



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

**Fakulta elektrotechnická
Katedra radioelektroniky**

Komprese stereoskopického videa

Compression of stereoscopic video

Bakalářská práce

Studijní program: Komunikace, multimédia a elektronika
Studijní obor: Multimediální technika

Vedoucí práce: Ing. Fliegel Karel Ph.D.

Jan Toman

Praha 2015

České vysoké učení technické v Praze
Fakulta elektrotechnická

katedra radioelektroniky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Jan Toman**

Studijní program: Komunikace, multimédia a elektronika
Obor: Multimediální technika

Název tématu: **Kompresa stereoskopického videa**

Pokyny pro vypracování:

Podejte přehled současných metod pro kompresi stereoskopického videa. Zaměřte se zejména na analýzu dopadu ztrátových kompresních metod na reprodukováný obraz, např. při použití nesymetrického kódování. Navrhněte experimentální řešení, které umožní posoudit vliv ztrátové komprese na kvalitu stereoskopického videa.

Seznam odborné literatury:

- [1] Ozaktas, H. M., Onural, L.: Three-Dimensional Television: Capture, Transmission, Display, Springer, 2008.
- [2] Javidi, B., Okano, F.: Three-Dimensional Television, Video, and Display Technologies, Springer, 2002.

Vedoucí: Ing. Karel Fliegel, Ph.D.

Platnost zadání: do konce letního semestru 2015/2016

L.S.

doc. Mgr. Petr Páta, Ph.D.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 10. 2. 2015

Abstrakt

Tato bakalářská práce se zabývá popisem metod zpracování a komprese stereoskopického videa, zejména se zaměřením na nesymetrické kódování. Dále bylo navrženo experimentální řešení umožňující posoudit vliv ztrátové komprese na kvalitu stereoskopického videa. V řešení byla využita metoda nesymetrického kódování za pomoci Gaussova rozostření, které bylo aplikováno s postupnou intenzitou rozostření od středu k okrajům na liché pruhy v levém obraze a sudé pruhy v pravém obraze. Účinnost metody byla ověřena prostřednictvím subjektivního hodnocení kvality obrazu a za pomoci dvou objektivních metrik PSNR a SSIM. Výsledky prokázaly, že námi použitá metoda nesymetrického kódování poskytuje efektivní způsob redukce datového toku stereoskopického videa při použití zpětně kompatibilní prostorové komprese ve formátu Side-by-side.

Klíčová slova

komprese stereoskopického videa, nesymetrické kódování, 3D, subjektivní a objektivní hodnocení kvality videa

Abstract

This bachelor thesis deals with the description of processing and compression methods of stereoscopic video and is particularly focused on asymmetric coding. Experimental solution allowing to assess the impact of lossy compression on the quality of the stereoscopic video was also proposed. In the solution we used asymmetric coding with the help of Gaussian blur. The Gaussian blur was applied gradually by raising the blurriness from the center to the sides in the right image to odd slices and in the left image to even slices. The efficiency was verified subjectively based on the quality of the picture and objectively by two metrics PSNR and SSIM. The results showed that method of asymmetrical coding, which we have used, provides an effective way to reduce bitrate of the stereoscopic video when using reverse compatible compression in Side-by-side format.

Key words

compression of stereoscopic video, asymmetric coding, 3D, subjective and objective video quality assessment

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne

.....

Jan Toman

Poděkování

Na tomto místě bych rád poděkoval vedoucímu práce Ing. Karlu Fliegelovi, PhD. za jeho čas, cenné rady a především trpělivost, kterou mi věnoval při psaní této práce. Dále bych chtěl poděkovat všem dobrovolníkům, kteří se zúčastnili subjektivního testování a Anetě Miklasové za pomoc s korekturou textu.

Obsah

| | | |
|-------|--|----|
| 1 | Úvod..... | 8 |
| 2 | Základy stereoskopie..... | 9 |
| 2.1 | Vnímání..... | 9 |
| 2.2 | Záznam..... | 11 |
| 2.3 | Zobrazování | 11 |
| 3 | Metody komprese obrazu..... | 14 |
| 3.1 | Bezeztrátová komprese | 14 |
| 3.2 | Ztrátová komprese..... | 15 |
| 4 | Metody kódování 3D videa..... | 16 |
| 4.1 | Prostorová komprese..... | 16 |
| 4.2 | 2D + Delta..... | 17 |
| 4.3 | 2D + Depth..... | 18 |
| 5 | Nesymetrické kódování..... | 19 |
| 5.1 | Teoretické základy | 19 |
| 5.2 | Praktická realizace | 21 |
| 6 | Návrh metody nesymetrického kódování..... | 23 |
| 6.1 | Gaussovo rozostření | 23 |
| 6.2 | Vývoj zvolené metodiky | 24 |
| 6.2.1 | Metoda rozostření celého obrazu | 24 |
| 6.2.2 | Metoda pruhů | 24 |
| 6.2.3 | Metoda pruhů s postupným rozostřením..... | 25 |
| 7 | Zpracování videosekvencí..... | 27 |
| 7.1 | Popis videosekvencí..... | 27 |
| 7.2 | Použité programy pro přípravu a zpracování videosekvencí | 28 |
| 7.3 | Příprava videosekvencí | 29 |
| 7.3.1 | Natočené záběry..... | 29 |
| 7.3.2 | Převzaté záběry | 30 |
| 8 | Subjektivní hodnocení..... | 31 |
| 8.1 | Realizace hodnocení se skupinou pozorovatelů..... | 31 |
| 8.1.1 | Popis testování | 31 |
| 8.1.2 | Metoda DSIS..... | 32 |

| | | |
|-------|---|----|
| 8.1.3 | Obsah testu..... | 32 |
| 8.1.4 | Otestování subjektů..... | 33 |
| 8.2 | Zpracování výsledků..... | 35 |
| 8.3 | Výsledky subjektivního testování..... | 35 |
| 9 | Objektivní hodnocení..... | 39 |
| 9.1 | Objektivní metriky..... | 39 |
| 9.2 | Použité programy pro výpočet objektivních metrik..... | 41 |
| 9.3 | Výsledky objektivního testování..... | 42 |
| 9.4 | Porovnání s výsledky subjektivního testování..... | 49 |
| 9.4.1 | Výpočet korelačních koeficientů..... | 49 |
| 9.4.2 | Grafické porovnání..... | 51 |
| 10 | Zhodnocení výsledků..... | 56 |
| 11 | Závěr..... | 59 |
| 12 | Použité zdroje a literatura..... | 60 |
| 13 | Seznam použitých obrázků..... | 63 |
| 14 | Přílohy..... | 65 |
| 14.1 | Příloha A - formulář pro subjektivní testování..... | 65 |
| 14.2 | Příloha B - obsah přiloženého DVD..... | 66 |

1 Úvod

Během několika posledních let jsme se stali svědky velikého nárůstu zájmu o 3D technologie. Z důvodu přidání třetího rozměru neboli hloubky obrazu se 3D stává populárním nejen v zábavním průmyslu, ale i mnoha jiných odvětvích jako je například medicína nebo strojírenství.

Jednou z nevýhod stereoskopického obsahu je nutnost dopravit k divákovi dvakrát více snímků než u klasického 2D obrazu, a tedy i jeho podstatně vyšší datový tok, který se například vylučuje s použitím současné infrastruktury televizního vysílání. To je jeden z mnoha důvodů, proč je komprese stereoskopického videa v současnosti tolik důležitá, a proč se jejím vývojem zabývá celá řada odborníků.

Naopak výhodou u stereoskopického obsahu může být to, že kromě klasických principů komprese obrazu se dají uplatnit i méně standardní postupy pomocí nesymetrického kódování, které je založeno na předzpracování jednoho z obrazů stereopáru a nedokonalosti lidského zraku.

Teoretická část se zabývá stručným popisem základů stereoskopie a současných metod pro kompresi stereoskopického videa, respektive jednotlivých způsobů kódování 3D obsahu. Dále práce obsahuje podrobnější popis metod nesymetrického kódování a jejich praktické realizace, které jsou stěžejní pro následující část práce.

Druhá, praktická část práce se zabývá návrhem, postupným vývojem a nakonec realizací konkrétní metody nesymetrického kódování. Dále jsou v práci popsány dva naprosto odlišné přístupy pro otestování a ověření účinnosti použité metody, konkrétně pomocí subjektivního ohodnocení kvality obrazu a prostřednictvím dvou objektivních metrik zvaných PSNR a SSIM.

2 Základy stereoskopie

V této kapitole budou popsány základy týkající se stereoskopie, jež jsou nutné pro pochopení celé problematiky. Nejdříve si vysvětlíme základní principy a pojmy, které jsou důležité pro prostorové vnímání a dále metody, jež nám napomáhají ke stanovení hloubky scény. Dále si stručně popíšeme základy záznamu a nakonec současné metody zobrazení stereoskopického obsahu.

2.1 Vnímání

Každý člověk má již od narození vybudovaný 3D pozorovací systém [15]. Ten je založený na binokulárním vidění, tzn. pomocí dvou očí, které jsou od sebe cca 6-7 cm vzdálené. V praxi to znamená, že každé oko vidí jednotlivé, nepatrně odlišné, obrazy (s horizontálním posunem), které se v mozku vyhodnotí jako jedna trojrozměrná scéna s danou informací o vzdálenosti jednotlivých objektů (hloubkou obrazu). Pro vnímání hloubky obrazu napomáhají také monokulární vodítka, která jsou přítomna i v klasické 2D fotografii a tak svým způsobem můžeme hloubku obrazu vnímat jedním okem. Jedná se spíše o mechanismy na základě zkušenosti, a patří mezi ně například osvětlení, perspektiva, hloubka ostrosti, pohybová paralaxa, překryv či relativní velikost známých objektů. Mnohem důležitější pro vnímání hloubky obrazu jsou však binokulární vodítka, která vycházejí z fyziologie zraku, a to především binokulární disparita, akomodace a vergence očí.

Akomodace a vergence

Akomodace je mechanický proces uvnitř našeho oka při němž se mění jeho optická mohutnost [27]. Hlavní funkcí akomodující čočky je upravit ohniskovou vzdálenost oka tak, aby bylo umožněno zaostření objektů, na které se soustředíme. Na krátkou vzdálenost jsou ciliární svaly ochablé a čočka tlustší, ovšem na větší vzdálenost se naopak čočka napíná do stran a rozšiřuje, tím dochází k lomu paprsků sledovaného objektu na jediné místo sítnice (žlutá skvrna).

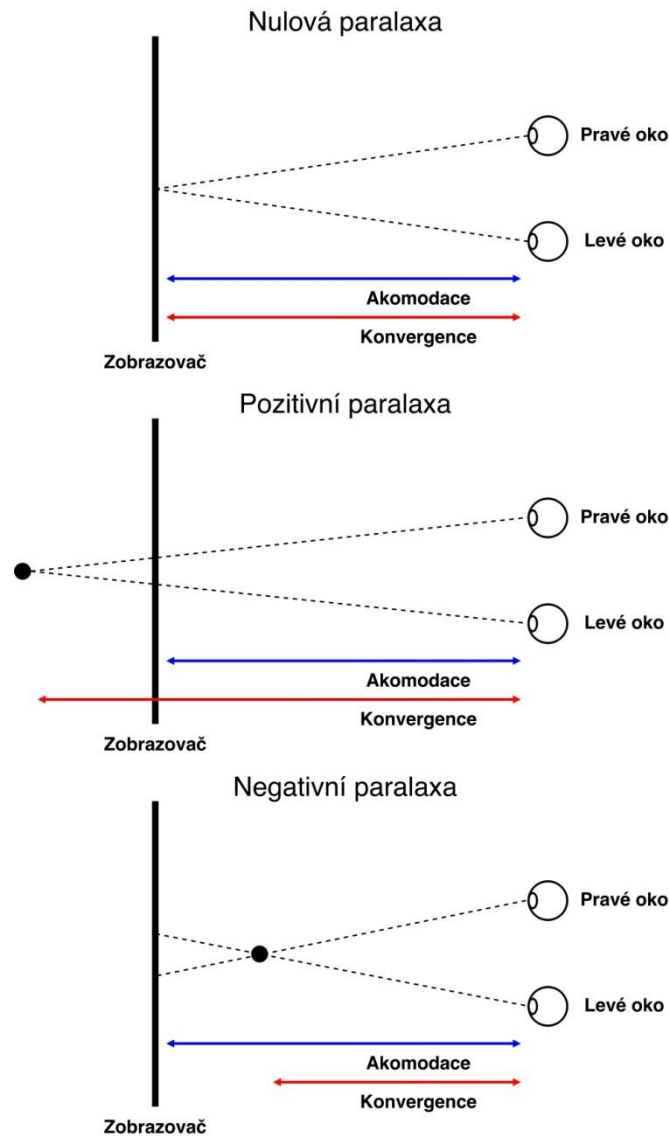
Vergence je obecně rotace očí, jež má za cíl, lepší vnímání bližšího a vzdálenějšího objektu [27]. Podle směru rotace ji dělíme na konvergenci a divergenci. Konvergence je rotace obou očí směrem dovnitř a je zajištěna prací okoohybných svalů. Jde o sbíhavost obou os očí při zaměření se na blízký předmět za cílem dopadu paprsků sledovaného předmětu do žluté skvrny. Konvergence se vždy účastní obě oči a je s akomodací úzce spjata. Opačným procesem je divergence, kdy při sledování velmi vzdálených předmětů dochází až k rovnoběžnosti os očí. Oba procesy jsou velmi namáhavé pro oční svalstvo a se podílí na celkové únavě očí.

Binokulární disparita

Zaměříme-li se na nějaký objekt v prostoru, promítne se nám jeho obraz na místo s nejhustším výskytem tyčinek a čípků, které se nazývá žlutá skvrna a dochází v něm k nejostřejšímu vidění (viz akomodace) [15]. Všechny ostatní objekty, které jsou v zorném poli blíže či dále než sledovaný objekt se nám promítnou v oku na jiné body sítnice. Rozdíl na sítnici mezi sledovaným objektem a ostatními objekty nazýváme disparita. Pokud je překročena určitá velikost disparity, tak nedojde v mozku ke spojení obrazu na trojrozměrný vjem, ale nastane dvojité vidění. Binokulární disparita se považuje za základní jev pro prostorové vidění, jež má dominantní hloubkový vjem a proto je důležité s ní umět v zobrazování stereoskopického obsahu pracovat.

Paralaxa

Posledním důležitým pojmem, který úzce souvisí právě s akomodací a konvergencí je paralaxa [42]. Jedná se o úhel či vzdálenost mezi optickými osami očí na projekční rovině. Rozdělujeme ji zejména na tři typy (viz obr. 1), a to podle pozice průtnutí optickým os očí vzhledem k projekční rovině. Pokud se osy protínají za projekční plochou, oči konvergují na větší vzdálenost než akomodují a vzniká dojem, že se předmět nachází v hloubce zobrazovacího zařízení. Naopak pokud se osy protínají před projekční rovinou, máme pocit, že předmět vystupuje z obrazovky.



Obr. 1. Typy paralax

Problémem sledování stereoskopického obsahu je tedy nepřirozený proces očí konvergovat na jiné místo než akomodují, z toho důvodu ovšem způsobuje zvýšenou únavu očí a může vést až k nevolnosti. Je tedy nutné na to brát zřetel a přizpůsobit tomu i jeho tvorbu. Například je dobré se vyhnout rychlým změnám v hloubce obrazu (při střihu) a extrémům, které se vylučují s přirozeným vnímáním člověka.

2.2 Záznam

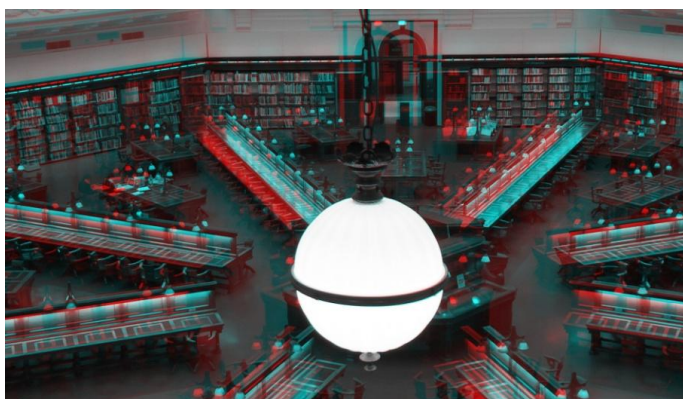
Chceme-li simulovat binokulární vjem, je nutné zachytit video či fotografii pomocí dvou zařízení, která snímají scénu s odlišným horizontálním posunem [22]. Výsledná hloubka obrazu je potom dána právě vzájemným horizontálním posuvem mezi zařízeními, tento rozdíl nazýváme stereobáze. Existují dvě metody natáčení, Toe-in a Off-axis, které se liší pouze tím, že u první metody dochází ke konvergenci kamer a jejich protínání optických os, zatímco u Off-axis metody jsou vždy osy kamer rovnoběžné. Zařízení, jež umožňuje precizní nastavení stereobáze a dalších parametrů nutných pro natáčení stereoskopického videa nazýváme stereoskopický rig. Nejjednodušším typem rigu je „Side-by-side“, kde jsou položeny obě kamery vedle sebe na jedné kolejnici. V profesionální sféře je pak spíše využíván zrcadlový rig, který je založený na kolmé montáži kamer vůči sobě a snímání scény skrz polopropustné zrcadlo. Oproti klasickému 2D filmu se musí při natáčení stereoskopického videa dodržovat několik důležitých pravidel pro zachování vjemu hloubky obrazu. Například pravidlo 1/30, které říká, že vzdálenost nejbližšího objektu od kamer musí být alespoň 30x větší než jejich stereobáze.¹

2.3 Zobrazování

Chceme-li vyvolat u diváka příslušný 3D vjem, musíme zaznamenané video pravou a levou kamerou zobrazit do příslušných očí diváka (levý obraz do levého oka a naopak). K dosažení takového efektu nám slouží hned několik metod, které si popíšeme níže [34].

Anaglyf

Jde o nejstarší a nejjednodušší technologii pro zobrazování 3D. Celý princip je založen na rozdílném obarvení dvou nepatrně odlišných obrazů jedné scény a jejich vložení do jednoho snímku. Za použití brýlí s barevnými filtry, které barevně odpovídají obarvení vašich snímků, se vám do oka propustí vždy jen příslušná složka obrazu a vzniká 3D vjem. V této technologii se používá různých kombinací obarvení a barev filtrů, ale mezi nejznámější patří červeno-azurový (viz obr. 2), červeno-modrý či červeno-zelený. Jednou z výhod této technologie je možnost realizace na papír a v podstatě jakýkoliv povrch. Nevýhodou zase může být znatelná deformace barev, která je pro mozek velmi náročná na zpracování.



Obr. 2. Ukázka snímku anaglyf

¹ Základní pravidla natáčení 3D : <http://www.dashwood3d.com/blog/beginners-guide-to-shooting-stereoscopic-3d/>

Pasivní technologie

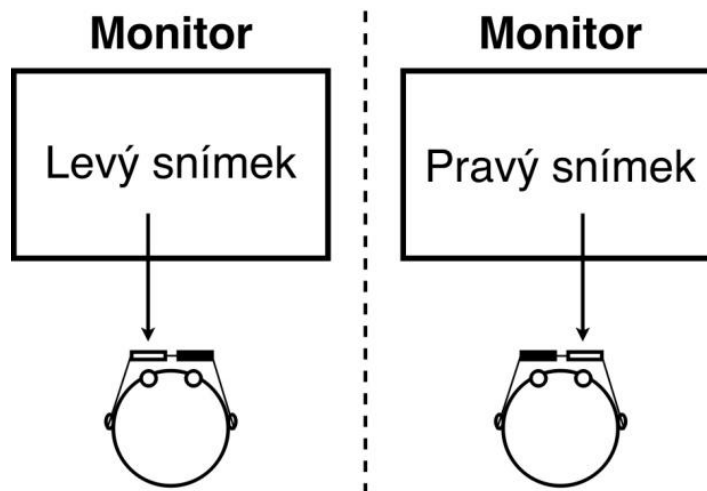
V současné době můžeme říct, že nejrozšířenější způsob zobrazování 3D je pasivní technologie. Ta je založena na principu lineární (vertikální a horizontální) či kruhové (levotočivá a pravotočivá) polarizaci světla a speciálních brýlí s polarizačními filtry. Obrazy jsou nejprve patřičně polarizovány a poté se vám před vstupem do oka propustí či odfiltrují pomocí speciálních brýlí s polarizačními filtry, které jsou nastavené pro každé oko jinak podle typu polarizace.

V kině je tento princip realizován pomocí dvou projektorů, které mají vždy před objektivem rozdílně polarizované filtry nebo jednoho projektoru, který disponuje měnícím se polarizačním filtrem. Tyto polarizované obrázky se odrazí od promítací plochy se speciální stříbrnou vrstvou, která zachovává polarizaci a putují k divákovi. Než vstoupí obrázky do oka, tak se jednoduše vyfiltrují pomocí brýlí jak již bylo řečeno a vznikne nám 3D vjem. Problém může nastat, pokud při sledování filmu nedržíte hlavu zcela vzpřímeně z důvodu použití dvou polarizačních rovin závislých na vzájemné poloze vyslaného obrázku z projektoru a filtru na brýlích. To může to mít za následek ztrátu 3D vjemu a velkého množství jasu. Z tohoto důvodu výrobci pasivních technologií přidávají navrch kruhovou polarizaci, na kterou nemá náklon hlavy téměř žádný vliv.

U monitorů je tato technologie o něco komplikovanější, protože máme jenom jeden zdroj obrazu. Využívá se tedy toho, že liché sloupce (resp. řádky) jsou polarizovány v jedné rovině a sudé sloupce (resp. řádky) jsou polarizovány v rovině druhé. Styl polarizace (řádky či sloupce) si určuje samotný výrobce. Takovéto polarizování má ale i svá technologická úskalí, protože polarizační filtry musí mít velmi jemnou strukturu, aby umožňovaly zakrýt přesně jeden sloupec (resp. řádek) pixelů. To znamená poloviční horizontální nebo vertikální rozlišení podle technologie výroby.

Aktivní technologie

Aktivní technologie se tak nazývá, protože k ní potřebujeme speciální brýle s baterií a elektronikou, které aktivně pracují, neboli dokážou střídavě zatmívat sklíčka brýlí, a tím do oka propouštět vždy jen potřebný snímek pro vytvoření 3D vjemu.

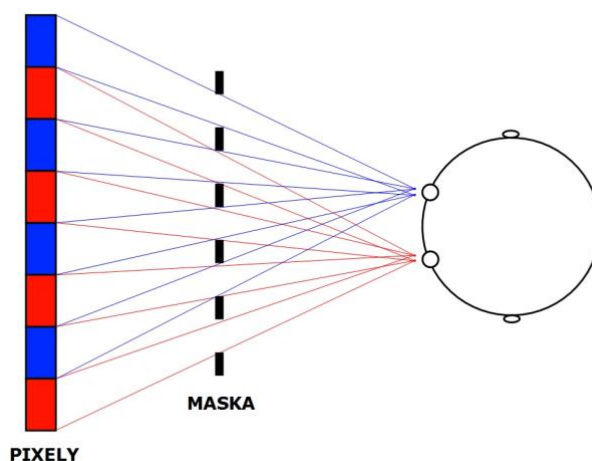


Obr. 3. Princip aktivní technologie

Princip je v kině i pro monitory (viz obr. 3) naprosto stejný a postačí nám vždy jeden zdroj, který pracuje na dvojnásobné frekvenci a vysílá střídavě obrázky pro levé a pravé oko. Snímky se odrazí od promítací plochy či zobrazí na monitoru a „putují“ k divákovi. Zde jsou jednoduše propuštěny jen levé snímky do levého oka a naopak pravé snímky do pravého oka pomocí zatemňování brýlí. Brýle a zdroj spolu musí být vzájemně synchronizovány. Nevýhodou tedy může být nutná synchronizace brýlí se zdrojem, složitější údržba a hlavně nepříjemné blikání, které je znatelné v nedostatečně zatemněném prostředí.

Auto-stereoskopie

Tato technologie se vyznačuje hlavně tím, že k ní není potřeba žádných speciálních brýlí. Základem všech auto-stereoskopických monitorů je speciální maska, která je umístěná před LCD displejem. Tato maska je buď vybavena lentikulárními čočkami, které vychylují různé sloupce pixelů do různých směrů nebo jednoduše utvořena s přesnými mezerami pro rozdělení obrazu pomocí paralaxní bariéry (viz obr. 4). Obrazy pro pravé i levé oko jsou tedy vtěsnány na obrazovku monitoru podle parametrů optické masky. Maska jednoduše způsobí to, že jsou obrazy pro jednotlivé oči viditelné pouze z určitých směrů. Pokud se tedy divák postaví do vhodné pozice, která může být pro každého trochu jiná podle jeho rozteče očí, uvidí levým okem pouze levý obraz a pravým okem obraz pravý.



Obr. 4. Princip auto-stereoskopického displeje s paralaxní bariérou

Z toho plyne nevýhoda, že rozlišení každého obrazu je pouze poloviční, oproti nativně horizontálnímu rozlišení monitoru. U monitoru pro více diváků navíc dochází ke snížení horizontálního rozlišení za každý sledující pár očí. Další velkou nevýhodou je pozorovací úhel, který je ve většině případů velmi nízký, takže máte předem stanovená místa pro 3D vjem, a poté existují místa slepá, kde dochází k tvoření „duchů“, tj. když se vám obraz pro pravé i levé oko slévá do jednoho. Tento problém se sice dá vyřešit za pomoci kamerového systému se sledováním očí, kdy se optická maska natáčí podle toho, kde sedíte, ale pokud máte před monitorem více diváků nebo nedostatečné světlo v místnosti, tak systém může jednoduše selhat.

3 Metody komprese obrazu

V současné době digitálních technologií hraje datový tok velmi důležitou roli, k jeho snížení a celkové redukci velikosti souborů nám slouží vhodné komprese. U kompresi si všímáme především kompresního poměru, který udává poměr dat na vstupu k datům výstupním. Komprese rozdělujeme na bezztrátové a ztrátové, přičemž ztrátové dosahují značně vyšších kompresních poměrů.

3.1 Bezeztrátová komprese

U bezztrátových kompresí dochází pouze k odstranění redundantních neboli nadbytečných dat. Nedochází tedy k žádné ztrátě informací a to co zakódujeme, tak to stejné i dekodujeme. Jedná se v podstatě o algoritmy, které jen vhodněji uspořádají a „zabalí“ data. Mezi ně patří například Huffmanovo kódování, které nejčastěji vyskytovaným symbolům přiřadí nejkratší kódové slovo, čímž se docílí komprese dat. Existuje mnoho různých formátů, které využívají bezztrátové komprese, ale mezi nejznámější například patří ZIP, RAR či námi v praktické části použité komprese statického snímku TIFF a RAW nebo AVI pro bezztrátovou kompresi videa.

TIFF

TIFF (Tag Image File Format), je bezztrátový formát používaný pro ukládání rastrové počítačové grafiky [13]. Původně byl vytvořen za účelem jednotného formátu pro stolní skenery. TIFF je velmi flexibilní a adaptivní formát, který nese v hlavičce „tagy“ o tom, jaká je jeho základní stavba obrazu (velikost, definice, uspořádání dat a aplikovaná komprese). Kvůli flexibilitě (mnoha typů) se v jednu dobu jeho zkratka překládala jako „Thousand of Incompatible File Formats“, a proto byl vytvořen jeho základní formát neboli Baseline TIFF (TIFF 6.0), který měl být jediným standardem. Tento standard umožňuje výstupní formát bez komprese, s modifikovaným Huffmanovým kódováním či PackBits kompresí (forma RLE), navíc má ještě rozšíření, které přidává LZW kompresi a mnoho dalších funkcí.

RAW

Tento bezkompresní formát, z překladu doslova jako „surový“, se přirovnává k digitálnímu negativu [11]. Důvodem je, že RAW soubory nejsou samotné fotografie, ale jsou to pouze minimálně zpracovaná data poskytnutá rovnou ze snímače. To samozřejmě nese svoje výhody, ale i úskalí. Mezi hlavní výhody tedy patří nejvyšší kvalita snímku po všech stránkách a nedestruktivní editace, kdy dochází ke změně instrukcí, které říkají, jak by měl vypadat finální výstup do jiného formátu (JPEG, TIFF, atd.). Ovšem značnou nevýhodou může být velikost souboru, neschopnost zobrazení snímku v jakémkoliv prohlížeči či nestandardizovanost formátu napříč výrobci.

AVI

Formát AVI (Audio Video Interleave), představený firmou Microsoft Windows v roce 1992, je multimediální kontejner, který může obsahovat jednu videostopu a několik zvukových stop naráz [12]. Používá RIFF (Resource Interchange File Format) strukturu ukládání a uspořádání dat, což znamená, že ho nelze přehrát, pokud není úplný. Další nevýhodou může být obrovský objem dat při jeho bezztrátovém ukládání, avšak vzhledem k jeho vysoké kompatibilitě s existujícími programy patří stále mezi jedny z nejpoužívanějších formátů.

3.2 Ztrátová komprese

Při ztrátové kompresi dochází nejen k odstranění redundantních dat, ale také irelevantních. To má za následek, že dekódovaná data po použití ztrátové komprese již nejsou stejná jako data zdrojová.

U audiovizuálního obsahu se převážně využívá psychovizuálního a psychoakustického modelu, zkrátka nedokonalosti lidských smyslů. Například pokud víme, že lidský sluch je schopný vnímat frekvence cca od 20 Hz do 20 kHz, je možné zbytek frekvencí odfiltrovat a ušetřit určitou část objemu dat. Na druhou stranu, obrazové formáty se ztrátovou kompresí dosahují redukce objemu dat, například odstraněním vyšších detailů v obraze a využitím vyšší citlivosti člověka na jas než na barvu. K dnes nejpopulárnějším ztrátovým kompresím patří především komprese statického snímku JPEG či hudebního obsahu Mp3. My se velmi stručně zaměříme nejprve na kompresní standard MPEG-2, který je předchůdcem a vede k pochopení druhého důležitého standardu H.264/MPEG-4 AVC, jenž byl použit pro kompresi videosekvencí v praktické části práce.

MPEG-2

MPEG-2 (Motion Pictures Experts Group) je standardem pro kódování pohyblivého obrazu [10]. Využívá ztrátové komprese videa založené na kompresi statických snímků JPEG a predikce pohybu. Neposílá tedy jenom celé snímky, ale pokud to jde, tak jenom informaci o změně v obraze. MPEG-2 využívá tři typů snímku: I (Intraframe) plné snímky, P (Predicted) závislé na předchozích I snímcích a snímky B (Bidirectionally predicted), které jsou závislé na předchozích i následujících snímcích I či P. Sekvence takových snímků se nazývá GoP (Group of Pictures). Struktura GoP je popsána dvěma čísly, první popisuje vzdálenost mezi nejbližšími I či P snímky a druhé udává vzdálenost mezi plnými snímky I.

H.264/MPEG-4 AVC

H.264 neboli MPEG-4 AVC (Advanced Video Coding), je jeden z momentálně nepoužívanějších kompresních standardů pro nahrávání, kompresi a distribuci video obsahu [37]. Je nástupcem staršího standardu MPEG-2 a vlastní celou řadu jeho principů. H.264 oproti svému předchůdci dosahuje polovičního bitového toku při zachování stejné subjektivní kvality, a dalšími vylepšeními, jako je například proměnná velikost bloků pro pohybovou kompenzaci, vnitrosnímková komprese, použití celočíselné transformace (obdoba diskrétní kosinové transformace) nebo vylepšené entropické kódování realizované pomocí Exponenciálního Golomova kódu, CAVLC (kódování s proměnnou délkou slova) a CABAC (adaptivní binární aritmetické kódování).

My se zaměříme pouze na část, ve které dochází ke ztrátě kvality, tedy již po samotném převodu dat z časové do spektrální oblasti pomocí celočíselné transformace, jež má na výstupu transformované koeficienty. Ty jsou dále po blocích kvantovány, tzn. že každý koeficient se vydělí celočíselnou hodnotou podle kvantizačního parametru (QP). Typickým výsledkem kvantizace po zaokrouhlení je blok s několika nenulovými hodnotami umístěnými v jeho levém horním rohu a zbylými nulami směrem dolů doprava. Právě nastavením parametru QP dochází k výslednému určení komprese a kvality obrazu. Vyšší hodnota znamená více nul po dělení a účinnější kompresi na úkor kvality obrazu.

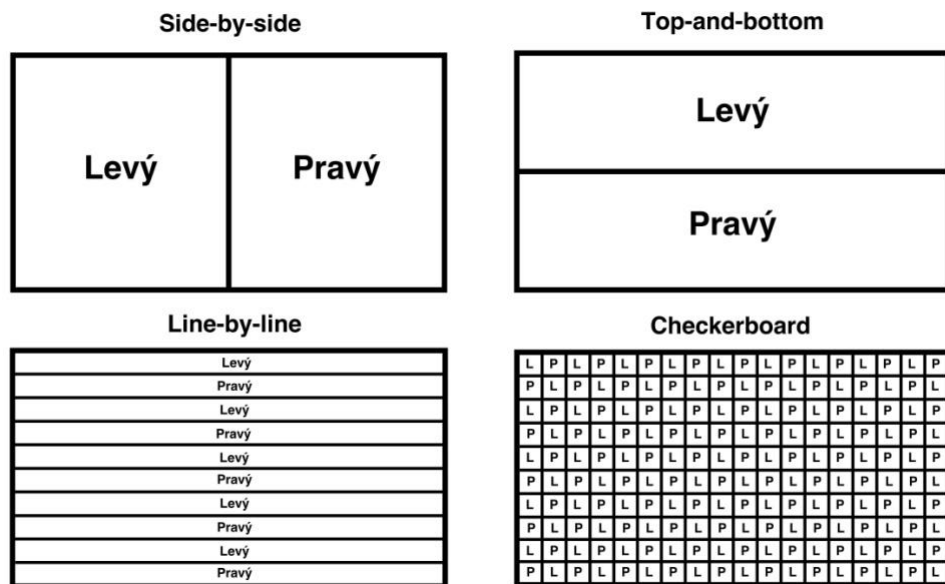
4 Metody kódování 3D videa

V současnosti můžeme 3D metody kódování, které mají za účel dopravit stereoskopický obsah do našich domovů, rozdělit na dva typy. Prvním je takzvaný zpětně kompatibilní formát pomocí prostorové komprese, který má výhodu v možnosti použití současné infrastruktury televizního vysílání (komprese, modulace, přenos a příjem) a neklade na její provozovatele ani uživatele žádné další náklady krom nákupu 3D zobrazovače. Druhým typem jsou o něco složitější metody využívající jednoho referenčního snímku a pomocných dat pro tvorbu stereoskopického vjemu, tzn. 2D + metadata, nevýhodou tohoto druhu kódování je nutnost nových dekodérů na straně příjemce. Do prvního typu ještě pro úplnost zapadá možnost využití principu Anaglyf, kdy jsou oba obrazy odlišně obarveny a následně vtěsnány do jednoho snímku. Tato metoda je ovšem nepoužitelná vzhledem k barevné destrukci a špatné kvalitě obsahu.

4.1 Prostorová komprese

Jedná se o podvzorkování obrazu pro levé i pravé oko a jejich následné prostorové přeskládání do jednoho obrazu [31]. Používají se normou definované dva typy uspořádání snímků, Side-by-side (SBS) a Top-and-Bottom (TAB), nazývané též jako „Frame Compatible“, jejichž jediným rozdílem je odlišné umístění snímků pro levý a pravý obraz. Nevýhodou těchto formátů je poloviční rozlišení po reprodukci na zobrazovači, kdy musí dojít k interpolaci chybějících pixelů. Další nevýhodou může být neschopnost příjmu 2D obrazu na obyčejné televizi, protože obraz se vám bude stále zobrazovat jen v daném formátu jako podvzorkovaná dvojice obrazů.

Existuje ještě celá řada realizací prostorové komprese (viz obr. 5), které většina stereoskopických zobrazovačů podporuje, ale vzhledem k jejich náročnějšímu dosažení se prakticky nepoužívají. Mezi ně patří například „Checkerboard“, kde je levý obraz reprezentován na lichých řádcích lichými pixely, a naopak na sudých sudými, nebo jednodušší princip „Line-by-line“, kde jsou oba obrazy vtěsnány po řádcích.



Obr. 5. Formáty prostorové komprese

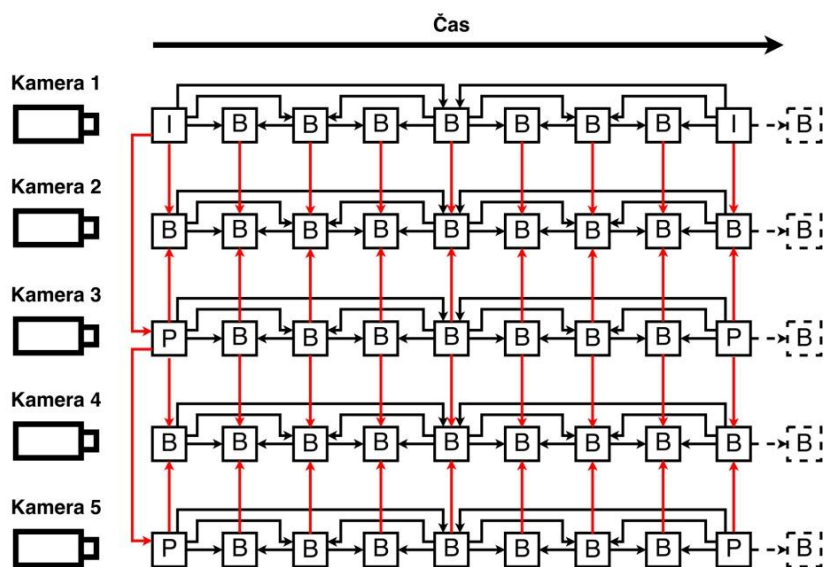
Pro úplnost uvádím i speciální TAB formát nazývaný „Frame Packing“ používaný Blu-ray disky, který má levý i pravý snímek poskládaný nad sebou v plném rozlišení se 45 px vysokou mezerou mezi snímky. Výsledné vertikální rozlišení je tak celých 2205 pixelů a vzhledem k vysokému datovému toku může být nevýhodou nutnost použití HDMI kabelu verze 1.4 a vyšší.

4.2 2D + Delta

Jde o metodu kódování, jež využívá vysoké korelace mezi obrazy pro levé a pravé oko. Jako základ využívá jeden pohled nazývaný *baseview* a k němu doplňující data ohledně ostatních pohledů. Nejedná se tedy už jen o využití mezi-snímkové komprese (viz. kapitola 3.2) jednotlivých obrazů, ale také o podobnosti mezi několika pohledy (mezi-pohledová komprese) z více kamer, které zachytávají téměř stejnou scénu z nepatrně odlišných úhlů. Praktickou realizací kompresního standardu využívajícího smyslu 2D + delta je Multiview Video Coding.

Multiview video coding

Multiview video coding (MVC) je doplněním kompresního standardu H.264 (MPEG-4 AVC) a jeho účelem je efektivní komprese vysokého množství (2 a více) diskrétních pohledů [37]. Jedním ze základních principů je právě mezi-pohledová komprese využívající minimálního rozdílu mezi jednotlivými pohledy zachycující scénu. MVC se v současnosti používá jen pro kódování 3D obsahu na Blu-ray discích, avšak díky zpětné kompatibilitě s H.264/MPEG-4 AVC umožňuje starším dekodérům a přehrávacím zařízením dekódovat video jako jeden pohled (2D) a nadbytečná metadata týkající se ostatních pohledů jednoduše „zahodit“, což by mohlo vést k budoucímu nasazení této metody kódování pro televizní vysílání. Dalším možným využitím by mohlo být ve FTV (Free Viewpoint Television), systému umožňující interaktivní přístup diváka k dynamickému sledování scény z mnoha úhlů. Nebo v technologii Tele-immersion, která slouží k zachycení a generování 3D virtuální scény v reálném čase a může být sdílena na dálku více účastníky najednou.



Obr. 6. Mezi-snímková/mezi-pohledová predikce struktury MVC

Obrázek 6 znázorňuje strukturu MVC, ve které je kromě klasické mezi-snímkové komprese (černé šipky) použita i komprese mezi-pohledová (červené šipky).

4.3 2D + Depth

2D + depth, někdy též zvaná jako 2D + Z, je metoda stereoskopického kódování podporovaná celou řadou autostereoskopických displejů, jejímž základem je jeden 2D pohled a speciální šedotónová hloubková mapa [30]. Hloubková mapa jednoduše zobrazuje, jakou hloubku obrazu budou mít jednotlivé pixely základního 2D obrazu. Pokud se objekty budou divákovi jevit „před obrazovkou“ (negativní paralaxa), budou v hloubkové mapě znázorněny pomocí nejsvětějšího odstínu šedi (bílá), naopak objekty umístěné „za obrazovkou“ (pozitivní paralaxa) budou do černa (viz obr. 7). Z toho je jasné, že právě objekty nacházející se v úrovni zobrazovače budou šedé. Toto obarvení není přesně definováno, a tak se můžeme setkat s opačným případem kdy objekty blíže k nám jsou zbarveny černě. Výhodou takového kódování je možnost konverze libovolného 2D obsahu na 3D, a možnost uživatelského nastavení hloubky obrazu, to vše bohužel za cenu výpočetně náročných procesů a složitých algoritmů. Na internetu existuje spousta stránek s volně dostupnými galeriemi, kde si můžete stáhnout 2D pohled a k němu příslušnou hloubkovou mapu. Podobně je to i se samotnými softwary na konverzi a tvorbu hloubkové mapy.²



Obr. 7. Ukázka 2D snímku (vlevo) a jeho hloubkové mapy (vpravo). Převzato z²

² Volně dostupnou galerii a software na výrobu hloubkové mapy naleznete na: <http://www.dofpro.com>

5 Nesymetrické kódování

Jednou z největších nevýhod stereoskopického videa je jeho objem dat, nutnost větší šířky pásma a datových úložišť jako takových. K redukci datového toku nám slouží patřičné komprese, které oproti klasickému 2D obrazu můžeme ještě doplnit o aplikaci nesymetrického kódování, které je založeno na předzpracování jednoho či obou kanálů ze stereopáru.

5.1 Teoretické základy

V této kapitole si popíšeme výzkum a poznatky ohledně nesymetrického kódování stereoskopického obrazu z minulých let. Zaměříme se na vnímanou kvalitu, hloubku a ostrost obrazu, jež jsou důležité parametry pro subjektivní kvalitu stereoskopického obsahu. Dalším velmi důležitým parametrem, který má značný dopad na subjektivní vjem je únava očí.

Vnímaná kvalita

Z teorie o oční supresi vychází, že binokulární vjem stereo páru převládá vždy u snímku s vyšší kvalitou [25]. Teoreticky to znamená, že pokud budeme mít jeden obrázek ve vysoké kvalitě a druhý zkomprimovaný, nemusíme vůbec zaznamenat nežádoucí artefakty způsobené kompresí méně kvalitního snímku.

Z toho také vychází koncept smíšeného rozlišení, který představil M. G. Perkins v roce 1992 [28], jenž předpokládá, že se binokulární vjem nezhorší, pokud je jeden z obrazů ve vysoké kvalitě a druhý v nižší. Tedy, že dominantní vjem převládá vždy u obrazu s vyšší kvalitou. Perkins dále zavedl aplikaci filtru typu dolní propusti (způsobuje rozostření), jako kompresní algoritmus na jeden z obrazů stereo páru. Autor dospěl k závěru, že metoda smíšených rozlišení je jednoduše implementovatelná a redukce datového toku je velmi znatelná.

L. Stelmach, Wa J. Tam a další, v roce 1998, ve své práci aplikovali rozdílné kompresní poměry na levý a pravý obraz stereoskopického videa pomocí kompresního standartu MPEG-2 (způsobuje kostičkování) a filtru dolní propusti (způsobuje rozostření) [33]. Výsledky ukázaly, že subjektivní kvalita nesymetricky kódovaného stereoskopického obrazu byla přibližným průměrem jednotlivých subjektivních kvalit samostatných obrazů při použitém MPEG-2. Při aplikaci rozostření odpovídala subjektivní kvalita vždy snímku s vyšší kvalitou obrazu.

D. Meegan a další v roce 2001 studovali kombinaci nesymetrického rozostření (filtrem typu dolní propust) a kostičkování (MPEG-2) na zhoršení kvality obrazu [28]. V případě aplikace nesymetrického rozostření byl dominantnější vjem u snímku s vyšší kvalitou. Při nesymetrickém použití komprese MPEG-2, byl subjektivní vjem stereo páru přibližným průměrem vjemů jednotlivých obrazů.

Výsledky z předchozích výzkumů jsou srovnatelné a je možné učinit závěr, že úspěch nesymetrického kódování závisí na typu kódovacích artefaktů.

Vnímaná hloubka

L. Stelmach a další, v roce 2000, zkoumali vliv nesymetrického prostorového a časového filtrování dolní propustí na vnímanou hloubku obrazu [32]. Výsledky ukazují, že prostorové ani časové

filtrování dolní propustí nemá na vnímanou hloubku žádný vliv. Vysvětlením může být, že filtrování dolní propustí odstraňuje vyšší frekvence v obraze (detaily), které nejsou pro vnímání hloubky příliš podstatné. Jejich studie tedy dokazuje velmi malou závislost mezi vnímáním kvality a hloubky obrazu.

Vnímaná ostrost

Problémem hodnocení ostrosti je, že může být ovlivněna několika parametry, jako je nezaostření kamery, kódováním nebo binokulární disparitou.

Roku 1997 A. Berthold zjistil, že stereo obraz s různě přidaným stupněm Gaussova rozostření byl vnímán ostřeji, oproti jednotlivým obrazům se stejným typem rozostření [28]. To naznačuje, že přidaná hloubka obrazu má dopad na vnímání hran, a tím i ostrost.

Wa J. Tam a další naopak při subjektivních testech v roce 1998 objevili, že ostrost byla subjekty hodnocena stejně, ně-li hůře u symetricky kódovaného stereoskopického videa pomocí MPEG-2 (způsobuje kostičkování), oproti stejně kódovaným jednotlivým obrazům [33]. Z obou studií vychází, že existuje vzájemná souvislost mezi vnímanou ostrotí a kvalitou obrazu podle typu kódovaných artefaktů.

V roce 2000, L. Stelmach a další vyšetřovali efekt smíšeného rozlišení na vnímanou ostrost a dospěli k závěru, že prostorové filtrování dolní propustí dává přijatelnou ostrost, která je vnímána vždy skrze ostřejší obraz ze stereo páru [32].

Vnímaná únava očí

Při prohlížení stereoskopického obsahu dochází k mnohem větší únavě očí nežli je tomu u klasického 2D, proto je také důležité počítat s jejím dopadem na subjektivní ohodnocení pozorovatelů.

V roce 1996, T. Mitsuhashi zjistil, že subjekty měly více unavené oči při sledování stereoskopického obsahu, nežli by tomu bylo při sledování klasického 2D obrazu [26]. K výsledkům došel pomocí objektivního měření, známém jako CCF (Critical Flicker Frequency) neboli měření kritického kmitočtu blikání. Kritický kmitočet blikání je největší frekvence, kdy pozorovatel stále vnímá blikání zdroje, pokud se tato frekvence překročí, pozorovatel již nadále vidí jen plynulé světlo. Z výsledků vyplývá, že během třiceti minut u subjektů došlo k významnému poklesu kritického kmitočtu při sledování stereoskopického obsahu. Ten je samozřejmě velmi úzce spjat s únavou očí.

V roce 2000, W. IJsselsteijn a další zkoumali účinek parametrů (stereobáze, ohnisková vzdálenost a konvergence kamer) při stereoskopickém natáčení a délce zobrazování stereoskopického obsahu na únavu očí testovaných osob [17]. Zprůměrované výsledky vykazují lineární závislost mezi únavou očí a zvyšující se disparitou.

5.2 Praktická realizace

Praktickou aplikací nesymetrického kódování je myšleno použití jakéhokoliv algoritmu či metody pro zpracování obrazu za účelem redukce datového objemu jednoho z kanálů stereopáru. Kompresi obrazu lze dosáhnout hned několika metodami, které by měly být ovšem prováděny tak, aby při výsledném zobrazování nedošlo k narušení vnímání subjektivní kvality a hloubky obrazu. V této kapitole si popíšeme čtyři základní způsoby předzpracování obrazu.

Smišené rozlišení

Metoda smíšeného rozlišení je jedna z prvních a nejjednodušších, jde v ní o podvzorkování jednoho z obrazů a tím i snížení jeho celkového rozlišení [2]. V důsledku přehrávání pak dochází ke zpětné interpolaci chybějících pixelů u pohledu s nižším rozlišení, což má za následek eliminaci detailů v obraze. Na obrázku 8 je znázorněné postupné snižování rozlišení originálního snímku (na polovinu, čtvrtinu a desetinu rozlišení) a následné přiblížení výsledků.



Obr. 8. Ukázka principu smíšeného rozlišení

Prostorová filtrace

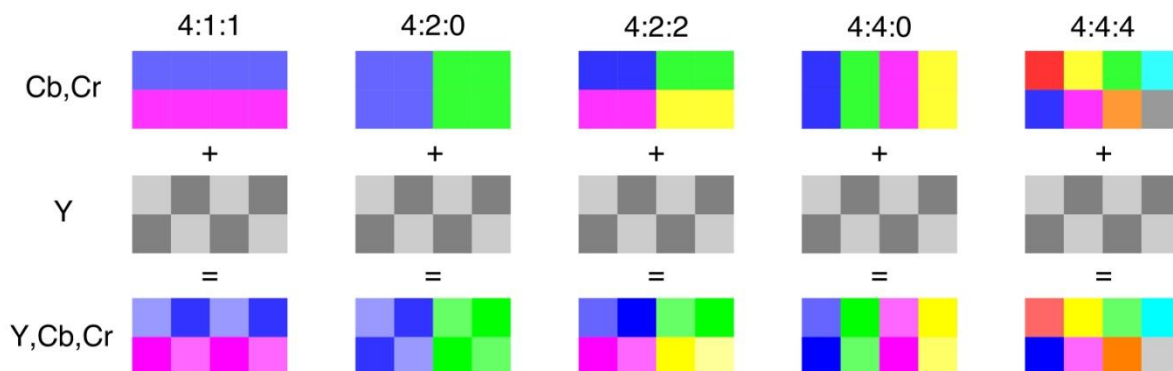
Prostorová filtrace je založena na rozostření jednoho z obrazů stereopáru, jedná se v podstatě o filtraci obrazu dolní propustí, jelikož dochází k odstraňování vysokých frekvencí (detaily) v obraze (viz obr. 9) a snížení celkového prostorového rozlišení [33]. K aplikaci rozostření nám slouží hned několik obrazových filtrů jako je například Mean (angl. průměr), Medián nebo Gaussův (viz kapitola 6.1), které se liší pouze v principu určení výsledné hodnoty rozostřeného pixelu. Zatímco u Mean filtru je výsledná hodnota vypočtena jako aritmetický průměr hodnot pixelů v určitém okolí, filtr Medián považuje za výslednou hodnotu právě medián z hodnot okolních pixelů.



Obr. 9. Ukázka rozostření pomocí Gaussova filtru s velikostí okna 15 a parametrem sigma 5

Podvzorkování barvosných složek

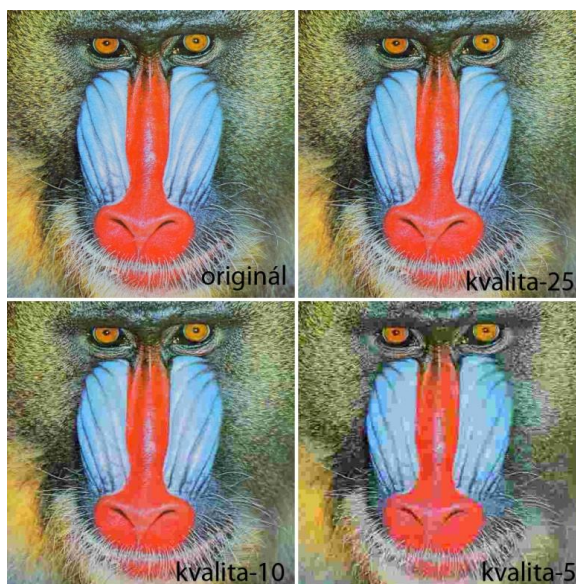
Jde o metodu, která využívá faktu, že jsme mnohem citlivější na jas nežli na barvu, a tak si můžeme dovolit informaci o barvě jednoho z obrazů uložit v nižším rozlišení [29]. Při podvzorkování dochází ke zprůměrování barevných hodnot jednotlivých pixelů podle předem daných poměrů. Obrázek 10 ilustruje podvzorkování sousedních pixelů chrominančních složek (Cb a Cr) podle různých poměrů a naopak ponechání původního rastru jasové složky (Y). Typické poměry jsou 4:2:2 (podvzorkování dvou bodů), 4:2:0 (čtyř bodů) či 4:4:4 (bez podvzorkování).



Obr. 10. Ukázka různých poměrů podvzorkování barvosných složek

Kvantizace transformovaných koeficientů

V této metodě dochází k rozdílnému kvantování transformovaných koeficientů levého a pravého obrazu [1]. K transformaci koeficientů využíváme různé algoritmy jako například DCT (Diskrétní Cosinova transformace), Vlnkovou transformaci či KLT (Karhunen-Loéveho transformace). Na tomto principu, za použití DCT, funguje například komprese statického snímku JPEG, kde s postupným zvyšováním kvantizace (zmenšování kvality) dochází k typickému nežádoucímu efektu „kostičkování“ (viz. obr 11).



Obr. 11. Ukázka změny kvality JPEG

6 Návrh metody nesymetrického kódování

Z teoretických základů lze usoudit, že nejpraktičtější metodou nesymetrického kódování je předzpracování pomocí jednoduchého rozostření. V námi navrhované metodě nesymetrického kódování použijeme právě Gaussova rozostření, které patří mezi nejčastěji používané metody rozostření obrazu. Tato metoda by měla být vzhledem k předchozím studiím velmi účinná a při správném použití by měla mít minimální vliv na subjektivní kvalitu obrazu.

6.1 Gaussovo rozostření

Gaussovo rozostření neboli Gaussian blur je metoda pro zpracování obrazu s efektem rozostření [16], která využívá Gaussovy funkce dané vztahem (1).

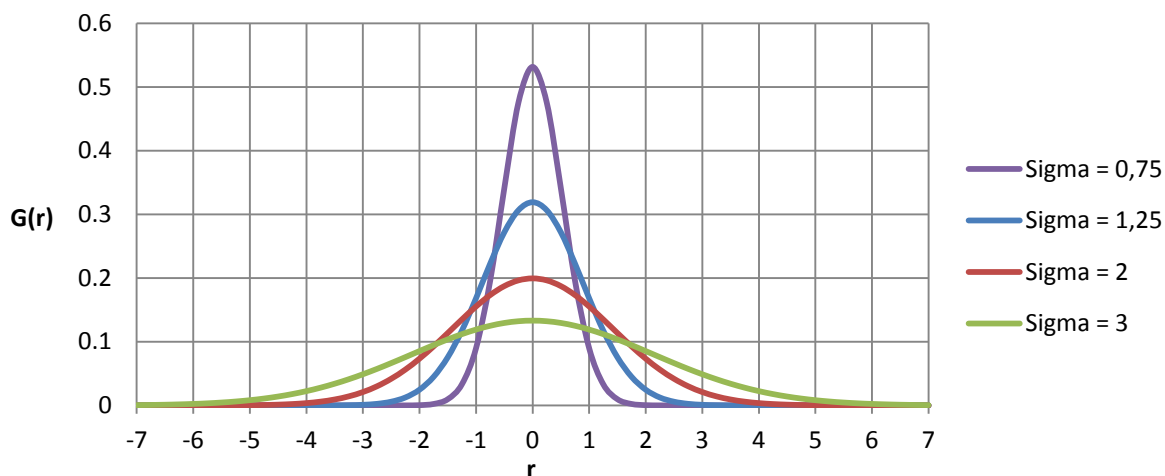
$$g(r) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (1)$$

Kde σ (sigma) je konstanta označovaná jako směrodatná odchylka a udává intenzitu rozostření (viz obr. 12). Pro aplikaci na fotografii je třeba funkci zobecnit pro dvojdimenzionální prostory (2).

$$g(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2)$$

Před samotnou aplikací rozostření je důležité definovat zaprvé parametr sigma a zadruhé velikost okna o rozměrech $m \times n$. Velikost okna udává oblast pixelů, která bude mít vliv na výslednou hodnotu rozostřeného pixelu. Proces se poté provádí pixel po pixelu přes celý obraz.

Problém ovšem nastává na okrajích fotografie, protože za okraji již nejsou definované pixely. Většina softwaru jednoduše dodefinuje chybějící pixely jako černé, a to ve výsledku způsobí ztmavení celé oblasti pixelů, která se nachází ve vzdálenosti menší než je poloviční velikost okna v daném směru. Tento problém se ovšem dá řešit vhodnějším dodefinováním pixelů, jako například periodickým prodloužením fotografie do všech stran (periodic extension) nebo podobně fungujícím periodickým prodloužením, ale tentokrát zrcadlově vzhledem k původnímu snímku (mirror extension).



Obr. 12. Ukázka Gaussovy křivky pro různé hodnoty parametru sigma

6.2 Vývoj zvolené metodiky

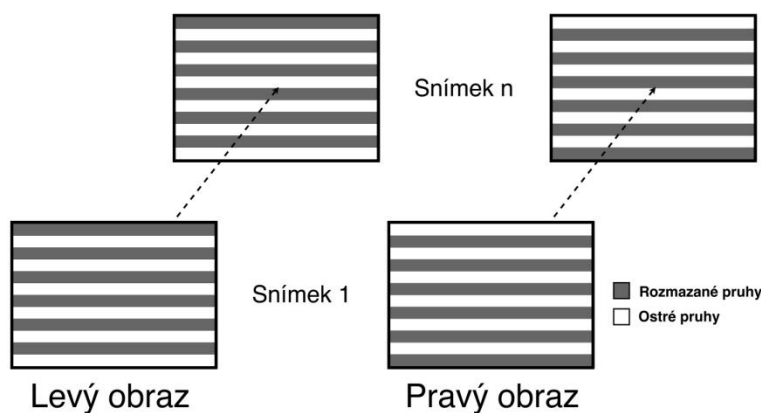
V této kapitole si podrobněji popíšeme možné způsoby aplikace Gaussova rozostření, jejich výhody či nevýhody a nakonec si rozebereme náš konkrétní přístup, který budeme používat v další části práce.

6.2.1 Metoda rozostření celého obrazu

Nejjednodušší realizací je rozostření jednoho celého obrazu ze stereopáru [26]. Tato metoda ovšem představuje velkou nevýhodu ohledně vrozené dominance očí člověka, která má za následek, že osoba, která má dominantní levé oko hodnotí subjektivní kvalitu stereoskopického obrazu hůře pro rozostřený levý snímek, nežli by to mu bylo naopak [5].

6.2.2 Metoda pruhů

Aby se předešlo diskriminaci velké části populace, je vhodnější metoda pruhů (viz obr. 13), během níž dochází u levého obrazu k rozostření lichých pruhů a u pravého sudých [36]. V levém i pravém obrazu je stejný počet pruhů daný celým sudým číslem, a tomu musí odpovídat i jejich výška. Při klasickém rozlišení FullHD, tedy vertikálním rozlišení 1080 pixelů, může být výška pruhu například 10, 30, 60, 108 nebo 270. My jsme zvolili pruh o výšce 108 pixelů, který se nám jevil subjektivně nejpříjemněji. Tato hodnota byla navíc doporučena v práci, která se věnovala velmi podobné problematice [3].



Obr. 13. Metoda pruhů

První otestování této metody prokázalo, že hrana mezi ostrou a rozostřenou částí obrazu je nadměrně rušivá a pro hodnotu sigma větší jedné se metoda prakticky nedá použít. Dalším postupem byla tedy snaha o odstranění těchto rušivých přechodů.



Obr. 14. Metody rozostření - celý obraz / pruhy / postupné pruhy

6.2.3 Metoda pruhů s postupným rozostřením

K odstranění viditelných hran byly realizovány a vyzkoušeny dohromady tři různé způsoby. Každý má své výhody i nevýhody a oblast využití vzhledem k parametru sigma.

Změna průhlednosti - Opacity

První metoda s názvem *Opacity* byla jednoduše realizována pomocí postupného prolnutí mezi rozostřeným pruhem a původním snímkem. Jedná se v podstatě o stejnou techniku, která je využívána pro plynulý přechod mezi dvěma záběry ve filmu (prolínačka). Nevýhodou je, že zprůměrování rozostřených pixelů s ostrými má za následek poměrně rychle ubývající intenzitu rozostření a proto i účinnost celé metody.

Lineární změna rozostření - Linear

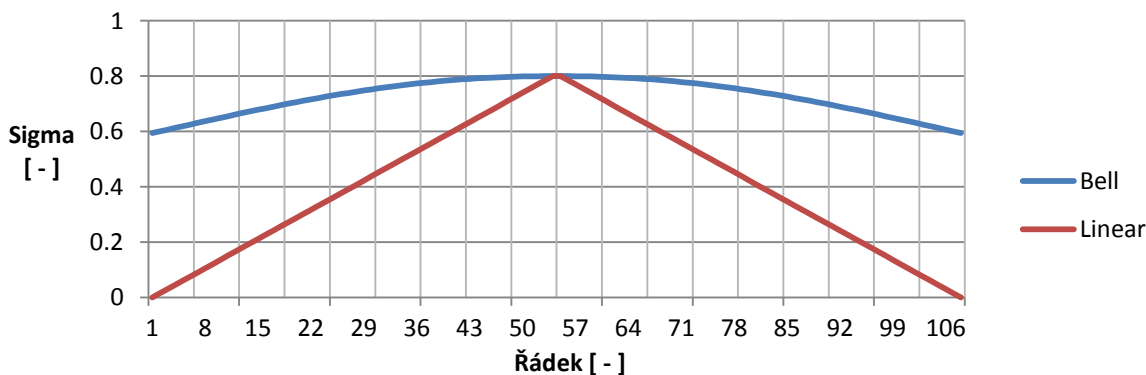
Jako další byla vyzkoušena metoda *Linear*, tato metoda je založena na lineárním snižování sigmy od středu k okrajům. Původní zvolená hodnota sigmy se tedy nachází jen ve středu pruhu a směrem k okrajům dochází ke snižování až na hodnotu 0. Nevýhodou této metody je, že na několik řádků velmi blízko okraje se aplikuje sigma, která se pohybuje pod mezní hodnotou 0,4, jež je hodnota kde nedochází prakticky k žádnému rozostření, a tedy i úspoře datového toku.

Nelineární změna rozostření - Bell

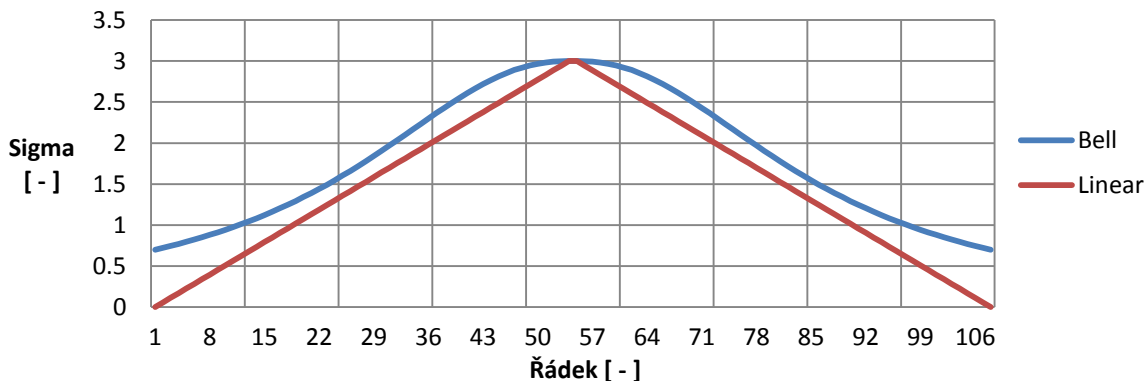
Poslední metoda s názvem *Bell*, je realizována pomocí bell-shaped funkce [7], která nelineárně snižuje hodnotu sigmy směrem od středu k okrajům a je dána vztahem (3).

$$f(x) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (3)$$

Tato funkce je definována pomocí tří parametrů, přičemž parametr c můžeme zanedbat neboť pouze definuje střed funkce. Důležitý je pro nás parametr a , který určuje vzdálenost poklesu funkce o polovinu od jejího středu a parametr b , jenž udává šířku vrcholu zvonu, oba tyto parametry spolu velmi úzce souvisí. V našem případě jsme zvolili parametr $b = 1,1$ a parametr a takový, aby pro zadanou sigmu do hodnoty 0,9 vždy krajní hodnota odpovídala o 0,2 menší než je sigma zadána (viz obr. 15) a pro parametru sigma vyšších hodnot než 0,9 je tato hranice nastavena vždy na hodnotu 0,7 (viz obr. 16).



Obr. 15. Průběh intenzity rozostření (sigma=0,8) jednoho pruhu o šířce 108 px



Obr. 16. Průběh intenzity rozostření (sigma=3) jednoho pruhu o šířce 108 px

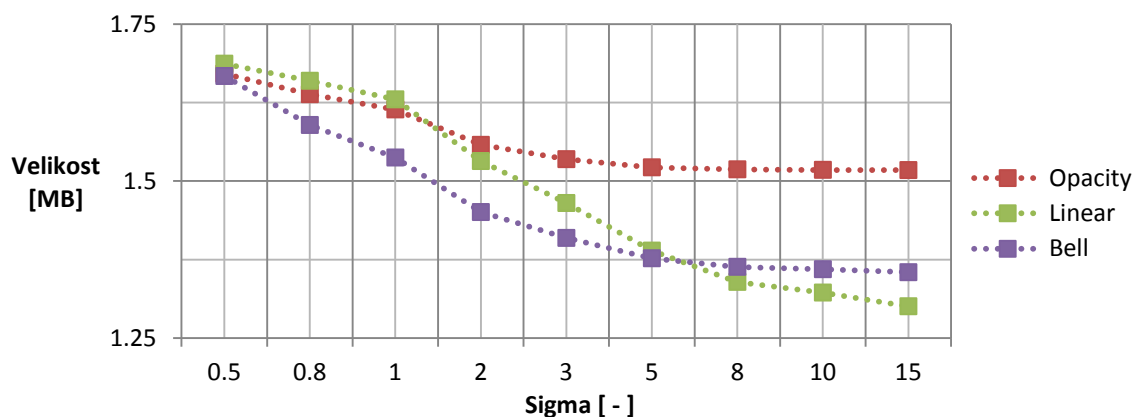
Porovnání metod

Všechny tři metody byly aplikovány na pět různých snímků (viz obr. 17) při devíti hodnotách sigma a stejné šířce pruhu (108 px), tyto snímky byly následně uloženy ve stejné kvalitě JPEG, aby mohlo dojít k porovnání datových objemů. V grafu jsou vyneseny zprůměrované křivky napříč všemi pěti snímky.



Obr. 17. Ukázka zdrojových snímků bez aplikace předzpracování

Vzhledem k faktu, že pod hodnotu sigma 5 je subjektivní vjem ze všech metod prakticky stejný a nad tuto hodnotu jsou pruhy v obraze již velmi patrné. Metoda *Bell* se zdá být pro dané rozmezí nejefektivnější (viz obr. 18) a my budeme pro tvorbu postupného rozostření pruhů používat právě ji. Výsledky ohledně subjektivní kvality metod jsou samozřejmě pouze orientační (4 pozorovatelé) a pro přesnější vyhodnocení by bylo třeba uskutečnit rozsáhlejší subjektivní ohodnocení.



Obr. 18. Závislost datového objemu fotografií na úrovni rozostření

7 Zpracování videosekvencí

V této kapitole si rozebereme celý postup zpracování videosekvencí. Nejdříve si popíšeme samotné videosekvence a jejich získání, poté se podíváme na programy použité pro jejich přípravu, která je důležitá například pro úpravu hloubky obrazu či zjednodušení následné práce, a nakonec si shrneme samotný postup přípravy.






7.1 Popis videosekvencí

Videosekvence byly vybrány s ohledem na rozmanitost prostředí, velikost a typ záběrů, pohybového a barevného obsahu či světelných podmínek.

První čtyři videosekvence jsme získali z veřejně dostupné databáze univerzity RMIT v Austrálii.³ Tyto videosekvence byly natočeny pomocí 3D kamery Panasonic AG-3DA1 s dvěma integrovanými objektivy, ve formátu Full HD, 10-bit, 4:2:2 YUV a 25 snímků za sekundu.

Vzhledem k absenci databáze na záběry se statickým detailem, byla pátá videosekvence natočena ve školní laboratoři pomocí zrcadlového rigu a dvou BMPCC (Blackmagic Pocket Cinema Camera) kamer ve Full HD nekomprimovaných RAW snímcích.⁴

Tab. 1. Parametry videosekvencí

| Číslo | Název | Sterobáze [cm] | Velikost záběru | Prostředí | Hloubka obrazu | Popis | Náhled |
|-------|-------|----------------|-----------------|------------------|----------------|---|---|
| 1 | Bikes | 6 | polocelek | exteriéry ve dne | střední | Kola v popředí, průjezd tramvaje v pozadí |  |
| 2 | Cube | 2 | detail | interiér ve dne | velká | Statická kostka |  |
| 3 | Fire | 6 | polocelek | exteriéry ve dne | velká | Plápolající oheň |  |
| 4 | Night | 6 | celek | exteriéry v noci | mírná | Zachycení noční dopravy |  |
| 5 | Train | 6 | velký celek | exteriéry ve dne | střední | Průjezd vlaku |  |

³ Databáze s 3D videi: <http://www.rmit3dv.com>

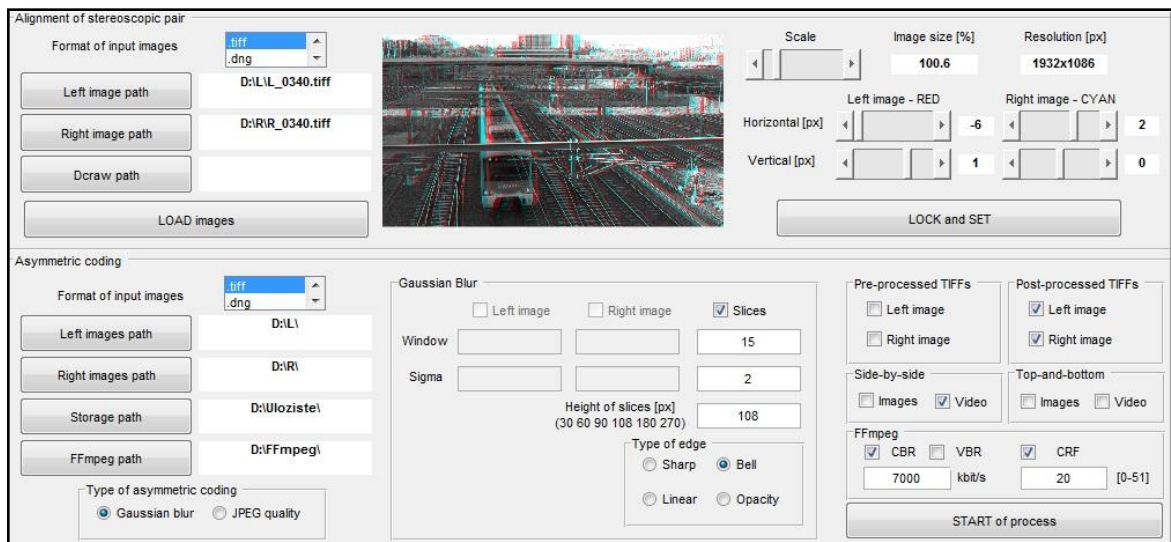
⁴ Podrobnější informace o kamerách: <https://www.blackmagicdesign.com/products/blackmagicpocketcinemacamera>

7.2 Použité programy pro přípravu a zpracování videosekvencí

Pro všechny kroky přípravy a zpracování videosekvencí nám postačí námi vytvořené GUI (grafické uživatelské rozhraní) v Matlabu s názvem ACT ve spolupráci s volně dostupnými nástroji Ddraw a FFmpeg, které budou popsány v následujících odstavcích.

ACT (Asymmetric Coding Tool)

GUI se skládá ze dvou částí (viz obr. 19), přičemž první má za úkol přípravu videosekvencí ve smyslu jejich zarovnání (hloubka obrazu), popřípadě archivaci a druhá slouží k aplikaci nesymetrického kódování s použitím Gaussova rozostření.



Obr. 19. Prostředí ACT

V první (horní) části s názvem *Alignment of stereoscopic pair* si můžeme všimnout levé části, která slouží k zadání cest pro zdrojové snímky ve formátu TIFF, a případně cestu k Ddraw za použití RAW vstupního formátu. Po stisknutí tlačítka *LOAD images* se vstupní snímky načtou do prostřední části GUI a nového *figure* ve formátu Anaglyf. Ten je velmi vhodný pro pouhý náhled na 3D scénu a její hloubku obrazu, protože tak můžeme učinit na jakémkoliv zobrazovacím zařízení jen za použití brýlí Anaglyf s barevnými filtry v kombinaci červená-azurová. Pokud není hloubka obrazu ideální, je zde možnost posunutí jednotlivých obrazů v horizontálním i vertikálním směru. Nevýhodou posunutí je nutný *scale* obou obrazů, při kterém dochází k interpolaci chybějících pixelů a znehodnocení původní kvality snímků. Po libovolných úpravách se použije tlačítko *LOCK and SET*, které znemožní další úpravy a zapamatuje si parametry zvětšení a posunutí obrazů pro využití v další části GUI.

Druhá část GUI (*Asymmetric coding*) zvládá masovou konverzi snímků, průběžné archivování po využití nesymetrického kódování a výstup do videa za použití nekomprimovaného AVI a kompresního formátu H.264. V levé části se začíná zadáním cest ke složkám levého a pravého obrazu se vstupními sekvencemi snímků, programu FFmpeg (při použití komprese) a do úložiště, kde se bude ukládat veškerý výstup.

Následuje možnost volby mezi dvěma typy nesymetrického kódování. Zaprvé námi používaná metoda *Gaussian Blur* a zadruhé metoda pomocí kvality kompresního standardu JPEG. Při zvolení jedné z možností nám prostřední okno nabídne detailnější nastavení parametrů konkrétní metody. U možnosti *Gaussian Blur* je to především možnost rozostření jednoho či obou z obrazů stereopáru, nastavení konkrétních parametrů rozostření a volba metody pruhů. Metoda *JPEG quality* nabídne pouze možnost volby obrazu, který bude komprimován pomocí kompresního standardu JPEG s určitou výslednou kvalitou od 0 (nejhorší) do 100. Po vyplnění konkrétních parametrů je nutné zvolit možné výstupy konverze, první možností je archivace nekomprimovaných TIFF snímků bez jakýchkoliv úprav, což se hodí zejména pro „přeuložení“ po zarovnání obrazů. Druhou možností je podobná archivace samotných TIFF snímků, ale tentokrát po aplikaci nesymetrického kódování. Dalšími možnostmi výstupu jsou TIFF snímky nebo AVI videa s využitím prostorové komprese SBS či TAB formátu. Poslední možností je pak využití komprese H.264/MPEG-4 AVC s nastavením konstantního, popřípadě variabilního bitového toku udávaného v kbit/s nebo pomocí udané kvality výstupního obrazu od 0 (nejlepší) do 51. Po nastavení potřebných možností a parametrů je tu pro spuštění tlačítko *START of process*.

Dcraw

Dcraw je volně dostupný program, který napsal David Coffin.⁵ Tento program je schopný číst a pracovat s celou řadou RAW formátů (.dng, .nef, cr2, atd.). Výstupem Dcraw jsou konvertované snímky do bezeztrátových formátů PPM nebo TIFF. Velmi jednoduše se ovládá a spouští přes příkazový řádek. Vždy je třeba zadat cestu k souboru a vybrat si z celé řady příkazů pro prvotní úpravy snímku, jako je například volba barevného prostoru, metody vyvážení bílé, prostorové manipulace (např. rotace a překlápění), saturace, gamma korekce a mnoho dalších.

FFmpeg

FFmpeg je volně dostupný nástroj, který je schopný pracovat s prakticky libovolným audio a video obsahem.⁶ Umí multimediální obsah nejen dekódovat a kódovat za účelem konverze, ale také přehrávat, grabovat, multiplexovat a mnoho dalších užitečných funkcí. Tento software je vyvíjen pomocí operačního systému Linux a jeho projekt založil vývojář Fabrice Bellard. Stejně jako software Dcraw se jednoduše spouští a ovládá přes příkazový řádek. My zejména využijeme schopnosti rozkládat video na jednotlivé snímky a kompresního standardu H.264/MPEG-4 AVC.

7.3 Příprava videosekvencí

Účelem přípravy videosekvencí bylo získat levý a pravý obraz všech pěti záběrů v nekomprimovaných a samostatných snímcích. Dalším důležitým krokem bylo zarovnání pravého a levého obrazu vzhledem k hloubce obrazu, která nebyla ve dvou videosekvencích z natáčení ideální.

7.3.1 Natočené záběry

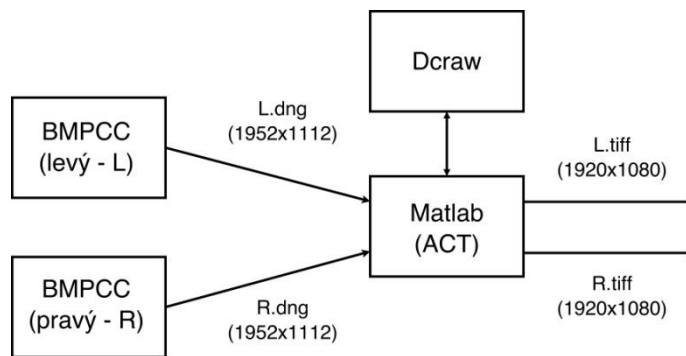
Při natáčení videosekvence byly pro nejvyšší výchozí kvalitu obrazu použity jako výstupní formát z kamer samotné RAW snímky. Ty jsme museli nejprve konvertovat na jiný formát, se kterým jsme

⁵ Program i s dokumentací ke stažení: <https://www.cybercom.net/~dcoffin/dcraw/>

⁶ Program i s dokumentací ke stažení: <https://www.ffmpeg.org/about.html>

byli dále schopni pracovat. Nejlepším formátem pro budoucí práci a zálohování zdrojových dat je bezztrátový TIFF. Ten nám oproti RAW formátu nabízí možnost prohlížení si snímků v jakémkoliv prohlížeči, což se nám hodilo zejména pro synchronizaci levého a pravého obrazu, kdy z důvodu manuálního spouštění kamer je jeden záznam vždy o pár snímků posunutý. Po zarovnání začátku a konce obou obrazů tak aby se shodovali v počtu snímků, jsme si oba kanály patřičně pojmenovali a vymazali nepotřebné části, což je vzhledem k velikosti samotných TIFF snímků velká úspora datového prostoru.

Ke konverzi jednotlivých RAW snímků do TIFF formátu byl potřeba program, který zvládá zejména masovou konverzi, což znamená, že vstupem může být celá složka se sekvencí snímků. Nejjednodušší možností je program Adobe Photoshop, který zvládá námi potřebné masové zpracování snímků, a má jednoduché ovládání. Bohužel z důvodů placené licence a neznalosti použitých algoritmů, které program využívá, není příliš vhodný.

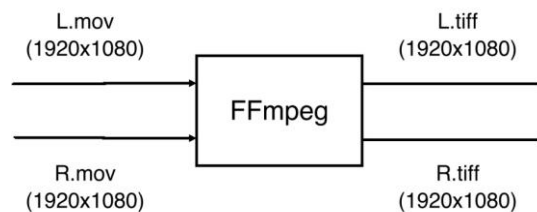


Obr. 20. Diagram zpracování RAW snímků

Další možností byl již dříve zmíněný software Dcraw, který má jeden velký nedostatek, a tím je neschopnost masové konverze. Abychom tento problém vyřešili, museli jsme do cesty zapojit ještě Matlab (GUI) (viz obr. 20), kde se nastavují samotné parametry změn a sám si dokáže vyvolávat Dcraw snímek po snímku. V GUI jsme navíc vyřešili potřebný ořez obrazu, k němuž muselo dojít, protože výstupní RAW snímky měly o něco větší rozlišení (1952x1112) než klasické FullHD (1920x1080), což je dáno tím, že obraz ze snímače obsahoval i okrajové pixely.

7.3.2 Převzaté záběry

Na rozdíl od námi natočených videosekvencí jsou ty převzaté již ve formě videa (.mov). To je pro naši další práci nevýhodné, a je třeba video rozložit na samotné snímky. K tomuto úkolu nám posloužil software FFmpeg za pomoci krátkého příkazu, kde se musí například definovat vstupní formát či barevný prostor: `-i input.mov -an -r 25 -pix_fmt rgb24 -vcodec tiff L_%04.tiff`.



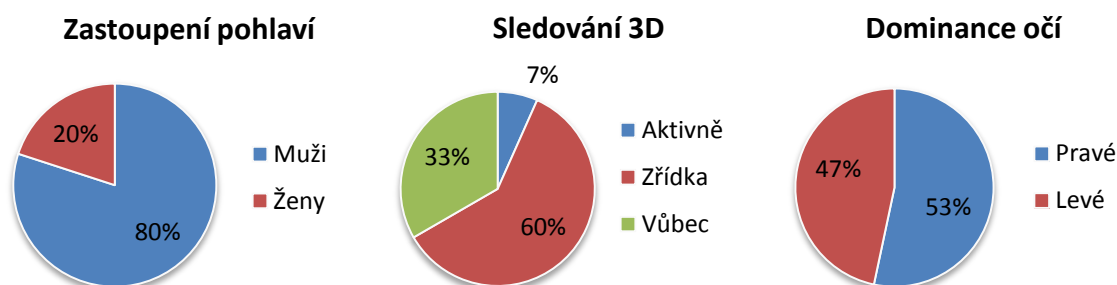
Obr. 21. Diagram zpracování převzatých záběrů

8 Subjektivní hodnocení

V této kapitole si popíšeme realizaci a vývoj subjektivního testování, dále se podíváme na způsob zpracování jeho výsledků, a nakonec si rozebereme konkrétní výsledky subjektivního hodnocení vlivu předzpracování videosekvencí a jejich následné komprese.

8.1 Realizace hodnocení se skupinou pozorovatelů

Subjektivní testování je způsob ohodnocení kvality videosekvencí pomocí určité skupiny lidí. Jelikož každá osoba vnímá kvalitu úplně jinak, je dobré testům podrobit co největší počet lidí obou pohlaví a s co nejširším věkovým rozpětím. Pro sjednocení podmínek testování nám slouží celá řada norem a doporučení. My jsme prováděli testování na základě doporučení ITU-R BT.2021-1 [18], které například definuje vzdálenost mezi subjekty a zobrazovačem v závislosti na jeho úhlopříčce nebo doporučuje metody průběhu testů pro hodnocení stereoskopického obsahu.



Obr. 22. Výsledky průzkumu subjektivního testování

8.1.1 Popis testování

Testování, kterého se zúčastnilo celkem 30 subjektů ve věku od 20 do 49 let bylo prováděno v laboratorních podmínkách na FEL ČVUT v Praze (viz obr. 23). K samotnému zobrazení videosekvencí byl použit aktivní displej Panasonic TX-P50VT20EA s úhlopříčkou 127 cm, který byl umístěn vždy před maximálně třemi subjekty najednou ve vzdálenosti cca 120 cm. Subjekty byly předem otestovány na zrakovou ostrost, barvocit a standardní vnímání hloubky obrazu, dále jim byla změřena rozteč a dominance očí. Celková doba testu činila 40 minut včetně trénovací sekvence a otestování subjektů. V testu byly hodnoceny tři parametry: celková kvalita zaměřená na vliv komprese, celková ostrost vůči intenzitě rozostření z předzpracování a hloubka obrazu.



Obr. 23. Uspořádání pracoviště pro subjektivní testy

8.1.2 Metoda DSIS

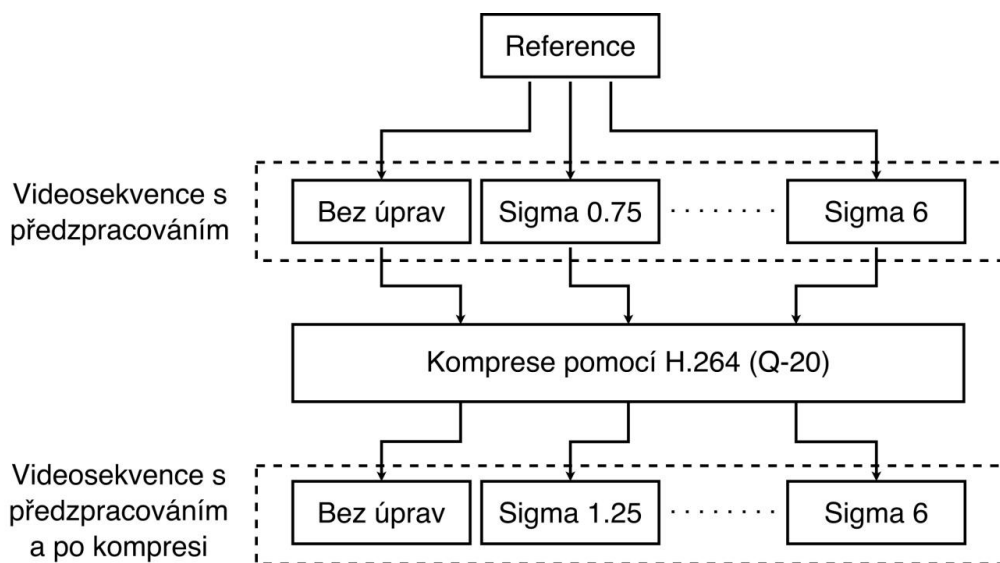
Testování bylo prováděno pomocí metody DSIS (Double-Stimulus Impairment Scale), která je vhodná zejména pro hodnocení vlivu zkreslení obrazu, při kterém zadavatel disponuje originální nezkreslenou předlohou (reference). Tato metoda předpokládá, že zkreslený obraz bude mít stejnou či horší kvalitu než známá reference. Hodnotí si v testu vždy první prohlédne referenci, tu porovná se zkreslenou videosekvencí, a poté ji ohodnotí pomocí pěti stupňové škály. Stupnice nabývá hodnot od 1 do 5 (viz tab. 2) a pro zjednodušení jsme oproti doporučení [18] otočili známkování. Průběh testu vypadá tak, že mezi referencí (T1) a porovnávanou videosekvencí (T3) je vždy tři sekundová pauza v podobě střední úrovně šedé (T2). Po zobrazení zkresleného záběru (T3) je znova šedá, tentokrát však sedmi sekundová (T4) a je určena pro samotné ohodnocení celkové kvality, ostrosti a hloubky obrazu do předem připraveného formuláře (viz kapitola 14.1).

Tab. 2. Stupnice pro subjektivní hodnocení a struktura prezentace testu

| Hodnota | Popis | Průběh testu |
|---------|------------------------------|--|
| 1 | nepostřehnutelné | <p>The diagram illustrates the test sequence. It shows a horizontal timeline with four segments labeled T1, T2, T3, and T4. T1 is a high-level pulse, T2 is a medium-level pulse, T3 is a high-level pulse, and T4 is a medium-level pulse. A dashed line with an arrow labeled 'Hlasování' (voting) spans the duration of T3 and T4. The timeline starts and ends with a break symbol (//).</p> |
| 2 | postřehnutelné, ale nerušivé | |
| 3 | lehce rušivé | |
| 4 | rušivé | |
| 5 | velmi rušivé | |

8.1.3 Obsah testu

Na každý z pěti záběrů bylo aplikováno předzpracování s pěti různými hodnotami sigma: 0,75, 1,25, 2, 3 a 6 (viz kapitola 6.1). Předzpracování bylo aplikováno pomocí metody pruhů s postupným rozostřením (*Bell* metoda), kdy v levém obraze byly rozostřené liché pruhy a v pravém obraze pruhy sudé. Oba obrazy byly následně zakódovány pomocí prostorové komprese (SBS formát). Tyto záběry byly dále komprimovány pomocí kompresního standardu H.264/MPEG-4 AVC s konstantním kvantizačním parametrem 20 (viz obr. 24), který byl zvolen vzhledem k výstupnímu bitovému toku, jenž se měl pohybovat okolo 6 Mbit/s podle doporučení pro streaming [41]. Zkomprimované videosekvence byly nakonec překonvertovány do bezztrátového *.avi* formátu, kvůli optimalizaci přehrávání. Na začátku testování byly zařazeny dvě trénovací porovnání, aby si subjekty mohly vyzkoušet způsob hodnocení a zkalibrovaly jeho stupnici. Nakonec test obsahoval 25 porovnávaných záběrů pouze s předzpracováním, dalších 25 předzpracovaných po kompresi a 10 kontrolních referencí od obou skupin, které měly být podle předpokladu hodnoceny stupněm 1 (nepostřehnutelné rušení).



Obr. 24. Blokové schéma postupu tvorby videosekvencí do testu

8.1.4 Otestování subjektů

Jak již bylo řečeno v popisu subjektivního testování, tak před samotným testem jsme provedli otestování subjektů ohledně jejich oční dominance, zrakové ostrosti a barvocitu. Dále jsme změřili jejich rozteč očí, která vyšla v průměru 6.25 cm. Poměrně zásadním a chybějícím otestováním byl test na přirozené vnímání hloubky obrazu, například pomocí Stereo Randot testu. Z tohoto důvodu byla subjektům položena pouze otázka ohledně jejich vnímání hloubky obrazu z předchozích zkušeností se 3D. Popřípadě jim byla puštěna testovací videosekvence, která potvrdila vnímání hloubky, které nakonec bylo u všech subjektů zcela v pořádku.

Oční dominance

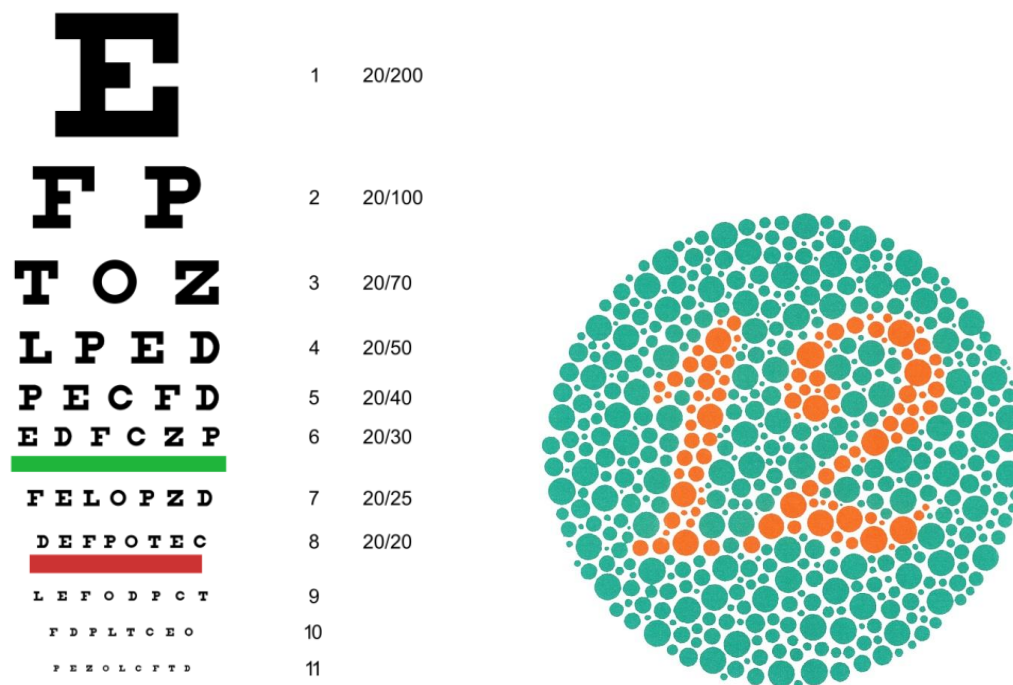
Prvním otestováním bylo určení oční dominance, které se dá uskutečnit podle celé řady metod [35]. Mezi nejznámější patří Porta test nebo Dolmanova metoda. Další známou metodou je Milesův test, který jsme použili my. Realizace testu je velmi jednoduchá, jde v něm o spojení dlaní tak, aby oba ukazováčky a palce mířily proti sobě, čímž vytvoří malý trojúhelník. Subjekt se po utvoření trojúhelníku skrze něj zaměří na nějaký objekt, zhruba ve výši očí a vzdálenosti 6 metrů. Pokud objekt po zakrytí pravého oka z trojúhelníku zmizí, subjekt je dominantní na pravé oko a naopak.

Zraková ostrost

Jako další byl test na zrakovou ostrost subjektů pomocí Snellenovy tabule (viz obr. 25), jenž je velmi rozšířená diagnostická pomůcka pro určení zrakové ostrosti [14]. Tabule běžně obsahuje jedenáct řádků písmen, které se postupně zmenšují od 88 mm až po písmena menší než půl centimetru. Vedle každé řádky se nachází Snellenův zlomek, který udává poměr mezi vaší vzdáleností od tabule (čitatel) a vzdáleností průměrného člověka (jmenovatel), který by dokázal písmena bez problému přečíst. Pro nás je stěžejní právě řádek označený zlomkem 6/6 s velikostí písmen 8,8 mm, který by měl být subjekty čitelný bez jakýchkoliv potíží ze vzdálenosti 6 metrů. Pro původní Snellenovu tabuli, udávanou ve stopách, by to byl řádek označený zlomkem 20/20.

Barvocit

Nakonec byl prověřen barvocit pomocí Ishiharových obrazců (viz obr. 25), resp. Ishiharova testu, který slouží k přesnému odhalení druhu a rozsahu poruch barvocitu [39]. Jedná se o jednoduchou identifikaci čísel, která se nacházejí v odlišně obarveném prostředí. Osoba s poškozeným barvocitem většinou identifikuje číslici špatně, nebo v nejhorším případě žádnou číslici nevidí.



Obr. 25. Ukázka Snellenovy tabule (vlevo) [14] a Ishiharova obrazce (vpravo) [39]

8.2 Zpracování výsledků

Vzhledem k faktu, že subjekty hodnotily celkově tři parametry: celkovou kvalitu, ostrost a hloubku obrazu. Bylo od každého subjektu vyplněno 180 ohodnocení. Z daných výsledků bylo vždy určeno průměrné skóre dle vzorce (4) napříč všemi hodnotícími, a ještě před samotným vynesemím do grafů, jim byly určeny intervaly spolehlivosti.

$$\bar{u}_{ijk} = \frac{1}{N} \sum_{i=1}^n u_{ijk} \quad (4)$$

kde u_{ijk} je hodnota vyplněná pozorovatelem i pro parametr j u sekvence k .

Interval spolehlivosti

Protože bodový odhad (aritmetický průměr) je ovlivněn, například výběrem daných subjektů, používá se ještě intervalový odhad. V něm nalezneme interval, ve kterém se bude skutečná hodnota populačního průměru nacházet s určitou pravděpodobností. Jako hranice mezi „je velmi nepravděpodobné“ a „ne tak nepravděpodobné“ se běžně používá pravděpodobnost 5%, tuto hodnotu nazýváme hladina významnosti a označujeme α . Potom hodnota $1-\alpha$ se nazývá koeficient spolehlivosti, který se po převedení na procenta dá považovat za $(1-\alpha) \cdot 100$ -ní interval spolehlivosti. V našem případě použijeme doporučený 95% interval spolehlivosti [19], který se vždy přičte a odečte od průměrného skóre u hodnocení (5).

$$[\bar{u}_{ijk} - \delta_{ijk}, \bar{u}_{ijk} + \delta_{ijk}] \quad (5)$$

kde δ_{ijk} se vypočítá podle vzorce (6).

$$\delta_{ijk} = 1,96 \cdot \frac{S_{ijk}}{\sqrt{n}} \quad (6)$$

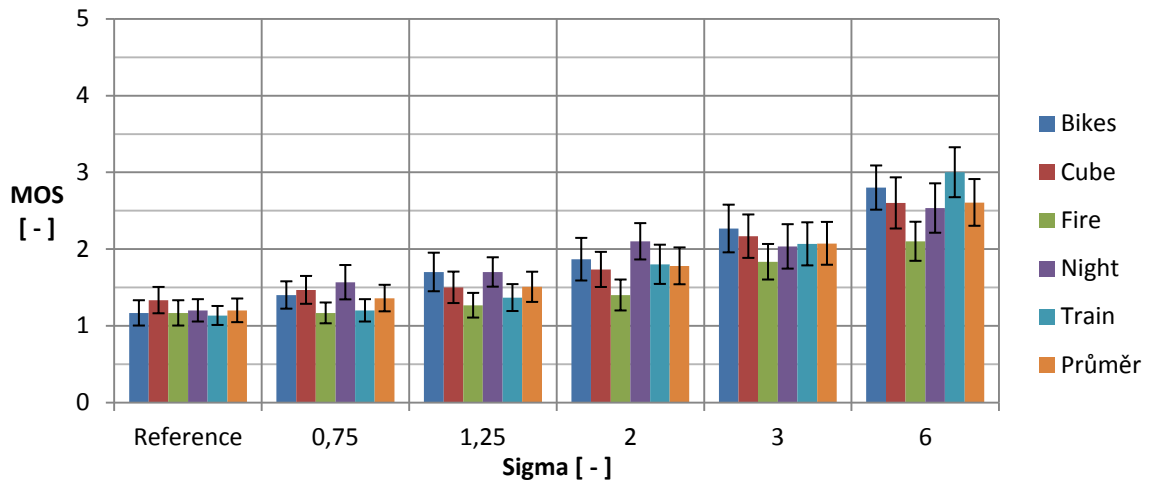
kde n je počet hodnotících a S_{ijk} je směrodatná odchylka daná vzorcem (7).

$$S_{ijk} = \sqrt{\sum_{i=1}^n \frac{(\bar{u}_{ijk} - u_{ijk})^2}{(n-1)}} \quad (7)$$

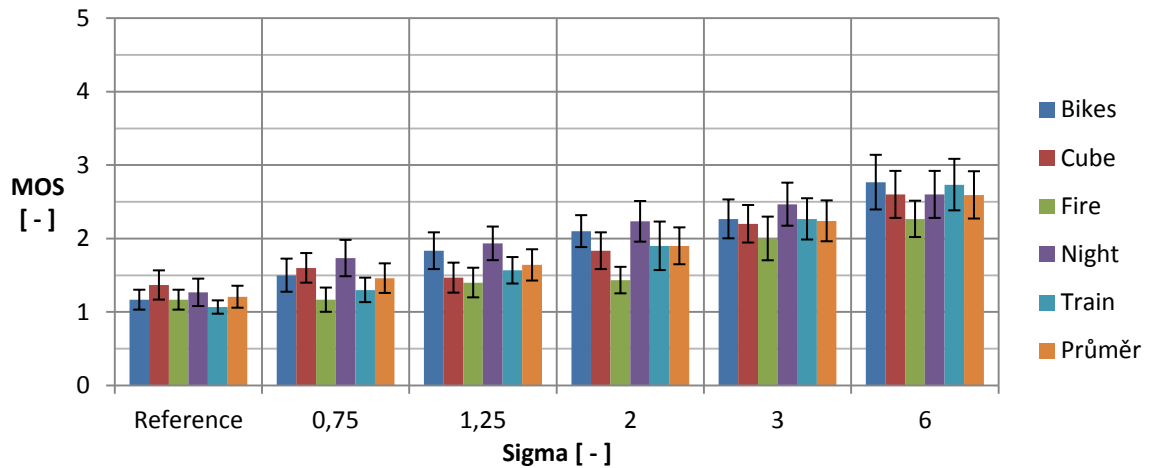
8.3 Výsledky subjektivního testování

Výsledky subjektivních testů byly vyneseny do grafů v závislosti na intenzitě předzpracování (parametr sigma). První tři grafy (viz obr. 26-28) znázorňují výsledky ohodnocení celkové kvality, ostrosti a hloubky obrazu jednotlivých videosekvencí po vlivu předzpracování. Druhé tři grafy (viz obr. 29-31) ukazují výsledky subjektivního hodnocení po aplikované kompresi (Q-20). V každém grafu je navíc přidán průběh, který je průměrem výsledků hodnocení jednotlivých videosekvencí.

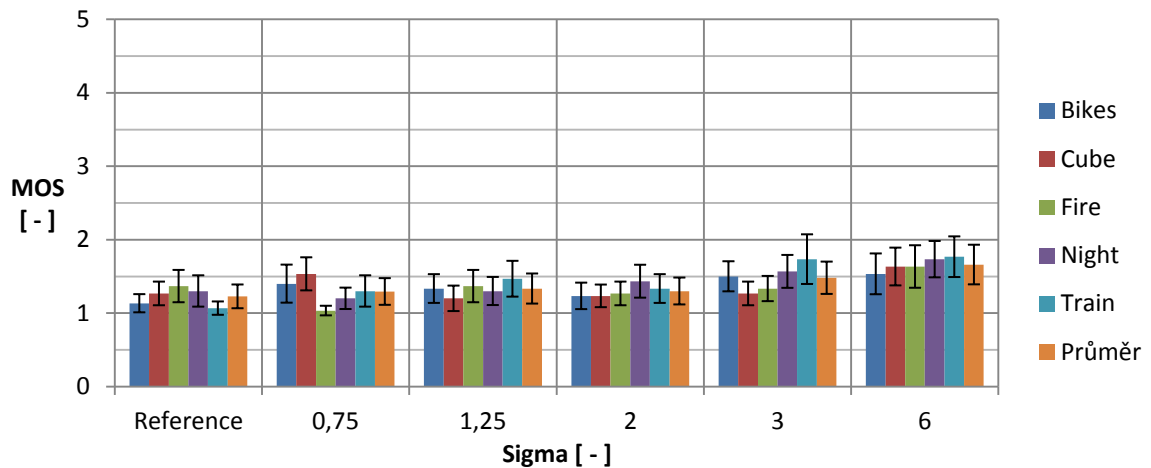
Hodnocení vlivu předzpracování



Obr. 26. Celková kvalita: Subjektivní hodnocení v závislosti na parametru sigma

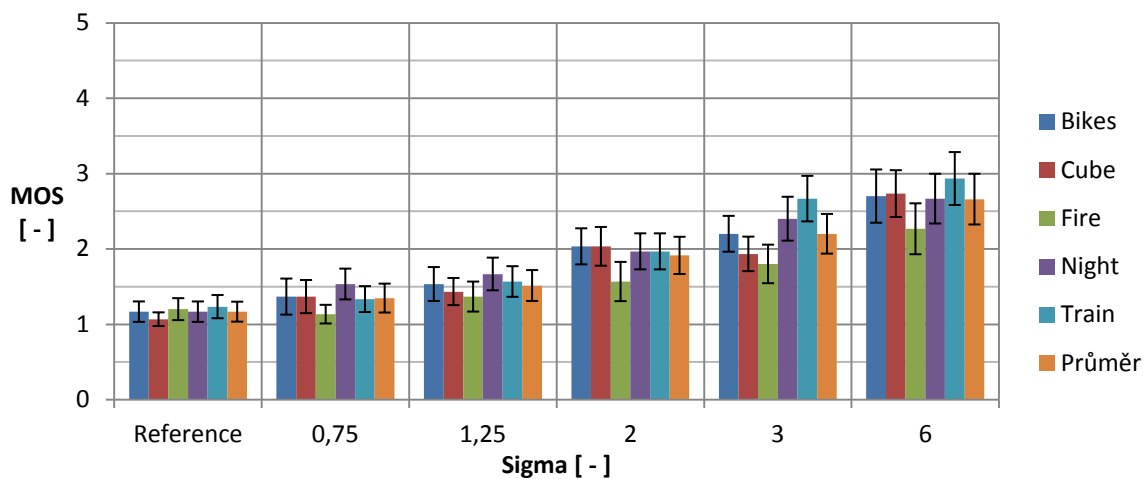


Obr. 27. Celková ostrost: Subjektivní hodnocení v závislosti na parametru sigma

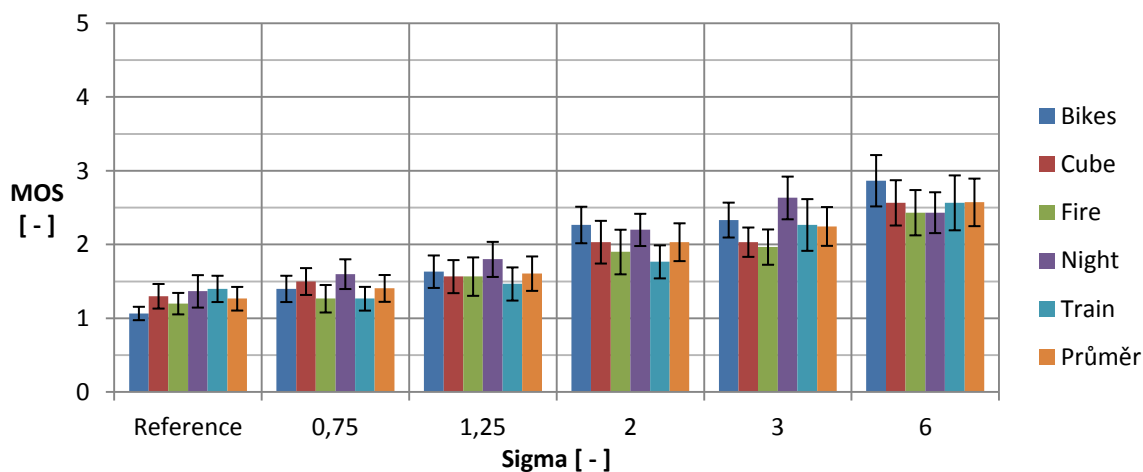


Obr. 28. Hloubka obrazu: Subjektivní hodnocení v závislosti na parametru sigma

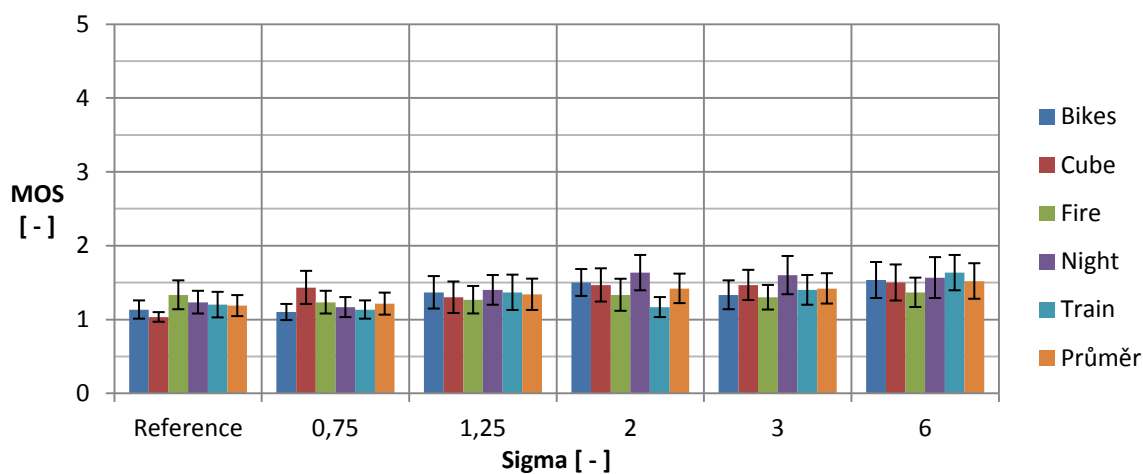
Hodnocení vlivu komprese



Obr. 29. Celková kvalita: Subjektivní hodnocení v závislosti na parametru sigma



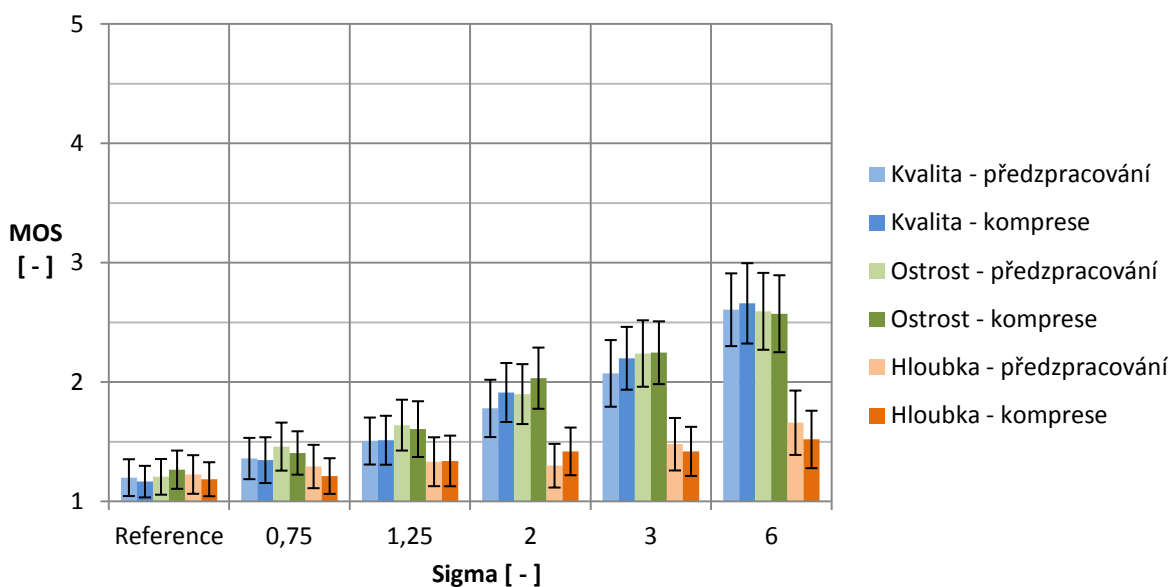
Obr. 30. Celková ostrost: Subjektivní hodnocení v závislosti na parametru sigma



Obr. 31. Hloubka obrazu: Subjektivní hodnocení v závislosti na parametru sigma

Shrnutí výsledků

Z výsledků subjektivního testování je patrné, že se celková kvalita a ostrost se zvyšující intenzitou rozostření (parametrem sigma) prakticky lineárně zhoršují pro všechny záběry, nehledě na aplikovanou kompresi. Je vidět, že mezi vnímáním kvality a ostrosti je velmi úzký vztah, dá se tedy říct, že intenzita rozostření se silně podílí na celkovém vnímání kvality. Mezi výsledky hodnocení vlivu předzpracování a komprese není praktický žádný rozdíl, to může být způsobeno nedostatečně viditelnými artefakty vznikajícími samotnou kompresí. Na druhou stranu lze říci, že zvyšování parametru sigma nemá na hloubku obrazu prakticky žádný vliv a proto se výsledné hodnoty pohybují vždy maximálně do skóre 2. Mírně zhoršené vnímání hloubky obrazu může být způsobené spíše metodou porovnávání, kdy si subjekt po zhlédnutí reference zvykne na její hloubku obrazu, takže na něj porovnávané video tolik nezapůsobí.



Obr. 32. Srovnání průměrů subjektivního hodnocení vlivu předzpracování a komprese

9 Objektívni hodnocení

Objektivní testování je založeno na matematických popisech vlastností obrazu, různých algoritmech a výpočtech z dostupných obrazových dat. Metriky rozdělujeme podle toho, zda máme k dispozici originální obrázek, který je nezkreslený nebo nějakým způsobem poškozený. My se budeme dále zabývat pouze případem, kdy máme jeden originální (nezkreslený) snímek, a k němu druhý upravený (rozostření či komprese). Výhodou objektivního testování může být rychlost vyhodnocení nebo finanční nenáročnost. Na druhou stranu je zřejmá nevýhoda v dosaženém ohodnocení, které se pouze přibližuje výsledku hodnocení skupinou pozorovatelů.

9.1 Objektívni metriky

Existuje celá řada objektivních metrik, ovšem my se zaměříme na jedny z nejpoužívanějších. PSNR a SSIM jsou dvě asi nejrozšířenější metriky vhodné pro srovnávání výsledků, které naleznete ve většině vědeckých pracích, které se zabývají objektivním hodnocením obrazu [24].

MSE

Jedna z nejjednodušších a nejčastěji používaných objektivních metrik se nazývá MSE (Mean Squared Error), jde o výpočet střední kvadratické chyby mezi srovnávanými snímky a je dána vztahem (8).

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X(i,j) - Y(i,j))^2 \quad [-] \quad (8)$$

kde $M \times N$ je rozlišení obrazu, X je vstupní testovaný snímek a Y snímek referenční (originál). Pokud by se snímky shodovaly, hodnota MSE by se rovnala nule, je tedy zřejmé, že se zvyšující se hodnotou MSE dochází k odchýlování referenčního snímku od porovnávaného.

PSNR

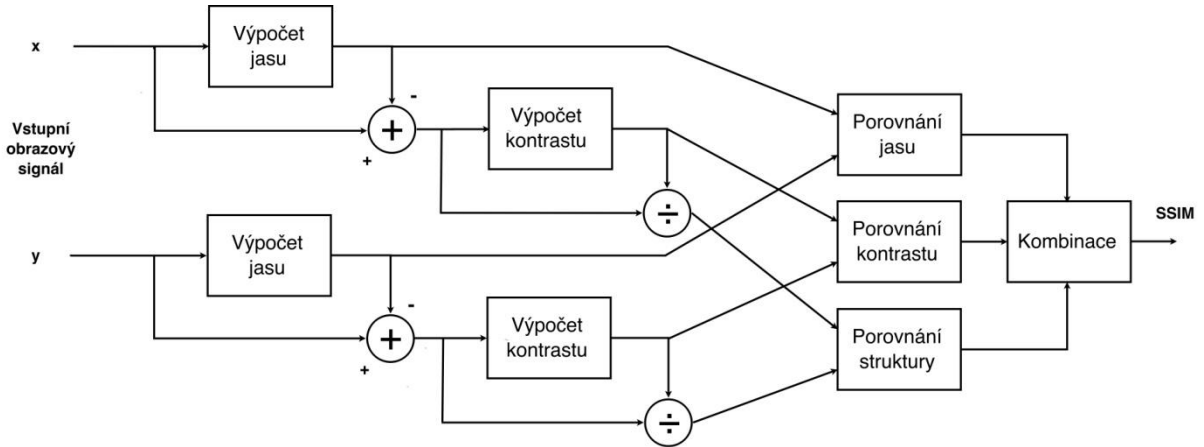
Metrika PSNR (Peak Signal-to-Noise Ratio) vychází z výpočtu MSE a zavádí se z důvodu možnosti srovnávání snímků s rozdílnými dynamickými rozsahy (bitovou hloubkou obrazu). Vyjadřuje špičkový poměr signálu k šumu a je dána vztahem (9).

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad [dB] \quad (9)$$

kde MAX je maximální možná hodnota pixelu ve snímku dána jeho bitovou hloubkou. Například pro 8 bitovou hloubku obrazu se bude MAX rovnat hodnotě 255.

SSIM

Jedná se o metodu porovnávání dvou obrazů, která se oproti klasickým metodám snaží brát v úvahu i psychovizuální model lidského vnímání [38]. SSIM (Structural Similarity) index vyjadřuje číselný rozdíl mezi dvěma obrazy, který může nabývat hodnot od -1 do 1 a skládá se ze tří částí porovnání: jasového, kontrastního a strukturálního. Metoda je symetrická, a tudíž nehlédě na pořadí porovnávaných snímků, vrací vždy stejnou hodnotu.



Obr. 33. Blokový diagram výpočtu SSIM indexu

Vzhledem k náročným principům, které metoda používá ji zde nebudu podrobněji rozepisovat. Náhled na metodu si lze udělat z blokového schématu (viz obr. 33), a pro ilustraci ještě uvádím její výslednou rovnici (10).

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (10)$$

Metriky pro hodnocení stereoskopického obsahu

Přestože v práci používáme pouze základní metriky PSNR a SSIM pro výpočet 2D obrazu, existuje ještě celá řada dalších metrik, které byly vyvinuty přímo za účelem hodnocení stereoskopického obsahu [9]. Naprostá většina těchto metrik vychází z využití základních metod (MSE, SSIM, atd.) pro posouzení kvality obrazu, které ovšem neberou v potaz specifika, která jsou zásadní pro stereoskopický obsah (disparita, crosstalk, hloubka obrazu, vizuální komfort, únava očí, atd.). Tyto základní metriky se proto doplňují algoritmy, které s těmito jevy počítají a výsledné hodnocení je potom kombinací těchto dvou posouzení.

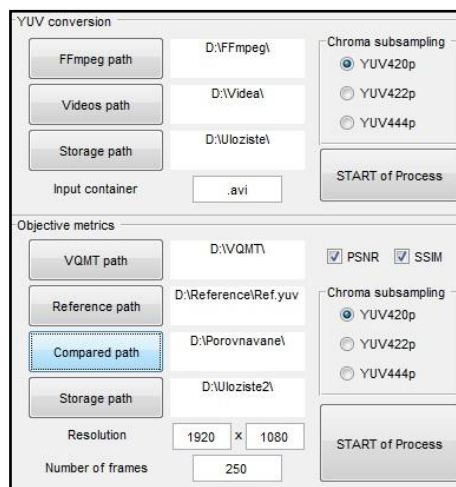
Každá metoda je sama o sobě velmi složitá a proto zde uvedu jen pár příkladů s příslušnými odkazy na literaturu. Například metody počítající zvlášť levý a pravý obraz, které využívají metriku SSIM a berou v potaz buď crosstalk [40], disparitu [6] nebo hloubku obrazu [23]. Lepší výsledky by ovšem měli vykazovat metody, které jsou založené na faktu, že při sledování 3D si mozek oba obrazy spojí v jeden, takzvaný kyklopský pohled, kterého je dosaženo v metrikách například pomocí 3D-DCT transformace. Na té jsou založeny metriky jako je například PHSV-3D [21], PHSD (pro mobilní aplikace) [20] nebo HV3D [4], která v současnosti patří podle dostupných parametrů mezi nejúčinnější metriky pro hodnocení stereoskopického obsahu.

9.2 Použité programy pro výpočet objektivních metrik

Pro výpočet objektivních metrik nám posloužilo druhé GUI vytvořené v Matlabu s názvem OMT ve spolupráci se softwarovými nástroji FFmpeg a VQMT.

OMT (Objective Metrics Tool)

GUI se skládá ze dvou částí (viz obr. 34). První část slouží pro konverzi zdrojových videosekvencí do YUV formátu, který je zcela nezbytný, jako vstupní formát pro výpočet objektivních metrik ve druhé části GUI. Pro konverzi je nejprve nutné zadat cestu k softwaru FFmpeg, zdrojovým videím a do samotného úložiště. Dále příponu, která slouží k identifikaci kontejneru zdrojových videí a poměr podvzorkování barvonosných složek na výstupu. Po převodu již můžeme využít druhé části GUI a dvou možných výpočtů objektivních metrik PSNR a SSIM. Před samotným výpočtem je však nutné zadat cestu k VQMT softwaru, referenčnímu videu, složce s porovnávanými videi, a nakonec do úložiště, kde se budou ukládat samotné výsledky v souborovém formátu .csv, jenž je formát čitelný například programem Microsoft Excel nebo Poznámkovým blokem. Dalšími potřebnými parametry jsou rozlišení vstupních videosekvencí, počet snímků k porovnání, a nakonec poměr podvzorkování barvonosných složek na vstupu.



Obr. 34. Prostředí OMT

VQMT

VQMT (Video Quality Measurement Tool) je volně dostupný software pro výpočet objektivních metrik, který byl vytvořený na základě implementace z Matlabu⁷ pro komerčnější užití a kratší dobu výpočtů. Nabízí možnost výpočtu několika objektivních metrik, jako je metrika PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity), MS-SSIM (Multi-Scale Structural Similarity), VIFp (Visual Information Fidelity, pixel domain version) a dalších dvou upravených verzí PSNR. Software pracuje pouze s YUV formátem videa a pro jeho činnost je potřeba mít nainstalovanou OpenCV library⁸ (jádro a modul imgproc).

⁷ Implementace z Matlabu: <http://www.cns.nyu.edu/lev/ssim/>

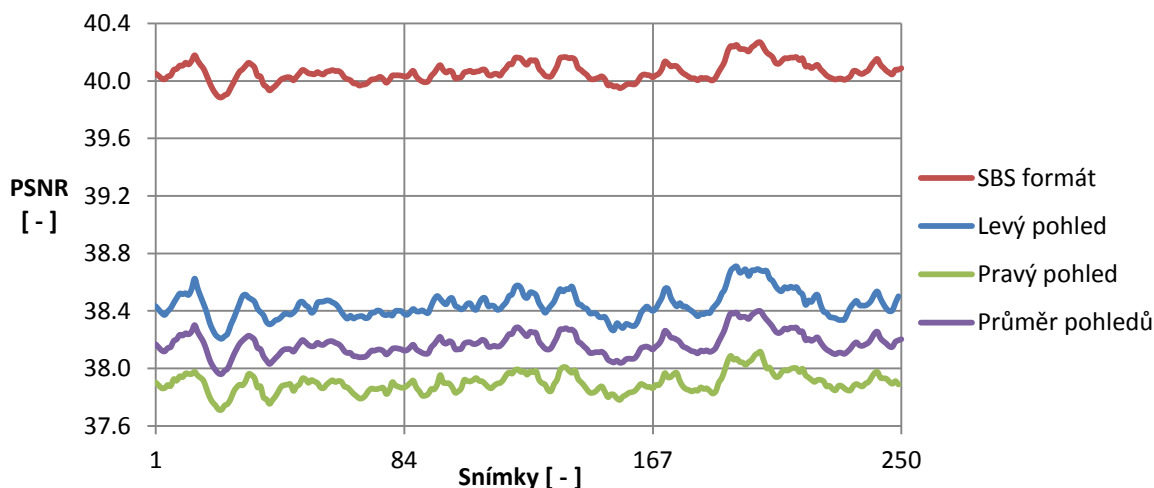
⁸ VQMT software s potřebnými knihovnami: <http://mmspg.epfl.ch/vqmt>

9.3 Výsledky objektivního testování

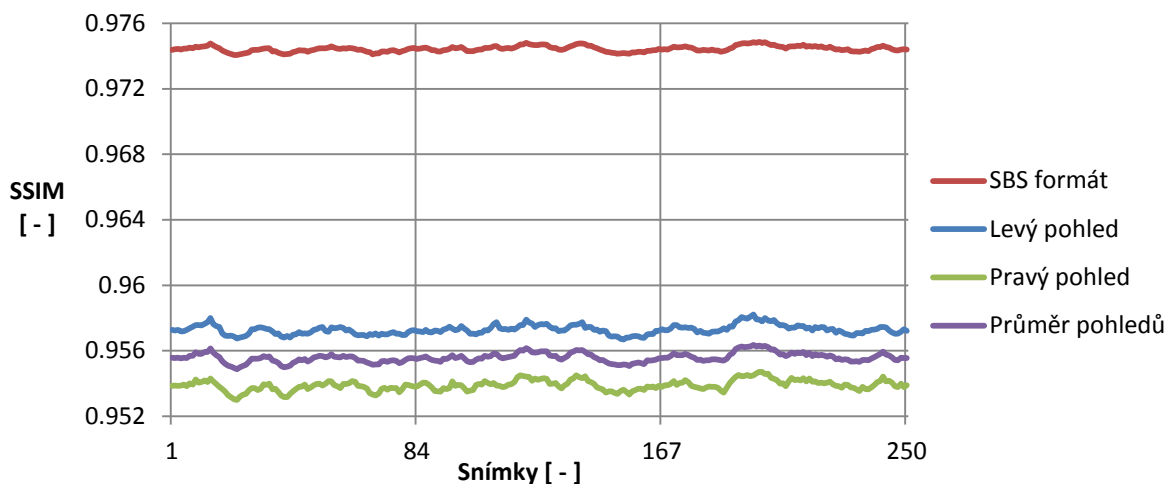
Při objektivním testování byla vždy porovnávána reference jednoho ze záběrů s jejími zkreslenými verzemi upravenými pomocí předzpracování (5 výpočtů) a následné komprese (6 výpočtů) (viz kapitola 8.1.3). Komprese byla provedena celkově ve třech úrovních s kvantizačním parametrem Q10, Q-20 a Q-30, přičemž stěžejní jsou pro nás výsledky s kvantizačním parametrem 20, který byl použit pro subjektivní hodnocení kvality obrazu. Ostatní dvě úrovně jsou spíše ilustrativní pro výsledné porovnání bitových toků. První úroveň (Q-10) demonstruje oblast komprese, která nemá na výslednou kvalitu obrazu žádný vliv („Lossless“). Druhá úroveň (Q-30) naopak oblast, kdy vlivem komprese dochází ke značné ztrátě kvality obrazu („Lossy“). Ve výsledných grafech jsou vždy průběhy parametru sigma v závislosti na jedné z metrik (PSNR, SSIM). V případě použitých kompresí je v grafu navíc hodnota odpovídající sigma 0, ta reprezentuje videosekvence bez použitého předzpracování, jde tedy pouze o porovnání reference s komprimovanou referencí.

Určení výsledné hodnoty metrik

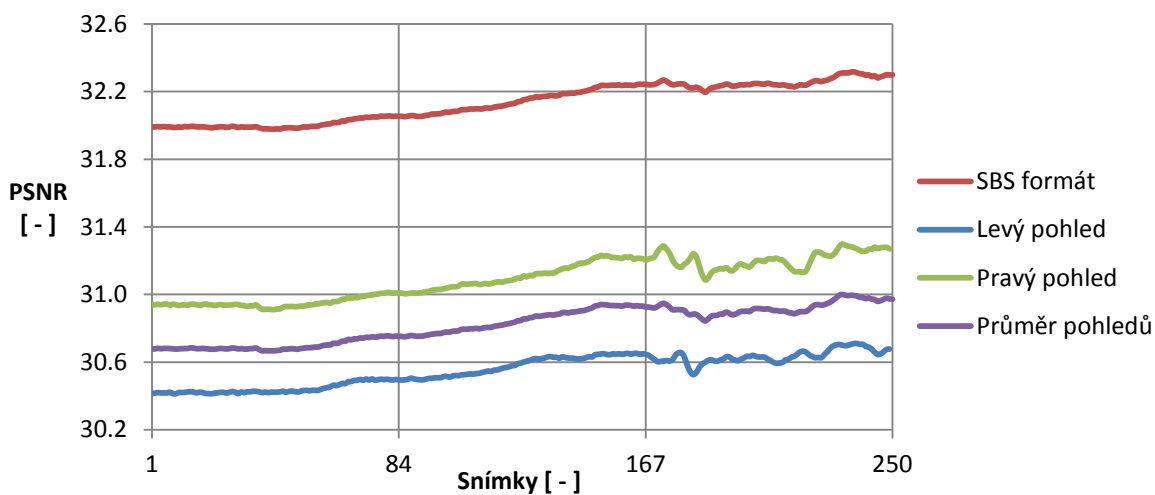
Ještě před samotným zobrazením výsledků je důležité definovat výslednou hodnotu z výpočtů obou metrik. Jelikož při výpočtu dochází k porovnávání snímků po snímku, tak se výsledné hodnoty mohou pohybovat v určitém rozsahu. My budeme za konečnou hodnotu považovat vždy aritmetický průměr jednotlivých výsledků v průběhu výpočtu celé videosekvence. Další důležitou věcí je, zdali se budou metriky počítat pro každý pohled stereopáru zvlášť a následně se udělá jejich průměrná hodnota, nebo se video ponechá prostorově komprimované ve formátu SBS. Obě možnosti byly vyzkoušeny a pro ilustraci vyneseny do grafů, kde jsou průběhy výpočtu obou metrik v závislosti na jednotlivých snímcích s aplikovaným předzpracováním (sigma = 1,25) a bez použité komprese. První čtyři grafy (viz obr. 35-38) obsahují srovnání výsledků jednotlivých pohledů a jejich průměrné hodnoty s formátem SBS u záběrů Fire a Train. Zbylé grafy (viz obr. 39-41) jsou již jen samotné výsledky pro SBS formát ostatních tří záběrů, kde červená křivka reprezentuje střední hodnotu konkrétního průběhu, která je výslednou hodnotou výpočtu objektivních metrik.



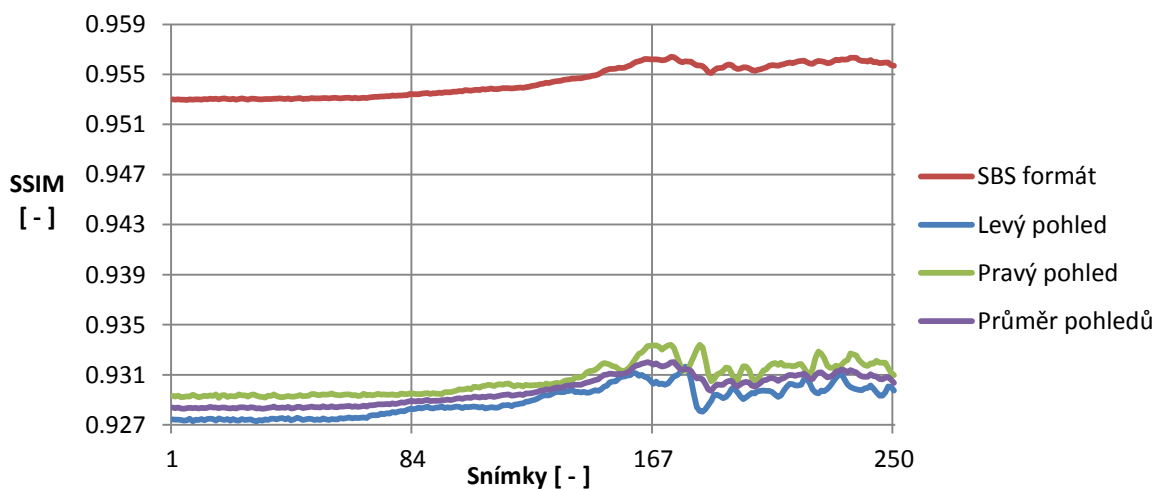
Obr. 35. Porovnání metriky PSNR jednotlivých pohledů stereopáru s formátem SBS záběru Fire



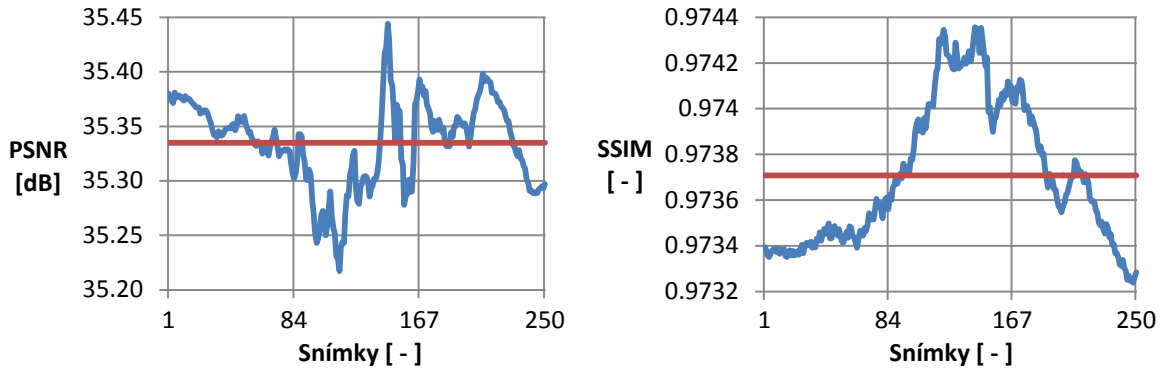
Obr. 36. Porovnání metriky SSIM jednotlivých pohledů stereopáru s formátem SBS záběru Fire



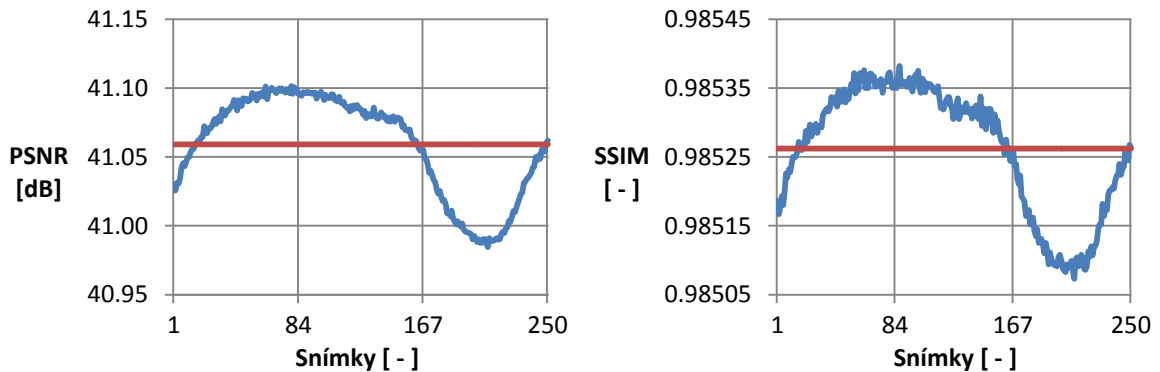
Obr. 37. Porovnání metriky PSNR jednotlivých pohledů stereopáru s formátem SBS záběru Train



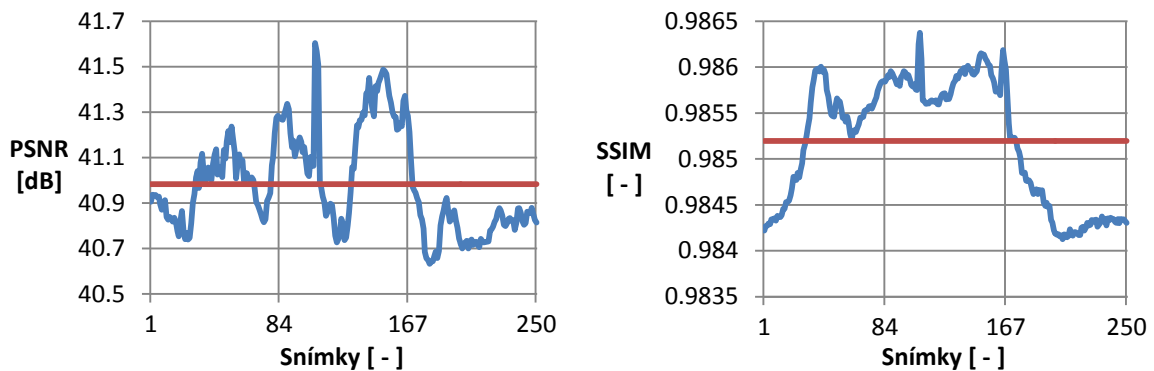
Obr. 38. Porovnání metriky SSIM jednotlivých pohledů stereopáru s formátem SBS záběru Train



Obr. 39. Průběh výpočtu PSNR a SSIM záběru Bikes



Obr. 40. Průběh výpočtu PSNR a SSIM záběru Cube

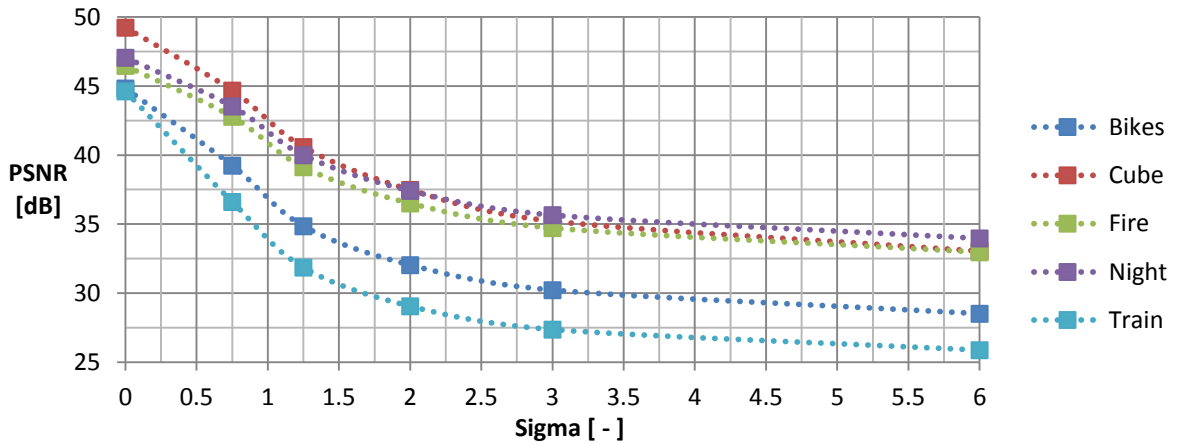


Obr. 41. Průběh výpočtu PSNR a SSIM záběru Night

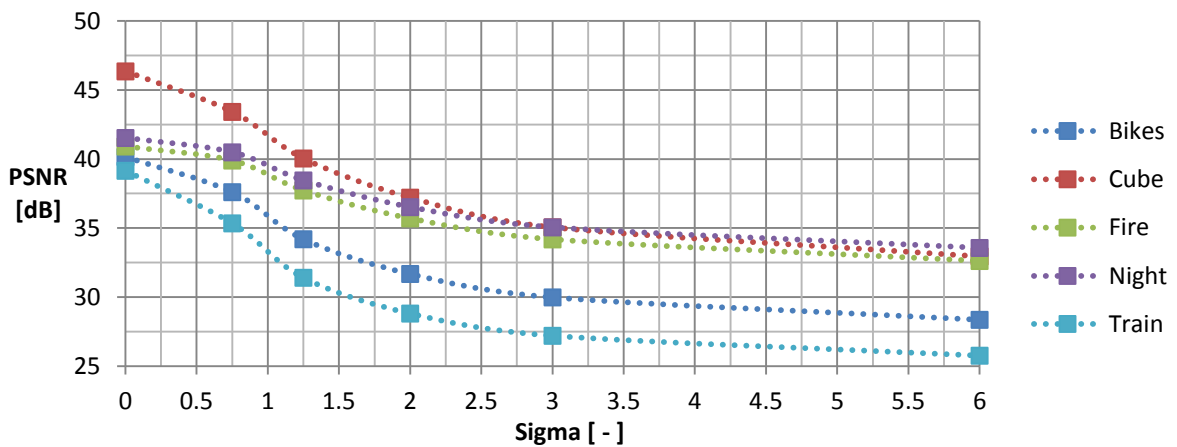
Z grafů je patrné, že průběhy jednotlivých pohledů (levý a pravý) jsou si velmi podobné, ale jejich výsledky jsou vždy trochu posunuté, což je dáno převážně prostorově závislým předzpracováním. Dále jsou tyto průběhy velmi podobné i SBS formátu, přičemž jejich průměrná hodnota je s ním prakticky identická. Z tohoto důvodu, budeme pro zjednodušení počítat objektivní metriky vždy pro prostorově komprimovanou videosekvenci ve formátu SBS. Z grafů je také vidět, že časový vývoj křivky je velmi závislý na obsahu videosekvence, a pro každou vypadá úplně jinak. Na druhou stranu se průběhy křivek pro obě metriky ve třech případech skoro neliší. Dále bylo ozkoušeno, že pro různé hodnoty sigma z předzpracování se výsledné tvary křivek prakticky nemění, stejně tak i následná komprese nemá na jejich tvar žádný výrazný vliv.

Výsledky vlivu komprese - metrika PSNR

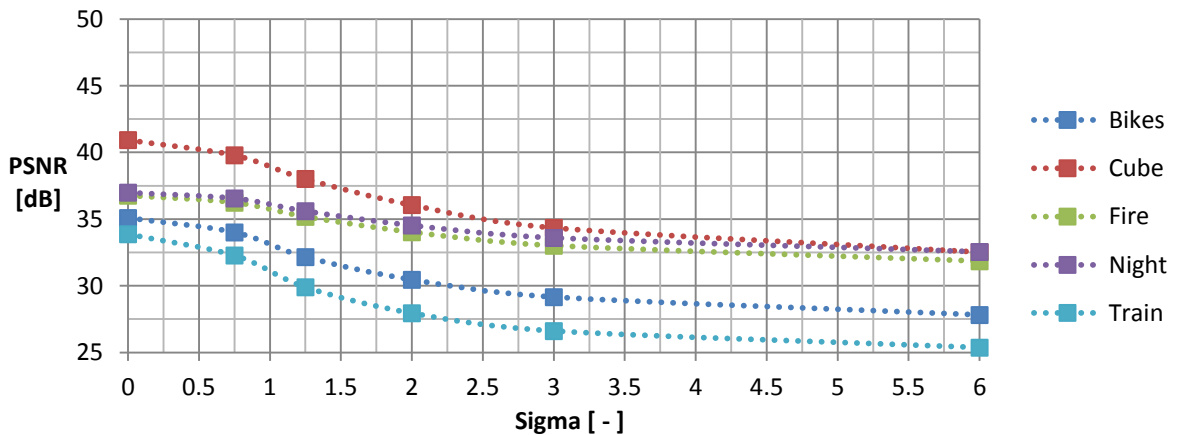
Grafy znázorňující výsledky objektivní metriky PSNR v závislosti na intenzitě rozmazání (parametr sigma) jednotlivých záběrů po aplikaci komprese ve třech úrovních (Q-10, Q-20 a Q-30).



Obr. 42. Výpočet PSNR v závislosti na parametru sigma s vlivem komprese (Q-10)



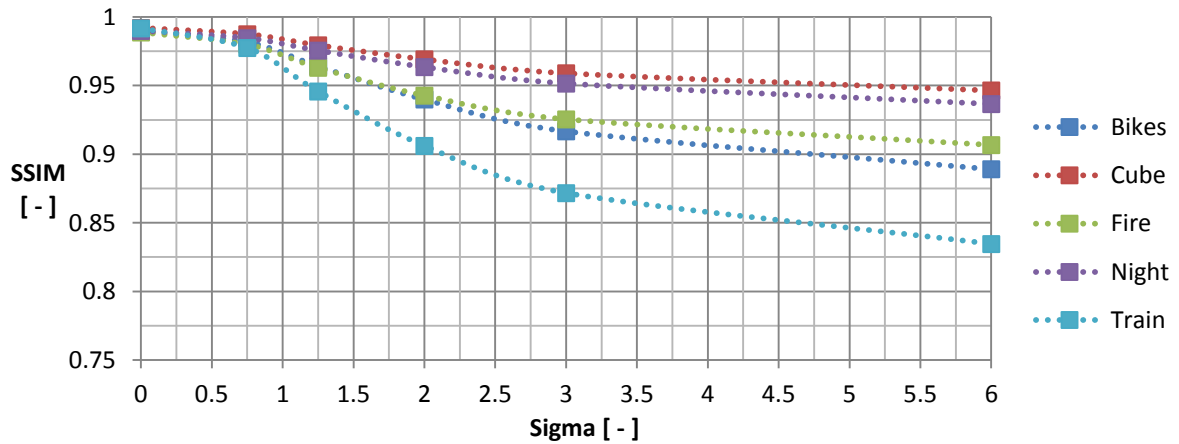
Obr. 43. Výpočet PSNR v závislosti na parametru sigma s vlivem komprese (Q-20)



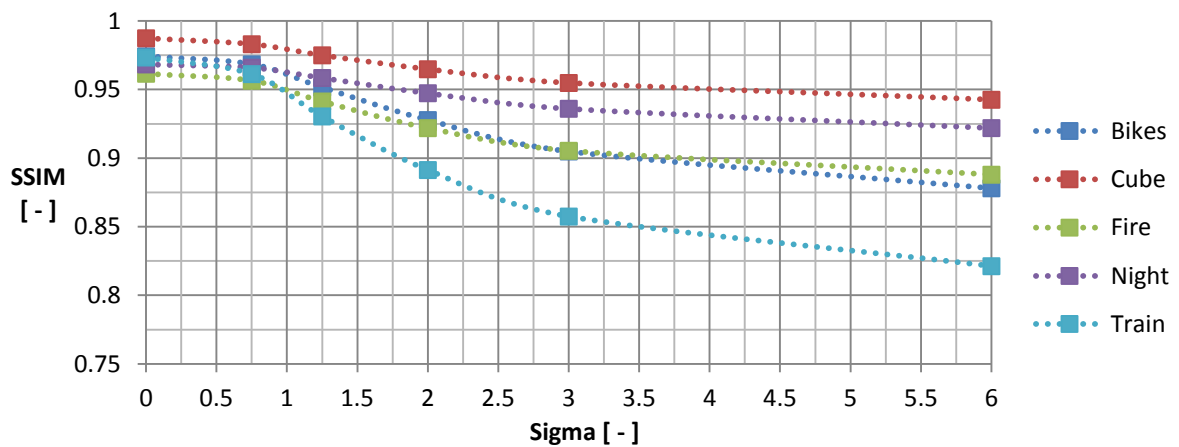
Obr. 44. Výpočet PSNR v závislosti na parametru sigma s vlivem komprese (Q-30)

Výsledky vlivu komprese - metrika SSIM

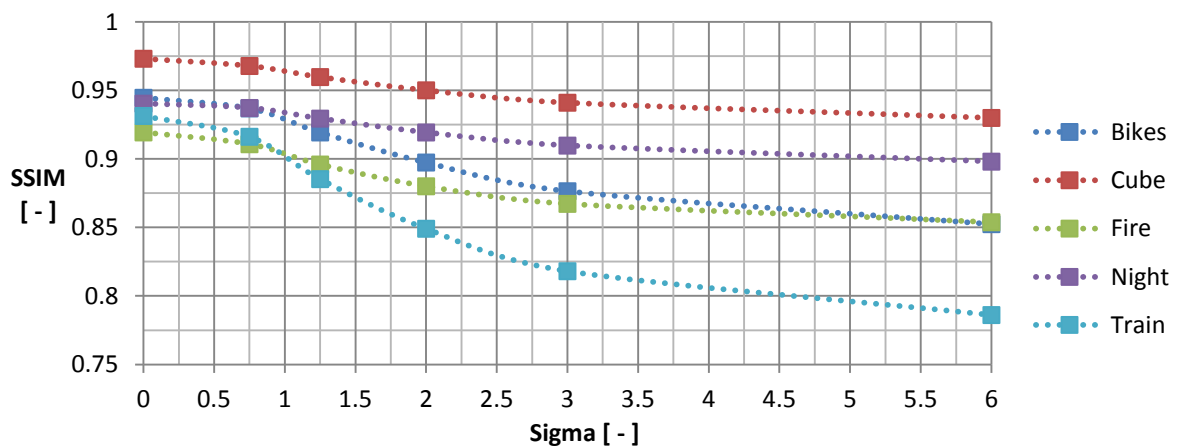
Grafy znázorňující výsledky objektivní metriky SSIM v závislosti na intenzitě rozmazání (parametr sigma) jednotlivých záběrů po aplikaci komprese ve třech úrovních (Q-10, Q-20 a Q-30).



Obr. 45. Výpočet SSIM v závislosti na parametru sigma s vlivem komprese (Q-10)



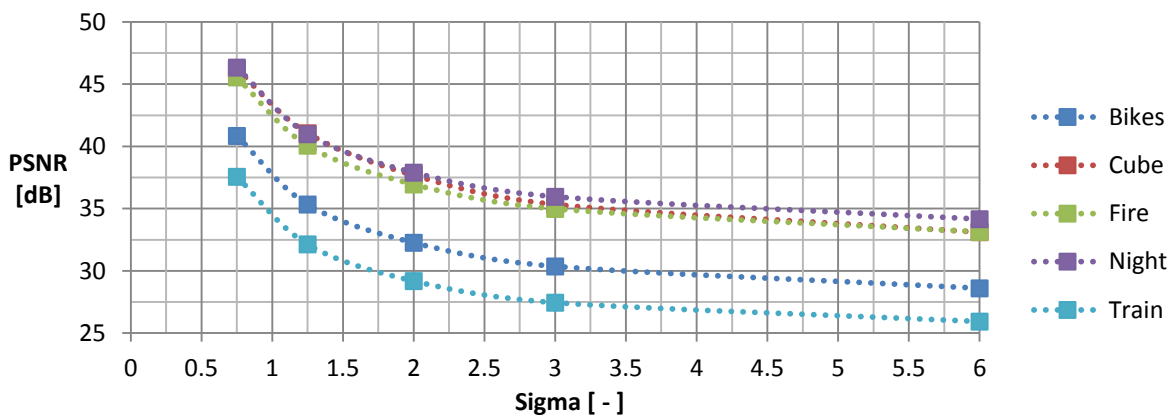
Obr. 46. Výpočet SSIM v závislosti na parametru sigma s vlivem komprese (Q-20)



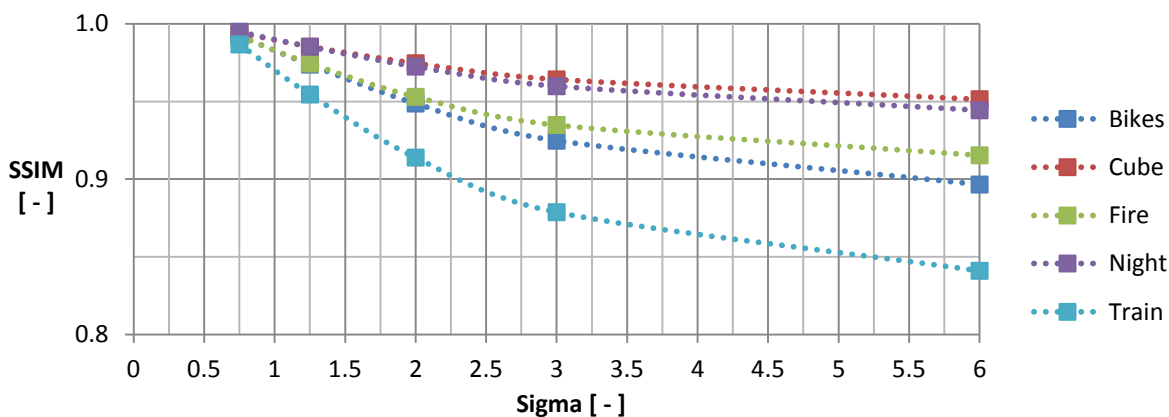
Obr. 47. Výpočet SSIM v závislosti na parametru sigma s vlivem komprese (Q-30)

Výsledky vlivu předzpracování - metrika PSNR i SSIM

Grafy znázorňující výsledky obou objektivních metrik (PSNR, SSIM) v závislosti na intenzitě rozmazání (parametr sigma) jednotlivých záběrů pouze po vlivu předzpracování.



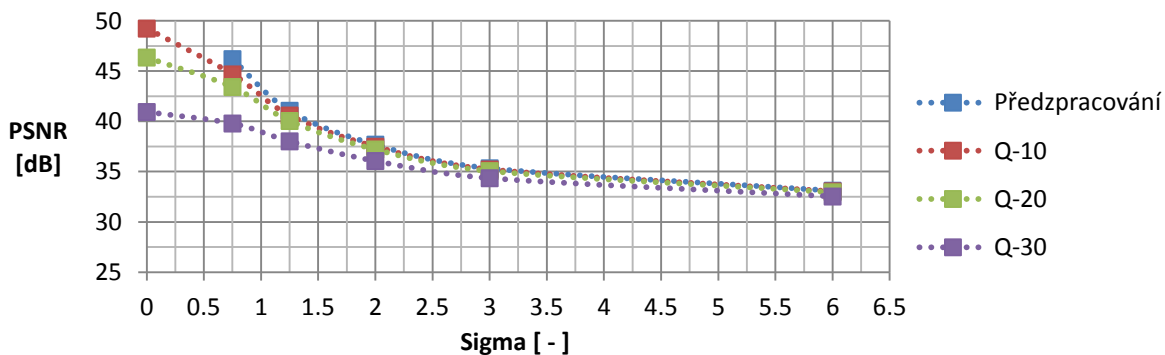
Obr. 48. Výpočet PSNR v závislosti na parametru sigma s vlivem předzpracování



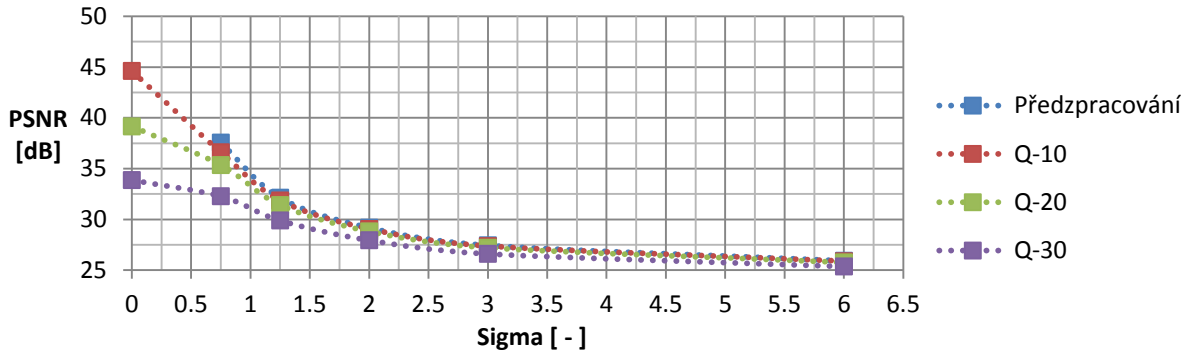
Obr. 49. Výpočet SSIM v závislosti na parametru sigma s vlivem předzpracování

Výsledky vlivu předzpracování a komprese - metrika PSNR i SSIM

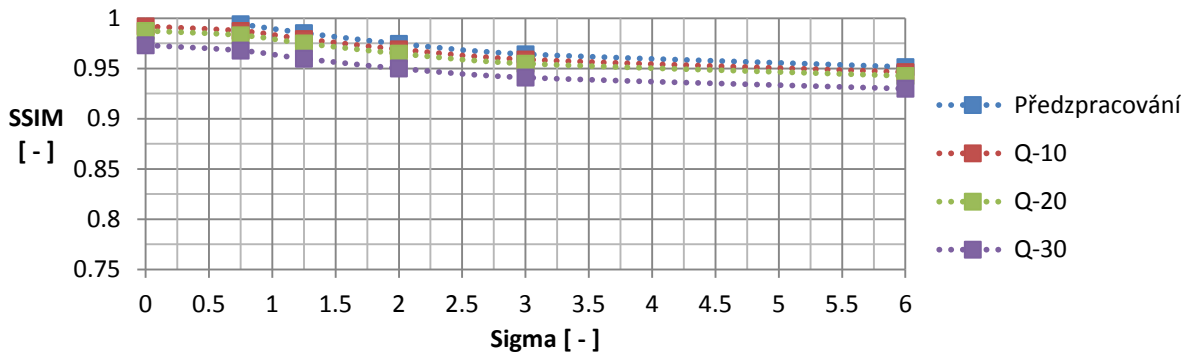
Porovnání výsledků vlivu předzpracování a komprese pro záběr Cube a Train.



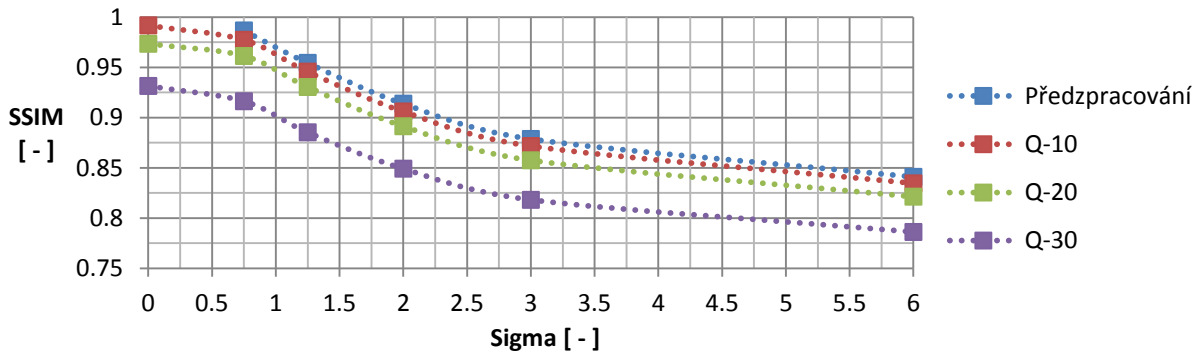
Obr. 50. Porovnání výpočtu PSNR s vlivem předzpracování a komprese záběru Cube



Obr. 51. Porovnání výpočtu SSIM s vlivem předzpracování a komprese záběru Cube



Obr. 52. Porovnání výpočtu PSNR s vlivem předzpracování a komprese záběru Train



Obr. 53. Porovnání výpočtu SSIM s vlivem předzpracování a komprese záběru Train

Shrnutí výsledků

Z výsledků objektivních metrik podle předpokladů vychází, že zvýšením intenzity rozostření dochází k větším odlišnostem mezi referenční a porovnávanou videosekvencí. Podobný vliv má i použití komprese, kdy zvyšováním kvantizačního parametru dochází k celkovému zhoršení kvality obrazu, a tudíž i samotných výsledků. Zajímavé jsou odlišnosti mezi jednotlivými záběry, kdy například záběr Train s vyšším obsahem pohybu a detailů dopadl hůře nežli statická kostka Cube, nehledě na použitou metriku nebo míru komprese. Z toho lze vyvodit, že výsledky objektivních metrik jsou velmi závislé na obsahu videosekvence. Přesto má každý graf svůj charakteristický průběh křivky, kterému odpovídají jednotlivé výsledky. To je pravděpodobně dáno velmi podobným vlivem předzpracování na výslednou kvalitu obrazu.

9.4 Porovnání s výsledky subjektivního testování

Výsledné hodnoty objektivních metrik jsou, bez patřičného porovnání s výsledky subjektivního testování, takřka nic neříkající, proto je důležité mezi oběma způsoby najít jejich vztah a souvislost. První možností srovnání je s využitím korelačních koeficientů. Druhým a jednodušším způsobem je vynesení výsledků do grafů, které budou v závislosti subjektivního ohodnocení (MOS) na odpovídajících výsledcích výpočtů obou metrik.

9.4.1 Výpočet korelačních koeficientů

Pro podrobnější porovnání a zjištění míry korelace dvou náhodných veličin nám slouží korelační analýza. Na jejím výstupu se obvykle nachází bezrozměrný korelační koeficient, který popisuje míru závislosti vstupních dat. Ten může nabývat hodnot od -1 do 1, přičemž hodnota 0 znamená neexistenci korelační závislosti veličin, naopak pokud se hodnota korelačního koeficientu přibližuje směrem k okrajům, začíná se jednat o její lineární závislost. Kladný korelační koeficient pak značí přímou korelaci mezi veličinami a záporná hodnota koeficientu zase nepřímou. Mezi nejčastěji používané korelační koeficienty patří Pearsonův a Spearmanův [8], které dále využijeme pro určení korelace mezi subjektivním testováním a objektivními metrikami. Koeficienty jsou spočítány pro obě metriky vzhledem k hodnoceným parametrům, typům záběrů a užití komprese (Q-20) po předzpracování.

Pearsonův korelační koeficient

Prvním spočítaným korelačním koeficientem je Pearsonův, který je daný vztahem (11).

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (11)$$

Tab. 3. Pearsonův korelační koeficient: závislost subjektivního ohodnocení všech parametrů s vlivem předzpracování a komprese na výpočtu objektivních metrik pro jednotlivé typy záběrů

| Pearson | | MOS | | | | | |
|---------|------------------|-----------------|--------|-----------------|--------|----------------|--------|
| | | Celková kvalita | | Celková ostrost | | Hloubka obrazu | |
| Metrika | Záběr / Komprese | bez | s | bez | s | bez | s |
| PSNR | Bikes | -0,913 | -0,961 | -0,949 | -0,977 | -0,375 | -0,874 |
| | Cube | -0,890 | -0,933 | -0,856 | -0,949 | -0,040 | -0,797 |
| | Fire | -0,884 | -0,949 | -0,892 | -0,984 | -0,853 | -0,523 |
| | Night | -0,906 | -0,984 | -0,975 | -0,958 | -0,948 | -0,893 |
| | Train | -0,859 | -0,910 | -0,931 | -0,852 | -0,787 | -0,694 |
| SSIM | Bikes | -0,990 | -0,993 | -0,980 | -0,979 | -0,731 | -0,813 |
| | Cube | -0,972 | -0,961 | -0,961 | -0,976 | -0,357 | -0,716 |
| | Fire | -0,957 | -0,975 | -0,958 | -0,988 | -0,646 | -0,577 |
| | Night | -0,957 | -0,981 | -0,951 | -0,915 | -0,963 | -0,819 |
| | Train | -0,961 | -0,988 | -0,997 | -0,971 | -0,900 | -0,801 |

Tab. 4. Pearsonův korelační koeficient: závislost subjektivního ohodnocení všech parametrů s vlivem předzpracování a komprese na výpočtu objektivních metrik pro všechny typy záběrů

| Pearson | Celková kvalita | | Celková ostrost | | Hloubka obrazu | |
|---------|-----------------|--------|-----------------|--------|----------------|--------|
| | bez | s | bez | s | bez | s |
| PSNR | -0,687 | -0,803 | -0,676 | -0,710 | -0,594 | -0,494 |
| SSIM | -0,786 | -0,773 | -0,752 | -0,691 | -0,694 | -0,452 |

Spearmanův korelační koeficient

Druhým použitým korelačním koeficientem je Spearmanův. Ten je oproti předchozímu jednodušší na výpočet (12) a vzhledem k faktu, že využívá při výpočtu pouze pořadí hodnot sledovaných veličin, je vhodný pro popis jakékoliv závislosti (lineární i nelineární).

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^n D_i^2}{n(n^2 - 1)} \quad (12)$$

Kde D_i je rozdíl mezi pořadím hodnot x_i a y_i příslušných korelačních dvojic a n je celkový počet korelačních dvojic.

Tab. 5. Spearmanův korelační koeficient: závislost subjektivního ohodnocení všech parametrů s vlivem předzpracování a komprese na výpočtu objektivních metrik pro jednotlivé typy záběrů

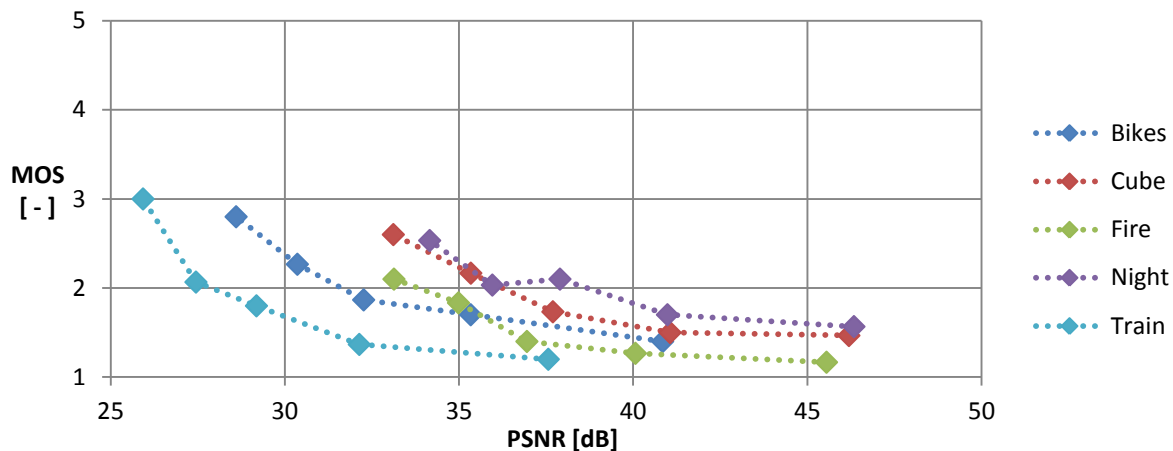
| Spearman | | MOS | | | | | |
|----------|------------------|-----------------|--------|-----------------|--------|----------------|--------|
| | | Celková kvalita | | Celková ostrost | | Hloubka obrazu | |
| Metrika | Záběr / Komprese | bez | s | bez | s | bez | s |
| PSNR | Bikes | -1 | -1 | -1 | -1 | -0,771 | -0,771 |
| | Cube | -1 | -0,943 | -0,943 | -0,957 | -0,243 | -0,900 |
| | Fire | -0,957 | -0,943 | -0,957 | -1 | -0,300 | -0,443 |
| | Night | -0,943 | -1 | -1 | -0,943 | -0,871 | -0,714 |
| | Train | -1 | -1 | -1 | -0,943 | -0,943 | -0,714 |
| SSIM | Bikes | -1 | -1 | -1 | -1 | -0,771 | -0,771 |
| | Cube | -1 | -0,943 | -0,943 | -0,957 | -0,243 | -0,900 |
| | Fire | -0,957 | -0,943 | -0,957 | -1 | -0,300 | -0,443 |
| | Night | -0,943 | -1 | -1 | -0,943 | -0,871 | -0,714 |
| | Train | -1 | -1 | -1 | -0,943 | -0,943 | -0,714 |

Tab. 6. Spearmanův korelační koeficient: závislost subjektivního ohodnocení všech parametrů s vlivem předzpracování a komprese na výpočtu objektivních metrik pro všechny typy záběrů

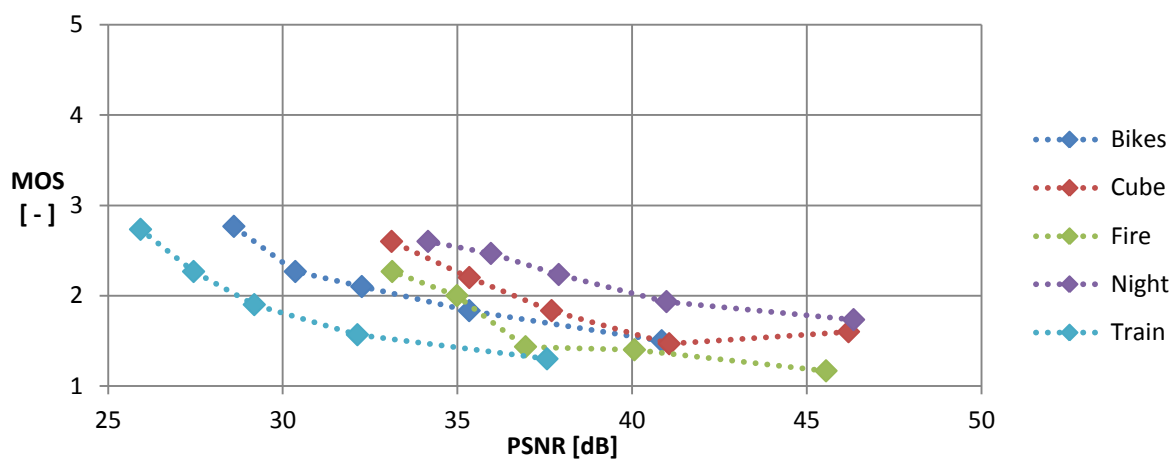
| Spearman | Celková kvalita | | Celková ostrost | | Hloubka obrazu | |
|----------|-----------------|--------|-----------------|--------|----------------|--------|
| | bez | s | bez | s | bez | s |
| PSNR | -0,785 | -0,825 | -0,785 | -0,730 | -0,610 | -0,495 |
| SSIM | -0,828 | -0,771 | -0,820 | -0,725 | -0,613 | -0,490 |

9.4.2 Grafické porovnání

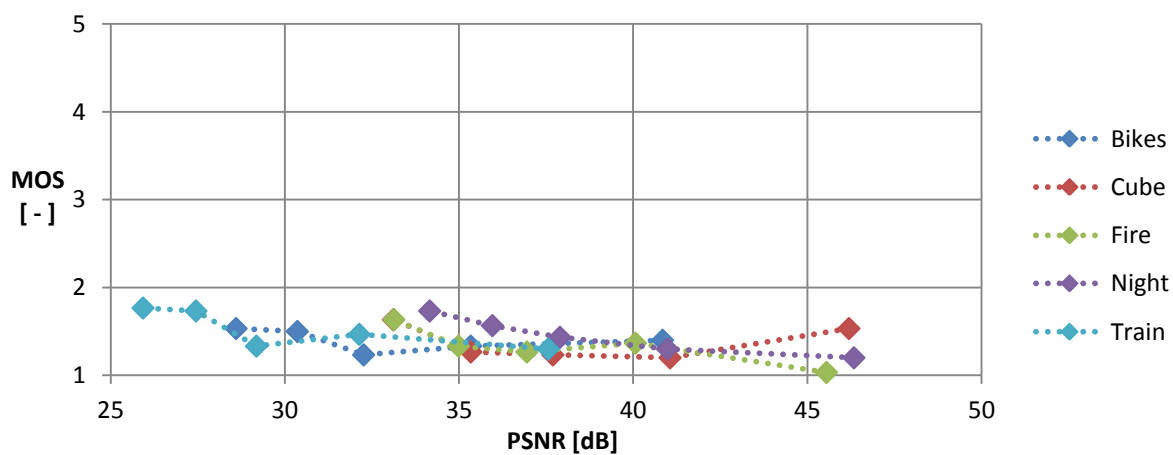
Vliv předzpracování - metrika PSNR



Obr. 54. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky PSNR

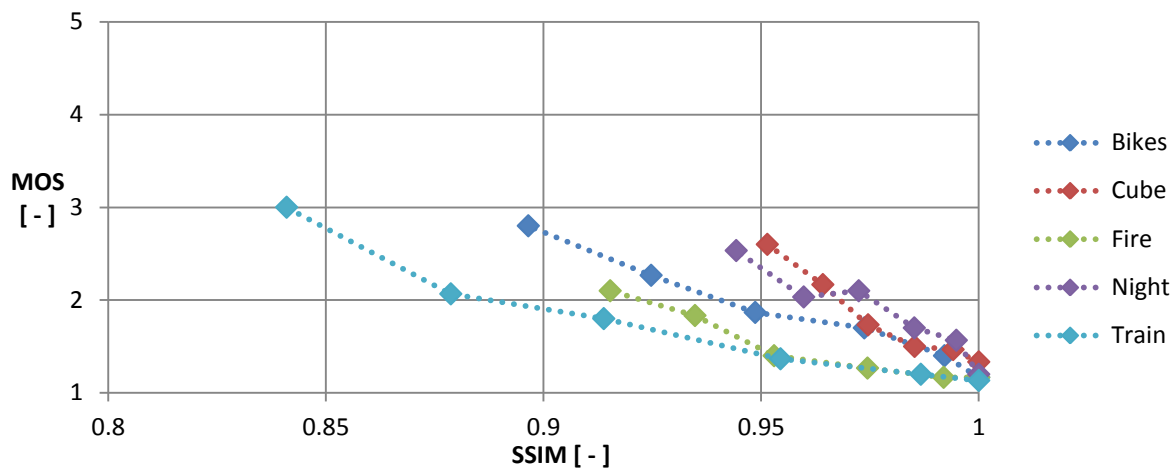


Obr. 55. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky PSNR

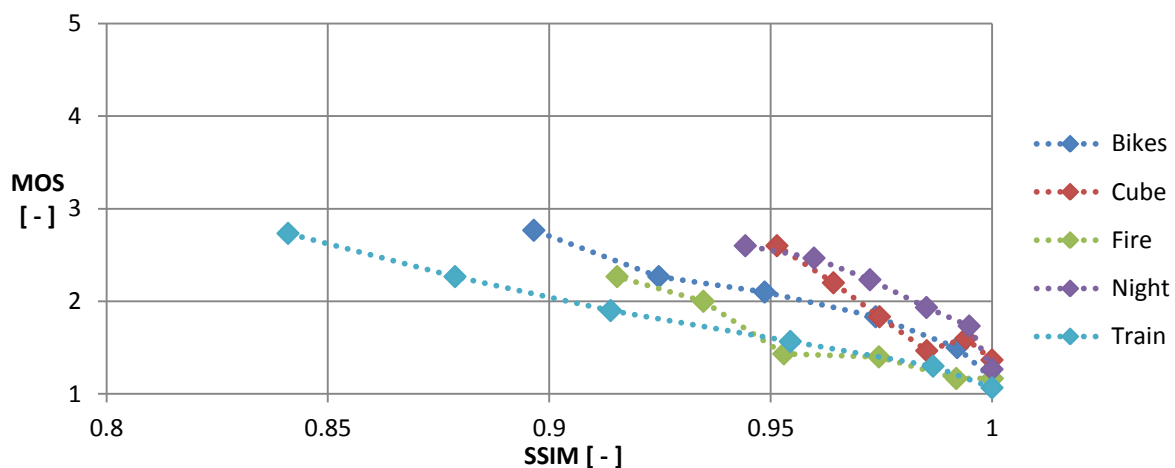


Obr. 56. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky PSNR

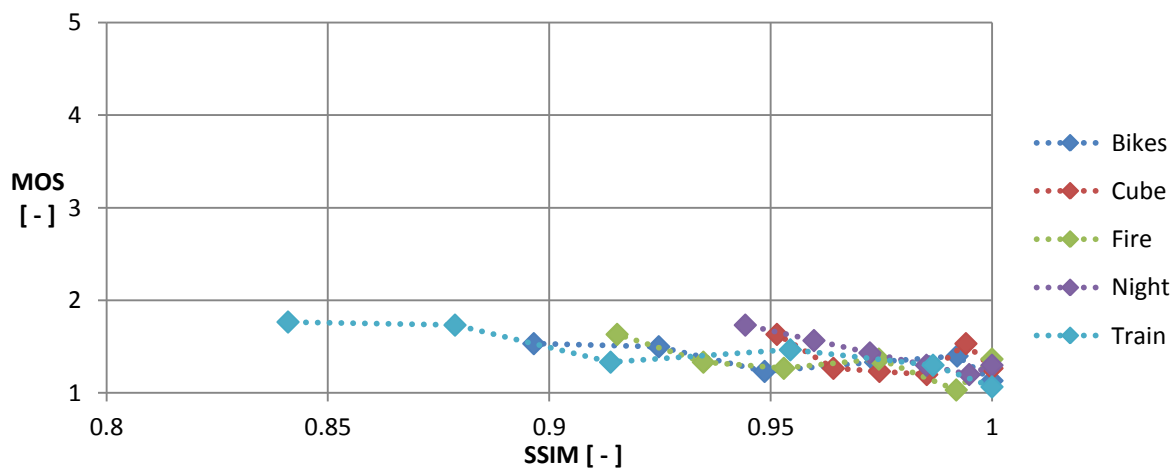
Vliv předzpracování - metrika SSIM



Obr. 57. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky SSIM

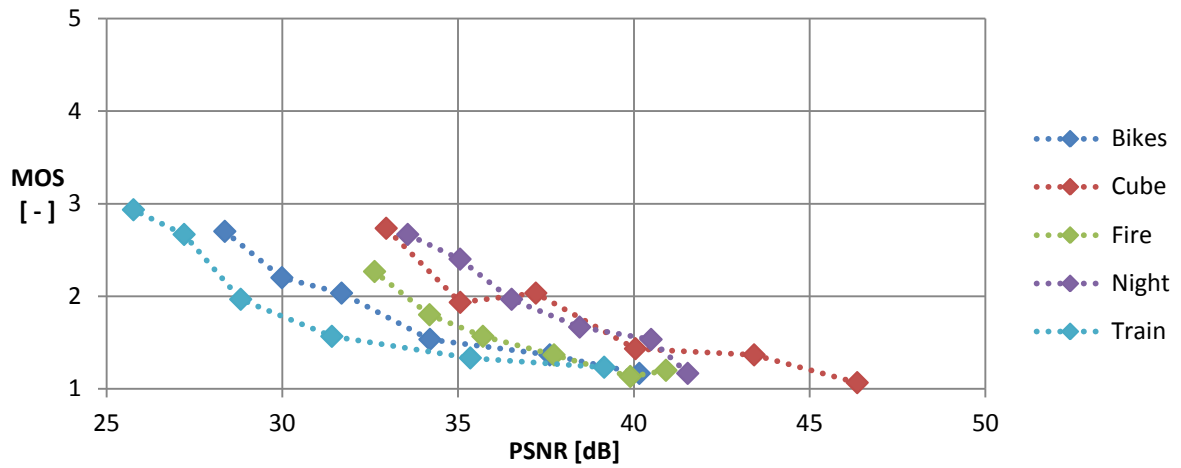


Obr. 58. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky SSIM

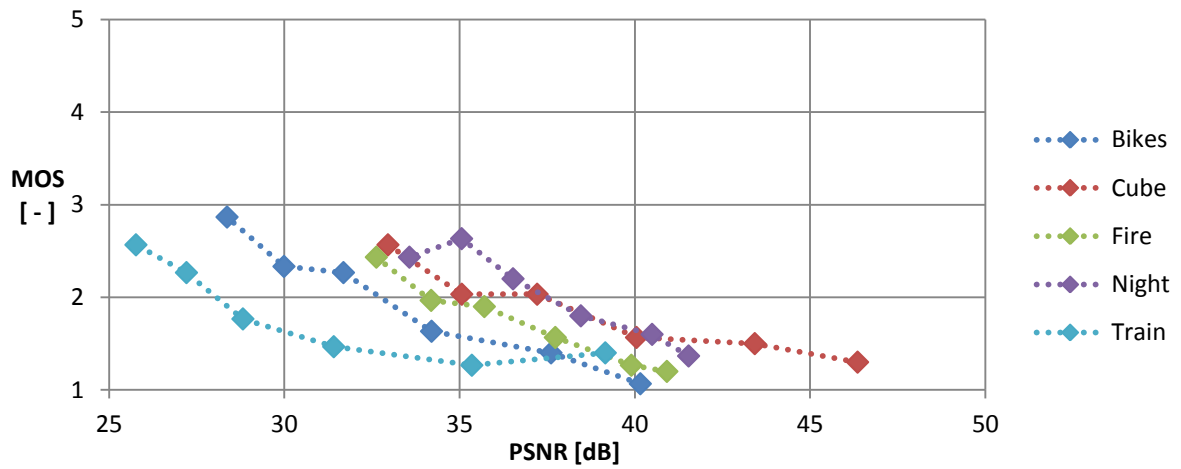


Obr. 59. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky SSIM

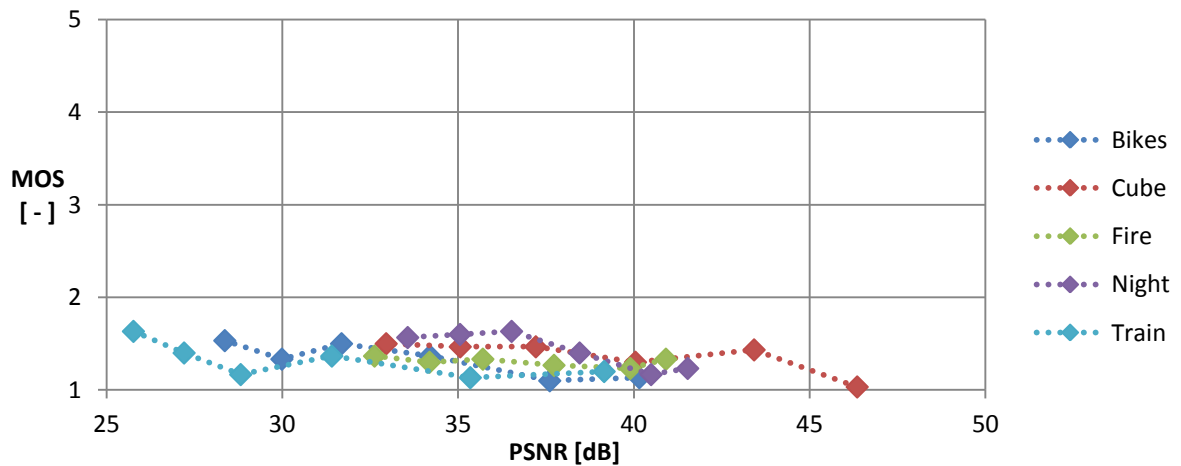
Vliv komprese (Q-20) - metrika PSNR



Obr. 60. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky PSNR

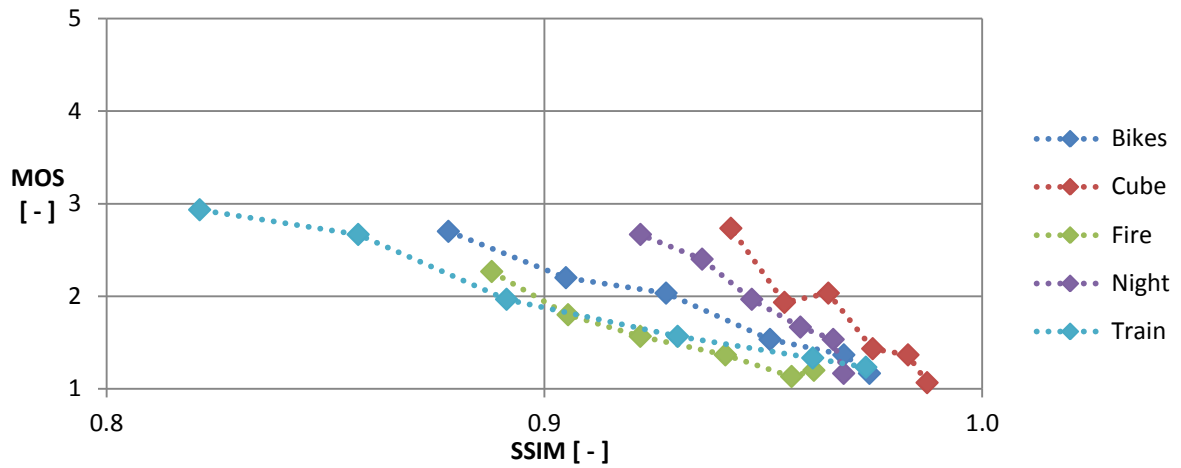


Obr. 61. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky PSNR

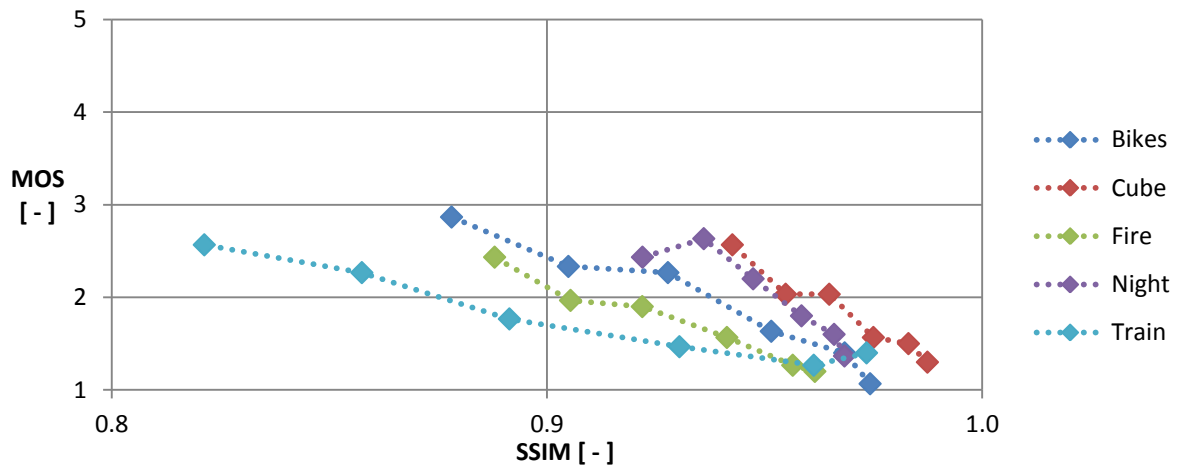


Obr. 62. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky PSNR

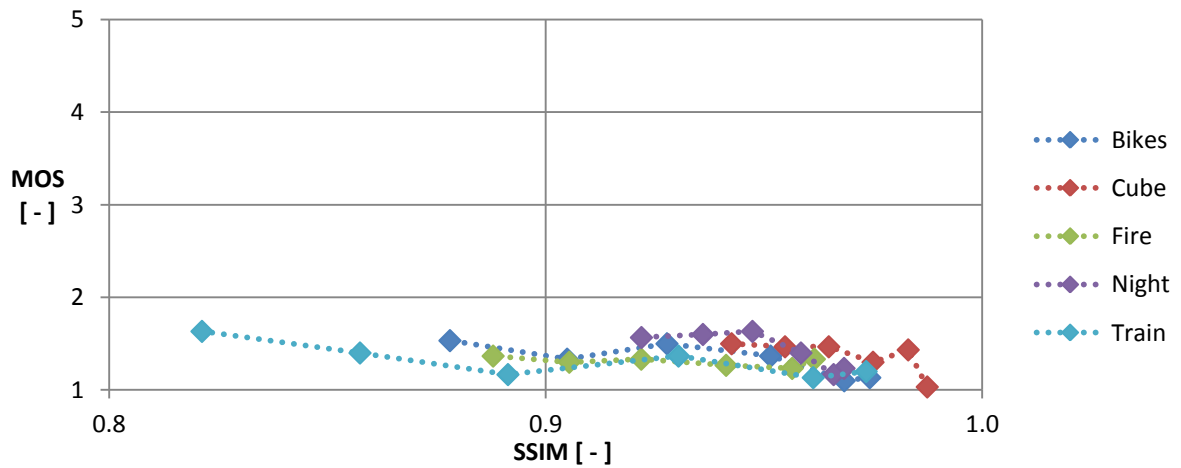
Vliv komprese (Q-20) - metrika SSIM



Obr. 63. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky SSIM



Obr. 64. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky SSIM



Obr. 65. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky SSIM

Porovnání subjektivního hodnocení a objektivních metrik

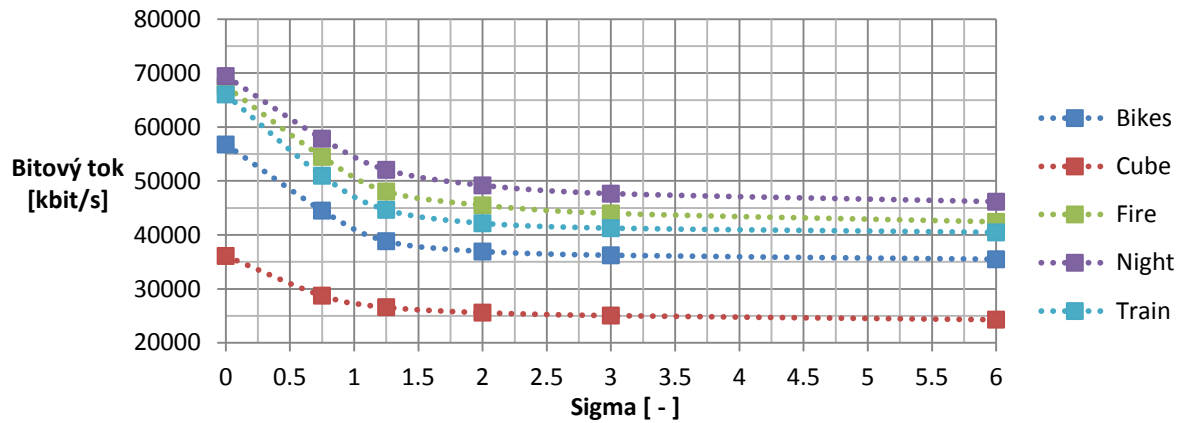
Z výpočtů korelačních koeficientů si můžeme povšimnout, že výsledné hodnoty vycházejí vždy záporně, to je dáno použitím opačného známkování, kdy ohodnocení 1 vypovídalo o nezkresleném parametru v obraze.

Z tabulky pro výpočet Pearsonova korelačního koeficientu je patrné, že metrika SSIM, ve valné většině případů, lépe koreluje s výsledky subjektivního hodnocení než jednodušší metrika PSNR. Výsledky koeficientu jsou ovšem pro obě metody velmi podobné, a pro subjektivní ohodnocení celkové kvality a ostrosti obrazu s nimi vždy vychází vysoká korelace. Naopak subjektivní hodnocení hloubky obrazu celkově nejméně koreluje s výsledky objektivních metrik, což může být dáno jejím nepravidelným ohodnocením, které se pohybovalo převážně mezi hodnotami jedna a dvě.

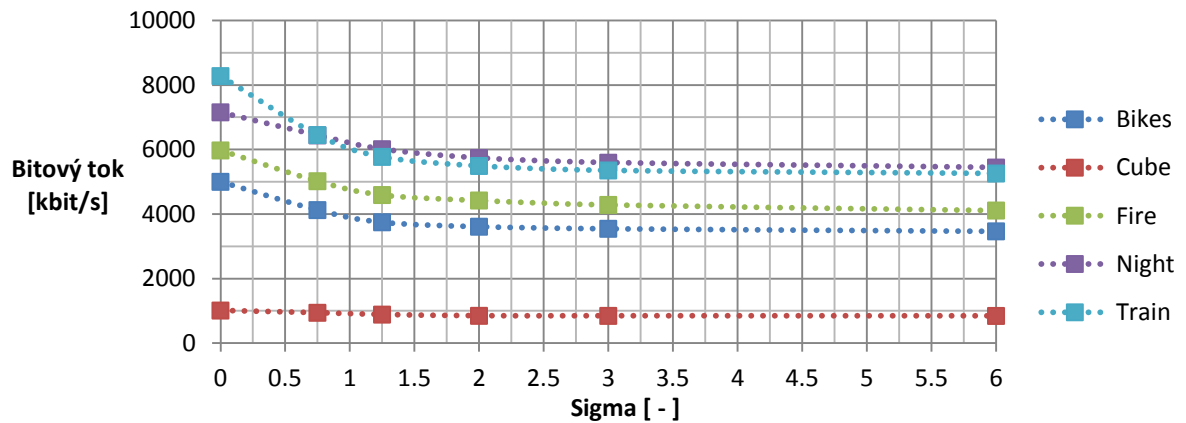
Co se týče Spearmanova korelačního koeficientu, tak ten je při srovnání subjektivního hodnocení celkové kvality a ostrosti prakticky nepoužitelný, protože dochází ke srovnání pouhých pěti či šesti hodnot, které jsou v případě obou veličin klesající. To má za následek, že většina výsledků pro hodnocení celkové kvality a ostrosti, nehledě na použité kompresi, vychází jako absolutní korelace mezi veličinami. Další nevýhodou Spearmanova koeficientu je nemožnost určení vhodnější objektivní metriky z důvodu stejných průběhů vstupních dat, a tudíž i výsledků.

10 Zhodnocení výsledků

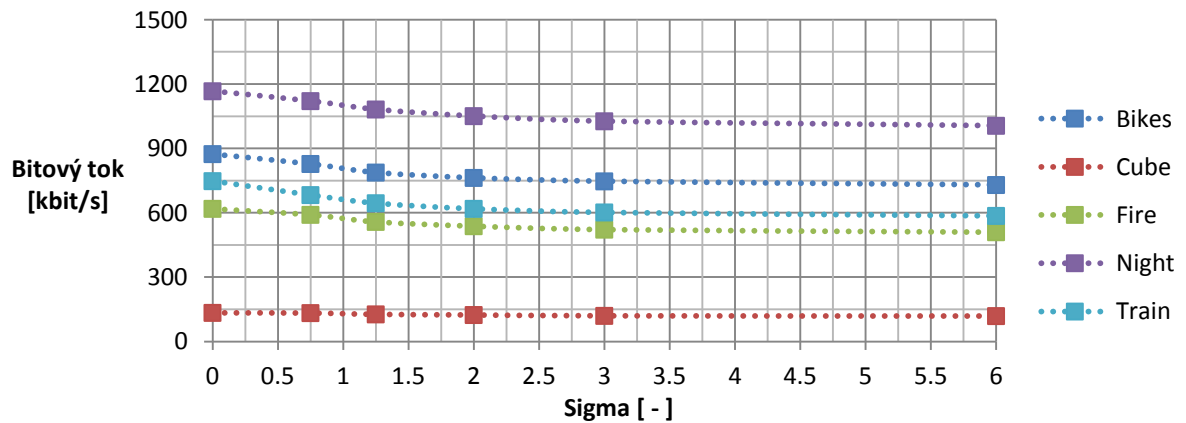
Pro zjištění účinnosti metody je důležité najít optimální hodnotu parametru sigma, které zaprvé odpovídá použitelné hodnotě z výsledků subjektivního testování, a zadruhé při něm dochází k podstatnější redukci datového toku.



Obr. 66. Závislost bitového toku na parametru sigma po kompresi videí (Q-10)



Obr. 67. Závislost bitového toku na parametru sigma po kompresi videí (Q-20)



Obr. 68. Závislost bitového toku na parametru sigma po kompresi videí (Q-30)

Tab. 7. Výsledné bitové toky a jejich datová úspora po kompresi videosekvencí ve všech třech úrovních

| Úroveň komprese | | Q-10 | | Q-20 | | Q-30 | |
|-----------------|-------|---------------------|------------|---------------------|------------|---------------------|------------|
| Záběr | Sigma | Bitový tok [kbit/s] | Úspora [%] | Bitový tok [kbit/s] | Úspora [%] | Bitový tok [kbit/s] | Úspora [%] |
| Bikes | 0 | 56768 | - | 5001 | - | 874 | - |
| | 0,75 | 44518 | 21,6 | 4125 | 17,5 | 828 | 5,3 |
| | 1,25 | 38838 | 31,6 | 3748 | 25,1 | 788 | 9,8 |
| | 2 | 36929 | 34,9 | 3611 | 27,8 | 763 | 12,7 |
| | 3 | 36227 | 36,2 | 3546 | 29,1 | 747 | 14,5 |
| | 6 | 35489 | 37,5 | 3467 | 30,7 | 730 | 16,5 |
| Cube | 0 | 36108 | - | 1013 | - | 134 | - |
| | 0,75 | 28750 | 20,4 | 943 | 6,9 | 133 | 0,7 |
| | 1,25 | 26610 | 26,3 | 887 | 12,4 | 127 | 5,2 |
| | 2 | 25605 | 29,1 | 849 | 16,2 | 124 | 7,5 |
| | 3 | 25059 | 30,6 | 847 | 16,4 | 120 | 10,4 |
| | 6 | 24303 | 32,7 | 846 | 16,5 | 119 | 11,2 |
| Fire | 0 | 67522 | - | 5974 | - | 619 | - |
| | 0,75 | 54480 | 19,3 | 5020 | 16 | 591 | 4,5 |
| | 1,25 | 48051 | 28,8 | 4592 | 23,1 | 558 | 9,9 |
| | 2 | 45491 | 32,6 | 4423 | 26 | 538 | 13,1 |
| | 3 | 43985 | 34,9 | 4286 | 28,3 | 522 | 15,7 |
| | 6 | 42425 | 37,2 | 4112 | 31,2 | 510 | 17,6 |
| Night | 0 | 69456 | - | 7155 | - | 1167 | - |
| | 0,75 | 57856 | 16,7 | 6441 | 10 | 1121 | 3,9 |
| | 1,25 | 52075 | 25,0 | 6013 | 16 | 1082 | 7,3 |
| | 2 | 49189 | 29,2 | 5739 | 19,8 | 1051 | 9,9 |
| | 3 | 47658 | 31,4 | 5594 | 21,8 | 1027 | 12 |
| | 6 | 46162 | 33,5 | 5443 | 23,9 | 1006 | 13,8 |
| Train | 0 | 66057 | - | 8278 | - | 748 | - |
| | 0,75 | 50983 | 22,8 | 6450 | 22,1 | 683 | 8,7 |
| | 1,25 | 44658 | 32,4 | 5774 | 30,2 | 645 | 13,8 |
| | 2 | 42162 | 36,2 | 5493 | 33,6 | 619 | 17,2 |
| | 3 | 41259 | 37,5 | 5353 | 35,3 | 602 | 19,5 |
| | 6 | 40476 | 38,7 | 5258 | 36,5 | 586 | 21,7 |

Z průběhu grafů a z tabulky pro srovnání bitových toků, lze vidět, že nejvýraznější pokles bitového toku, a s tím i spojená jeho úspora, nastává zhruba do hodnoty parametru sigma 1,25 nehlédě na úroveň komprese. Pro vyšší hodnoty sigma jsou již rozdíly v bitovém toku podstatně menší, takže nedochází k již tak výrazné kompresi. To je dáno převážně nelinearitou intenzity rozostření se vzrůstajícím parametrem sigma.

Dále si lze povšimnout, že pokles bitového toku je velmi rozdílný podle typu záběru a použité úrovně komprese. Co se týče typu záběru, největší pokles nastává vždy pro záběr *Train* a nejmenší pro záběr *Cube*. To je dáno samotnými principy komprese, která si lépe poradí se statickým obsahem a jednoduchým pozadím. Záběr *Cube* má tedy již sám o sobě velmi malý bitový tok, což znamená, že aplikace předzpracování na něj nemá ve výsledku takový vliv. Podobně to funguje i pro samotné úrovně komprese, kdy se zvyšujícím se kvantizačním parametrem, a tedy zhoršující se výslednou kvalitou obrazu, klesá i vliv předzpracování na výsledný bitový tok.

Porovnání s výsledky subjektivních testů

Z výsledků subjektivních testů můžeme považovat vliv předzpracování do hodnoty parametru sigma 3 za použitelný, do této hodnoty by prakticky nemělo dojít ke zpozorování jakýchkoliv změn v obraze, ať už se jedná o celkovou kvalitu, ostrost či jeho hloubku. Z tabulky výsledných bitových toků jsme ovšem zjistili, že nejúčinnější komprese (Q-20) lze dosáhnout do parametru sigma 2, a proto by byla míra předzpracování o hodnotě sigma 3 zbytečně veliká.

Výsledný parametr míry předzpracování se tedy nachází někde mezi hodnotami sigma 1,25 a 2. V tomto rozmezí je vliv předzpracování prakticky nezpozorovatelný a úspora datového objemu podle obsahu videosekvence až 33,6 %.

11 Závěr

Cílem této práce bylo podání přehledu současných metod pro kompresi, respektive kódování stereoskopického videa. Dále navrhnout experimentální řešení pomocí nesymetrického kódování, které umožní posoudit vliv předzpracování a komprese na kvalitu stereoskopického videa.

Po stručném shrnutí základů vnímání, záznamu a zobrazování stereoskopického obsahu, důležitých pro pochopení celé problematiky byly v teoretické části popsány tři současné metody kódování 3D obsahu. Z jejich rozboru lze usoudit, že pro současnou infrastrukturu televizního vysílání je nejpříhodnější způsob kódování pomocí prostorové komprese. To by se ovšem mohlo změnit s příchodem UHD vysílání, které bude potřebovat pro svou činnost výměnu starých dekodérů za nové na straně přijímače. Nové dekodéry by měly být kompatibilní s MVC způsobem kódování, což by mohlo vést k jeho budoucímu využití pro vysílání ve 3D. Ovšem problémem stále zůstává samotná poptávka po 3DTV vysílání, které se nezdá být pro diváky příliš atraktivní, a tak se momentálně vývoj 3D technologií soustředí na oblast mobilních aplikací a internetového streamingu.

Ve druhé části teoretické přípravy jsme se zaměřili na metody nesymetrického kódování a jejich související publikované práce, ze kterých plyne, že předzpracování pomocí rozostření obrazu patří mezi nejefektivnější metody nesymetrického kódování. Otázkou tedy zůstává, jaký konkrétní způsob aplikace rozostření je nejvhodnější a současně nejefektivnější. V této práci jsme popsali dosavadní použité způsoby včetně toho, kterým jsme se nechali inspirovat do praktické části práce.

Praktická část práce je věnována realizaci konkrétní metody nesymetrického kódování a jejímu otestování pomocí subjektivního hodnocení kvality obrazu a dvou objektivních metrik. V rámci přípravy a zpracování videosekvencí bylo vytvořeno GUI v Matlabu (ACT), které může posloužit nejen za účelem aplikace nesymetrického kódování, ale i jako vhodný nástroj pro archivaci či úpravu hloubky obrazu stereoskopického videa. Druhé GUI (OMT) bylo vytvořeno na základě práce s objektivními metrikami PSNR a SSIM, které se počítají prostřednictvím softwaru VQMT.

Ze subjektivního hodnocení kvality obrazu vyšlo, že použitá metoda nesymetrického kódování má minimální dopad na reprodukováný obraz, a prakticky ji lze použít v celém testovaném rozsahu. Se zvyšující se intenzitou rozostření se celková kvalita a ostrost mírně zhoršovaly, což se nedá říct o hloubce obrazu, jež zůstala prakticky beze změn. Po porovnání s výslednými bitovými toky jsme ovšem dospěli k závěru, že metoda je efektivní pouze do hodnoty parametru sigma 2.

Přestože metoda pruhů lépe eliminuje problém s oční dominancí člověka, tak výsledek není stoprocentní. Jednoduše by se mohlo stát, že pruhy v jednom z obrazů se treťi vždy do zajímavých míst scény, což znamená, že se v podstatě vracíme ke stejnému problému metody, kde dochází k rozostření celého obrazu stereopáru. Proto by pokračováním této práce mohl být například inteligentní systém pro rozostření neatraktivních oblastí v obraze. Další možností by mohl být systém, který by reagoval na paralaxu očí a následně by rozostřoval místa, kde je méně znatelná hloubka obrazu (pozitivní paralaxa), což by se v nejhroším případě mohlo jevit jako zvýšená hloubka ostrosti. Možností pokračování práce ohledně stereoskopického videa je opravdu mnoho, otázkou ovšem zůstává, která oblast by byla pro jeho současný rozvoj nejpříhodnější.

12 Použité zdroje a literatura

- [1] AFLAKI, Payman, Miska M. HANNUKSELA a Moncef GABBOUJ. Subjective quality assessment of asymmetric stereoscopic 3D video. *Signal, Image and Video Processing* [online]. 2013, vol. 9, issue 2, s. 331-345
- [2] AFLAKI, Payman, Miska M. HANNUKSELA, Jussi HAKALA, Jukka HAKKINEN a Moncef GABBOUJ. Estimation of subjective quality for mixed-resolution stereoscopic video. In: *2011 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)* [online]. 2011, s. 1-4
- [3] AZIMI, Maryam, Sima VALIZADEH, Xiaokang LI, Lino E. CORIA a Panos NASIOPOULOS. Subjective study on asymmetric stereoscopic video with low-pass filtered slices. In: *2012 International Conference on Computing, Networking and Communications (ICNC)* [online]. 2012, s. 719-723
- [4] BANITALEBI-DEHKORDI, Amin, Mahsa T. POURAZAD a Panos NASIOPOULOS. An efficient human visual system based quality metric for 3D video. *Multimedia Tools and Applications* [online]. 2015.
- [5] BANITALEBI-DEHKORDI, Amin, Mahsa T. POURAZAD a Panos NASIOPOULOS. Effect of eye dominance on the perception of stereoscopic 3D video. In: *2014 IEEE International Conference on Image Processing (ICIP)*[online]. 2014, s. 3469-3473
- [6] BENOIT, Alexandre, Patrick Le CALLET, Patrizio CAMPISI a Romain COUSSEAU. Using disparity for quality assessment of stereoscopic images. In: *2008 15th IEEE International Conference on Image Processing* [online]. 2008.
- [7] BindiChen [online]. [vid. 16.05.2015]. Dostupné z: <http://www.bindichen.co.uk/post/Fundamentals/bell-shaped-function.html>
- [8] CHOK, Nian Shong. *Pearson's Versus Spearman's and Kendall's Correlation Coefficients for Continuous Data*. Master's Thesis. University of Pittsburgh, 2010.
- [9] CAMPISI, Patrizio, P. L. CALLET a E. MARINI. Stereoscopic images quality assessment. Proceedings of 15th European Signal Processing Conference (EUSIPCO'0). 2007.
- [10] DigiZone [online]. [vid. 12.05.2015]. Dostupné z: <http://www.digizone.cz/clanky/dohlubin-formatu-mpeg-2/>
- [11] Digimanie [online]. [vid. 16.05.2015]. Dostupné z: <http://www.digimanie.cz/vse-o-formatu-raw-1dil/2182>
- [12] EZR8 [online]. [vid. 15.05.2015]. Dostupné z: <http://www.ezr8.com/avi.html#2>
- [13] FileFormat.info [online]. [vid. 22.04.2015]. Dostupné z: <http://www.fileformat.info/format/tiff/egff.htm>
- [14] Fyzmatik [online]. [vid. 10.05.2015]. Dostupné z: <http://fyzmatik.pise.cz/270-optotypy-vysetreni-ostrosti-zraku.html>

- [15] HANUS, Rostislav. *Binokulární vidění*. Brno, 2010. Bakalářská práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií.
- [16] HOŠEK, Lukáš. *Filtrace šumu z digitálních fotografií*. Praha, 2007. Bakalářská práce. Univerzita Karlova v Praze. Matematicko-fyzikální fakulta.
- [17] IJSELSTEIJN, W.A., H. de RIDDER a J. VLIEGEN. Subjective evaluation of stereoscopic images: effects of camera parameters and display duration. *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2000, vol. 10, issue 2, s. 225-233
- [18] ITU-R BT.2021-1. *Subjective methods for assessment of stereoscopic 3DTV systems*. Geneva: International Telecommunication Union, 2015.
- [19] ITU-R BT.500-13. *Methodology for the subjective assessment of the quality of television pictures*. Geneva: International Telecommunication Union, 2012.
- [20] JIN, Lina, Atanas BOEV, Atanas GOTCHEV a Karen EGIAZARIAN. Validation of a new full reference metric for quality assessment of mobile 3DTV content. In: *2011 19th European Signal Processing Conference*
- [21] JIN, Lina, Atanas BOEV, Atanas GOTCHEV a Karen EGIAZARIAN. 3D-DCT based perceptual quality assessment of stereo video. In: *2011 18th IEEE International Conference on Image Processing* [online].
- [22] KAHL, Jan. *Snímání a reprodukce stereoskopického obrazu*. Praha, 2012. Bakalářská práce. České vysoké učení technické v Praze. Fakulta elektrotechnická.
- [23] KIM, Donghyun, Seungchul RYU a Kwanghoon SOHN. Depth perception and motion cue based 3D video quality assessment. In: *2012 IEEE international Symposium on Broadband Multimedia Systems and Broadcasting* [online].
- [24] LUKES, T., K. FLIEGEL a M. KLIMA. Objective image quality assessment of multiframe super-resolution methods. In: *2013 23rd International Conference Radioelektronika (RADIOELEKTRONIKA)* [online]. 2013, s. 267-272
- [25] LEVELT, W. *On binocular rivalry*. Soesterberg, Netherlands: Institute for Perception RVO-TNO, National Defence Research Organization TNO. 1965.
- [26] MITSUHASHI, TETSUO. Evaluation of stereoscopic picture quality with CFF. *Ergonomics* [online]. 1996, vol. 39, issue 11, s. 1344-1356
- [27] RANDULOVÁ, Jana. *Akomodace a konvergence*. Brno, 2008. Bakalářská práce. Masarykova univerzita v Brně. Lékařská fakulta.
- [28] SEUNTIENS, Pieter, Lydia MEESTERS a Wijnand IJSELSTEIJN. Perceived quality of compressed stereoscopic images. *ACM Transactions on Applied Perception* [online]. 2006, vol. 3, issue 2, s. 95-109

- [29] SHAO, Feng, Gangyi JIANG, Xu WANG, Mei YU a Ken CHEN. Stereoscopic video coding with asymmetric luminance and chrominance qualities. *IEEE Transactions on Consumer Electronics* [online]. 2010, vol. 56, issue 4, s. 2460-2468
- [30] SMOLIC, Aljoscha, Karsten MUELLER, Philipp MERKLE, Peter KAUFF a Thomas WIEGAND. An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution. In: *2009 Picture Coding Symposium* [online]. 2009, s. 1-4
- [31] SPERANZA, Filippo, Ron RENAUD, Andre VINCENT a Wa J. TAM. Perceived Picture Quality of Frame-Compatible 3DTV Video Formats. In: *2012 IEEE International Conference on Multimedia and Expo* [online]. 2012, s. 640-645
- [32] STELMACH, L., Wa James TAM, D. MEEGAN a A. VINCENT. Stereo image quality: effects of mixed spatio-temporal resolution. *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2000, vol. 10, issue 2, s. 188-193
- [33] STELMACH, Lew B. a W. James TAM. Stereoscopic image coding: Effect of disparate image-quality in left- and right-eye views. *Signal Processing: Image Communication* [online]. 1998, vol. 14, 1-2, s. 111-117
- [34] TOMAN, Jan. *Systém pro zpracování a reprodukci stereoskopického videa*. Praha, 2015. Projekt II. České vysoké učení technické v Praze. Fakulta elektrotechnická.
- [35] Topendsports [online]. [vid. 11.05.2015]. Dostupné z: <http://www.topendsports.com/testing/tests/eye-dominance.htm>
- [36] VALIZADEH, Sima, Maryam AZIMI a Panos NASIOPOULOS. Bitrate reduction in asymmetric stereoscopic video with low-pass filtered slices. In: *2012 IEEE International Conference on Consumer Electronics (ICCE)* [online]. 2012, s. 170-171
- [37] VETRO, A, T. WIEGAND a G. J. SULLIVAN. Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard. *Proceedings of the IEEE* [online]. 2011, vol. 99, issue 4, s. 626-642
- [38] WANG, Z., A.C. BOVIK, H.R. SHEIKH a E.P. SIMONCELLI. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing* [online]. 2004, vol. 13, issue 4, s. 600-612
- [39] Wikipedia [online]. [vid. 12.05.2015]. Dostupné z: http://en.wikipedia.org/wiki/Ishihara_Test
- [40] XING, Liyuan, Junyong YOU, Touradj EBRAHIMI a Andrew PERKIS. Assessment of Stereoscopic Crosstalk Perception. *IEEE Transactions on Multimedia* [online]. 2012, vol. 14, issue 2, s. 326-337
- [41] YouTube [online]. [vid. 13.05.2015]. Dostupné z: <https://support.google.com/youtube/answer/2853702?hl=en>
- [42] 3D Forums [online]. [vid. 12.05.2015]. Dostupné z: <http://www.3d-forums.com/threads/stereoscopic-parallax.4/>

13 Seznam použitých obrázků

| | |
|--|----|
| Obr. 1. Typy paralax | 10 |
| Obr. 2. Ukázka snímku anaglyf..... | 11 |
| Obr. 3. Princip aktivní technologie | 12 |
| Obr. 4. Princip auto-stereoskopického displeje s paralaxní bariérou | 13 |
| Obr. 5. Formáty prostorové komprese | 16 |
| Obr. 6. Mezi-snímková/mezi-pohledová predikce struktury MVC | 17 |
| Obr. 7. Ukázka 2D snímku (vlevo) a jeho hloubkové mapy (vpravo). Převzato z^2 | 18 |
| Obr. 8. Ukázka principu smíšeného rozlišení | 21 |
| Obr. 9. Ukázka rozostření pomocí Gaussova filtru s velikostí okna 15 a parametrem sigma 5..... | 21 |
| Obr. 10. Ukázka různých poměrů podvzorkování barvonosných složek..... | 22 |
| Obr. 11. Ukázka změny kvality JPEG | 22 |
| Obr. 12. Ukázka Gaussovy křivky pro různé hodnoty parametru sigma | 23 |
| Obr. 13. Metoda pruhů..... | 24 |
| Obr. 14. Metody rozostření - celý obraz / pruhy / postupné pruhy | 24 |
| Obr. 15. Průběh intenzity rozostření (sigma=0,8) jednoho pruhu o šířce 108 px | 25 |
| Obr. 16. Průběh intenzity rozostření (sigma=3) jednoho pruhu o šířce 108 px | 26 |
| Obr. 17. Ukázka zdrojových snímků bez aplikace předzpracování | 26 |
| Obr. 18. Závislost datového objemu fotografií na úrovni rozostření | 26 |
| Obr. 19. Prostředí ACT | 28 |
| Obr. 20. Diagram zpracování RAW snímků | 30 |
| Obr. 21. Diagram zpracování převzatých záběrů..... | 30 |
| Obr. 22. Výsledky průzkumu subjektivního testování..... | 31 |
| Obr. 23. Uspořádání pracoviště pro subjektivní testy | 31 |
| Obr. 24. Blokové schéma postupu tvorby videosekvencí do testu..... | 33 |
| Obr. 25. Ukázka Snellenovy tabule (vlevo) [14] a Ishiharova obrazce (vpravo) [39] | 34 |
| Obr. 26. Celková kvalita: Subjektivní hodnocení v závislosti na parametru sigma..... | 36 |
| Obr. 27. Celková ostrost: Subjektivní hodnocení v závislosti na parametru sigma | 36 |
| Obr. 28. Hloubka obrazu: Subjektivní hodnocení v závislosti na parametru sigma | 36 |
| Obr. 29. Celková kvalita: Subjektivní hodnocení v závislosti na parametru sigma..... | 37 |
| Obr. 30. Celková ostrost: Subjektivní hodnocení v závislosti na parametru sigma | 37 |
| Obr. 31. Hloubka obrazu: Subjektivní hodnocení v závislosti na parametru sigma | 37 |
| Obr. 32. Srovnání průměrů subjektivního hodnocení vlivu předzpracování a komprese | 38 |
| Obr. 33. Blokový diagram výpočtu SSIM indexu..... | 40 |

| | |
|--|----|
| Obr. 34. Prostředí OMT | 41 |
| Obr. 35. Porovnání metriky PSNR jednotlivých pohledů stereopáru s formátem SBS záběru Fire . | 42 |
| Obr. 36. Porovnání metriky SSIM jednotlivých pohledů stereopáru s formátem SBS záběru Fire.. | 43 |
| Obr. 37. Porovnání metriky PSNR jednotlivých pohledů stereopáru s formátem SBS záběru Train | 43 |
| Obr. 38. Porovnání metriky SSIM jednotlivých pohledů stereopáru s formátem SBS záběru Train | 43 |
| Obr. 39. Průběh výpočtu PSNR a SSIM záběru Bikes..... | 44 |
| Obr. 40. Průběh výpočtu PSNR a SSIM záběru Cube | 44 |
| Obr. 41. Průběh výpočtu PSNR a SSIM záběru Night..... | 44 |
| Obr. 42. Výpočet PSNR v závislosti na parametru sigma s vlivem komprese (Q-10) | 45 |
| Obr. 43. Výpočet PSNR v závislosti na parametru sigma s vlivem komprese (Q-20) | 45 |
| Obr. 44. Výpočet PSNR v závislosti na parametru sigma s vlivem komprese (Q-30) | 45 |
| Obr. 45. Výpočet SSIM v závislosti na parametru sigma s vlivem komprese (Q-10) | 46 |
| Obr. 46. Výpočet SSIM v závislosti na parametru sigma s vlivem komprese (Q-20) | 46 |
| Obr. 47. Výpočet SSIM v závislosti na parametru sigma s vlivem komprese (Q-30) | 46 |
| Obr. 48. Výpočet PSNR v závislosti na parametru sigma s vlivem předzpracování | 47 |
| Obr. 49. Výpočet SSIM v závislosti na parametru sigma s vlivem předzpracování | 47 |
| Obr. 50. Porovnání výpočtu PSNR s vlivem předzpracování a komprese záběru Cube..... | 47 |
| Obr. 51. Porovnání výpočtu SSIM s vlivem předzpracování a komprese záběru Cube | 48 |
| Obr. 52. Porovnání výpočtu PSNR s vlivem předzpracování a komprese záběru Train..... | 48 |
| Obr. 53. Porovnání výpočtu SSIM s vlivem předzpracování a komprese záběru Train | 48 |
| Obr. 54. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky PSNR | 51 |
| Obr. 55. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky PSNR..... | 51 |
| Obr. 56. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky PSNR..... | 51 |
| Obr. 57. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky SSIM | 52 |
| Obr. 58. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky SSIM | 52 |
| Obr. 59. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky SSIM..... | 52 |
| Obr. 60. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky PSNR | 53 |
| Obr. 61. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky PSNR..... | 53 |
| Obr. 62. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky PSNR..... | 53 |
| Obr. 63. Celková kvalita: závislost subjektivního hodnocení na výpočtech metriky SSIM | 54 |
| Obr. 64. Celková ostrost: závislost subjektivního hodnocení na výpočtech metriky SSIM | 54 |
| Obr. 65. Hloubka obrazu: závislost subjektivního hodnocení na výpočtech metriky SSIM..... | 54 |
| Obr. 66. Závislost bitového toku na parametru sigma po kompresi videí (Q-10)..... | 56 |
| Obr. 67. Závislost bitového toku na parametru sigma po kompresi videí (Q-20)..... | 56 |
| Obr. 68. Závislost bitového toku na parametru sigma po kompresi videí (Q-30)..... | 56 |

14 Přílohy

14.1 Příloha A - formulář pro subjektivní testování

Jméno: _____ Příjmení: _____
 Věk: _____ Pohlaví: _____
 Dominantní oko: _____ Rozteč očí: _____ [cm]
 Problémy: zraková ostrost / barvocit / vnímáním hloubky obrazu
 Sledování 3D: aktivně / zřídka / vůbec

| Vzorek | Kvalita | Ostrost | Hloubka | Vzorek | Kvalita | Ostrost | Hloubka |
|--------|---------|---------|---------|--------|---------|---------|---------|
| Test 1 | | | | Test 2 | | | |
| 1 | | | | 31 | | | |
| 2 | | | | 32 | | | |
| 3 | | | | 33 | | | |
| 4 | | | | 34 | | | |
| 5 | | | | 35 | | | |
| 6 | | | | 36 | | | |
| 7 | | | | 37 | | | |
| 8 | | | | 38 | | | |
| 9 | | | | 39 | | | |
| 10 | | | | 40 | | | |
| 11 | | | | 41 | | | |
| 12 | | | | 42 | | | |
| 13 | | | | 43 | | | |
| 14 | | | | 44 | | | |
| 15 | | | | 45 | | | |
| 16 | | | | 46 | | | |
| 17 | | | | 47 | | | |
| 18 | | | | 48 | | | |
| 19 | | | | 49 | | | |
| 20 | | | | 50 | | | |
| 21 | | | | 51 | | | |
| 22 | | | | 52 | | | |
| 23 | | | | 53 | | | |
| 24 | | | | 54 | | | |
| 25 | | | | 55 | | | |
| 26 | | | | 56 | | | |
| 27 | | | | 57 | | | |
| 28 | | | | 58 | | | |
| 29 | | | | 59 | | | |
| 30 | | | | 60 | | | |

14.2 Příloha B - obsah přiloženého DVD

Součástí přílohy je datový disk, který obsahuje především implementace z Matlabu (GUI) a programy důležité pro jejich chod. Dále disk obsahuje samotnou práci ve formě PDF, výsledky subjektivních testů a zdrojové videosekvence.

\BAP

\Komprese_stereoskopickeho_videa.pdf

\Implementace

\ACT.m

\ACT.fig

\OMT.m

\OMT.fig

\Software

\Dcraw

\FFmpeg

\VQMT

\Subjektivni_testovani

\Vysledky_subjektivnich_test.xlsx

\Zdrojove_videosekvence

\Bikes_L.avi

\Bikes_R.avi

\Cube_L.avi

\Cube_R.avi

\Fire_L.avi

\Fire_R.avi

\Night_L.avi

\Night_R.avi

\Train_L.avi

\Train_R.avi