

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ
Fakulta elektrotechnická
katedra radioelektroniky

**Metody pro kompresi obrazu s vysokým dynamickým
rozsahem**

Diplomová práce

Vedoucí: Ing. Karel Fliegel, Ph.D.

Vypracoval: Bc. Petr Rautenkranc

duben 2014

Zadání DP

Abstrakt

Diplomová práce se zabývá procesem snímání a komprese obrazových dat s vysokým dynamickým rozsahem. Práce vysvětluje pojem vysoký dynamický rozsah a popisuje metody mapování jasu, které jsou důležitou součástí metod komprese obrazových dat s vysokým dynamickým rozsahem. V práci jsou popsány zpětně kompatibilní metody pro kompresi obrazových dat s vysokým dynamickým rozsahem. Součástí práce jsou algoritmy vytvořené v programovém prostředí Matlab a jejich stručný popis. Zpracované kompresní algoritmy jsou v závěru práce posouzeny pomocí objektivních metod hodnocení kvality obrazu.

Klíčová slova

vysoký dynamický rozsah, HDR, mapování jasu, komprese, hodnocení kvality obrazu

Abstract

The thesis pursues high dynamic range image creation and data compression. The thesis explains term high dynamic range and describes tone mapping methods, which are important in process of high dynamic range compression. In the thesis are described important backward-compatible compression methods. These methods are implemented in programming environment Matlab and these algorithms are compared using image quality evaluation.

Keywords

high dynamic range, HDR, tone mapping, compression, image quality evaluation

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.“

Prohlašuji, že diplomovou práci Metody pro kompresi obrazu s vysokým dynamickým rozsahem jsem vypracoval samostatně a použil k tomu literaturu, kterou uvádím v seznamu přiloženém k práci.

Nemám námitky proti půjčování, zveřejnění a dalšímu využití práce, pokud s tím bude souhlasit katedra radioelektroniky.

V Praze dne 11. května 2014

.....

podpis studenta

Poděkování

Tímto bych chtěl poděkovat vedoucímu mé diplomové práce, Ing. Karlu Fliegelovi, Ph.D., za pomoc a čas, který mi věnoval při vypracování této diplomové práce.

Obsah:

1	VYSOKÝ DYNAMICKÝ ROZSAH	1
2	HDR SNÍMÁNÍ	3
2.1	SNÍMÁNÍ SKLÁDÁNÍM	3
2.2	PŘÍMÉ SNÍMÁNÍ	4
2.3	HDR DATA - FORMÁTY	5
3	HDR MAPOVÁNÍ JASU	7
3.1	GLOBÁLNÍ METODY MAPOVÁNÍ JASU	8
3.1.1	<i>Lineární globální operátor</i>	8
3.1.2	<i>Logaritmický globální operátor</i>	8
3.1.3	<i>Exponenciální globální operátor</i>	9
3.1.4	<i>Millerův globální operátor</i>	9
3.1.5	<i>Globální fotografický operátor</i>	10
3.1.6	<i>Adaptivní logaritmický operátor</i>	10
3.1.7	<i>Schlickův kvantizační globální operátor</i>	12
3.2	LOKÁLNÍ METODY MAPOVÁNÍ JASU	13
3.2.1	<i>Lokální fotografický operátor</i>	13
4	HDR KOMPRESSE	15
4.1	KONVENČNÍ METODA KOMPRESSE DLE WARDA A SPAULDINGA	16
4.2	ZPĚTNĚ KOMPATIBILNÍ METODA JPEG-HDR	16
4.3	METODA KOMPRESSE VYUŽÍVAJÍCÍ KÓDOVÁNÍ ZÁKLADNÍ MAPY	18
4.4	LOGLUV METODA S VYUŽITÍM JPEG2000	19
4.5	METODA KOMPRESSE JPEG2000 VYUŽÍVAJÍCÍ HVS-MODEL	21
4.6	METODA PŘÍMÉHO KÓDOVÁNÍ NA SNÍMACÍM ČIPU	24
4.7	DVOUSTUPŇOVÁ ZPĚTNĚ KOMPATIBILNÍ METODA	29
4.7.1	<i>Odhad exponentu snímku (E)</i>	30
4.7.2	<i>Odhad mantisy (RGB)</i>	31
4.7.3	<i>Dekompresní postup</i>	31
5	OBJEKTIVNÍ METODY PRO HODNOCENÍ KVALITY OBRAZU	32
5.1	PSNR	32
5.2	UNIVERZÁLNÍ INDEX KVALITY OBRAZU	33
5.3	SSIM.....	34
5.4	HDR-VDP-2	39
5.5	INDEX KVALITY MAPOVÁNÍ JASU.....	42
6	POPIS ALGORITMŮ V PROGRAMOVÉM PROSTŘEDÍ MATLAB.....	43
7	HODNOCENÍ VÝSLEDKŮ A KVALITY SNÍMKŮ	45
7.1	KOMPRESNÍ METODA HDR-JPEG2000	47
7.2	ZPĚTNĚ KOMPATIBILNÍ KOMPRESNÍ METODA HDR-JPEG	50
7.3	DVOUSTUPŇOVÁ KOMPRESNÍ METODA.....	54
7.4	GRAFICKÉ POROVNÁNÍ KOMPRESNÍCH METOD	57
8	ZÁVĚR.....	61
9	SEZNAM POUŽITÉ LITERATURY	62
	OBSAH DATOVÉHO NOSIČE	65

1 Vysoký dynamický rozsah

Dnešní fotografická technika má jeden zásadní problém. Nedokáže na jednom snímku zachytit dostatečně velký dynamický rozsah. Při zachycení snímku obvykle volíme clonové číslo a dobu expozice. Při volbě těchto parametrů však musíme počítat s tím, že ztratíme část dynamického rozsahu snímané scény. Hodnoty jasu se většinou ukládají v rozsahu 0 až 255. Tento rozsah nedokáže pokrýt všechny rozdíly a detaily scény.

Formát High Dynamic Range (data s vysokým dynamickým rozsahem, HDR) je takový formát, který ukládá jasové hodnoty ve větším rozsahu a tím dokáže lépe zachytit celou reálnou scénu. V průběhu posledního desetiletí zaznamenal tento formát obrovský rozkvět. Využití našel jak ve fotografii, tak v dalších oblastech audio-vizuální techniky (např. virtuální realita, počítačové hry, vizuální efekty v kinematografii). Tento formát dokáže zachytit jak nejsvětlejší, tak i nejtmaší oblasti scény bez ztráty informace v podobě přexponování nebo podexponování scény. Tradiční zobrazovací metody jsou často limitovány technologií, která dokáže zaznamenat pouze 8 bitů na jeden obrazový bod pro jeden barevný kanál. Taková zobrazovací technologie bývá také nazývána jako Low Dynamic Range (nízký dynamický rozsah, LDR).

Snímání obrazu s vysokým dynamickým rozsahem mění všechny části přístupu k zachycení a zobrazení daného obrazu. V první fázi je nutné určitým způsobem získat obrazová data použitelná pro formát HDR - za použití klasického přístroje složením z více obrazů nebo při použití moderních zobrazovacích zařízení, která vytváří snímky s vysokým dynamickým rozsahem přímo při ukládání dat. Poslední možností jsou přístroje, které dokáží získat HDR data jako jeden snímek, nicméně tato data nikdy nejsou tak robustní jako data složená z více snímků. Tyto metody a použité algoritmy pro výpočet obrazových dat jsou popsány v dalších částech této práce.

HDR obrazová data zabírají až 4x více paměti než data s nízkým dynamickým rozsahem. Jasová složka každého bodu se neuchovává jako 8-bitové číslo, ale jako 3 číslice s plovoucí desetinnou čárkou (floating point number). To ovlivňuje všechny další fáze zpracování obrazu včetně rychlosti zpracování a dalších parametrů a projevuje se také v rámci standardů JPEG a MPEG. Kvůli zrychlení tohoto procesu je žádoucí použít určité formy efektivnějšího ukládání čísel s plovoucí desetinnou čárkou.

Jakmile máme HDR data důsledně a správně uložena, je možné začít s těmito daty pracovat v dalších částech zobrazovacího řetězce. Problém nastane ve chvíli, kdy chceme tato

data zobrazit na klasickém CRT nebo LCD monitoru. Dynamický rozsah těchto zařízení není dostatečně vysoký (typicky kolem 200 : 1). Pro správné zobrazení je nutné tato data určitým způsobem upravit. Tento proces úpravy se nazývá „tone mapping“ (mapování tónů) a podrobněji se jím zabývá další část této práce. V nedávné době se také na trhu objevily monitory, které dokáží zobrazit HDR data přímo (bez nutnosti další úprav).

2 HDR snímání

Možností, jak pořizovat HDR data, je více. Ideální variantou jsou moderní zařízení, která umí snímat a ukládat data přímo v požadovaném formátu. Pokud takové zařízení nemáme k dispozici, pak je možné použít klasický digitální fotoaparát a vyfotit několik snímků s předem definovanými parametry a pomocí definovaného algoritmu složit z jednotlivých snímků výsledná HDR data.

2.1 Snímání skládáním

Postup, kterým se dají snímky skládat, je obdobný jak pro černobílé snímky, tak pro barevné. Při barevném snímání je nutné tento postup provádět pro jednotlivé barevné kanály zvlášť. Princip činnosti moderních CCD snímačů je takový, že výsledná hodnota expozice v daném bodě snímače odpovídá počtu fotonů, které na tento bod dopadly v určitém čase. Pokud tedy zdvojnásobíme dobu expozice a zajistíme poloviční hodnotu ozáření, dostaneme shodnou hodnotu expozice. Z toho vychází, že i z omezeného rozsahu expozice, kterou je schopen čip zpracovat, dokážeme vytvořit velký dynamický rozsah skládáním ze snímků s různou dobou expozice. Expozici je také možné ovlivnit změnou clonového čísla, tam však dochází ke změně hloubky ostrosti a změně vnímání celé scény.

Na výsledné sérii snímků pořízené digitálním fotoaparátem neodpovídají hodnoty pixelů přesně hodnotám expozice na čipu (dochází k nelineárnímu převodu z čipu do výsledných digitálních hodnot). Popis této nelinearity zajišťuje charakteristická křivka zařízení a pro správné skládání do výsledného HDR snímku je nutné tuto křivku odhalit.

Nejjednodušší metodou je přímý výpočet z pořízených snímků. Tato metoda dokáže určit hledanou křivku z hodnot HDR dat, která se pohybují v intervalu od 0 do 1. U digitální fotografie si můžeme vypomoci použitím známé křivky gamma 2,2. Digitální zpracování nemá tak výrazné nelineární zkreslení a tato křivka je v celém procesu použita uměle a tím se vytváří efekt, který napodobuje klasickou metodu pořizování fotografie na fotografický film. Navíc není nutné řešit vyvážení bílé, pokud použijeme stejnou křivku pro všechny barevné kanály. Další otázka je, kolik LDR snímků je nezbytných pro vytvoření HDR snímku složením z těchto LDR snímků a jaké rozložení doby expozice je vhodné pro celou sérii snímků. Ideální pro tuto konstrukci je střední část charakteristické křivky, kde je srovnatelná změna hodnot expozice a hodnot jasu na pixelu (tedy křivka se v této oblasti blíží lineární křivce).

Pokrytí celého dynamického rozsahu závisí na velikosti lineární části charakteristické funkce čipu a na dynamickém rozsahu snímané scény. Pro pokrytí celého dynamického rozsahu scény je nutné nastavit rozdíl doby expozice tak, že se rozsahy hodnot jasu nepřekrývají. Tím je zajištěno, že hodnota jasu v každém pixelu bude v lineární části charakteristické křivky právě na jednom snímku z celé série a tato hodnota pak bude použita pro výpočet hodnoty ozáření vybraného pixelu.

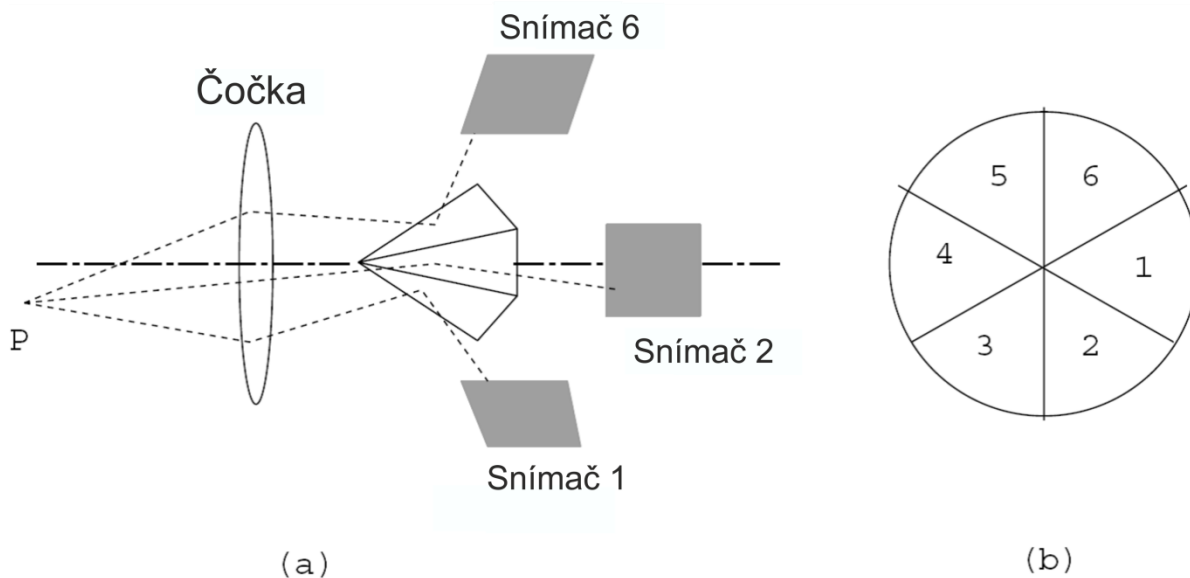
Při použití této metody hrozí zvýšení šumu. Pak je tedy lepší použít metodu, kdy se přepočítaná hodnota jasu pohybuje v lineární části charakteristické křivky na více snímcích a použitím průměrování nebo váženého průměrování eliminujeme značnou část šumu.

2.2 Přímé snímání

Metoda přímého snímání má oproti metodě skládání několik významných výhod. Asi nejdůležitější změnou je nižší časová náročnost. U skládání snímáním bylo nutné nejdříve vyfotografovat několik snímků s různou dobou expozice a následně je pomocí jednoduchých nebo složitějších algoritmů složit dohromady. Problémy také mohly nastat při pořizování více snímků u scény, která nebyla statická. Všechny tyto problémy odpadnou, pokud použijeme zařízení schopné snímat HDR snímky přímo.

V dnešní době existuje již několik typů kamer, které umožňují snímat HDR data v reálném čase. Princip těchto zařízení si můžeme představit takto: kamera má jeden společný objektiv, kterým vstupují paprsky světla. Za tímto objektivem je optický hranol, který rozděluje svazky paprsků na jednotlivé snímače. Pokud je celá soustava dobře seřízená, není poloha snímaného bodu ve scéně v místě dopadu na jednotlivé senzory závislá na vzdálenosti tohoto bodu od objektivu a tím pádem se celé zařízení chová jako klasická kamera s jedním snímačem.

Různých časů expozice můžeme dosáhnout například použitím barevně neutrálních filtrů, které budou mít různý přenos. Při správné konfiguraci s optickým hranolem můžeme dosáhnout správný poměr expozic pro vytvoření HDR snímku. Podobného výsledku je možné docílit také rozdílnou integrační dobou jednotlivých snímačů. Tato metoda je však použitelná pouze pro statické scény, u nestatických by vedla k nepatrným rozdílům v obrazech na jednotlivých snímačích.



Obrázek 1 - princip kamery pro přímé snímání HDR dat. Převzato z [1].

2.3 HDR data - formáty

Pro ukládání HDR obrazových dat nepostačuje klasický formát barevného snímku (JPEG, BMP a další), který využívá maximálně 8 bitovou barevnou hloubku pro jednotlivé barevné kanály (tedy hodnoty 0–255). HDR data je nutné ukládat ve formátu, který umožňuje pracovat s číselnou hodnotou ve formátu „float“, tedy s plovoucí desetinnou čárkou. To umožňuje například formát TIFF, v jedné ze svých variant určené pro práci s plovoucí desetinnou čárkou. Existují však také formáty speciálně určené pro ukládání HDR dat. Tyto formáty totiž zohledňují i případnou kompresi dat v rámci specifických parametrů HDR snímků.

Jedním z těchto formátů je Radiance RGBE. Tento formát využívá vysoké pravděpodobnosti toho, že hodnoty v jednotlivých kanálech mají stejný nebo podobný řád. Z tohoto důvodu je možné řád, tedy exponent, ukládat a přenášet pouze jednou v rámci úspory dat. Pro uchování a uložení snímků také není nutná absolutní přesnost dat. Data pro jednotlivé obrazové body se tedy ukládají následovně: 8 bitů mantisa pro každý barevný kanál zvlášť a 8 bitů jako společný exponent. Každý obrazový bod má tedy informaci velkou 32 bitů.

Tento formát má množství výhod a také množství nevýhod. Problém nastává v momentě, kdy se hodnoty v jednotlivých barevných kanálech výrazně liší. Další nevýhodou je rovnoměrné rozdělení 16 bitů na 8 bitů jako mantisa a 8 jako exponent. Výhodou je naopak celkové množství dat, která se ukládají pro jednotlivé obrazové body – 32 bitů na jeden obrazový bod. To zjednodušuje následné úpravy a práci s HDR daty a po provedení

jednoduchých kompresních algoritmů také docílíme poměrně přijatelné velikosti ukládaných dat pro jeden snímek.

Další používaný formát pro ukládání HDR dat je OpenEXR. Tento formát prošel velkým vývojem a nabízí široké možnosti nastavení přesnosti ukládaných dat. Zároveň také umožňuje kompatibilitu mezi jednotlivými možnostmi. Nejčastěji se používá buď 32-bitová nebo 16-bitová přesnost ukládaných dat. 16 bitový typ přesnosti („half“) podporuje také práci s daty, která jsou uložena ve tvaru s 5 bity exponentu a 10 bity jako mantisa. Takové nastavení umožňuje zachytit vysoký dynamický rozsah a používá se v moderních grafických kartách a nedochází tak k žádné úpravě při zobrazení pomocí grafických karet.

Na tento formát je možné aplikovat množství ztrátových a bezztrátových kompresních algoritmů (RLE, ZIP, PIZ – bezztrátová vlnková komprese, a další). Formát také není omezen počtem snímaných kanálů. Kromě jasového kanálu a barevných kanálů je možné uložit v tomto formátu i libovolné doplňkové kanály, které nesou dodatečné informace o snímané scéně. Tyto informace se nemusí ukládat do jiných souborů a přenášet zvlášť. Každý kanál pak může být uložen v jiném datovém typu v závislosti na požadované přesnosti. Na výběr je ze tří typů: „half“ jako 16 bitový desetinný typ pro obrazová data, „float“ jako 32 bitový desetinný typ pro data z bufferů a „uint“ jako 32 bitový celočíselný formát pro uložení identifikátorů.

Přes všechny komplikované možnosti nastavení tento formát nabízí také zjednodušený přístup. Ten je dostačující pro základní použití při HDR snímání. Lze zapisovat obyčejná RGBA data s přesností 16 bitů na obrazový bod do souboru formátu OpenEXR a tento postup přináší dobré výsledky pro statické snímky.

3 HDR mapování jasu

Postupy, které vedou ke snímání a uložení HDR obrazových dat, jsou použitelné na snímací straně. V další fázi, tedy zobrazování, dochází k objevu celé řady dalších problémů. Jak například zobrazit data s vysokým dynamickým rozsahem na běžných zobrazovacích zařízeních s nízkým dynamickým rozsahem. Běžná zařízení mají dynamický rozsah zpravidla na úrovni dvou, maximálně tří řádů. HDR data se však pohybují v jiných úrovních dynamického rozsahu.

Nejjednodušší metodou pro zobrazení HDR dat na běžném zařízení je nezobrazit všechna data, ale pouze vybrat určitou část pro jeden nebo dva expoziční kroky. V tomto případě se pak určitá část hodnot s plovoucí desetinnou čárkou dokáže vejít nebo transformovat do rozsahu 0 až 255 a je možné ji zobrazit jako běžný snímek. Tuto metodu bychom mohli použít jako náhled na HDR data, pro přesnější zobrazení však není použitelná. Pokud chceme zobrazit data v celém rozsahu, je nutné najít metodu, jak celý dynamický rozsah snímané scény (tedy data s plovoucí desetinnou čárkou) „namapovat“ do rozsahu hodnot 0 až 255.

Uvažujme nyní jen černobílá obrazová data, kde hodnoty jasu v obrazovém bodě přímo odpovídají osvětlení daného bodu. Pro jednotlivé barevné kanály se pak celý proces odehrává stejně jako u černobílého snímku, jen pro každý kanál zvlášť, nebo se barevné složky převedou na jasovou složku podle následujícího vztahu:

$$L = 0,3R + 0,59G + 0,11B , \quad (1)$$

kde L je hodnota jasu a R,G a B jsou hodnoty pro jednotlivé barevné kanály. Po provedení všech potřebných matematických operací se pak jasová složka převede zpět na jednotlivé barevné kanály pomocí vztahu:

$$C_{out} = \left(\frac{C_{in}}{L_{in}} \right)^s \cdot L_{out} , \quad (2)$$

kde C značí jednotlivé barevné kanály a tento vzorec je nutné použít pro každý barevný kanál zvlášť. Exponent „s“ je zodpovědný za výslednou barevnou sytost a nejčastěji se pohybuje v intervalu 0,4 až 0,6.

Převedení dat s vysokým dynamickým rozsahem do běžného obrázku, který zvládne zobrazit i klasické zobrazovací zařízení s nízkým dynamickým rozsahem je možné nazvat jako komprese dynamického rozsahu, častěji se však používá termín mapování jasu.

Mapování jasu má dva hlavní proudy, které se liší přístupem k jednotlivým obrazovým bodům a jejich souvislostem ve snímané scéně.

První skupina metod se označuje jako globální metody mapování jasu a jedná se o skupinu, která pracuje s obrazovými body nezávisle na jejich poloze ve snímané scéně. Je to skupina, která se vyznačuje nízkou výpočetní náročností a jednoduchou implementací. Nevýhodou je často nedostatečná úroveň komprese, obzvláště pro HDR data pokrývající takřka celý dynamický rozsah, protože v tomto případě se vytrácí lokální kontrast.

Druhá skupina metod mapování jasu využívá principy, jež určitým způsobem zohledňují polohu obrazového bodu v celém snímku, jeho okolí a souvislosti s okolními body. Tuto skupinu označujeme jako lokální metody mapování jasu. Metody v této skupině vykazují daleko lepší výsledky pro lokální kontrast, protože pro jednu vstupní hodnotu obrazového bodu může být na jednotlivých místech snímku různá výstupní hodnota [2]. Nevýhody těchto metod jsou převážně vyšší výpočetní náročnost a možnost vzniku nelinearit a artefaktů způsobených variancí v odlišném převodu lokálního zobrazení.

3.1 Globální metody mapování jasu

3.1.1 Lineární globální operátor

Globální metody mapování jasu se používají pro celou snímanou scénu rovnoměrně. První z nich je lineární uniformní komprese. Vstupní data jsou rovnoměrně namapována na interval 0 až 255. Z principu této metody vychází jednoznačná nevýhoda – vysoké potlačení lokálního kontrastu a celý snímek je tak často velmi tmavý.

HDR data většinou obsahují malé množství míst, která jsou velmi světlá (nejčastěji zdroje světla ve scéně) a to způsobí, že části snímku, obsahující velké množství detailů jsou mapována do tmavé oblasti výsledných hodnot. Dochází tedy k relativně velké ztrátě informace.

3.1.2 Logaritmický globální operátor

Druhá metoda vychází z výše uvedeného lineárního operátoru. Jedná se o logaritmickou kompresní funkci (též **logaritmický globální operátor**). Využívá prostého faktu, že lidské oko vnímá intenzitu světla také přibližně logaritmicky. Pokud na vstupní data nejdříve aplikujeme logaritmus, pak se dostaneme do situace, že velmi světlá místa jsou mapována do malé oblasti ve světlé části rozsahu a zbývá tedy poměrně velká část pro ty hodnoty jasu, které obsahují podstatné informace o detailech. I přes toto významné zlepšení je stále problém s potlačením lokálního kontrastu [1].

Kompresní funkce má následující tvar:

$$L_d(x, y) = \frac{\log(1 + L_w(x, y))}{\log(1 + L_{wmax})}, \quad (3)$$

kde L_d jsou hodnoty jasu na snímku s nízkým dynamickým rozsahem (LDR) a L_w jsou hodnoty jasu na snímku s vysokým dynamickým rozsahem (HDR).

3.1.3 Exponenciální globální operátor

Obdobou logaritmické metody je exponenciální metoda (nebo též **exponenciální globální operátor**). Jedná se o úpravu logaritmické kompresní funkce použitím exponenciály. Je to výpočetně nenáročná metoda, která se obdobně jako logaritmická metoda hojně využívá jako reference. Kompresní funkce je v následujícím tvaru:

$$L_d(x, y) = 1 - \exp\left(-\frac{L_w(x, y)}{L_{avlog}}\right), \quad (4)$$

kde L_{avlog} je aritmetický průměr logaritmu hodnot jasu v každém bodě HDR snímku a z něj vypočítaná exponenciální funkce:

$$L_{avlog} = \exp\left(\frac{1}{N} \sum_{x,y} \log(L_w(x, y))\right). \quad (5)$$

Pro jednotlivé barevné složky se pak použije následující vztah:

$$L_w = 0,2126 \cdot L_R + 0,7152 \cdot L_G + 0,0722 \cdot L_B. \quad (6)$$

3.1.4 Millerův globální operátor

Globální operátory mohou být také založeny na psychovizuálních vlastnostech lidského zraku. Toho využívá např. **Millerův globální operátor**. Kompresní funkce tohoto operátoru zachycuje vztah mezi jasem zobrazovacího LDR zařízení a původním jasem scény:

$$L_{di}(x, y) = \frac{Q_i(x, y)}{Q_{i max}} \cdot L_{d max}, \quad (7)$$

a Q_i potom získáme takto:

$$Q_i = (-1,5 \cdot \log(L_i) + 6,1) \cdot L_i^{0,338 \cdot L_i^{0,034}}, \quad (8)$$

3.1.5 Globální fotografický operátor

Další přístup je inspirován fotografickou stránkou této problematiky. Pomocí měření expozice a dalších nástrojů při fotografování vznikl **globální fotografický operátor**. Konstanta „key“ se pohybuje v rozmezí 0,18 až 0,72 a závisí na jasů snímané scény.

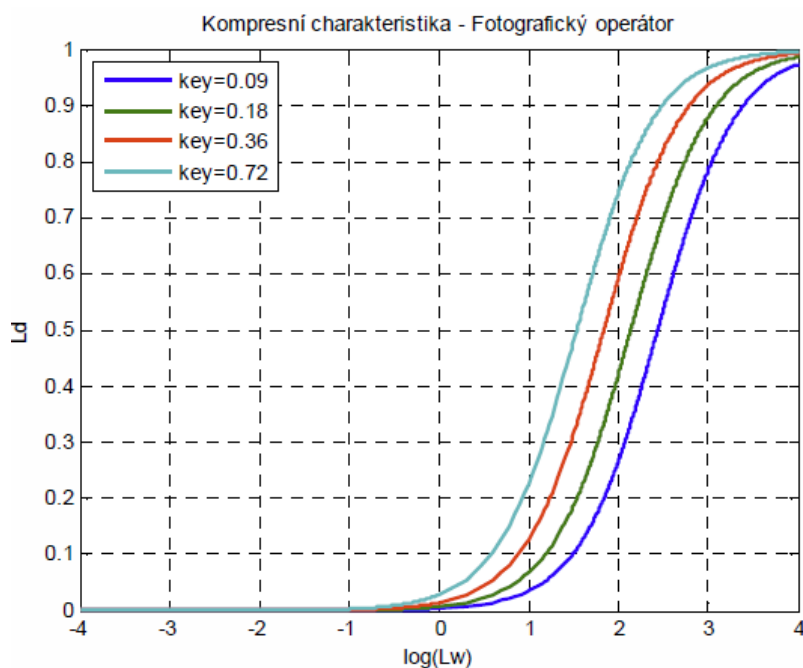
V první fázi této metody je potřeba pomocí této konstanty přepočítat hodnoty jasu:

$$L(x, y) = \frac{k}{L_{avlog}} \cdot L_w(x, y), \quad (9)$$

kde L_{avlog} vychází ze vztahu (5) a „k“ je konstanta popisující celkovou světlost snímané scény.

V další fázi je nutné vypočítat hodnotu jasu pro LDR zobrazovací zařízení:

$$L_d(x, y) = \frac{L(x, y) \cdot \left(1 + \frac{L(x, y)}{L_{white}^2}\right)}{1 + L(x, y)}, \quad (10)$$



Obrázek 2 – Kompresní funkce globálního fotografického operátoru. Převzato z [4]

3.1.6 Adaptivní logaritmický operátor

Úpravou logaritmického operátoru je možné získat **adaptivní logaritmický operátor**. Podrobný postup je popsán v článku [3]. Algoritmus mapování jasu pomocí logaritmu je upraven pro výstupní zobrazovací zařízení (v základní verzi se využívá jas 100 cd/m^2).

Základní vztah pro logaritmický operátor se upraví nejdříve tak, že se použije jiný základ logaritmu a to je možné vyjádřit takto:

$$\log_b(x) = \frac{\log_{10}(x)}{\log_{10}(b)}. \quad (11)$$

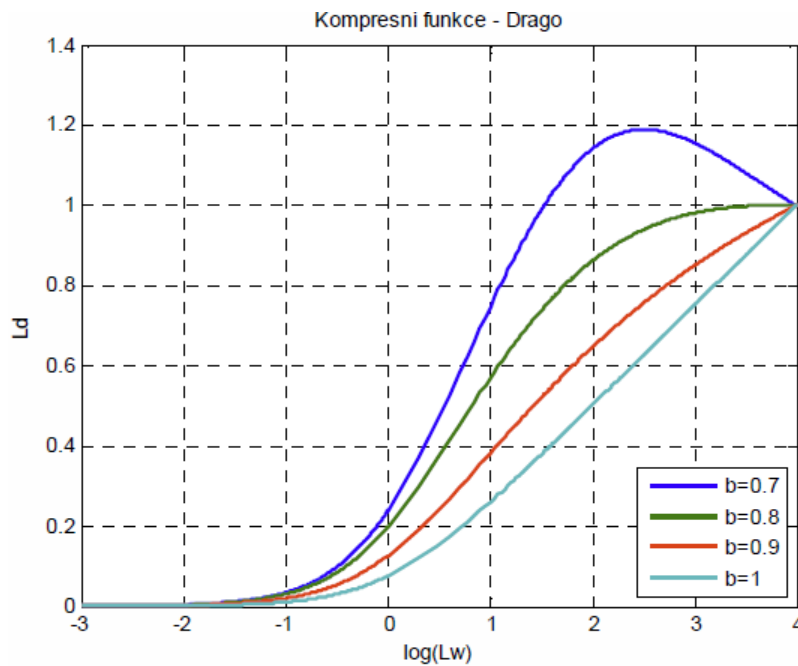
Základní kompresní funkce pak má tedy tvar:

$$L_d(x, y) = \frac{\log_b(1 + L_w(x, y))}{\log_b(1 + L_{w, \max})}. \quad (12)$$

Kombinací těchto vztahů vznikne výsledná funkce, která popisuje adaptivní logaritmický operátor:

$$L_d(x, y) = \frac{L_{d, \max} \cdot 0,01}{\log(1 + L_{w, \max})} \cdot \frac{\log(1 + L_w(x, y))}{\log \left\{ 2 + 8 \cdot \left[\left(\frac{L_w(x, y)}{L_{w, \max}} \right)^{\frac{\log(b)}{\log(0,5)}} \right] \right\}}, \quad (13)$$

kde b je parametr, který ovlivňuje vyhlazení křivky a ideálně se pohybuje v rozmezí 0,7 až 0,9.



Obrázek 3 – Kompresní funkce adaptivního logaritmického operátoru. Převzato z [4].

3.1.7 Schlickův kvantizační globální operátor

Další globální operátor je Schlickův uniformní kvantizační operátor. Jedná se o empiricky vylepšený operátor založený na logaritmickeém operátoru, který obsahuje prvky gamma korekce. Kompresní funkce obsahuje pouze dva parametry, z nichž první se vypočítá z daného snímku s vysokým dynamickým rozsahem a druhý se vypočítá podle parametrů zobrazovacího systému.

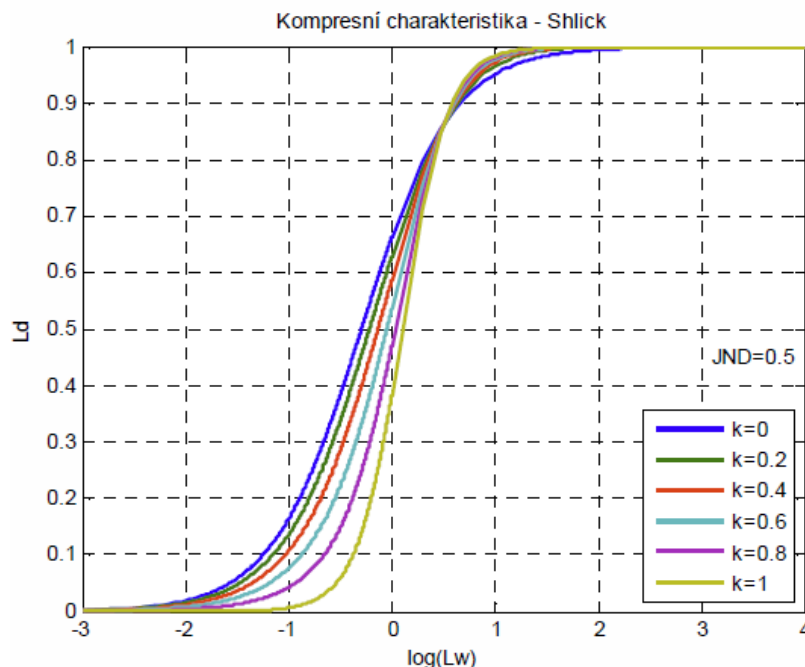
Parametr p se získá takto:

$$p = \frac{JND}{N} \cdot \frac{L_{w,max}}{L_{w,min}} \left(1 - k + k \cdot \frac{L_w(x,y)}{\sqrt{L_{w,min}L_{w,max}}} \right), \quad (14)$$

kde N je počet hodnot jasu, které dokáže výstupní zařízení zobrazit (tedy pro 8-bitové zařízení je tato hodnota 256) a JND je „just noticeable difference“, tedy minimální hodnota kterou je pozorovatel schopen rozlišit. Parametr „ k “ se pohybuje od 0 do 1, nejlépe s krokem 0,1 a ovlivňuje strmost křivky – viz obrázek.

Výsledná křivka má potom následující tvar:

$$L_d(x,y) = \frac{p \cdot L_w(x,y)}{(p-1) \cdot L_w(x,y) + L_{w,max}}. \quad (15)$$



Obrázek 4 – Kompresní funkce Schlickova kvantizačního operátoru. Převzato z [4]

3.2 Lokální metody mapování jasu

Lokální operátory pro mapování jasu jsou výpočetně náročnější než globální metody. Inspiraci získaly přímo z postupů, které využívá lidské zrakové ústrojí. Lidské oko dokáže adaptovat svůj dynamický rozsah na právě sledované místo a jasovou úroveň v něm obsaženou. Tyto metody jsou účinnější i pro velké dynamické rozsahy a vykazují přesnější a lepší výsledky. Problém však nastává na ostrých přechodech (vznikají artefakty atd.). Tomu se dá zabránit u některých metod vhodným nastavením vstupních parametrů.

3.2.1 Lokální fotografický operátor

Nejpoužívanější metodou lokálního mapování jasu je tzv. **lokální fotografický operátor** (jinak též **Reinhardův operátor** [5]). Tato metoda mapování jasu napodobuje klasické efekty, které fotografové používají již více než sto let – „dodging“ a „burning“ (zastiňování a nadržování určitých míst na fotografii).

Princip vychází z globálního fotografického operátoru. V první fázi se použije stejný postup, který je na závěr doplněn o lokální úpravy. Nejdříve tedy lineární úprava jasu:

$$L(x, y) = \frac{k}{L_{avlog}} \cdot L_w(x, y), \quad (16)$$

kde L_{avlog} vychází ze vztahu [5] a „k“ je konstanta popisující celkovou světlost snímané scény. Rozdělení na světlou, normální a tmavou část je důležité k určení, jak se dále pracuje s danými zónami snímku a k predikci, jak bude vypadat výsledná mapa. Ze vztahu pro mapování jasu vyplývá, že vyšší hodnoty jsou více komprimovány, zatímco ostatní hodnoty se blíží lineárnímu rozdělení. Na rozdíl od běžné fotografie však příliš jasné části fotografie nevedou k přesaturování hodnoty v daném obrazovém bodě při procesu zvyšování kontrastu celé fotografie. Předchozí rovnici je možné upravit takto:

$$L_d(x, y) = \frac{L(x, y) \cdot \left(1 + \frac{L(x, y)}{L_{white}^2}\right)}{1 + L(x, y)}, \quad (17)$$

kde L_{white} je nejmenší hodnota jasu, kterou mapujeme jako bílou a je rovna L_{max} . Pokud je hodnota $L_{white} > L_{max}$, hodnoty větší než L_{white} jsou ořezány (metoda „burning“ v klasické fotografii).

Pro začátek byla definována globální metoda. Nyní se použije lokální metoda, při které je nutné pro každý bod najít v jeho okolí další body, které nemají ostrý přechod kontrastu. K tomu použijeme Gaussovké filtry a porovnáme hodnotu obrazového bodu

uprostřed dané oblasti s Gaussovsky váhovaným průměrem celé uvažované oblasti. Pokud není v dané oblasti výrazný rozdíl v kontrastu, pak je výsledná hodnota blízká nule a naopak pokud je výrazný rozdíl v kontrastu, pak je i tato hodnota vysoká [1],[5].

Gaussovsky rozmazaný snímek s měřítkem s potom vypadá takto:

$$L_s^{blur}(x, y) = L(x, y) * R_s(x, y), \quad (18)$$

kde je použit symbol *, což je matematický operátor konvoluce, v tomto případě diskrétní dvourozměrná konvoluce. Metoda pro výpočet okolí daného bodu je následující:

$$V_s(x, y) = \frac{L_s^{blur} - L_{s+1}^{blur}}{2^\Phi \frac{a}{s^2} + L_s^{blur}}, \quad (19)$$

kde Φ je parametr popisující ostrost přechodu kontrastu ve výsledném snímku a nejčastěji se pohybuje mezi hodnotou 4 až 32 (nejlépe v hodnotách mocniny čísla 2).

V dalším kroku je také nutné nastavit práh epsilon, který definuje maximální povolenou odchylku kontrastu (porovnání středního bodu s jeho okolím, viz výše). Matematicky to lze popsat takto:

$$\frac{s_{max}}{|V_{s,max}(x, y)|} < \epsilon, \quad (20)$$

kde s_{max} popisuje maximální možnou velikost okolí bodu, kde změna kontrastu nepřesáhne parametr epsilon. Výsledný vztah pro tuto metodu je potom následující:

$$L_d(x, y) = \frac{L(x, y)}{1 + L_{s,max}^{blur}(x, y)}. \quad (21)$$

Tato metoda je v další části práce popsána při realizaci v programovém prostředí Matlab a je součástí přílohy této práce v digitální podobě – funkce „tmoReinhard.m“.

4 HDR komprese

Rostoucí kvalita a obliba HDR snímků má za následek nutnost komprimovat používaná obrazová data. Výpočetní náročnost operací s HDR daty a také jejich velikost jsou negativním projevem jinak pozitivního zlepšení snímaných a zobrazovaných dat. 16-bitové ukládání dat v jednotlivých barevných kanálech je natolik náročné na ukládací prostor a přenosové cesty, že se přistupuje při přenášení a ukládání těchto snímků ke kompresi.

Dynamický rozsah lidského oka a celé soustavy, která zpracovává obrazové informace, může dosahovat více než 200 dB v logaritmické míře. Poměr mezi světlými a tmavými místy scény se pohybuje až kolem 100 až 120 dB v běžné reálné scéně, v rámci testování se nejčastěji pracuje s dynamickým rozsahem okolo 130 dB. Naproti tomu dynamický rozsah většiny moderních fotoaparátů pro běžného spotřebitele vykazuje hodnoty okolo 80 dB a nemůže tak zachytit věrně celou snímanou scénu [6]. Popis výroby HDR snímku je popsán výše a tak se pro potřeby komprese spokojíme s podstatnou informací. HDR snímky mají enormní datovou náročnost jak pro datová úložiště, tak pro přenosové cesty a práci s nimi.

Komprese HDR snímků má několik možností, jak k ní přistoupit. V této kapitole jsou vybrané metody popsány a porovnány. První skupinou metod jsou metody, které využívají dvoustupňové kódování, kde se nejdříve zkomprimuje LDR snímek a v separátní větvi se pracuje na kódování a kompresi rozdílového nebo detailního snímku (v tomto snímku jsou obsaženy všechny významné informace o HDR detailech a tento snímek je v podstatě rozdíl mezi LDR a HDR snímkem, proto „rozdílový“).

Další skupina metod vychází ze standardu JPEG2000 a využívá ke kompresi vlnkovou transformaci nebo diskretní kosinovou transformaci. Standard JPEG2000 není sám o sobě vhodný pro HDR snímky a jejich kompresi, nicméně vybrané metody upravují tento standard tak, že vykazují dobré výsledky i u snímků s vysokým dynamickým rozsahem.

Třetí možný přístup je komprese s využitím známých vlastností lidského zrakového ústrojí. Nazývají se též „perceptuálně bezztrátové“, čímž je myšleno to, že i když jsou to ztrátové komprese z hlediska definice, tak lidské oko nedokáže tuto ztrátu informace rozpoznat. Tyto metody mívají obecně vyšší kompresní poměr při zachování stejné kvality reprodukováného obrazu.

4.1 Konvenční metoda komprese dle Warda a Spauldinga

Tato metoda se zaměřuje na dvoustupňovou kompresi HDR snímku. Ideální se jeví její použití pro snímky typu RAW, tedy surová data přímo ze snímacího zařízení. Podle Spauldinga [7] rozšiřuje tato metoda dynamický rozsah a barevné spektrum digitálního fotoaparátu pomocí nalezených nelinearit v LDR snímcích a zároveň zakódováním rozdílu mezi rozšířeními LDR snímky a originálním RAW snímkem. Později tuto metodu vylepšil Ward [8] tím, že ji použil přímo na HDR data. LDR snímky se v této metodě nejdříve komprimují pomocí klasického JPEG kodéru a následně se zjistí rozdílová informace mezi LDR snímkem a výsledným HDR snímkem v každém obrazovém bodě. Tato rozdílová informace se přenáší a používá v JPEG kodéru a dekodéru jako dodatečná informace pro dopočítání HDR dat.

Rozdílová informace se získá z HDR a LDR snímků následovně:

$$I_r = \frac{I_{HDR}}{I_{LDR}}. \quad (22)$$

Na straně dekodéru se potom snímek získá inverzní operací takto:

$$I'_{HDR} = I'_{LDR} \cdot I_r. \quad (23)$$

Konvenční metoda komprese používá ke kódování I_{LDR} a I_r (tedy LDR snímku a rozdílového snímku) kódér JPEG. Tato data se kódují nezávisle na sobě a zpět se skládají až v dekodéru. To přináší některé výrazné nevýhody. Efektivita komprese výrazně klesá na vyšších prostorových frekvencích, protože rozdílový snímek v nich nabývá vysokých hodnot energie a jasů. Druhá důležitá vlastnost a nevýhoda zároveň spočívá v saturaci odstupů signál-šum (SNR). Ta vzniká v případě, že LDR snímek I_{LDR} je použit k výpočtu HDR snímku I_{HDR} .

Proto se v následující kapitole 4.3 používá dekodovaný snímek I'_{LDR} tak, aby nevznikal drift ve výsledném snímku. Použitím dekodovaného snímku však přineseme do systému jinou chybu, kvantizační. Ta se u dekodovaného snímku projevuje převážně ve vyšších frekvencích a u nižších naopak ne. Proto tedy u konvenční metody vzniká saturace SNR. Tato metoda je výpočetně nenáročná, nedosahuje však příliš dobrých výsledků a proto je nutné ji vylepšit.

4.2 Zpětně kompatibilní metoda JPEG-HDR

Metoda komprese podle Warda a Simmonse [8] byla dále rozvíjena s cílem dosáhnout lepších výsledků. Jednoduchá myšlenka vedla k vytvoření této kompresní metody [9]. Provede se metoda mapování jasů na HDR snímky a pak se zkomprimují pomocí standardu JPEG.

Kompresi začíná v prvním kroku mapování jasu HDR snímku. Pak je vytvořen rozdílový (nebo též doplňkový) snímek. Ten se vytváří tak, že je původní snímek vydělen snímkem po provedení mapování jasu. Takto vytvořený rozdílový snímek budeme dále využívat. Je možné tento snímek podvzorkovat a přenášet ho jako zkomprimovanou informaci, protože lidský zrakový systém má omezenou schopnost vnímat vysoké změny prostorové frekvence jasové složky. Pokud však provedeme toto podvzorkování, musíme také provést korekci tohoto snímku tak, aby při následných operacích nevznikaly nežádoucí efekty okolo ostrých přechodů a hran.

Původní metoda využívala korekce ve snímku po provedení mapování jasu. Rozdílový snímek byl převzorkován a tímto snímkem byl vydělen původní HDR snímek. Takto vznikl snímek po provedení mapování jasu s korekcemi, který se dal použít pro další operace. Tato novější metoda však využívá řízené převzorkování [10], ale je pak výpočetně náročnější než původní metoda. Převzorkovaný rozdílový snímek je diskrétně uložen v 8 bitovém logaritmicky rozděleném prostoru aplikační vrstvy formátu JPEG a je tedy přenášen přímo jako součást snímku ve formátu JPEG. Snímek po provedení mapování jasu je nutné ještě upravit pro přenos tak, aby byly lépe a věrně zachovány barevné složky.

Je nutné použít dvě techniky, které kombinovaně vedou potřebným výsledkům - komprese barevné škály a využití barevného prostoru YCbCr. Je nutné provést „desaturaci“ pro kompresi barevné škály. Pokud uvažujeme definici saturace takto,

$$S(x) = 1 - \frac{\min[R(x), G(x), B(x)]}{L_w(x)}, \quad (24)$$

Pak „desaturaci“ jednotlivých barevných kanálů můžeme definovat následovně:

$$\begin{bmatrix} R_c(x) \\ G_c(x) \\ B_c(x) \end{bmatrix} = (1 - S(x)') \begin{bmatrix} L_w(x) \\ L_w(x) \\ L_w(x) \end{bmatrix} + S(x)' \begin{bmatrix} R_c(x) \\ G_c(x) \\ B_c(x) \end{bmatrix}, \quad (25)$$

a pro $S(x)'$ platí:

$$S(x)' = \alpha S(x)^{\beta-1}, \quad (26)$$

kde $\alpha \leq 1$ je parametr, který řídí úroveň saturace v průběhu kódování a β určuje výsledný barevný kontrast.

Druhý krok je převod do barevného prostoru YCbCr, protože tento barevný prostor má větší použitelnou škálu než prostor RGB. Z toho je patrné, že nevyužité koeficienty v prostoru YCbCr mohou být využity k uchování dodatečných informací z originálního HDR snímku.

Toho dosáhneme mapováním hodnot originálního snímku přesně do volných míst, například pro červený kanál vypadá mapování takto:

$$R'(x) = \begin{cases} 1,055R_c(x)^{0,42} - 0,055 & \text{pro } R_c(x) > 0,0031308, \\ 12,92 R_c(x) & \text{pro } |R_c(x)| \leq 0,0031308, \\ -1,055(-R_c(x))^{0,42} + 0,055 & \text{pro } R_c(x) < -0,0031308. \end{cases} \quad (27)$$

Tento postup je obdobný pro modrý a zelený barevný kanál.

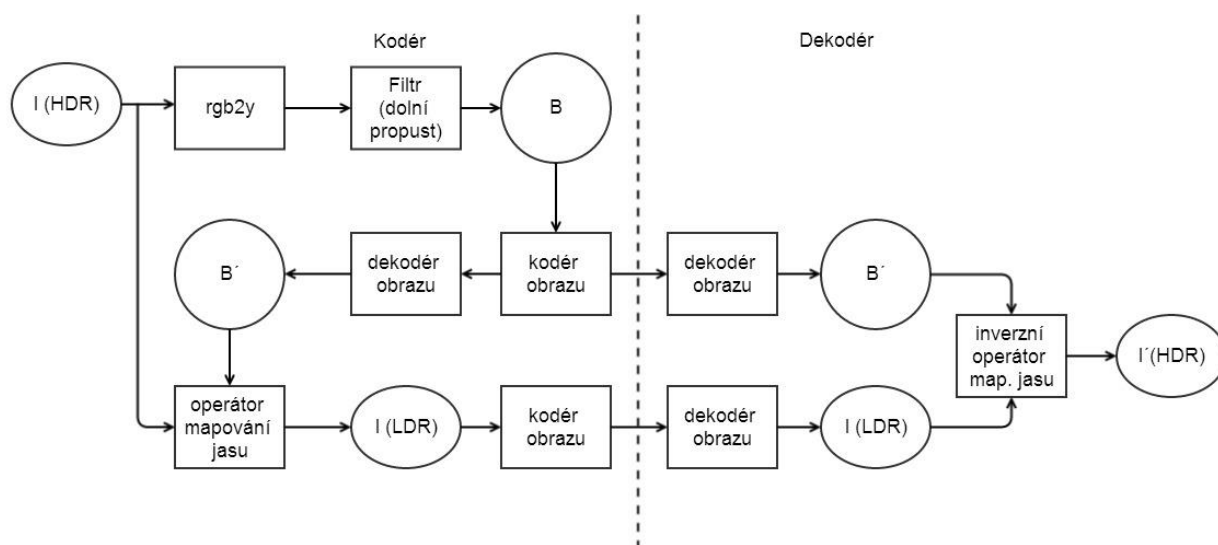
Dekódování probíhá v několika krocích. Nejprve je snímek po provedení mapování jasu dekomprimován pomocí JPEG dekodéru a barevná škála je zvětšena operací inverzní k operaci ve vztahu (25). Poté je převzorkovaný rozdíllový snímek dekodován, rozšířen z logaritmické škály do lineární a nakonec převzorkován zpět do rozlišení snímku po provedení mapování jasu. V posledním kroku je pak provedeno vynásobení obou těchto snímků.

4.3 Metoda komprese využívající kódování základní mapy

Metoda komprese podle Warda [8] je neefektivní z hlediska výpočetního výkonu, protože kvantizační chyba LDR snímku přináší vyšší a nepřesné hodnoty energie a jasu v rozdíllovém snímku (zvláště na vyšších frekvencích). Tato metoda je popsána v následujícím článku [6]. Principem metody je použití globálního fotografického operátoru mapování jasu – podle Reinharda [1]. Ten je použit v obou fázích dvoustupňového kódování, tedy jak na LDR snímek, tak na mapu jasů. Z HDR snímku je nejprve vytvořena základní mapa jasu B („base map“) a to tak, že HDR snímek projde dolno-propustním filtrem. LDR snímek se získává mapováním jasů dle Reinharda s pomocí základní mapy B. Oba tyto snímky jsou poté kódovány na straně kodéru.

Na straně dekodéru se provede dekodování obou přenášených snímků a poté se provede inverzní operace mapování jasu a tím se získá výsledný HDR snímek. Metody mapování jasu jsou popsány v předcházejících kapitolách, a proto se spokojíme s informací, že je použita metoda globálního fotografického operátoru. Rozdíl mezi konvenční metodou a metodou s mapou jasu je ten, že při použití mapy jasu se neprojevuje ani saturace SNR ani kvantizační chyba v takové míře jako u předchozí metody. Inverzní mapa jasu B je použita pro rekonstrukci HDR snímku a tento postup je popsán následovně:

$$I_{HDR}(x) = I_{LDR}(x) \left(1 + B'(x, s(x)) \right). \quad (28)$$



Obrázek 5 – Schéma metody kódování a komprese HDR snímku pomocí „Base Map Coding“ algoritmu (stran kodéru i dekodéru).

Na obrázku 5 je potom vidět schematické znázornění této metody. $I(HDR)$ a $I(LDR)$ jsou snímky, se kterými se pracuje buď v HDR formátu nebo LDR. B a B' zde reprezentují základní mapy jasů – B je po průchodu dolní propustí, B' je potom základní mapa po kódování a dekódování (obsahuje kvantizační chybu). $Rgb2y$ je potom operátor, který převádí snímek z barevných kanálů RGB na jasovou složku Y a rozdílové kanály (vhodné pro další práci se snímkem).

4.4 LogLuv metoda s využitím JPEG2000

Kompresní metody založené na vlnkové transformaci jsou povětšinou výhodné pro klasické LDR snímky složené z maximálně 24 bitů pro všechny barevné kanály. Jedná se například o JPEG2000. Nedokáží tak efektivně zvládnout HDR snímky. Zatímco lidské oko si dokáže s vysokým dynamickým rozsahem poradit, protože má vysokou schopnost lokální adaptace na rozdíly jasu, tak klasické kompresní algoritmy toto nedokážou provádět se stejnou účinností. Další problém nastává při metodě, využívající matematické operace v každém barevném kanálu zvlášť. Proto je nutné upravit klasické kompresní standardy typu JPEG pro potřeby snímku s vysokým dynamickým rozsahem.

Vylepšená metoda s vlnkovou transformací HDR snímků využívá tzv. *LogLuv* barevný prostor (viz [11]). Tento barevný prostor značně vylepšuje výhody vlnkové transformace u HDR snímků a dokáže odstranit problémy, které toto přináší. Tento formát barevného prostoru byl navrhnout pro bezztrátovou kompresi HDR snímků ve formátu TIFF.

Uvažujme nyní situaci, kdy jako vstupní HDR data používáme snímek v XYZ barevném formátu. V rámci této metody se jako nejvýhodnější jeví použití 32-bitového kódování.

Je nutné provést určité úpravy vstupního HDR snímku. Nejprve se vstupní data převedou z XYZ barevného prostoru do CIE(x,y):

$$x = \frac{X}{X + Y + Z} \quad (29)$$

$$y = \frac{Y}{X + Y + Z} \quad (30)$$

Tato reprezentace je nutná pro další postup celé komprese. Můžeme z ní určit hodnoty veličin v reprezentaci $LogLuv$, která koresponduje s kvantifikovanou verzí chromatického prostoru u' , v' (zde se používá také logaritmus jasu, $LogL$).

Celá tato operace využívá následující vztahy:

$$L_e = (256 \cdot (\log_2 Y + 64)) \quad (31)$$

$$u_e = 410 \cdot u' \quad (32)$$

$$v_e = 410 \cdot v' \quad (33)$$

kde veličiny u' a v' jsou popsány takto:

$$u' = \frac{4x}{-2x + 12y + 3} \quad (34)$$

$$v' = \frac{9y}{-2x + 12y + 3} \quad (35)$$

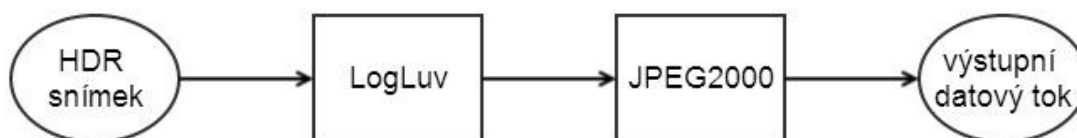
Jak je uvedeno v předchozích odstavcích, tak tato komprese využívá kódování pomocí 32 bitů a ty jsou rozděleny mezi jednotlivé složky takto: 16 bitů pro L_e a 8 bitů pro jednotlivé kanály u' a v' . Toto rozdělení dokáže zachytit celý dynamický rozsah snímané scény. Barevná reprezentace pomocí $LogLuv$ je velice podobná reprezentaci $YCbCr$. Ta se využívá u ztrátových kompresních algoritmů typu JPEG a MPEG, Y představuje jasový kanál a Cb , Cr potom reprezentují rozdílové složky Y-B a Y-R).

V našem případě tedy $LogLuv$ představuje pouze barevný prostor a nelze v tomto formátu ukládat přímo celý snímek, je potřeba rozšířit tento barevný prostor na 48-bitovou reprezentaci. Tím se změní rovnice (31), (32) a (33) takto:

$$L_e = (256.(\log_2 Y + 64)) \quad (36)$$

$$u_e = 32768.u' \quad (37)$$

$$v_e = 32768.v' \quad (38)$$



Obrázek 6 – Princip komprese LogLuv s pomocí vlnkové transformace

Koeficienty u_e a v_e se změni z 8 na 16 bitů. Princip komprese je potom naznačen na obrázku 6. Klasický kompresní princip využívající JPEG2000 nevykazuje dostatečně dobré výsledky s HDR snímky. V tomto případě je možné tyto výsledky značně vylepšit změnou barevného prostoru vstupních dat. To představuje blok *LogLuv* a princip je uveden v minulých odstavcích. Blok *LogLuv* nahrazuje klasický „preprocessing“ u běžné komprese JPEG2000. K tomu je pak nutné upravit vstupní data do barevného prostoru CIE XYZ, což však není nutné ve všech případech – často se HDR snímky již v tomto formátu vyskytují. Toto jsou tedy operace na straně kodéru.

Na straně dekodéru je pak situace podobná. Celý postup se zopakuje v opačném pořadí, tedy nejdříve se dekóduje JPEG2000 a potom se provede převod z *LogLuv* barevného prostoru. Z výše uvedených vztahů je možné vyjádřit zpětné vztahy pro výpočet koeficientů a hodnot jednotlivých obrazových bodů a rekonstruovat tak celý snímek. V tomto případě se jedná o bezztrátovou kompresi.

4.5 Metoda komprese JPEG2000 využívající HVS-model

Vedle bezztrátových kompresních metod, které vykazují nižší kompresní poměr, jsou také metody ztrátové. Tato ztrátovost je však relativní. Při správném využití vlastností lidského zrakového ústrojí je možné dospět k vyšším kompresním poměrům i při zachování kvality snímků, kdy lidské oko nedokáže rozpoznat rozdíl mezi bezztrátovou a ztrátovou kompresí. Takto ztrátové komprese je možné nazvat jako „percepčně bezztrátové“ [12].

Tato metoda v sobě skrývá několik kroků, které vedou k uspokojivým výsledkům. Je založena na standardu JPEG2000. V minulé kapitole se využíval *LogLuv* barevný prostor se

16 bity pro každý kanál (LogL, u, v). Ke snížení objemu přenášených dat se v této metodě aplikuje DWPT (Discrete Wavelet Packet Transformation), která využívá vlastností lidského zrakového ústrojí a v tomto případě vede k úspoře dat, jejichž nepřítomnost nedokáže lidské oko rozpoznat. Dále se také využívá CSF (Contrast Sensitivity Function), tedy váhovací funkce, která se uplatní v doméně vlnkových sub-pásem. Tyto operace vedou k signifikantnímu zvýšení obrazové kvality při zachování vysoké úspory přenášených dat.

Během posledních desítek let bylo vyvinuto několik modelů popisujících vlastnosti lidského zrakového ústrojí. Jeden z nejdůležitějších popisuje vztah mezi citlivostí lidského oka na kontrast a prostorovou frekvencí. To popisuje CSF, kterou můžeme získat experimentálně a podrobně se tomu věnují následující publikace [13] a [14]. Pro naše potřeby jsou důležité informace, že lidské oko je citlivější na nižší prostorové frekvence. U modelu pro šedotónové snímky se využívá CSF jako nelineární transformace s využitím modulační přenosové funkce MTF. Vzorkovací prostorovou frekvenci (v obrazových bodech na stupeň) získáme takto:

$$f_s = \frac{2\vartheta \cdot \tan(0,5^\circ)\gamma}{0,0254}, \quad (39)$$

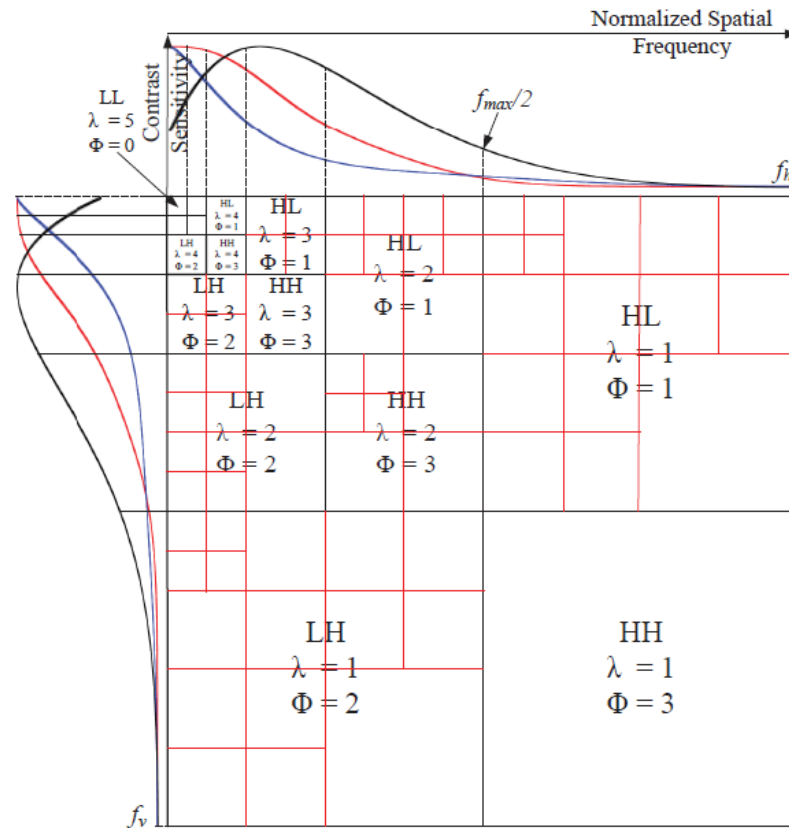
kde ϑ je pozorovací vzdálenost (nejčastěji 1 metr) a γ je rozlišení obrazovky v obrazových bodech na palec. Maximální frekvence uvedená v CSF je definována jako f_{max} podle Nyquistova teorému, tedy $f_{max} = 0,5 f_s$. CSF se rozděluje na jasový a dva chrominanční kanály:

$$CSF_L(f) = 2,6 \cdot (0,0192 + 0,114f \cdot \exp(-0,114f^{1,1})) \quad (40)$$

$$CSF_u(f) = 5,623 \cdot \exp(-0,00001f^{3,4066}) + 41,936 \cdot \exp(-0,083f^{1,368}) \quad (41)$$

$$CSF_v(f) = 91,228 \cdot \exp(-0,0003f^{2,803}) + 74,907 \cdot \exp(-0,0038f^{2,601}) \quad (42)$$

Na obrázku 7 je vidět použití diskrétní vlnkové transformace v kombinaci s CSF modelem (využití vlastností lidského zrakového ústrojí). Vlnková transformace se jeví jako nejlepší a nejúčinnější při kompresi obrazových dat a zvláště vhodná je při použití s modely využívajícími lidské zrakové ústrojí. Samotná vlnková transformace však nemusí být postačující kvůli nízkému frekvenčnímu rozlišení. Jednou z možností jak vylepšit výsledky je použití spojité vlnkové transformace (Continuous wavelet transformation – CWT). CWT vykazuje v teoretické podobě skvělé parametry včetně neomezených možností flexibility a rozlišení. Bohužel je však enormně náročná na implementaci a realizaci, takže se hledala ještě jiná varianta s diskrétním přístupem.



Obrázek 7 – Vztah mezi jasovou složkou CSF (černá křivka), chrominančními složkami CSF (modrá a červená křivka) a 2D-WPT (wavelet packet transformation se stupněm 5). Převzato z [12].

Lepší rozlišení při zachování stejné úrovně výpočetní složitosti vykazuje WPT – Wavelet Packet Transform (paketová vlnková transformace). Vlnkové pakety jsou funkce definované spojitě, ale dá se s nimi pracovat jako s diskretními funkcemi podobně jako v DWT. Hlavním přínosem této transformace je možnost signály adaptivně rozkládat podle nastavených podmínek, např. podle prahové hodnoty.

Diskretní nastavení této transformace se tedy nazývá DWPT (Discrete Wavelet Packet Transformation). Na obrázku 7 je znázorněno rozdělení po provedení této transformace. Písmena L a H v jednotlivých segmentech popisují, jestli byl použit filtr typu dolní propust (L) nebo horní propust (H). Parametr λ popisuje úroveň provedené DWPT a parametr Φ popisuje orientaci daného stupně DWPT. f_h a f_v jsou hodnoty horizontální a vertikální frekvence a odpovídají sub-pásmům DWPT.

Pro spojení těchto dvou kroků, tedy CSF a DWPT, je nutné provést další úpravy detailně popsané v [12]. Výsledná váhová funkce pro celou operaci komprese je potom definovaná jako váhová funkce DWPT koeficientů průměrovaných podle jednotlivých sub-pásem a získá se z křivky CSF v oblasti prostorových frekvencí:

$$\overline{W}_{CSF}(f_h, f_v) = \frac{\sqrt{\frac{\sum_{f_h} CSF^2(f_h)}{card(f_h)} + \frac{\sum_{f_v} CSF^2(f_v)}{card(f_v)}}}{2} \quad (43)$$

Celá tato komprese využívá 5-úrovňovou diskretní 2D paketovou vlnkovou transformaci. Jasové i chromatické funkce CSF jsou implementovány do této transformace a váhovací funkce jsou využity pro každé sub-pásmo. Rozklad do 5 úrovní využívá vlastností lidského zrakového ústrojí tak, že lidské oko je nejcitlivější v jasovém kanálu na úroveň 3 a citlivost klesá s prostorovou frekvencí v úrovních 1,2,4 a 5. Citlivost na chrominanční kanál se značně snižuje s rostoucí prostorovou frekvencí. Výsledný rozklad je potom uveden na obrázku 7.

4.6 Metoda přímého kódování na snímacím čipu

Moderní přístup ke kompresi HDR dat se provádí přímo na snímacím čipu [15]. Komprese se stejně jako v předchozích případech provádí ve dvou stupních. Prvním je optimalizace kódování hodnoty jednotlivých obrazových bodů a druhým je komprese založená na diskretní kosinové transformaci (DCT). Optimalizace v prvním kroku využívá kódovací techniku s plovoucí desetinnou čárkou se společným exponentem pro pixely ve stejném bloku. Ve druhém kroku je pak na každý blok aplikována diskretní kosinová transformace. Tento moderní přístup ke kompresi HDR obrazových dat se vyznačuje jednoduchou výpočetní strukturou a vysokou kvalitou výsledného komprimovaného snímku. Kompresní poměr se pohybuje u běžných snímků kolem hodnoty 0,75 a odstup signál šum dosahuje až 40 dB.

První krok dvouступňového kódování je následující. Vstupní obrazová data se rozdělí na makrobloky (nejčastěji 8x8 obrazových bodů) a pak se použije kódovací metoda s plovoucí desetinnou čárkou (v anglické literatuře označovaná jako „Mantissa-Exponent technique“). Tato metoda zvyšuje dynamický rozsah snímače pomocí úpravy integračního času podle množství dopadajících fotonů. Integrační čas se vybírá z předem definované řady diskretních hodnot. Vybraný optimální integrační čas reprezentuje celý exponent, zatímco diskretní hodnoty výstupního napětí reprezentují mantisu tohoto formátu s plovoucí desetinnou čárkou.

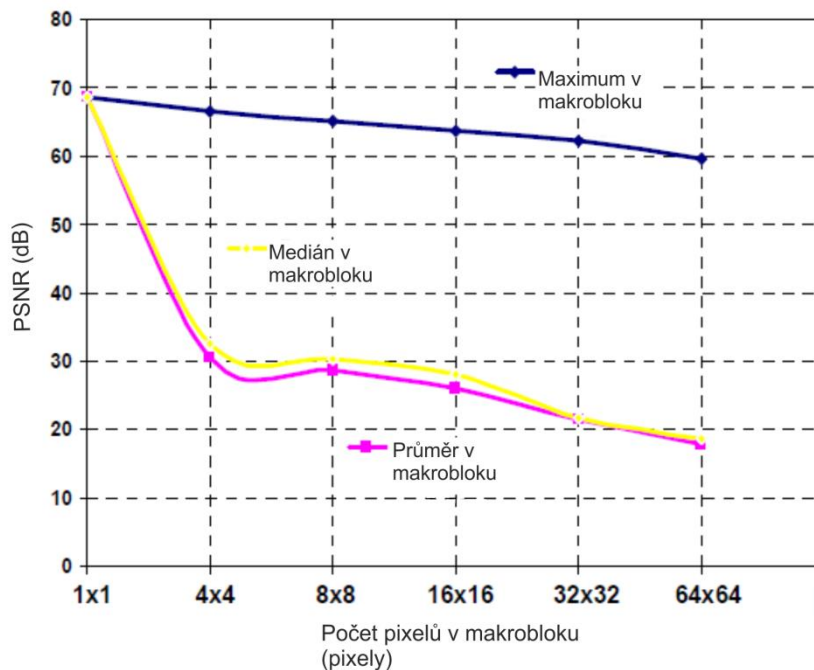
To lze zapsat následovně:

$$L(x, y) = M \cdot 2^E, \quad (44)$$

kde E je hodnota exponentu společná pro celý makroblok a M je hodnota mantisy, tedy přesná hodnota ve formátu s plovoucí desetinnou čárkou.

V tomto formátu je kódování velice účinné. Mantisa nabývá hodnot od 0 do $2^{N_m}-1$, kde N_m je počet bitů, který je potřeba ke kódování mantisy. Podobně exponent nabývá hodnot od 0 do $2^{N_e}-1$, kde N_e je počet bitů, který je potřeba ke kódování exponentu. Hodnoty jasu v obrazových bodech pak tedy nabývají hodnoty v následujícím rozsahu:

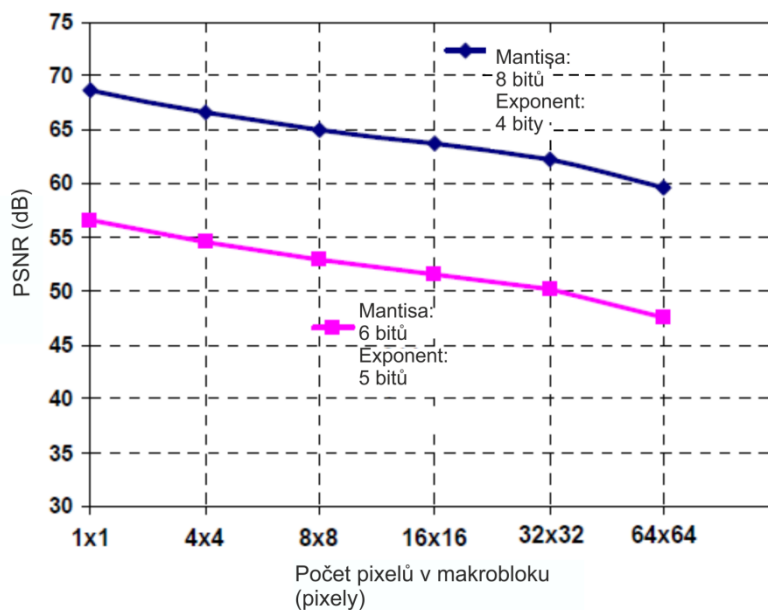
$$L(x, y) \in \{0; 2^{(N_m+2^{N_e})} - 1\}, \quad (45)$$



Obrázek 8 – Graf závislosti PSNR (dB) na velikosti makrobloků pro jednotlivé metody – modře je metoda maxima v makrobloku, žlutě je metoda mediánu v makrobloku a růžově metoda průměru v makrobloku. Převzato z [15].

Pokud tedy potřebuje v HDR snímku 22 bitů k popsání celého dynamického rozsahu scény, pak po provedení této metody kódování postačuje 8 bitů pro mantisu a 4 bity pro exponent. To představuje vysokou úsporu bitů pro každý obrazový bod.

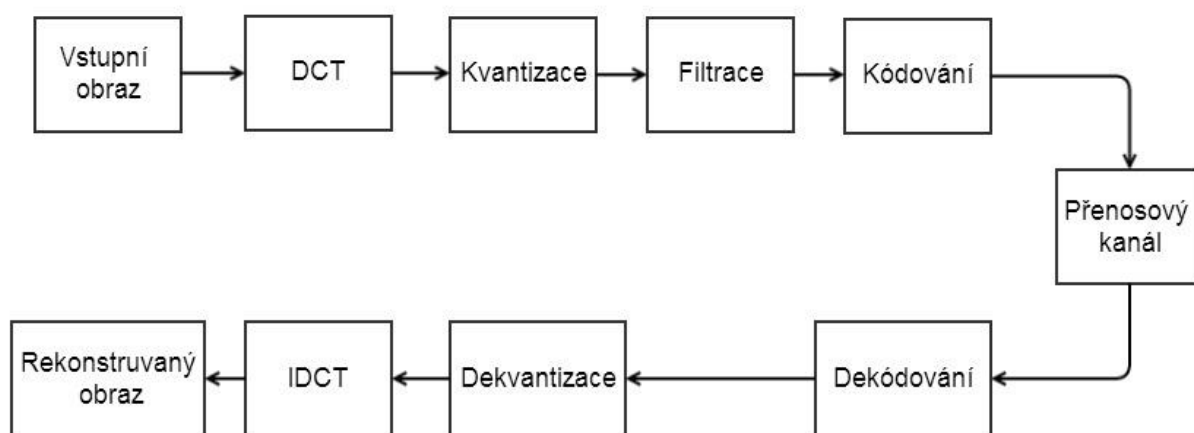
Další úsporu lze získat tak, že pro celý makroblok využijeme stejný exponent. To by znamenalo, že exponent stačí uložit a přenášet pouze jednou pro každý makroblok, nejčastěji tedy pro 64 obrazových bodů). Této úspory lze docílit několika metodami, například pomocí mediánu makrobloku, maxima v makrobloku nebo průměru v makrobloku. Všechny tři metody jsou použitelné, nicméně nejlepších výsledků dosahuje metoda maxima v makrobloku a to především u makrobloků 8x8 a větších (16x16, 32x32 atd.). Jednotlivé metody jsou porovnány pomocí PSNR, tedy špičkového odstupů signál-šum. Výsledky porovnání jednotlivých metod jsou v následujícím grafu, ze kterého vyplývá, že dále už se pracuje pouze s metodou maxima v makrobloku.



Obrázek 9 – Graf závislosti PSNR (dB) na velikosti makrobloků pro jednotlivé metody – modře metoda s 8 bity pro mantisu a 4 pro exponent, růžově metoda s 6 bity pro mantisu a 5 pro exponent. Převzato z [15].

Pro kódér bude tedy použita metoda maxima v makrobloku a optimalizace hodnot jasu v jednotlivých obrazových bodech. Pro realizaci bude postačující jednoduchá architektura s 8 až 10 bitovými A/D převodníky. Buď se využije výše popsaná metoda s 8 bity pro mantisu a 4 pro exponent nebo druhá metoda s 6 bity pro mantisu a 5 pro exponent (u této metody je nutný pátý bit exponentu, jinak by mohlo dojít k saturaci snímku a významné ztrátě informace. Obě metody jsou postačující pro uvažované modelové snímky s dynamickým rozsahem okolo 130 dB, jak vyplývá z následujícího obrázku. Podle očekávání se zde projevuje rozdíl přibližně 12 dB mezi oběma metodami, nicméně tato hodnota odpovídá úrovni kvantizační chyby, respektive kvantizačního zkreslení.

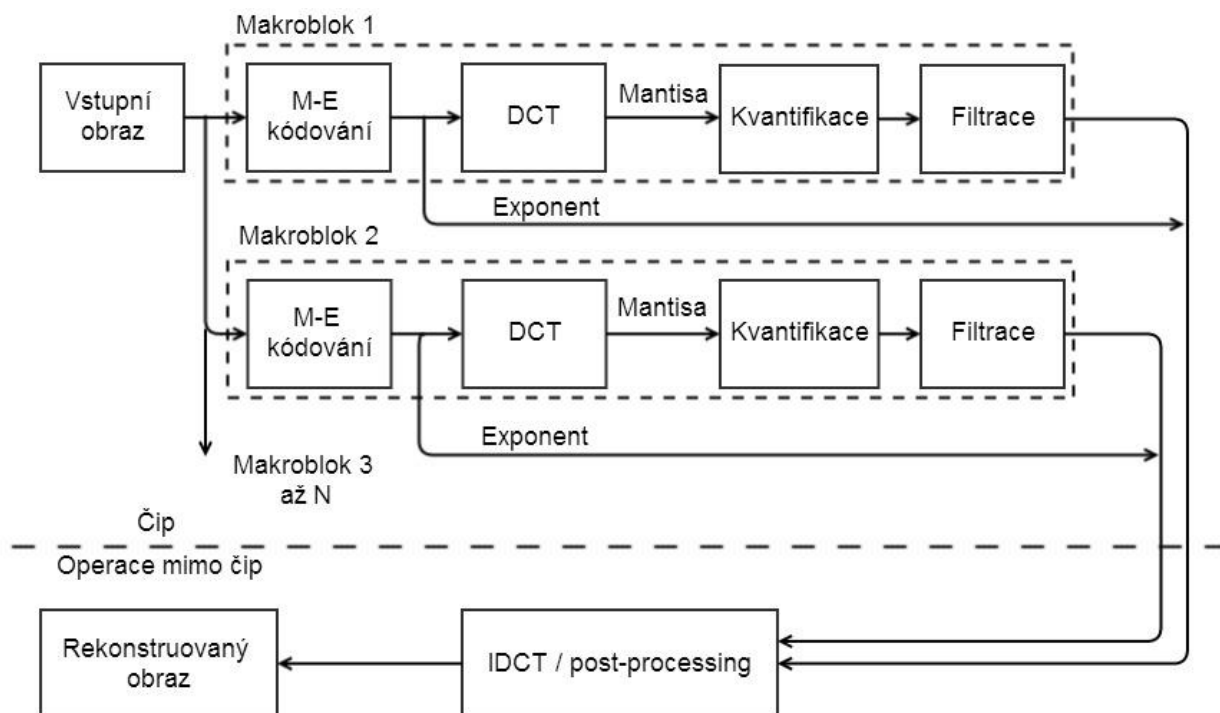
Druhý stupeň komprese využívá diskretní kosinovou transformaci. Jedná se o obdobu Fourierovy transformace, protože rovněž převádí signál z prostorové domény do frekvenční domény. U obrazových dat se využívá dvourozměrná diskretní kosinová transformace (2D-DCT), což se dá vyložit jako dvě oddělené jednorozměrné DCT, první pro data v řádcích, druhá pro data ve sloupcích. DCT se dá použít na celý snímek jako celek, bohužel však výpočetní náročnost dramaticky narůstá. Z tohoto důvodu se jako nejvhodnější jeví rozdělení celého snímku na makrobloky 8x8 obrazových bodů a použití DCT na jednotlivé makrobloky.



Obrázek 10 – Schéma postupu při kompresi a rekonstrukci pomocí DCT. Jedná se o princip, který je použit na každý makroblok zvlášť.

Výhoda této transformace spočívá v tom, že energie je koncentrována v několika koeficientech na nízkých frekvencích. To vede ke značné úspoře dat pro přenos a na tomto principu je založen kodér typu JPEG. Po filtraci se pak použijí pouze nenulové koeficienty v každém makrobloku. Princip této metody je potom znázorněn obrázkem 10. Použití tohoto postupu bez dalších úprav přímo v chipu by bylo neúnosně nákladné a výpočetně náročné. V kombinaci s prvním stupněm této metody se však dá dosáhnout uspokojivých výsledků i nižší výpočetní a prostorové náročnosti.

Každý makroblok se zpracovává nezávisle na ostatních, proto je 2D-DCT ideální ke kompresi dat v jednotlivých makroblocích. V prvním stupni této kompresní metody se vytvoří makroblok se společným exponentem a jediné rozdíly jsou v mantisách jednotlivých obrazových bodů. Navíc je tato forma komprese velice účinná při zachování všech informací potřebných k rekonstrukci obrazu. Společný exponent má značný přínos pro snížení výpočetní náročnosti při 2D-DCT. Postačí k tomu jediná podmínka, makrobloky v prvním a druhém stupni musí mít stejnou velikost a musí spolu korespondovat. DCT se provádí pouze u makrobloků obsahujících mantisy, neboť exponent se přenáší pouze jeden společný pro celý makroblok a tím dochází k obrovskému snížení množství bitů pro přenos (matematické operace s mantisou mohou probíhat nezávisle na exponentu a na konci dochází k opětovnému složení). V rámci DCT je také možnost uspořít velikost přenášených dat pomocí přenosu pouze významných koeficientů po DCT transformaci – viz obrázek 11.



Obrázek 11 – Schéma postupu při kompresi ve druhém stupni metody. Rozdělení na makrobloky, výpočetní algoritmy se společným exponentem (DCT probíhá pouze u makrobloků s mantisami).

Celý proces se tedy skládá ze dvou hlavních částí. Nejprve je provedeno kódování mantisa – exponent, kdy se použije pro nejběžnější HDR snímek (22 bitů pro jeden obrazový bod a kanál) komprese na 8 bitů pro mantisu a 4 bity pro exponent. Ve druhém kroku se provede komprese pomocí DCT. Ta se uplatní pouze na mantisy v makroblocích 32x32 obrazových bodů a exponent se přenáší pouze jednou pro celý makroblok.

Na obrázku 11 je naznačen postup komprese pomocí DCT, kdy se při DCT transformaci zvyšuje délka kódového slova. To není žádoucí jev pro další zpracování a tak je nutné snížit délku tak, aby odpovídala délce na vstupu. K tomu se používá blok „kvantifikace“. Kvantifikace po provedení DCT znamená, že se použijí pouze koeficienty 8 nejvýznamnějších bitů. Dalším krokem je filtrace, kde se zaměříme na výběr nízkofrekvenčních koeficientů po provedení DCT (v ideálním případě se použije entropický kodér, který vybere pouze nenulové koeficienty). To je hlavní rozdíl mezi touto metodou komprese a kompresí pomocí JPEG – metoda On-Chip komprese má stálý datový tok, zatímco u standardu JPEG se datový tok neustále mění tak, aby byl co nejmenší a zároveň zachoval kvalitu snímků nad prahovou hodnotou.

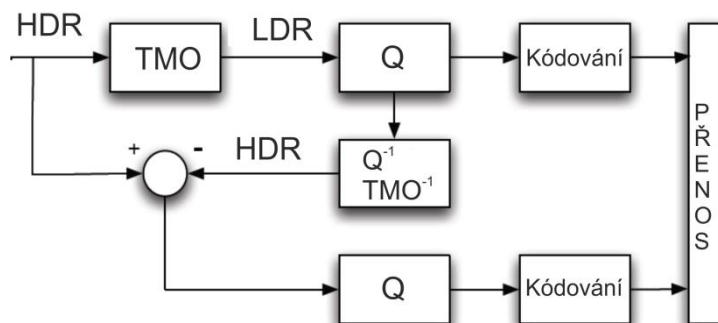
Vztah mezi počtem komponent, které prošly celým kompresním procesem, a počtem vstupních komponent se nazývá kompresní poměr. Do přenosového kanálu jsou poslány

pouze komponenty, které nesou potřebnou informaci (viz výše) a jeden exponent, ostatní části se nepřenáší. Rekonstrukce snímku pak probíhá v inverzním režimu. Nejdříve se rekonstruuje matice koeficientů a pak se provede inverzní DCT, čímž získáme zpět celý makroblok mantis. Ty se vynásobí společným exponentem a získáme tak rekonstruovaná data HDR snímku.

4.7 Dvoustupňová zpětně kompatibilní metoda

Dvoustupňová metoda komprese ([16], [17]) využívá převod do vlastního formátu podobného RGBE a přenáší dodatečné informace potřebné k přesnému dekompresnímu algoritmu. Originální informace je uložena jako LDR data po použití vybraného operátoru mapování jasu – v tomto případě buď tzv. Reinhardův operátor mapování jasu nebo tzv. Gamma operátor. Tato přepočítaná originální data jsou uložena ve formátu vycházejícím z RGBE, který využívá společný exponent pro popis jednotlivých bodů ve formátu čísla s plovoucí desetinnou čárkou. V tomto případě je mantisa složená z aproximace LDR snímku a společný exponent reprezentuje vrstvu dodatečných informací. To umožňuje rozdělení HDR snímku na barevný snímek (mantisu), a snímek ve škále odstínů šedi (exponent). Odstíny šedi dovedou zobrazit velmi jemné odchylky původního snímku a je možné je velice výrazně komprimovat pomocí klasických metod komprese. Výše uvedený postup vede ke dvěma důležitým vlastnostem komprese – vysoké efektivitě výpočtů a možnosti zobrazit snímek po kompresi pomocí klasických prohlížečů obrazových dat (není nutné snímek dále převádět nebo dekomprimovat pro potřeby zobrazení).

Dvoustupňová kompresní metoda nejprve provede mapování jasu na vstupní HDR snímek, což vede k vytvoření LDR snímku. Dále je pak tento proces proveden obráceně, čímž vznikne prediktivní HDR snímek. Nakonec vznikne residuální snímek, který nese dodatečné detailní informace a je velice důležitý při rekonstrukci snímku.



Obrázek 12 – Schema dvoustupňové komprese s využitím residuálního snímku. TMO je operátor mapování jasu, Q je kvantizace, HDR a LDR jsou naznačení dynamického rozsahu obrazových dat v daném bodě

Vlastní postup dvoustupňové komprese popisuje Boschetti [18]. Základní myšlenkou je odvození formátu pro kompresi dat z formátu RGBE, kde HDR data jsou reprezentována ve třech barevných kanálech RGB doplněných o společný exponent E :

$$E = \lceil \log_2 \max(\text{HDR}(r, g, b)) \rceil \quad (46)$$

kde HDR jsou obrazová data a E je společný exponent a

$$RGB = \left\lfloor \frac{2^N \cdot \text{HDR}}{2^E} \right\rfloor \quad (47)$$

Formát RGBE umožňuje kompaktní popis (4 byty na obrazový bod) původní informace, která byla ve formátu s plovoucí desetinnou čárkou (12 bytů na obrazový bod). Chybovost způsobená snížením počtu bytů pro každý obrazový bod je zanedbatelná. Na rozdíl od klasického formátu RGBE tato metody využívá v části RGB snímek po provedení mapování jasu (tedy LDR snímek). Takto mapovaný snímek může být na straně dekodéru rekonstruován pomocí prvních 3 kanálů RGBE reprezentace. HDR rekonstruovaný snímek je pak získán metodou podobnou metodě RGBE. Není možné použít pouze jednoduché vnoření LDR snímku do RGB matice, je nutné odhadnout exponent.

4.7.1 Odhad exponentu snímku (E)

První krok celého algoritmu je vytvoření LDR snímku ze vstupního HDR snímku s použitím jakéhokoli operátoru mapování jasu. Vstupní snímek nazveme I_{HDR} a následný LDR snímek označíme I_{LDR} a víme, že vzniká takto: $I_{LDR} = TMO(I_{HDR})$, kde TMO je zvolený operátor mapování jasu. Matici exponentu E pak získáme pomocí následujícího vztahu:

$$E = \text{mean}_{r,g,b} \left[\log_2 \left(\varepsilon + \frac{I_{HDR}}{1 + I_{LDR}} \right) \right], \quad (48)$$

kde operátor *mean* je průměr přes všechny 3 barevné kanály a ε je takové nenulové číslo, které zajistí, že argument logaritmu nikdy nedosáhne nuly. Tento vztah je tzv. vyhlazovací operátor, který uvažuje vztah mezi jasovou složkou a primárními barevnými kanály formátu RGBE. Výsledkem jsou hodnoty exponentu, které jsou nejčastěji rozloženy na relativně malém intervalu a je nutné provést odpovídající kvantizaci tak, aby nedocházelo k příliš vysoké chybě kvůli zaokrouhlování. Počet bitů reprezentujících exponent nazvěme N_B . Pak je exponent E dynamicky rozdělen do intervalu $\langle 0, 2^{N_B} - 1 \rangle$. Všechny hodnoty jsou zaokrouhleny na nejbližší celé číslo a zkomprimovány pomocí kodéru JPEG2000 při kompresním poměru R_E . Rozšířený a kvantizovaný exponent je nutný pro redukci kvantizační chyby vložené při zaokrouhlení:

$$E_q = \text{round} \left[(E - \min(E)) \frac{2^{N_B} - 1}{\max(E)} \right], \quad (49)$$

4.7.2 Odhad mantisy (RGB)

Snaha o minimalizaci chyby při rekonstrukci snímku vede k tomu, že RGB snímek je přepočítán a odhadnut s využitím hodnot exponentu E_q z kapitoly 4.7.1. Tento kvantizovaný exponent je využit k nejlepší aproximaci \hat{E} matice exponentu tak, že je přiřazen k původní matici exponentu E . Výsledkem je získání LDR snímku vydělením hodnot HDR snímku v jednotlivých obrazových bodech exponentem \hat{E} pomocí následujícího vztahu:

$$RGB = \frac{I_{HDR}}{2^{\hat{E}}} - 1 \quad (50)$$

Jasně patrný je vliv vztahu kvality LDR snímku na velikosti chyby vzniklé kvantizací exponentu E . Čím větší bude počet bitů N_B , tím lepší kvality bude dosahovat aproximovaný LDR snímek. Z toho se dá vyvodit také vztah mezi celkovou účinností komprese a kvalitou výsledného snímku. Celý proces pak končí kvantizací RGB snímku, jeho kompresí pomocí JPEG2000 kodéru při kompresním poměru R_{LDR} .

4.7.3 Dekompresní postup

Proces dekomprese opakuje inverzně kroky v procesu komprese. Nejprve je dekódován LDR snímek s kompresním poměrem R_{LDR} . Poté i snímek s exponenty projde dekompresí při dodržení podmínky kompresních poměrů $R \leq R_E$ a přenesení se do původního rozsahu hodnot. Použitím aproximovaných snímků \widehat{LDR} a \hat{E} na rekonstruovaný LDR snímek získáme zpět původní HDR snímek:

$$HDR = (1 + \widehat{LDR}) \cdot 2^{\hat{E}} \quad (51)$$

5 Objektivní metody pro hodnocení kvality obrazu

5.1 PSNR

Základním matematickým parametrem, který dokáže porovnat kvalitu dvou rozdílných obrazů je PSNR („Peak Signal-to-noise Ratio“, špičkový odstup signál-šum). Popisuje poměr mezi maximální možnou hodnotou signálu a hodnotou šumu, který snižuje celkovou kvalitu obrazu. Protože obrazové signály s vysokým dynamickým rozsahem dosahují širokého spektra hodnot, tento poměr se většinou definuje v logaritmické stupnici. Kvalita obrazu je velice subjektivní pojem, každý pozorovatel může stejnou kvalitu vnímat odlišně. Proto jsou důležité matematické popisy kvality obrazu, protože jsou objektivním měřítkem pro kvalitu obrazu.

Pro obrazová data je nutné nejdříve definovat střední kvadratickou chybu (MSE, „Mean Square Error“). Rozměry obou porovnávaných snímků musí být stejné. Pak je tedy MSE definována takto:

$$MSE = \frac{1}{mn} \sum_0^{m-1} \sum_0^{n-1} \|f(i, j) - g(i, j)\|^2 \quad (52)$$

kde $f(i, j)$ reprezentuje originální snímek, $g(i, j)$ komprimovaný snímek, m a n potom reprezentují počet řádků a sloupců matice obrazových dat.

Parametr PSNR je potom definován takto:

$$PSNR = 20 \log_{10} \left(\frac{max_f}{\sqrt{MSE}} \right) \quad (53)$$

kde max_f je maximální hodnota signálu, která se vyskytuje v originálním snímku. V tomto případě tedy MSE dokáže porovnat pravou hodnotu v jednotlivých obrazových bodech obou snímků. MSE reprezentuje průměrnou kvadratickou chybu mezi původním a komprimovaným snímkem. Velikost této chyby je rovna odchylce hodnot komprimovaného snímku od původních hodnot. Čím vyšší je hodnota PSNR, tím lépe lze rekonstruovat komprimovaný snímek. Tedy cílem je to, aby byla MSE co nejmenší.

Pokud se MSE počítá pro dva identické obrazy, pak je rovna 0. Tím pádem je hodnota PSNR nekonečná. Nevýhodou této metody pro objektivní hodnocení kvality obrazu je to, že uvažuje pouze srovnání hodnot v jednotlivých obrazových bodech a neuvažuje vlastnosti lidského zrakového ústrojí. Tato metoda je nevhodná pro hodnocení kvality HDR snímků a proto bylo nutné zvolit jiné níže popsané metody.

5.2 Univerzální index kvality obrazu

Matematické metody pro hodnocení kvality obrazu ve výše uvedených kapitolách jsou založeny na globálním přístupu k obrazu. Další metody se snaží využít vlastností lidského zrakového ústrojí k věrnějšímu hodnocení kvality obrazu. Tato metoda, jak ji popsal Wang [19], patří do druhé skupiny. Matematické metody jsou velmi používané z několika důvodů: mají jednoduchou implementaci, nízkou výpočetní náročnost a jsou nezávislé na pozorovacích podmínkách.

Je známo, že při subjektivních metodách mají pozorovací podmínky zásadní vliv na výsledky, ale málokdy jsou podmínky při testování kvality obrazu uvedeny ve výsledcích. Proto je nutné vytvořit objektivní popis nezávislý na pozorovacích podmínkách. V názvu metody je slovo univerzální – v tomto případě zastupuje univerzalitu ve smyslu nezávislosti této metody na testovaném obrazu, pozorovacích podmínkách nebo jednotlivých pozorovatelích. U metod PSNR a MSE jsou výsledky stále diskutabilní. Porovnání výsledků jednotlivých metod je uvedeno v samostatné kapitole.

Index kvality obrazu označíme písmenem Q . Potom uvažujme, že $\mathbf{x} = \{x_i | i = 1, 2, \dots, N\}$ a $\mathbf{y} = \{y_i | i = 1, 2, \dots, N\}$ jsou vzorky originálního obrazu a komprimovaného signálu. Navrhovaný index kvality má pak následující tvar:

$$Q = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)[(\bar{x})^2 + (\bar{y})^2]} \quad (54)$$

kde jsou jednotlivé prvky signálu a statistické parametry definovány následovně:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (55)$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (56)$$

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (57)$$

$$\sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2 \quad (58)$$

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \quad (59)$$

Rozsah hodnot Q je od -1 do +1, přičemž nejlepší je hodnota 1. Této hodnoty index nabývá pouze, pokud jsou oba obrazy identické ve všech obrazových bodech. Hodnoty -1 nabývá v případě, že $y_i = 2\bar{x} - x_i$ pro všechna $i = 1, 2, \dots, N$.

Pro lepší pochopení přepíšeme původní vztah takto:

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (60)$$

Vznikly tři separátní části, z nichž každá popisuje jinou měřenou vlastnost. První je korelační koeficient mezi x a y , který popisuje stupeň lineární korelace mezi x a y a jeho dynamický rozsah je v intervalu $\langle -1, 1 \rangle$. Hodnoty 1 dosahuje pouze, pokud platí $y_i = ax_i + b$ pro všechna $i = 1, 2, \dots, N$, kde „ a “ a „ b “ jsou konstanty a $a > 0$. I pokud jsou x a y lineárně závislé, stále mezi nimi mohou být drobné odchylky, které uvažujeme v rámci druhého a třetího členu. Druhý člen nabývá hodnot v intervalu $\langle 0, 1 \rangle$ a popisuje jak blízko je průměrný jas obou porovnávaných obrazů. Hodnoty jedna dosahuje v případě, že $\bar{x} = \bar{y}$. Třetí člen má podobnou úlohu, dosahuje hodnot ve stejném intervalu a hodnoty 1 dosáhne v případě, že $\sigma_x = \sigma_y$.

Pro hodnocení kvality obrazu je lepší počítat a porovnávat snímky lokálně a výsledek zkombinovat dohromady. Lokální hodnocení probíhá pro matici o definované velikosti $B \times B$ bodů, která se posouvá po jednotlivých obrazových bodech v řádcích a sloupcích. V každém kroku je vypočítán lokální index kvality Q_j . Pro celkový počet M kroků je pak celkový index kvality vypočítán takto:

$$Q = \frac{1}{M} \sum_{j=1}^M Q_j \quad (61)$$

Nejčastěji se využívá plovoucí okno o velikosti 8×8 obrazových bodů, což je obdoba výpočtu pro standard JPEG. Tuto metriku je možné ještě vylepšit, což popisuje následující kapitola.

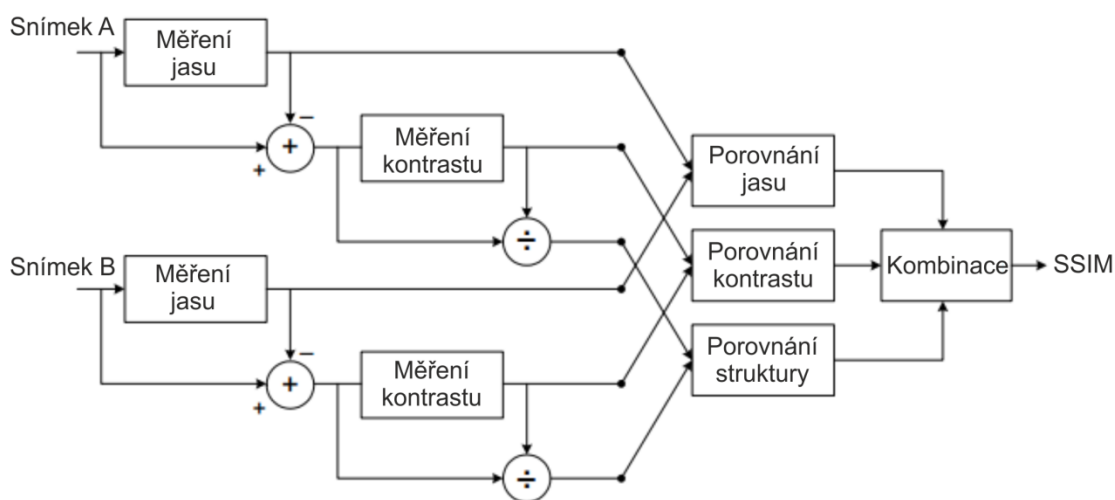
5.3 SSIM

Index strukturální podobnosti ([20], „Structural Similarity Index“) je matematický popis podobnosti dvou porovnávaných snímků s využitím informací o chování lidského zrakového systému. Jedná se o vylepšení předchozího univerzálního indexu kvality obrazu. Jas povrchu snímaného objektu je produktem daného osvětlení scény a odrazivosti povrchu objektu. Struktura objektu je však nezávislá na osvětlení snímané scény. Pokud tedy chceme zjistit

strukturální informaci snímku, musíme oddělit vliv osvětlení scény. Strukturální informaci snímku definujeme jako atributy reprezentující strukturu objektů ve scéně nezávisle na průměrném osvětlení a kontrastu. Protože se jas a kontrast v jednotlivých místech scény výrazně liší, v této definici budeme uvažovat lokální jas a lokální kontrast.

Tento matematický popis není příliš vhodný pro srovnání HDR snímků, ale je možné ho využít pro porovnání LDR snímků ve formátu JPEG. Vhodnou úpravou výpočtu indexu SSIM pro HDR data lze i tuto metodu považovat za dostatečně vypovídající o kvalitě snímku, což je vidět v kapitole srovnávající výsledky jednotlivých objektivních metod porovnání kvality obrazu.

Schéma postupu hodnocení kvality snímku je uvedeno na obrázku 16. Uvažujme nyní dva nezáporné obrazové signály x a y , které jsou k sobě přesně zarovnané. Pokud uvažujeme jeden signál jako perfektní (originál), pak se při měření podobnosti jedná o kvantitativní popis kvality druhého signálu. Celý proces je rozdělen do tří částí: porovnání jasu, porovnání kontrastu a porovnání struktury.



Obrázek 13 – Schéma principu výpočtu indexu strukturální podobnosti. Překresleno z[20].

Nejprve je tedy porovnán jas obou signálů. Pokud uvažujeme diskrétní hodnoty, pak lze vypočítat průměrnou hodnotu intenzity jasu:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i \quad (62)$$

Funkce pro porovnání jasu $l(x, y)$ je pak funkcí μ_x a μ_y .

Dále je nutné odebrat průměrnou intenzitu ze signálů. V diskretní podobě bude výsledný signál $\mathbf{x} - \mu_x$ korespondovat s projekcí vektoru \mathbf{x} do roviny definované takto:

$$\sum_{i=1}^N x_i = 0 \quad (63)$$

Použitím standardní odchylky určíme odhad kontrastu signálu. Nevážený odhad v diskretní podobě vypadá následovně:

$$\sigma_x = \sqrt{\left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2\right)} \quad (64)$$

Pro porovnání kontrastu $c(\mathbf{x}, \mathbf{y})$ se využije porovnání σ_x a σ_y .

V rámci třetího kroku je signál normován svojí vlastní standardní odchylkou tak, že oba porovnávané signály mají svojí standardní odchylku. Strukturální porovnání $s(\mathbf{x}, \mathbf{y})$ je provedeno na normovaných signálech $(\mathbf{x} - \mu_x)/\sigma_x$ a $(\mathbf{y} - \mu_y)/\sigma_y$.

Výsledná funkce v sobě zahrnuje všechny tři komponenty:

$$S(\mathbf{x}, \mathbf{y}) = f(l(\mathbf{x}, \mathbf{y}), c(\mathbf{x}, \mathbf{y}), s(\mathbf{x}, \mathbf{y})) \quad (65)$$

Jednotlivé složky jsou na sobě relativně nezávislé. Například změna jasu a kontrastu neovlivní strukturu snímku. Pro přesný matematický popis vztahu 59 musíme definovat jednotlivé funkce (\mathbf{x}, \mathbf{y}) , $c(\mathbf{x}, \mathbf{y})$, $s(\mathbf{x}, \mathbf{y})$ i výslednou funkci $f(l(\mathbf{x}, \mathbf{y}), c(\mathbf{x}, \mathbf{y}), s(\mathbf{x}, \mathbf{y}))$. Pro tyto definice platí následující podmínky:

- Symetrie: $S(\mathbf{x}, \mathbf{y}) = S(\mathbf{y}, \mathbf{x})$;
- Ohraničení: $S(\mathbf{x}, \mathbf{y}) \leq 1$;
- Unikátní maximum: $S(\mathbf{x}, \mathbf{y}) = 1$ pro $\mathbf{x} = \mathbf{y}$ (v diskretní reprezentaci $x_i = y_i$ pro všechny $i = 1, 2, \dots, N$);

Pro srovnání jasu definujeme vztah takto:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (66)$$

kde C_1 je konstanta, která zabraňuje nestabilitě výrazu při $\mu_x^2 + \mu_y^2$ blízcím se nule. Stanovení konstanty:

$$C_1 = (K_1 L)^2 \quad (67)$$

kde L je dynamický rozsah obrazových bodů (255 pro 8 bitové šedotónové snímky) a $K_1 \ll 1$.

Podobné podmínky platí pro porovnání kontrastu i struktury. Rovnice 66 kvalitativně odpovídá Weberovu zákonu, který se hojně využívá pro modelování světelné adaptace lidského zrakového ústrojí („luminance masking“). Podle tohoto zákona je magnituda nejmenší jasové změny, kterou je lidské oko schopno zaregistrovat, přibližně stejná jako jasové pozadí škály použitých hodnot jasu. Jinými slovy lidské zrakové ústrojí vnímá relativní změny jasu, nikoli absolutní změny jasu. Pokud R reprezentuje velikost změny jasu v porovnání s jasovým pozadím, pak můžeme napsat jasový signál takto: $\mu_y = (1 + R)\mu_x$. Použitím na vztah (66) dostaneme následující vztah:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2(1 + R)}{1 + (1 + R)^2 + C_1/\mu_x^2} \quad (68)$$

Navíc je možné uvažovat, že C_1 je dostatečně malé v porovnání s μ_x^2 a tak celý člen můžeme ignorovat. Pak $l(\mathbf{x}, \mathbf{y})$ je pouze funkcí R a kvalitativně odpovídá Weberovu zákonu.

Porovnání kontrastu má podobnou formu:

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (69)$$

kde $C_2 = (K_2L)^2$ a $K_2 \ll 1$. Tato definice opět vyhovuje všem třem výše uvedeným, nutným a postačujícím podmínkám. Zajímavou vlastností této funkce je to, že množství změny kontrastu $\Delta\sigma = \sigma_y - \sigma_x$ vykazuje menší sensitivitu pro případy vysokého kontrastu základního snímku. To odpovídá vlastnosti lidského zrakového ústrojí známé jako „contrast masking“.

Porovnání struktury je poslední dílčí krok celé operace a provádí se s pomocí normování dvou předchozích parametrů. Sjednotíme dva vektory $(\mathbf{x} - \mu_x)/\sigma_x$ a $(\mathbf{y} - \mu_y)/\sigma_y$ tak, že je položíme do stejné roviny, kterou využívají výše uvedené vztahy. Korelace mezi těmito dvěma vektory je jednoduchým a efektivním postupem pro objektivní porovnání kvality strukturální podobnosti.

Ze vztahů je jasné patrné, že korelace mezi $(\mathbf{x} - \mu_x)/\sigma_x$ a $(\mathbf{y} - \mu_y)/\sigma_y$ je srovnatelná s korelačním koeficientem mezi \mathbf{x} a \mathbf{y} . Potom tedy můžeme odvodit následující funkci pro popis strukturálního porovnání:

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (70)$$

Stejně jako v předchozích případech je i zde nutnost použít konstantu v čitateli i jmenovateli celého vztahu. V diskrétní formě můžeme σ_{xy} určit takto:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (71)$$

Geometricky lze popsat korelační koeficient jako kosinus úhlu mezi vektory $\mathbf{x} - \mu_x$ a $\mathbf{y} - \mu_y$. Funkce $s(\mathbf{x}, \mathbf{y})$ také může nabývat záporných hodnot.

Kombinací tří základních funkcí pro popis jednotlivých srovnávaných prvků získáme funkci nazvanou „index strukturální podobnosti“ (SSIM) signálů \mathbf{x} a \mathbf{y} :

$$SSIM(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha \cdot [c(\mathbf{x}, \mathbf{y})]^\beta \cdot [s(\mathbf{x}, \mathbf{y})]^\gamma \quad (72)$$

kde $\alpha > 0$, $\beta > 0$ a $\gamma > 0$ jsou parametry použité k úpravě relativní váhy jednotlivých složek. V rámci zjednodušení dalšího procesu můžeme v našem případě uvažovat $\alpha = \beta = \gamma = 1$ a $C_3 = C_2/2$. Pak dostaneme zvláštní případ indexu strukturální podobnosti, který vypadá následovně:

$$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (73)$$

Praktické využití parametru SSIM je popsáno v kapitole zabývající se výsledky a zhodnocením kvality výsledných obrazů. Pro potřeby porovnání obrazů je ve většině případů výhodnější a přesnější využít index SSIM lokálně. Poruchy v obraze mohou záviset na lokálních vlastnostech obrazu, ale mohou být také prostorově variantní. Dále také při optimální pozorovací vzdálenosti může lidské oko pozorovat pouze lokální změny kontrastu a struktury v celém snímku s vysokým dynamickým rozsahem.

Lokální hodnocení kvality může přinést prostorově variabilní mapu kvality obrazu, která nese více informace o snížení kvality v rámci komprese obrazu. Ve výše uvedených vztazích se využívá statistických prvků μ_x , σ_x a σ_{xy} , s využitím lokální matice 8×8 obrazových bodů. Tato matice se posouvá vždy o jeden obrazový bod a takto se projde celý snímek. V každém bodě jsou vypočítány statistické parametry a index SSIM. To se v některých případech jeví jako nevýhodné pro lokální popis kvality – vznikají různé nežádoucí chyby na přechodu bloků apod. Proto se jako vhodnější jeví použití kruhově symetrické Gaussovské váhovací funkce $\mathbf{w} = \{w_i | i = 1, 2, \dots, N\}$ o rozměrech 11×11 , se standardní odchylkou 1,5 vzorků, normovanou na jednotkovou sumu $\sum_{i=1}^N w_i = 1$.

Odhady jednotlivých statistických parametrů μ_x , σ_x a σ_{xy} se upraví následovně:

$$\mu_x = \sum_{i=1}^N w_i x_i \quad (74)$$

$$\sigma_x = \sqrt{\left(\sum_{i=1}^N w_i (x_i - \mu_x)^2 \right)} \quad (75)$$

$$\sigma_{xy} = \sum_{i=1}^N w_i (x_i - \mu_x)(y_i - \mu_y) \quad (76)$$

Tento váhovaný proces vede k lepším výsledkům a lokální mapa kvality obrazu je přesnější. V praxi je často potřebný také celkový popis kvality obrazu vyjádřený pomocí jednoho čísla. Používá se průměrný index SSIM nazvaný „MSSIM“ (Mean Structural Similarity Index):

$$MSSIM(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{j=1}^M SSIM(\mathbf{x}_j, \mathbf{y}_j) \quad (77)$$

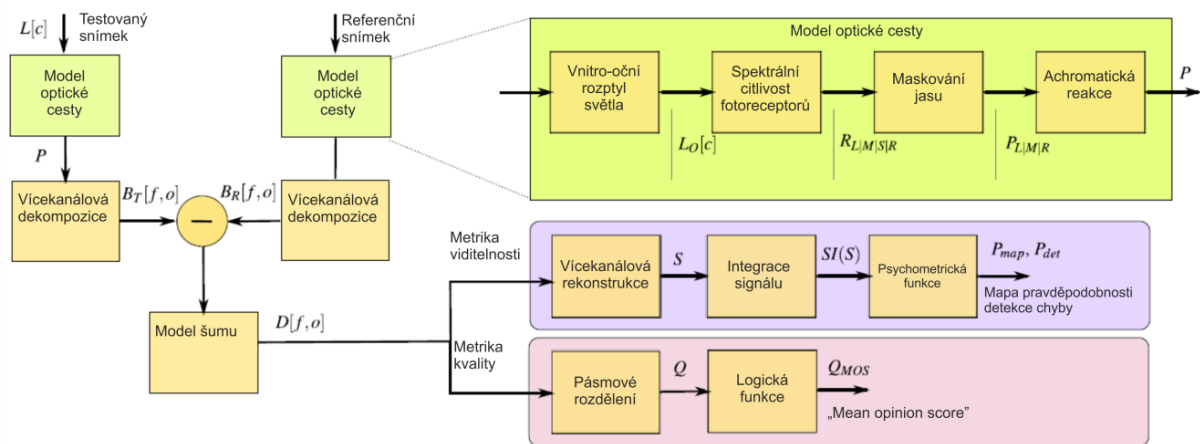
kde \mathbf{X} a \mathbf{Y} jsou originální a komprimovaný snímek, \mathbf{x}_j a \mathbf{y}_j jsou části obou snímků v j -tém lokálním okně v rámci operace výpočtu MSSIM a M je počet lokálních oken použitých při výpočtu. Také je možné vypočítat váhovaný průměr různých vzorků z mapy kvality obrazu. Toho se využívá například, pokud je část obrazu výrazně odlišná od ostatních a lidské oko ji tak vnímá výrazněji než zbylé části. V praxi je však nejpoužívanější rovnoměrné váhování, tedy klasický průměr.

5.4 HDR-VDP-2

HDR-VDP-2 („Visual Difference Predictor“, [22]) je metrika popisující odlišnosti dvou porovnávaných snímků (referenčního a testovaného). Tato metrika je založená na modelu lidského zrakového ústrojí a snaží se jej co nejvěrněji napodobit. HDR-VDP-2 se skládá ze dvou metrik popisujících kvalitu obrazu. Jsou to viditelnost a kvalita. Viditelností se rozumí pravděpodobnost, s jakou je průměrný pozorovatel schopen rozpoznat odlišnost mezi oběma snímky. Kvalitou se rozumí popis, jak moc je testovaný obraz odlišný (degradovaný) od originálu.

Na obrázku 14. je naznačen postup výpočtu obou metrik. Vstupem jsou dva snímky, jeden jako referenční a druhý jako testovaný. Oba snímky projdou nejprve blokem, který simuluje průchod světla lidským okem (zde je nazván „optická cesta“). V dalším bloku je

provedena vícerozměrná dekompozice a oba snímky jsou od sebe odečteny. Na tento rozdílový snímek se použije šumový model, který obsahuje dva druhy šumu (šum závislý na signálu a šum nezávislý na signálu). Další část je pro každou metriku rozdílná. Pro viditelnost se nejprve určí pravděpodobnost všech dále prováděných operací a provede se vícerozměrná rekonstrukce. Velké objekty a vzory se snadno detekují díky bloku prostorové integrace. Posledním krokem je použití psychometrické funkce, které vede k výsledné hodnotě P_{det} (střední pravděpodobnosti detekce rozdílu mezi oběma snímky) a mapě této pravděpodobnosti. Pro určení kvality je nutné vycházet z mapy pravděpodobnosti (vzniklé při určení viditelnosti) a také využít rozdělovací funkce.



Obrázek 14 – Blokové schéma dvoukanálové metriky pro určení viditelnosti a kvality obrazu (HDR-VDP-2). Převzato z [22].

Blok optické cesty se skládá z několika dílčích bloků, které jsou detailně popsány v článku [22]. První simuluje vnitro oční rozptyl světelných paprsků. Určitá část paprsků světla se rozptýlí při průchodu rohovkou, čočkou, ale také přímo v komoře a na sítnici. Rozptyl se projevuje nejvíce u vyšších prostorových frekvencí a výsledkem je nižší kontrast světla dopadajícího na sítnici. Celý tento blok je pak popsán modulační přenosovou funkcí:

$$MTF = \sum_{k=1}^4 a_k e^{-b_k \rho} \quad (78)$$

kde ρ je prostorová frekvence a a_k a b_k jsou parametry modelující vlastnosti rozptyl a jsou podrobně definovány v příloze článku [22].

Druhý blok popisuje spektrální citlivost fotoreceptorů, tedy pravděpodobnost s jakou fotoreceptor zaznamená foton s příslušnou vlnovou délkou. Křivky spektrální citlivosti jsou definovány pro všechny typy čípků a pro tyčinky. Příklad použitých křivek je uveden v článku [22]. Poslední dva bloky jsou nazvány „maskování jasu“ a „achromatická reakce“. Blok maskování jasu vyjadřuje vlastnost fotoreceptorů regulovat citlivost v závislosti na

intenzitě dopadajícího světla. Fotoreceptory využívají tuto vlastnost lokálně, takže pro přesné hodnocení kvality obrazu je potřeba použít lokální model dle Mantiuka [24] (model mechanismu pro přizpůsobení fotoreceptorů dle vstupní úrovně jasu). Blok achromatická reakce je pouze součtem příspěvků od tyčinek a čípků.

Blok „vícekanálová dekompozice“ napodobuje děj, probíhající v části mozku zodpovědné za interpretaci podnětů z fotoreceptorů. Informace jsou na maximální možný počet prostorových frekvencí podle rozlišení snímku. Blok „model šumu“ se skládá ze dvou částí. Jednou je signálově nezávislý šum (reprezentovaný neurální CSF, tedy funkcí kontrastní citlivosti) a druhou je signálově závislý šum (maskování kontrastu). Podrobný popis obou šumových modelů popisuje Mantiuk [22].

Dále se pracuje v kanálech určených přímo pro danou metriku. Část, která vede k výpočtu metriky viditelnost, využívá nejprve celkovou pravděpodobnost pro výše uvedené mechanismy na všech prostorových frekvencích. Tuto pravděpodobnost definoval Daly [25]:

$$P_{map} = 1 - \prod_{(f,o)} (1 - P(f, o)) \quad (79)$$

kde $P(f, o)$ je pravděpodobnost vypočtená dle psychometrické funkce podle Dalyho [25]:

$$P(f, o) = 1 - \exp(\log(0,5)D^\beta(f, o)) \quad (80)$$

kde β je sklon psychometrické funkce a v tomto případě je tento sklon $\beta = 1$ a $D(F, o)$ je kontrast na jednotlivých prostorových frekvencích.

Kanál, který vede ke zjištění kvality obrazu (MOS, „Mean opinion score“), se snaží vytvořit hodnocení kvality podobné subjektivnímu hodnocení velkého počtu pozorovatelů [26]. Vstupní prvky celého procesu jsou velikost zobrazovacího zařízení a pozorovací vzdálenost. Všechny obrazové body je nutné rozdělit podle rozdělovací funkce:

$$Q = \frac{1}{F \cdot O} \sum_{f=1}^F \sum_{o=1}^O w_f \log \left(\frac{1}{I} \sum_{i=1}^I D_i^2(f, o) + \varepsilon \right) \quad (81)$$

kde i je index popisující pořadí obrazového bodu, ε je konstanta, která zabraňuje nulovým hodnotám kontrastu $D(f, o)$. Dále I je celkový počet obrazových bodů ve snímku a w_f váhová funkce jednotlivých pásem prostorových frekvencí. Výsledné hodnocení kvality obrazu dle [26] je potom definováno takto:

$$Q_{MOS} = \frac{100}{1 + \exp(q_1(Q + q_2))} \quad (82)$$

kde q_1 a q_2 jsou váhovací koeficienty definované v normě ITU [26].

5.5 Index kvality mapování jasu

Index kvality mapování jasu (TMQI) je metrika, která je speciálně navržena pro porovnání dvou snímků, které nemají stejný nebo podobný dynamický rozsah. Je tedy vhodná pro srovnání zpětně kompatibilních metod komprese HDR obrazových dat, při které je nutné srovnat originální HDR snímek s komprimovaným LDR snímkem. Protože všechny kompresní metody, které jsou v rámci této práce zpracovány prostředím Matlab, obsahují operátory mapování jasu, je vhodné využít i tuto metriku pro hodnocení kvality obrazu.

Tato metrika vychází z indexu strukturální podobnosti (kapitola 5.3) a používá se pro celý snímek jako lokální. Pokud uvažujeme, že x a y jsou lokální úseky z HDR snímku a z LDR snímku po provedení mapování jasu, pak definujeme strukturální podobnost takto:

$$S_{local}(x, y) = \frac{2\sigma_x\sigma_y + C_1}{\sigma_x^2 + \sigma_y^2 + C_1} \cdot \frac{\sigma_{xy} + C_2}{\sigma_x\sigma_y + C_2} \quad (83)$$

kde σ_x , σ_y a jsou standardní odchylky a σ_{xy} je vzájemná korelace mezi oběma úseky HDR a LDR snímku. C_1 a C_2 jsou konstanty, které zabraňují, aby čitatel nebo jmenovatel byl nulový. Pokud tento vztah srovnáme s indexem SSIM, pak chybí člen, který srovnává jasovou složku a strukturální člen je stejný. Jediný rozdíl je ve srovnání hodnot, kterých mohou oba snímky nabývat. Je nutné rozlišit, jestli je lokální odchylka obou snímků významná nebo nevýznamná (tedy taková, že neovlivní výsledek). K tomu se používá nelineární mapování standardní odchylky tak, že významné složky se mapují na hodnotu 1 a nevýznamné na hodnotu 0 s plynulým přechodem mezi nimi. Celý tento postup je detailně popsán v článku [27].

Výpočet lokální strukturální podobnosti je pak pomocí plovoucího okna použit na celý snímek. Dále se také využívá víceúrovňová metoda podobná jako u indexu strukturální podobnosti. Při každém průchodu se změní měřítko tak, aby bylo možné odhalit co možná největší množství odchylek mezi oběma snímky. Výsledkem je mapa, která popisuje lokální odchylky strukturální podobnosti obou snímků. Pro každé měřítko je mapa průměrována:

$$S_l = \frac{1}{N_l} \sum_{i=1}^{N_l} S_{local}(x_i, y_i) \quad (84)$$

kde x_i a y_i jsou i -té části porovnávaných snímků a N_l je počet částí v l -té vrstvě. Celková strukturální podobnost je pak vypočtena takto:

$$S = \prod_{l=1}^L S_l^{\beta_l} \quad (85)$$

kde L je celkový počet měřítek a β_l je váha přiřazená každému měřítku.

6 Popis algoritmů v programovém prostředí Matlab

Celý program je uveden v příloze na digitálním nosiči, který je součástí této práce. Hlavní ovládací algoritmy jsou v základní složce celého programu. Jedná se o skripty *Main.m*, *MainHDRVDP2.m*, *MainTMQI.m* a *MainSSIM.m*. Tyto skripty pak využívají ke své činnosti další funkce uvedené v této složce. Skript *Main.m* je hlavní částí celého programu. Jeho součástí je načtení vstupního snímku, vytvoření dalších upravených názvů pro výstupní snímky, vlastní provedení kompresních algoritmů a zapsání výstupních dat v požadovaném formátu po kompresi. Jednotlivé části jsou odděleny a v další fázi je pak provedena zpětná dekomprese pomocí převrácených postupů komprese. Jednotlivé kroky mohou v závislosti na testovaném snímku a výkonu hardwarového vybavení trvat minuty až desítky minut, proto je po každém kroku uveden očíslovaný stav provedené části algoritmu.

Skripty *MainHDRVDP2.m*, *MainTMQI.m* a *MainSSIM.m* jsou ovládací prvky pro testování kvality obrazových dat po kompresi. Každá použitá objektivní metoda má svoji vlastní funkci. Skript *MainHDRVDP2.m* využívá funkce převzaté z [22]. Pro správnou funkci tohoto skriptu je nutné mít v programovém prostředí Matlab implementován doplněk „*matlab Pyr Tools*“, dostupný z [28]. Jedná se o doplňkové programové prostředí a funkce vytvořené v jazyce C++. Tento doplněk je nutné nakompilovat do jádra prostředí Matlab, v UNIXových systémech to jde snadno pomocí jednoduchého skriptu, v systémech typu Windows je kompilace složitější. Pokud není toto prostředí nakompilované do systému, pak je výpočetní náročnost daleko vyšší (každý soubor typu .mex si musí Matlab přeložit sám pro každý výskyt v dané funkci). Výpočetní náročnost u snímků HDR je vysoká i při běžných operacích. Skripty *MainTMQI.m* a *MainSSIM.m* jsou obdobně navržené pro načtení jednotlivých snímků a provedení výpočtu objektivních parametrů kvality obrazu.

Nejdůležitější funkce celého programu jsou funkce pro vybrané kompresní metody (popsané v kapitolách 4.2, 4.4 a 4.7). První z nich je nazvaná kompreseJPEGHDR.m, vycházející z principu uvedeném v kapitole 4.2 a dále v [8],[9] a [10]. Část tohoto kódu je uvedena v digitální příloze knihy [9] a pro potřeby této práce byl kód upraven a rozšířen. Její vstupy jsou testovaný snímek HDR, název výstupního snímku a nastavení kvality komprese JPEG. Funkce nejprve načte vstupní HDR snímek, provede mapování jasu pomocí Reinhardova operátoru a vytvoří si poměrový snímek „ratio“. Ten je důležitý při celé fázi komprese. Funkce k němu opakovaně přistupuje a provedení každého cyklu ho přepíše na stále menší a menší měřítko. Výsledný snímek je pak upraven podle operátoru mapování jasu, parametrů tohoto operátoru (např. gamma snímku). Důležitou součástí této funkce je ošetření

extrémních hodnot tak, že každý bod výsledného snímku může nabývat pouze hodnot v daném rozmezí. Výstupem funkce jsou dva soubory ve formátu JPEG. Jeden je samotný komprimovaný snímek a druhý je zbytkový snímek „ratio“.

Funkce *kompreseHDRJPEG2000.m* je variantou ztrátové komprese s pomocí kodéru typu JPEG2000 tak, jak je naznačeno v kapitole 4.4 a dále uvedeno v [9] a [11]. Tato komprese může být při určitém nastavení kompresního poměru považována za bezztrátovou. Vstupem je testovaný snímek, název výstupního souboru a kompresní poměr pro kompresi JPEG2000. Funkce nejprve zredukuje dynamický rozsah tak, aby se vešel do 16 bitového formátu JPEG2000 a pak přepočítá celý snímek v jednotlivých kanálech dle vztahů uvedených v popisu této metody. Výstupem je pak snímek ve formátu JPEG2000.

Funkce *kompreseDvoustupnova.m* je funkcí, která provede kompresi dle Boschettiho [18]. Vstupem je testovaný snímek, název výstupního snímku, kompresní poměr pro exponent, kompresní poměr pro mantisu a dále nepovinně také počet bitů pro JPEG2000 a operátor mapování jasu. Funkce nejprve zjistí, jestli má všechny potřebné vstupní parametry a pokud ne, tak použije předdefinované hodnoty. Pak provede mapování jasu, v tomto případě pomocí Reinhardova operátoru a operátoru Gamma. Provede kvantizaci, provede výpočet odhadu exponentu, zkomprimuje výsledný exponent, uloží jako soubor s příponou „_bos_E.jp2“. Pak se exponent dekóduje, vypočítá se mantisa a uloží se jako snímek s příponou „_bos_RGB_.jp2“.

Funkce dekomprese mají obdobné názvy jako ty kompresní a jejich součástí je proces inverzní k těm, které byly provedeny při kompresi. Dále pak program využívá řadu podpůrných funkcí. Jedná se například o funkce *tmoReinhard.m*, *tmoReinhardAlpha.m*, *tmoReinhardFiltr.m*, *tmoReinhardGauss.m* a *tmoReinhardBila.m*. Jsou to funkce, které provádí mapování jasu dle Reinhardova lokálního operátoru. Funkce *tmoGamma.m* provádí mapování jasu pomocí Gamma operátoru.

Další podpůrné funkce jsou např. *zmenaJasu.m*, která změní hodnotu jasového kanálu v celém snímku na novou hodnotu dle vstupních parametrů. Dále pak *odebratExtremy.m*, což je důležitá funkce, která zabrání zápisu hodnoty „NaN“ a „Inf“ číselného formátu s plovoucí desetinnou čárkou do snímku. Funkce *ssim.m* je určena pro výpočet hodnoty indexu strukturální podobnosti. Důležité funkce pro zápis a čtení HDR a LDR snímků jsou *hdrimread.m*, *hdrimwrite.m* a *ldimread.m* převzaté z digitální přílohy [9]. Jedná se o upravené funkce pro čtení a zápis HDR dat v Matlabu. Jsou vhodné pro tyto metody kompresí a také jsou výpočetně efektivnější. Všechny funkce mají v hlavičce popsány své vlastnosti a vstupní a výstupní hodnoty.

7 Hodnocení výsledků a kvality snímků

Komprese a hodnocení kvality probíhalo na sadě testovacích snímků. Jako zdroj posloužily databáze HDR snímků Empa [21] a Fairchildova databáze [22]. Vybrané snímky jsou různorodou směsí snímků s vysokým dynamickým rozsahem, zvláštní důraz byl kladen na obsah snímků – hlavně na to, aby bylo možné testovat a posuzovat kvalitu komprese. Názvy snímků jsou použity z obou databází a názvy souborů na datovém nosiči jsou stejné jako názvy uvedené v této práci.



Obrázek 15 – „Cafeteria“ originální snímek (pro zobrazení provedeno mapování jasu pomocí Reinhardova operátoru). Převzato z [21] **Chyba! Nenalezen zdroj odkazů.**



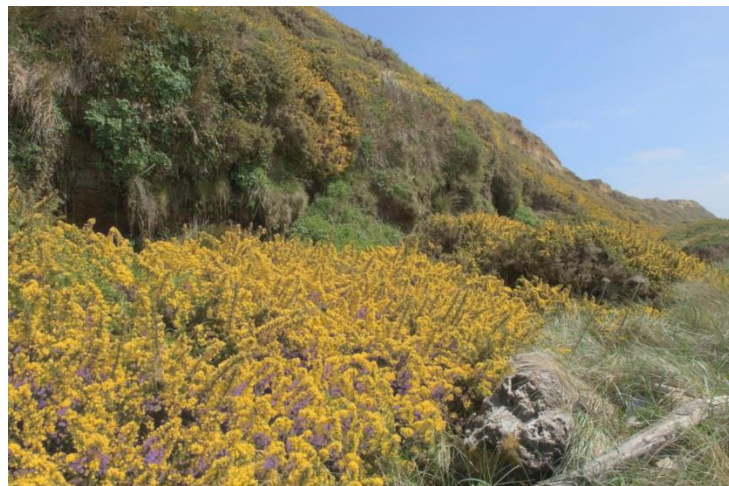
Obrázek 16 – „Market“ originální snímek (pro zobrazení provedeno mapování jasu pomocí Reinhardova operátoru). Převzato z[21].



Obrázek 17 – „Mont Saint Michel“ originální snímek (pro zobrazení provedeno mapování jasu pomocí Reinhardova operátoru). Převzato z [21].



Obrázek 18 – „Zurich“ originální snímek (pro zobrazení provedeno mapování jasu pomocí Reinhardova operátoru). Převzato z [21].



Obrázek 19 – „BloomingGorse“ originální snímek (pro zobrazení provedeno mapování jasu pomocí Reinhardova operátoru). Převzato z [22].

7.1 Kompresní metoda HDR-JPEG2000

Metoda komprese využívající formát JPEG2000, je vysoce efektivní metodou pro ztrátovou i bezztrátovou kompresi. Rekonstruované snímky v bezztrátové verzi identické s originálním snímkem na vstupu kompresní metody. Z následujících tabulek a grafů je vidět zajímavý jev. Výsledné skóre po kompresi pomocí metriky HDR-VDP-2 je nečekaně nízké i přesto, že provedení dekomprese jsou snímky rekonstruovány s vysokou přesností (pro kompresní poměr menší než 20 jsou snímky identické).



Obrázek 20 – Snímek „MontSaintMichel“ po provedení komprese HDR-JPEG2000 s kompresním poměrem 5.



Obrázek 21 – Snímek „MontSaintMichel“ po provedení dekomprese metodou HDR-JPEG2000 a pro potřeby zobrazení provedeno mapování jasu pomocí Reinhardova operátoru.

Tato metoda využívá vlastní kompresní mechanismus, při kterém snímek po kompresi má menší kontrast a neposkytuje příliš věrné podání barev. Při dekompresi se však barevné podání inverzní operací vrátí zpět. Na dalším obrázku je ukázka snímku „MontSaintMichel“ po kompresi s kompresním poměrem 5 a pak také výsledný snímek po zpětné dekompresi se stejným poměrem. Výsledné snímky jsou přiloženy v digitální příloze této práce. Pro porovnání byla zvolena škála kompresního poměru JPEG2000 (2, 5, 20, 50, 200, 500, 999).

Tabulka 1: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG2000 (fotografie „Market“)

Popis veličin: *mSSIM* je průměrná hodnota indexu strukturální podobnosti v celém snímku, *MOS* je „Mean Opinion Score“ a *P_{det}* pravděpodobnost detekce rozdílu mezi dvěma snímky podle metody HDR-VDP-2 [24], *TMQI* je index kvality mapování jasu a *TMQI-s* je strukturální složka indexu kvality mapování jasu podle metody [27].

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P _{det} (-)	TMQI	TMQI-s
2	24,00	0,999	66,91	1	0,793	0,966
5	9,60	0,999	66,90	1	0,793	0,965
20	2,40	0,998	66,88	1	0,787	0,944
50	0,96	0,994	66,61	1	0,779	0,911
200	0,24	0,968	63,84	1	0,755	0,824
500	0,10	0,923	57,42	1	0,729	0,733
999	0,05	0,868	52,12	1	0,702	0,667

Tabulka 2: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG2000 (fotografie „Cafeteria“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P _{det} (-)	TMQI	TMQI-s
2	24,00	0,999	75,41	1	0,785	0,935
5	9,60	0,999	75,41	1	0,784	0,933
20	2,40	0,993	75,31	1	0,771	0,882
50	0,96	0,985	74,36	1	0,762	0,847
200	0,24	0,963	67,29	1	0,742	0,779
500	0,10	0,938	59,84	1	0,724	0,716
999	0,05	0,909	53,40	1	0,702	0,649

Tabulka 3: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG2000 (fotografie „BloomingGorse“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI	TMQI-s
2	24,00	0,999	72,48	1	0,791	0,958
5	9,60	0,997	72,47	1	0,788	0,946
20	2,40	0,979	71,56	1	0,774	0,893
50	0,96	0,965	67,95	1	0,768	0,871
200	0,24	0,928	60,12	1	0,75	0,807
500	0,10	0,903	51,99	1	0,731	0,74
999	0,05	0,865	44,49	1	0,707	0,664

Tabulka 4: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG2000 (fotografie „MontSaintMichel“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI	TMQI-s
2	24,00	0,999	55,26	1	0,792	0,965
5	9,60	0,999	55,29	1	0,792	0,964
20	2,40	0,999	55,17	1	0,789	0,952
50	0,96	0,998	54,53	1	0,784	0,929
200	0,24	0,985	48,49	1	0,761	0,844
500	0,10	0,962	42,47	1	0,735	0,754
999	0,05	0,936	36,95	1	0,715	0,688

Tabulka 5: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG2000 (fotografie „Zurich“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI	TMQI-s
2	24,00	0,999	90,54	1	0,797	0,983
5	9,60	0,999	90,94	1	0,797	0,982
20	2,40	0,992	90,53	1	0,783	0,926
50	0,96	0,975	90,11	1	0,75	0,854
200	0,24	0,951	88,27	1	0,734	0,75
500	0,10	0,908	84,10	1	0,711	0,673
999	0,05	0,864	80,29	1	0,686	0,601

Z tabulek a grafů je dále patrné, že metoda HDR-JPEG2000 využívá vlastnosti výsledného formátu. Pro nízké kompresní poměry ji považujeme za bezztrátovou (v některých případech až do kompresního poměru 50, tedy pro počet bitů na pixel v našem případě cca 1. Pro nižší počet pixelů je pak znatelný pokles kvality. To je dáno ztrátovým charakterem komprese.

Nejvhodnější metrika použitá v této práci je metrika HDR-VDP-2. Výstupem této metriky je index kvality fotografie označovaný jako MOS („Mean Opinion Score“) a střední pravděpodobnost detekce chyby nebo rozdílu ve dvou porovnávaných fotografiích (P_det), případně tato pravděpodobnost jako mapa pravděpodobnosti pro celý snímek.

Metoda SSIM vykazovala podobné výsledky pro tuto kompresní metodu. Není však vhodná pro přímé použití mezi formáty radiance RGB (.hdr) a JPEG2000 (.jp2). Proto bylo nutné nejprve originální HDR snímek převést pomocí mapování jasu do LDR formátu (např. JPEG2000) a porovnávat snímky takto upravené. Z grafu je zřejmé, že pro vysoký počet bitů na pixel (nízký kompresní poměr) je průměrný index strukturální podobnosti blízký hodnotě 1 (identické snímky). S klesajícím počtem bitů na pixel pak též klesá tento index.

Zajímavé srovnání nabízí TMQI – index kvality mapování jasu. To je specifická metrika, která je speciálně navržena pro srovnání dvou snímků v odlišných formátech (jeden snímek HDR, druhý snímek LDR). I tento index vykazoval podobné výsledky jako dva předchozí. Výstupem z tohoto hodnocení kvality obrazu jsou dva hlavní parametry. Jeden nazvaný jako TMQI a druhý jako TMQI-s (strukturální index kvality mapování jasu). V našem případě byl zajímavější strukturální index, který více vypovídá o kompresních ztrátách kvality snímku. Pro nízký kompresní poměr se hodnoty tohoto indexu blížily 1, pro vyšší kompresní poměr naopak klesaly.

7.2 Zpětně kompatibilní kompresní metoda HDR-JPEG

Tato kompresní metoda využívá pro ukládání dat široce rozšířený ztrátový formát JPEG. Pro porovnání kvality obrazu byla použita škála komprese odpovídající kvalitě JPEG komprese (1 až 95, rozdělena takto: 95, 90, 80, 65, 50, 25, 10, 1). Úspora dat i při nastavení kvality na úrovni kvality JPEG 95 byla oproti HDR originálu vysoká. Tato úspora byla dána především kompresí typu JPEG, kde i při nejvyšší kvalitě byl výsledný počet bitů na pixel cca okolo hodnoty 2. Na snímcích se projevovaly klasické negativní jevy, které se vyskytují u komprese JPEG obzvláště při nižším nastavení kvality. Všechny výsledné snímky jsou jako příloha této práce na digitálním nosiči. Pro představu zde uvedu pouze dva vybrané snímky. Zajímavější informace nám poskytne mapa pravděpodobnosti detekce rozdílu dle metriky HDR-VDP-2 tak, jak je uvedena jako příklad na následujících dvou snímcích. Pro nastavení kvality 95 je průměrná hodnota indexu HDR-VDP-2 97,01 a průměrná pravděpodobnost detekce je 0,003. Pro nastavení kvality je průměrná hodnota indexu HDR-VDP-2 96,04 a průměrná pravděpodobnost detekce je 0,023. Pro nižší nastavení kvality již mapa pravděpodobnosti postrádá smysl, neboť je v celém snímku rovna hodnotě 1.



Obrázek 22 – snímek „Cafeteria“ po provedení komprese HDR-JPEG s nastavením kvality komprese na hodnotu 95.



Obrázek 23 - snímek „Cafeteria“ po provedení komprese HDR-JPEG s nastavením kvality komprese na hodnotu 25.



Obrázek 24 – Snímek „Cafeteria“, kvalita 95, mapa pravděpodobnosti detekce chyby



Obrázek 25 - Snímek „Cafeteria“, kvalita 25, mapa pravděpodobnosti detekce chyby

V následujících tabulkách jsou uvedeny výsledné hodnoty jednotlivých parametrů a indexů kvality obrazu. Jednotlivé metriky vykazují podobné výsledky v závislosti na množství detailů obsažených v testovacích snímcích. Například u snímků „MontSaintMichel“ a „Market“ je index kvality nižší než u ostatních snímků. To je dáno rozložením objektů a obsahem snímků – např. u „MontSaintMichel“ je hlavním motivem řada sloupů.

Tabulka 6: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG (fotografie „Market“).

Popis veličin: $mSSIM$ je průměrná hodnota indexu strukturální podobnosti v celém snímku, MOS je „Mean Opinion Score“ a P_{det} pravděpodobnost detekce rozdílu mezi dvěma snímky podle metody HDR-VDP-2 [24], $TMQI$ je index kvality mapování jasu a $TMQI-s$ je strukturální složka indexu kvality mapování jasu podle metody [27].

kvalita	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
95	1,71	0,931	90,4	0,609	0,845	0,85
90	1,13	0,931	90,4	0,615	0,848	0,847
80	0,74	0,93	90,36	0,614	0,841	0,838
65	0,53	0,928	90,23	0,605	0,837	0,824
50	0,43	0,927	89,93	0,663	0,833	0,807
25	0,29	0,92	88,22	0,621	0,816	0,75
10	0,19	0,89	80,95	0,948	0,777	0,626
1	0,14	0,76	60,49	0,913	0,683	0,401

Tabulka 7: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG (fotografie „Cafeteria“)

kvalita	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
95	2,20	0,959	97,01	0,003	0,85	0,838
90	1,46	0,959	97	0,003	0,849	0,835
80	0,96	0,959	97,03	0,003	0,846	0,827
65	0,70	0,958	96,97	0,003	0,843	0,816
50	0,56	0,957	96,82	0,003	0,84	0,807
25	0,37	0,951	96,04	0,023	0,832	0,778
10	0,22	0,922	93,31	0,584	0,804	0,686
1	0,14	0,828	74,79	0,999	0,715	0,466

Tabulka 8: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG (fotografie „BloomingGorse“)

kvalita	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
95	3,50	0,982	99,26	0	0,965	0,876
90	2,35	0,982	99,27	0	0,965	0,874
80	1,56	0,981	99,28	0	0,963	0,867
65	1,14	0,979	99,25	0	0,961	0,86
50	0,93	0,977	99,12	0	0,96	0,857
25	0,62	0,969	98,71	0	0,958	0,848
10	0,34	0,933	96,45	0,118	0,945	0,802
1	0,16	0,803	81,29	0,998	0,87	0,585

Tabulka 9: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG (fotografie „MontSaintMichel“)

kvalita	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
95	2,06	0,925	87,33	0,015	0,818	0,883
90	1,36	0,925	87,37	0,016	0,817	0,88
80	0,89	0,924	87,36	0,015	0,814	0,872
65	0,64	0,923	87,39	0,016	0,81	0,858
50	0,52	0,922	87,09	0,021	0,806	0,841
25	0,34	0,915	86,88	0,036	0,79	0,785
10	0,21	0,869	80,79	0,039	0,742	0,632
1	0,14	0,779	53,23	0,999	0,631	0,375

Tabulka 10: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro kompresní metodu HDR-JPEG (fotografie „Zurich“)

kvalita	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
95	1,54	0,995	99,78	0	0,757	0,707
90	0,98	0,995	99,78	0	0,754	0,699
80	0,59	0,994	99,77	0	0,746	0,678
65	0,41	0,993	99,74	0	0,735	0,649
50	0,33	0,969	99,68	0	0,729	0,632
25	0,23	0,979	99,41	0,003	0,715	0,592
10	0,16	941	98,19	0,133	0,659	0,451
1	0,13	0,886	81,58	0,998	0,549	0,251

Grafické znázornění výsledků je uvedeno na následujících grafech. I pro tuto kompresní metodu platí, že pro vysoké nastavení výstupní kvality jsou odpovídající výsledné hodnoty jednotlivých metrik. Nastavení kvality větší než 65 vede k tomu, že výsledné snímky po provedení dekomprese jsou zdařilými kopiemi vstupních snímků. Pro nižší nastavení převládnu negativní jevy známé u komprese typu JPEG. Tato metoda při dekompresi rekonstruuje HDR snímek na jeho původní velikost. Velikost se však může lišit, pokud je na vstupu snímek ve formátu *OpenEXR*, pak výstupní snímek *RadianceRGBE* může být větší než původní snímek. To je způsobeno odlišným kódováním a formátem při čtení dat ze souboru a zapisování do souboru, zvláště při použití upravených funkcí pro čtení a zápis (viz kapitola 6). Při použití vysokého stupně komprese je však rekonstruovaný snímek zatížený řadou JPEG artefaktů.

7.3 Dvoustupňová kompresní metoda

Dvoustupňová kompresní metoda byla na testování složitější. Zpětná rekonstrukce u této bezztrátové komprese je zbytečná, takže porovnávat výsledné HDR snímky s originálem nebudeme. Porovnávat budeme zkomprimované snímky pomocí této metody a výše popsaných objektivních metrik. Výstupem této kompresní metody jsou dva snímky nazvané RGB a E. RGB je snímek, popisující mapu odhadu mantisy přes celý snímek a E je podobný jednokanálový snímek s mapou odhadu exponentu pro formát s plovoucí desetinnou čárkou. Celý proces při kompresi tohoto snímku je popsán v kapitole 4.7. Zdálo by se, že pro potřeby objektivního hodnocení je nutné hodnotit oba snímky. Protože však šedotónový snímek E má pouze odstíny šedi (popis exponentu), stačí pro objektivní hodnocení pomocí zvolených metrik pouze snímek RGB. Strukturální vlastnosti obou snímků RGB a E jsou stejné.



Obrázek 26 - snímek „Market“ (RGB) po provedení dvoustupňové komprese s poměrem 4.



Obrázek 27 - snímek „Market“ (E) po provedení dvoustupňové komprese s poměrem 4.

Tabulka 11: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro dvoustupňovou kompresní metodu (fotografie „Market“)

Popis veličin: *mSSIM* je průměrná hodnota indexu strukturální podobnosti v celém snímku, *MOS* je „Mean Opinion Score“ a *P_det* pravděpodobnost detekce rozdílu mezi dvěma snímky podle metody HDR-VDP-2 [24], *TMQI* je index kvality mapování jasu a *TMQI-s* je strukturální složka indexu kvality mapování jasu podle metody [27].

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
2	24,00	0,998	75,98	1	0,781	0,919
4	12,00	0,979	74,81	1	0,780	0,917
8	6,00	0,952	74,67	1	0,778	0,909
16	3,00	0,921	74,55	1	0,775	0,895
32	1,50	0,893	74,52	1	0,767	0,868
64	0,75	0,841	71,86	1	0,751	0,849

Tabulka 12: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro dvoustupňovou kompresní metodu (fotografie „Cafeteria“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
2	24,00	0,999	64,53	1	0,764	0,856
4	12,00	0,994	64,52	1	0,763	0,852
8	6,00	0,978	64,51	1	0,762	0,850
16	3,00	0,968	64,49	1	0,758	0,833
32	1,50	0,955	63,38	1	0,747	0,785
64	0,75	0,926	61,28	1	0,724	0,693

Tabulka 13: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro dvoustupňovou kompresní metodu (fotografie „BloomingGorse“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
2	24,00	0,993	75,98	1	0,771	0,883
4	12,00	0,976	74,81	1	0,767	0,866
8	6,00	0,946	74,67	1	0,761	0,845
16	3,00	0,934	74,55	1	0,747	0,794
32	1,50	0,920	74,52	1	0,739	0,769
64	0,75	0,876	71,86	1	0,712	0,721

Tabulka 14: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro dvoustupňovou kompresní metodu (fotografie „MontSaintMichel“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
2	24,00	0,999	79,55	0,98	0,777	0,907
4	12,00	0,999	79,53	0,98	0,777	0,906
8	6,00	0,998	78,94	0,97	0,776	0,901
16	3,00	0,995	78,12	0,98	0,771	0,881
32	1,50	0,987	76,12	0,98	0,762	0,849
64	0,75	0,951	73,58	0,98	0,741	0,798

Tabulka 15: Výsledné hodnoty objektivních parametrů pro měření kvality obrazu pro dvoustupňovou kompresní metodu (fotografie „Zurich“)

kompresní poměr JPEG2000	počet bitů na pixel (bpp)	mSSIM (-)	MOS (-)	P_det (-)	TMQI (-)	TMQI-s (-)
2	24,00	0,999	79,58	0,98	0,776	0,902
4	12,00	0,998	79,12	0,98	0,774	0,894
8	6,00	0,995	78,65	0,98	0,761	0,847
16	3,00	0,981	77,54	0,99	0,762	0,850
32	1,50	0,953	73,92	0,99	0,768	0,850
64	0,75	0,921	71,18	0,99	0,759	0,846

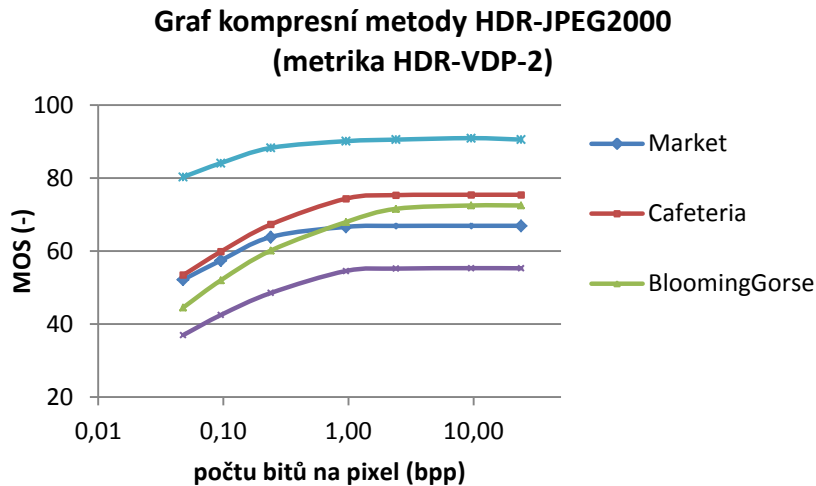
Pouze v tomto případě metrika HDR-VDP-2 vykazovala odlišné výsledky od ostatních metod. Pravděpodobně není tato metoda vhodná pro porovnání pouze mantisy (RGB snímku) dvoustupňové metody komprese.

V ostatních případech je tato metrika nejpřesnější a nejpokročilejší. Výsledná číslo MOS je pouze průměrné objektivní skóre této metody pro celý snímek. V rámci této práce na digitálním nosiči je možné tuto metriku využít jako celek – pak jsou dostupné také mapy pravděpodobnosti detekce chyby a další funkce obsažené v programovém prostředí v rámci této metriky.

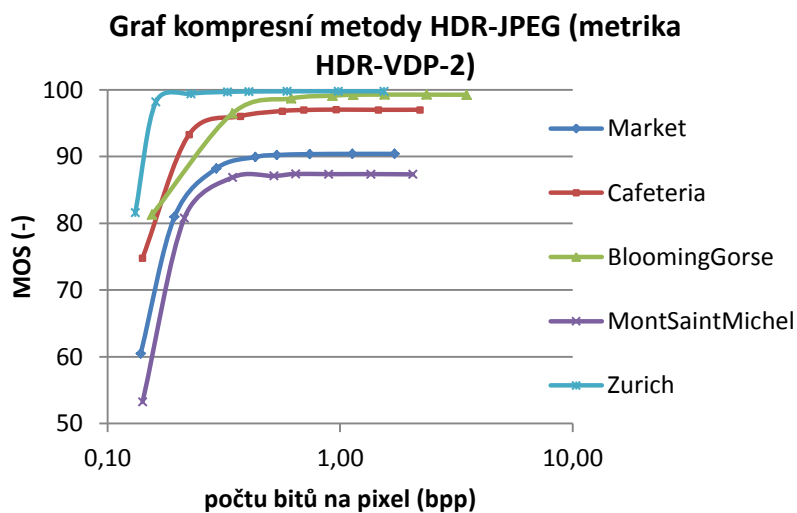
7.4 Grafické porovnání kompresních metod

Následující obrázky poskytují srovnání jednotlivých kompresních metod. Nejprve jsou kompresní metody porovnávány podle parametrů získaných při metrice HDR-VDP-2 [24]. Zde je vidět vysoká kvalita pro metody HDR-JPEG2000 a HDR-JPEG. U dvoustupňové metody se projevuje vliv nastavení jasu, takže snímky se jeví jako odlišné v rámci této metriky.

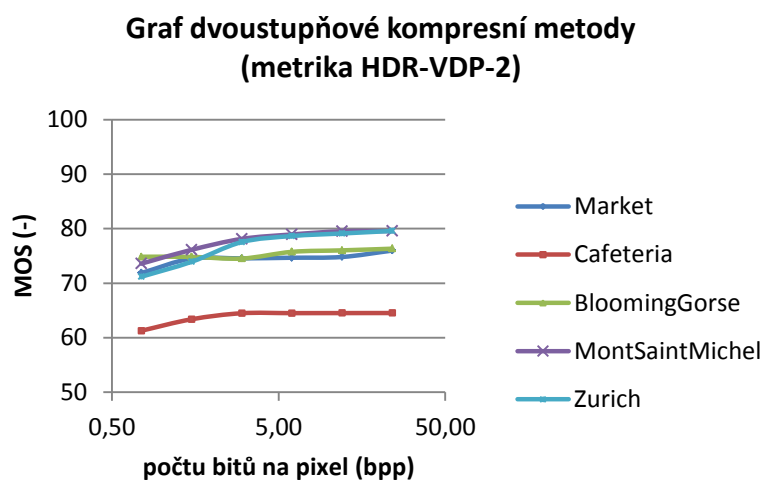
Druhá část poskytuje porovnání kompresních metod podle indexu strukturální podobnosti [20]. V této části již vidíme srovnatelné výsledky u všech kompresních metod. Podobný vliv je vidět i u poslední metriky – indexu kvality mapování jasu [27]. V tomto případě je zobrazena strukturální složka tohoto indexu.



Obrázek 28 – Závislost MOS metriky HDR-VDP-2 na počtu bitů na pixel. Metoda komprese HDR-JPEG2000

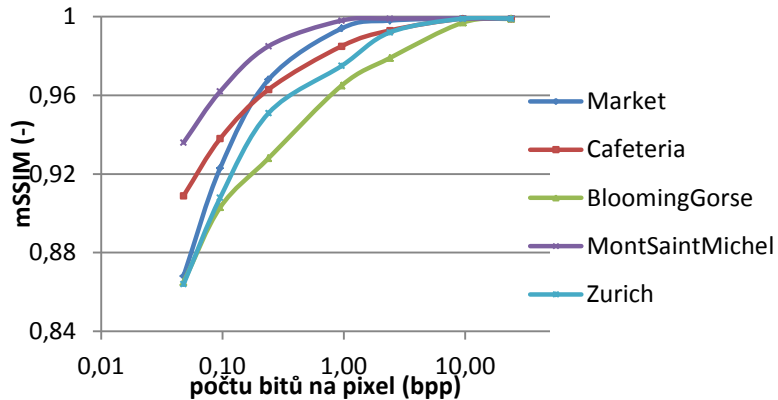


Obrázek 29 - Závislost MOS metriky HDR-VDP-2 na počtu bitů na pixel. Metoda zpětně kompatibilní komprese HDR-JPEG



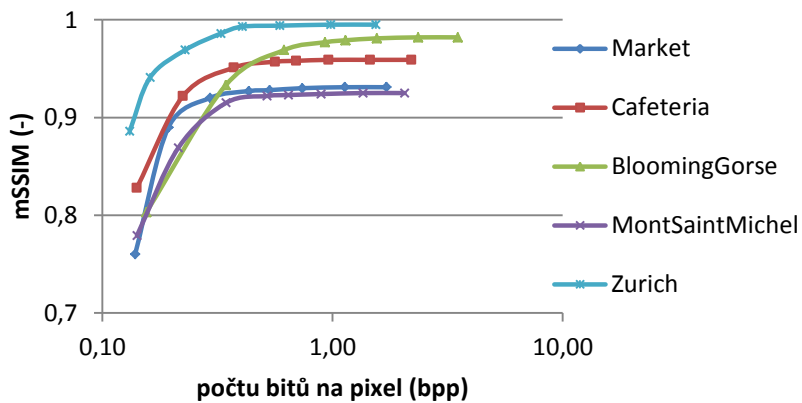
Obrázek 30 - Závislost MOS metriky HDR-VDP-2 na počtu bitů na pixel. Metoda dvoustupňové komprese

**Graf kompresní metody HDR-JPEG2000
(metrika SSIM)**



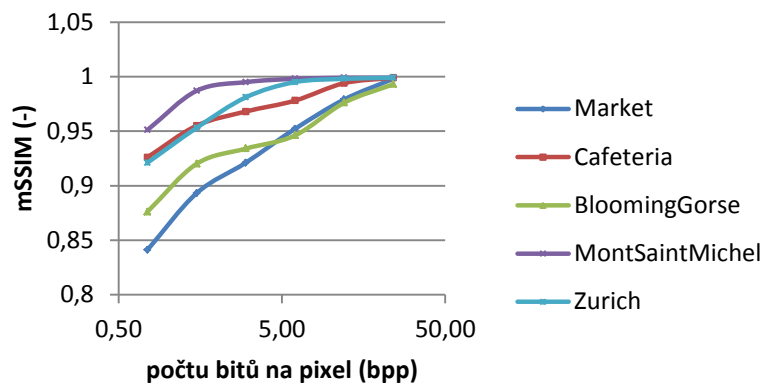
Obrázek 31 - Závislost středního indexu strukturální podobnosti na počtu bitů na pixel. Metoda komprese HDR-JPEG2000.

**Graf kompresní metody HDR-JPEG
(metrika SSIM)**

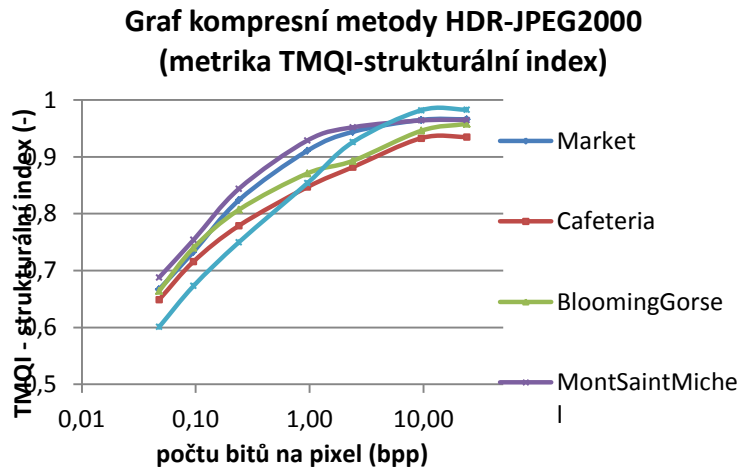


Obrázek 32 - Závislost středního indexu strukturální podobnosti na počtu bitů na pixel. Metoda zpětně kompatibilní komprese HDR-JPEG.

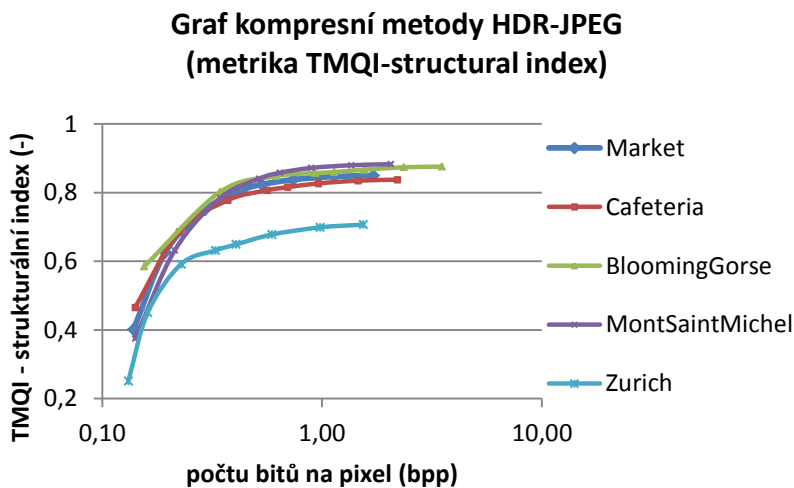
**Graf dvoustupňové kompresní metody
(metrika SSIM)**



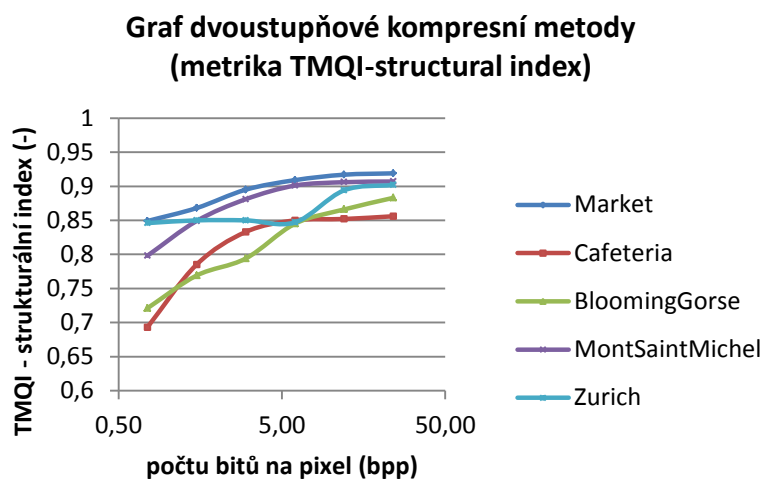
Obrázek 33 - Závislost středního indexu strukturální podobnosti na počtu bitů na pixel. Metoda dvoustupňové komprese



Obrázek 34 - Závislost strukturálního indexu metriky TMQI na počtu bitů na pixel. Metoda komprese HDR-JPEG2000.



Obrázek 35 - Závislost strukturálního indexu metriky TMQI na počtu bitů na pixel. Metoda zpětně kompatibilní komprese HDR-JPEG.



Obrázek 36 - Závislost strukturálního indexu metriky TMQI na počtu bitů na pixel. Metoda dvoustupňové komprese.

8 Závěr

Cílem práce bylo popsat zpětně kompatibilní metody pro kompresi dat s vysokým dynamickým rozsahem. Zároveň bylo vytvořeno a prozkoumáno několik vybraných algoritmů v programovém prostředí Matlab. Tyto algoritmy byly porovnány pomocí objektivních metod pro hodnocení kvality obrazu.

Pro pochopení problematiky bylo nutné vysvětlit základní pojmy, postupy a metody vytváření snímků s vysokým dynamickým rozsahem a dále také popsat vybrané metody pro mapování jasu (kapitola 3), které jsou nutné při kompresi obrazových dat. Popis samotných metod komprese je uveden v samostatné kapitole 4. Jednotlivé metody vykazovaly podobné výsledky, takže nelze přesně určit nejvhodnější metodu pro kompresi. Metoda JPEG-HDR je metodou, která vede k nejvyššímu kompresnímu poměru – komprimované snímky mají nejmenší velikost. Při snížení kvality JPEG se pak projeví typické problémy formátu JPEG, artefakty apod. Další dvě metody (HDR-JPEG2000 a dvoustupňová metoda) využívají formát JPEG2000 pro ukládání komprimovaných dat. Tyto metody vytvářejí větší soubory po kompresi, ale při dostatečně vysoké úrovni nastavení kvality se dají považovat za bezztrátové. Všechny metody mají také pro srovnání dekompresní část, kde se rekonstruuje vstupní snímek.

Hodnocení kvality komprimovaných obrazů probíhalo pomocí vybraných objektivních metrik. Protože pro data s vysokým dynamickým rozsahem není vhodné používat klasické metody (odstup signál šum, střední kvadratická chyba atd.), bylo nutné využít metriky přímo určené pro tyto účely. Pro srovnání pomocí indexu strukturální integrity (SSIM) a metriky HDR-VDP-2 bylo nutné referenční snímky převést pomocí operátorů mapování jasu na nízký dynamický rozsah. Pro přímé srovnání HDR a LDR snímků byl použit index kvality mapování jasu, hlavně jeho strukturální složka. Testování kompresních metod probíhalo na sadě testovacích fotografií vybraných z veřejně dostupných databází. V programovém prostředí je vytvořena struktura funkcí, které jsou podrobně popsány v kapitole 6 a jsou přiloženy na datovém nosiči, který je součástí práce.

Kompresní metody popsané v této práci jsou efektivní, nicméně vysoká výpočetní náročnost neumožňuje jejich použití v reálném čase a proto jsou vhodné pouze pro statické snímky. Některé z nich je možné dále rozvíjet, což povede k vyšší účinnosti a ke snížení výpočetní náročnosti. Zvláště vhodné jsou pak bezztrátové metody, které jsou vhodné pro ukládání HDR dat na datové nosiče.

9 Seznam použité literatury

- [1] REINHARD E., WARD G., PATTANAİK S., DEBEVEC P.: *High dynamic range imaging – acquisition, display and image-based lighting*. Morgan Kaufmann Publishers, San Francisco, USA. 2006, 502 s., ISBN 978-0-12-585263-0
- [2] DEBEVEC P., and MALIK J.: *Recovering High Dynamic Range Radiance Maps from Photographs*, Proceedings of SIGGRAPH 1997, ACM Press / ACM SIGGRAPH, 1997.
- [3] DRAGO F., MYSZKOWSKI K., ANNEN T. and CHIBA N.: *Adaptive logarithmic mapping for displaying high contrast scenes*. In Computer Graphics Forum 22(3), 2003.
- [4] Speciální metody pro snímání a reprodukci obrazu (HDR a superresolution). Prezentace a výukový materiál předmětu A0M37MOT. www.multimediatech.cz, ČVUT FEL, 2012.
- [5] REINHARD E., STARK M., SHIRLEY P., FERWERDA J.: *Photographic Tone Reproduction for Digital Images*. Proceedings of SIGGRAPH 2002, s. 267-276. 2002.
- [6] FUJIKI Takuya, ADAMI Nicola, JINNO Takao, OKUDA Masahiro: *High dynamic range image compression using base map coding*. The University of Kitakyushu, Brescia University, Toyohashi University of Technology, 2012, ISBN 978-1-4673-4863-8
- [7] SPAULDING K. E., JOSHI R. L., WOOLFE G. J.: *Using a residual image formed from a clipped limited color gamut digital image to represent an extended color gamut digital image*, United States Patent 6301393.
- [8] WARD, Greg., and SIMMONS Maryann.: *JPEG-HDR: A Backwards-Compatible, High Dynamic Range Extension to JPEG*, Proceedings of the Thirteenth Color Imaging Conference, 2005.
- [9] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced High Dynamic Range Imaging: Theory and Practice*. AK Peters (CRC Press), Natick, USA. 2011, 278 s., ISBN 978-156881-719-4
- [10] KOPF J., COHEN M. F., LISCHINSKI D., UYTTENDAELE M.: *Joint Billateral Upsampling*. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2007), vol. 26, str. , 2007

- [11] OKUDA M., ADAMI N.: Effective color space representation for wavelet based compression of HDR images. 14th International Conference on Image Analysis and Processing, ICIAP, 2007, ISBN 0-7695-2877-5
- [12] ZHANG Y., REINHARD E., BULL D. R.: *Perceptually lossless high dynamic range image compression with JPEG 2000*. 19th IEEE International Conference on Image Processing (ICIP), 2012, ISBN 978-1-4673-2532-5
- [13] BARTEN P.G.J.: *Contrast sensitivity of the human eye and its effects on image quality*, SPIE-International Society for Optical Engineering, 1999.
- [14] MULLEN K.T.: *The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings*. The Journal of Physiology, vol. 359, s. 381–400, 1985.
- [15] GUEZZI MESSAOUD F., PEIZERAT A., DUPRET A., BLANCHARD Y.: *On-Chip Compression for HDR Image Sensors*, Paris-Est Universitym ESYCOM, ESIEE Paris, CEA LETI, L3I – MINATEC, France, 2010, ISBN 978-1-4244-8734-9
- [16] CHRISTOPOULOS C., SKODRAS A., EBRAHIMI T.: *The JPEG2000 Still Image Coding System: An Overview*. IEEE Transactions on Consumer Electronics 46:4, s. 1103-1127, 2000.
- [17] XU R., PATTANAIAK S. N., HUGHES C. E.: *High-Dynamic-Range Still-Image Encoding in JPEG 2000.*, IEEE Computer Graphics and Applications 25:6, s. 57-64, 2005
- [18] BOSCHETTI A., ADAMI N., LEONARDI R.: *Flexible and Effective High Dynamic Range Image Coding*. 17th IEEE International Conference on Image Processing (ICIP), str. 3145-3148, 2010
- [19] WANG Z., BOVIK A. C.: *A universal image quality index*. IEEE Signal Processing Letters, vol. 9, str. 81-84, 2002
- [20] WANG Z., BOVIK A. C., SHEIKH H. R., SIMONCELLI E. P.: *Image Quality Assessment: From Error Visibility to Structural Similarity*. IEEE Transactions on Image Processing, vol. 43, str. 600-612, 2004
- [21] *Empamedia HDR database* [online]. Empamedia Media Technology. [vid. 11.5.2014]. Dostupné z: <http://empamedia.ethz.ch/hdrdatabase/index.php>
- [22] *The HDR Photographic Survey* [online]. Mark Fairchild. [vid. 11.5.2014]. Dostupné z: <http://www.cis.rit.edu/fairchild/HDRPS/HDRthumbs.html>
- [23] MANTIUK R., KIM K. J., REMPEL A. G., HEIDRICH W.: *HDR-VDP-2: A Calibrated Visual Metric for Visibility and Quality Predictions in All Luminance*

Conditions. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2011). vol. 30, 2011

- [24] MANTIUK R., DALY S., MYSZKOWSKI K., SEIDEL H.: Predicting visible differences in high dynamic range images: model and its calibration. Proceedings of SPIE 2005, vol. 5666, str. 204-214, 2005
- [25] DALY S.: *Digital Images and Human Vision*. MIT Press, str. 179-206, 1993
- [26] ITU-R-BT.500-13, Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union, 2012
- [27] YEGANEH H., WANG Z.: Objective Quality Assessment of Tone-Mapped Images. IEEE Transactions on Image Processing, vol 22, str. 657-667, 2013
- [28] *Matlab Pyr Tools* [online]. Laboratory for Computational Vision. [vid. 11.5.2014]. Dostupné z: <http://www.cns.nyu.edu/~lcv/software.php>

Obsah datového nosiče

1. Text diplomové práce – soubor *rautepet_DP.pdf*
2. Zpracované kompresní algoritmy včetně pomocných funkcí – složka *Program*
3. Originální HDR snímky – složka *Program/Photo*
4. Komprimované fotografie pomocí algoritmů v programovém prostředí Matlab – složka *Program/Vysledky*
5. Podpůrné programy pro výpočet objektivního hodnocení kvality obrazu – složka *Program/HDRVDP* a složka *Program/MatlabPyrTools*