



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

VYUŽITÍ POKROČILÝCH OBJEKTIVNÍCH KRITÉRIÍ HODNOCENÍ PŘI KOMPRESI OBRAZU

ADVANCED OBJECTIVE MEASUREMENT CRITERIA APPLIED TO IMAGE COMPRESSION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. JOSEF ŠIMEK

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. JAN MALÝ

BRNO 2010



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Diplomová práce

magisterský navazující studijní obor
Telekomunikační a informační technika

Student: Bc. Josef Šimek

ID: 78319

Ročník: 2

Akademický rok: 2009/2010

NÁZEV TÉMATU:

Využití pokročilých objektivních kritérií hodnocení při kompresi obrazu

POKYNY PRO VYPRACOVÁNÍ:

Nastudujte princip objektivních metod hodnocení obrazu na bázi Structural Similarity Index (SSIM) a jejich využití při kompresi obrazu. Navrhněte ztrátové kompresní schéma na bázi vlnkové transformace, které funguje na minimalizaci chyby kritéria SSIM mezi zdrojovým a výsledným obrazem. Řešení implementujte v MATLABu.

DOPORUČENÁ LITERATURA:

[1] Wang Z., Bovik A.C.: Mean Squared Error: Love It or Leave It?, IEEE Signal Processing Magazine, Vol. 26, No. 1, Jan 2009, pp. 98-117, ISSN: 1053-5888

[2] Wang Z., Li Q., Shang X.: Perceptual image coding based on a maximum of minimal structural similarity criterion, Proceedings of IEEE Int. Conf. Image Processing, Vol 2., Sept. 2007, pp. 121-124

Termín zadání: 29.1.2010

Termín odevzdání: 26.5.2010

Vedoucí práce: Ing. Jan Malý

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

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

ABSTRAKT

Tato diplomová práce pojednává o problematice využití metod objektivního hodnocení kvality při kompresi obrazových dat. Ztrátová komprese vždy zavádí do zpracovávaných dat určité zkreslení, které je příčinou degradace kvality u výsledné podoby obrazu. Míra tohoto zkreslení může být hodnocena pomocí subjektivních nebo objektivních metod. Pro možnost optimalizace kompresních algoritmů jsou využívána objektivní kritéria. V práci je představen index SSIM jako vhodný nástroj pro popis kvality komprimovaných obrazů. Ztrátové kompresní schéma je realizováno s využitím vlnkové transformace a algoritmu SPIHT. Implementována byla varianta s členěním vlnkových koeficientů do samostatných bloků a následným nezávislým kódováním vhodná zejména pro účely paralelního zpracování. Pro daný kompresní poměr je řešen tradiční problém – jak rozdělit dostupný počet bitů mezi jednotlivé bloky, abychom dosáhli co nejvyšší kvality obrazu. Prodiskutovány byly možné přístupy řešení a navrženy některé metody využívající k rozdělení bitů kritéria MSSIM. K testování účinnosti těchto metod bylo použito programové prostředí MATLAB.

Klíčová slova: SSIM, objektivní hodnocení kvality, vlnková transformace, ztrátová komprese, SPIHT, MATLAB

ABSTRACT

This diploma thesis deals with the problem of using an objective quality assessment methods in image data compression. Lossy compression always introduces some kind of distortion into the processed data causing degradation in the quality of the image. The intensity of this distortion can be measured using subjective or objective methods. To be able to optimize compression algorithms the objective criteria are used. In this work the SSIM index as a useful tool for describing the quality of compressed images has been presented. Lossy compression scheme is realized using the wavelet transform and SPIHT algorithm. The modification of this algorithm using partitioning of the wavelet coefficients into the separate tree-preserving blocks followed by independent coding, which is especially suitable for parallel processing, was implemented. For the given compression ratio the traditional problem is being solved – how to allocate available bits among the spatial blocks to achieve the highest possible image quality. The possible approaches to achieve this solution were discussed. As a result, some methods for bit allocation based on MSSIM index were proposed. To test the effectivity of these methods the MATLAB environment was used.

Keywords: SSIM, objective quality assessment, wavelet transform, lossy compression, SPIHT, MATLAB

ŠIMEK, J. *Využití pokročilých objektivních kritérií hodnocení při kompresi obrazu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2010. 74 s. Vedoucí diplomové práce Ing. Jan Malý.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „VYUŽITÍ POKROČILÝCH OBJEKTIVNÍCH KRITÉRIÍ HODNOCENÍ PŘI KOMPRESI OBRAZU“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č.121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne

.....

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu mé diplomové práce Ing. Janu Malému za vstřícné jednání, metodické vedení a odbornou pomoc při zpracování diplomové práce. Rovněž děkuji Mgr. Pavlu Rajmicovi, Ph.D. za poskytnutí užitečných myšlenek a rad.

V Brně dne

.....

OBSAH

ÚVOD	12
1 METRIKY OBJEKTIVNÍHO HODNOCENÍ KVALITY OBRAZU	13
1.1 Mean Squared Error	13
1.1.1 Nedostatky kritéria MSE	14
1.2 Model hodnocení kvality obrazu založený na lidském vnímání	18
1.2.1 Předzpracování.....	19
1.2.2 Rozklad na subpásma.....	19
1.2.3 Normování chyby.....	20
1.3 Hodnocení kvality obrazu na základě strukturální podobnosti.....	20
1.3.1 SSIM Index	20
1.3.2 Použití SSIM pro hodnocení kvality obrazu.....	23
1.3.3 Implementace výpočtu indexu SSIM v MATLABu	24
1.3.4 Porovnání MSE a MSSIM	25
2 OBJEKTIVNÍ HODNOCENÍ KVALITY U KOMPRESI OBRAZŮ	28
2.1 Použití indexu SSIM při ztrátové kompresi obrazu	28
2.1.1 Vlnková transformace obrazu.....	29
2.1.2 Algoritmus SPIHT	30
3 VYUŽITÍ KRITÉRIA SSIM U KOMPRESI OBRAZU ZALOŽENÉ NA SPIHT	33
3.1 Zvyšování kvality části obrazu s minimální hodnotou indexu SSIM	33
3.2 Kódování koeficientů algoritmem SPIHT v nezávislých blocích.....	34
3.3 Kompres barevného obrazu založená na algoritmu SPIHT s odděleným zpracováním kanálů	36
3.4 Shrnutí	37
4 IMPLEMENTACE KOMPRESNÍHO SCHÉMATU	39
4.1 Vlastnosti kodéru	39
4.2 Kompresní schéma	40
4.2.1 Rozdělení koeficientů do bloků.....	41
4.3 Výpočet metriky MSSIM.....	42
4.3.1 Výpočet MSSIM celého obrazu.....	42
4.3.2 Výpočet MSSIM bloku.....	43

5	NÁVRH A TESTOVÁNÍ METOD PRO ROZDĚLENÍ BITŮ	44
5.1	Problém rozdělení bitů	44
5.1.1	Využití kritéria MSSIM.....	44
5.1.2	Závislost hodnoty MSSIM na počtu přidělených bitů	45
5.2	Návrh metod.....	49
5.2.1	Obecné poznatky	49
5.2.2	Rozdělení bitů podle funkční hodnoty	52
5.2.3	Rozdělení bitů podle velikosti směrnice.....	53
5.2.4	Kombinovaná metoda rozdělení bitů.....	54
5.3	Testování navržených metod	55
5.3.1	Rozdělení bitů podle funkční hodnoty	56
5.3.2	Rozdělení bitů podle velikosti směrnice.....	57
5.3.3	Kombinovaná metoda rozdělení bitů.....	57
5.4	Shrnutí a interpretace dosažených výsledků	59
	ZÁVĚR	60
	SEZNAM POUŽITÉ LITERATURY	62
	PŘEHLED POUŽITÝCH ZKRATEK.....	63
	SEZNAM PŘÍLOH	64
A	SUBJEKTIVNÍ KVALITA TESTOVANÝCH OBRÁZKŮ	65
B	OBSAH CD	74

SEZNAM OBRÁZKŮ

Obr. 1.1: Nezávislost MSE na vztahu mezi chybovým a obrazovým signálem (hodnoty metrik MSE a MSSIM se vztahují ke zkresleným obrázkům).....	15
Obr. 1.2: Nezávislost MSE na vztahu mezi jednotlivými pixely v obraze (hodnoty metrik MSE a MSSIM se vztahují ke zkresleným obrázkům)	16
Obr. 1.3: Vliv znaménka vzorků chybového signálu na hodnotu MSE (hodnoty metrik MSE a MSSIM se vztahují ke zkresleným obrázkům)	17
Obr. 1.4: Rozložení chybového signálu při výpočtu MSE (vpravo) a MSSIM (vlevo) ..	18
Obr. 1.5: Blokový diagram hodnocení kvality obrazu	19
Obr. 1.6: Blokový diagram výpočtu indexu SSIM.....	21
Obr. 1.7: Srovnání hodnot MSE a MSSIM pro různé druhy zkreslení	26
Obr. 1.8: Srovnání SSIM mapy s rozdílovým obrázkem	27
Obr. 2.1: Algoritmus 2D-DTWT pro první úroveň dekompozice	29
Obr. 2.2: Dekompoziční obrazec s hloubkou dekompozice $n = 3$	30
Obr. 2.3: Stromová struktura používaná algoritmem SPIHT	31
Obr. 4.1: Stromová struktura algoritmu dSPIHT	40
Obr. 4.2: Postup komprese obrazu.....	40
Obr. 4.3: Způsob tvorby bloků ze stromové struktury dekompozičního obrazce	42
Obr. 4.4: Princip výpočtu indexu MSSIM pro jednotlivé bloky.....	43
Obr. 5.1: Průběh závislostí hodnot MSSIM bloků obrázku na počtu přidělených bitů (64 bloků, rozměry obrázku 512×512 pixelů, graf má pouze ilustrativní charakter)	46
Obr. 5.2: Závislost MSSIM na počtu přidělených bitů u vybraných charakteristických bloků	47
Obr. 5.3: Průběh závislostí hodnot PSNR bloků obrázku na počtu přidělených bitů (64 bloků, rozměry obrázku 512×512 pixelů, graf má pouze ilustrativní charakter)	48
Obr. 5.4: Detail počátku průběhu křivek z obr. 5.2	50
Obr. 5.5: Průběhy chybových funkcí mezi původními (obr. 5.2) a interpolovanými hodnotami	51
Obr. 5.6: Testované obrázky	55
Obr. 5.7: Závislost hodnoty MSSIM testovaných obrázků na počtu přidělených bitů při rovnoměrném rozdělení (plná čára) a rozdělení podle funkční hodnoty (přerušovaná čára)	56

Obr. 5.8: Závislost hodnoty MSSIM testovaných obrázků na počtu přidělených bitů při rovnoměrném rozdělení (plná čára) a rozdělení podle velikosti směrnice (přerušovaná čára).....	57
Obr. A.1: Obrázek <i>kodak02.bmp</i> pro $\text{bpp}=1$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,9291$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,9357$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	65
Obr. A.2: Obrázek <i>kodak02.bmp</i> pro $\text{bpp}=0,25$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,7979$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,8048$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	66
Obr. A.3: Obrázek <i>kodak02.bmp</i> pro $\text{bpp}=0,125$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,7325$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,7399$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	67
Obr. A.4: Obrázek <i>kodak03.bmp</i> pro $\text{bpp}=1$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,9549$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,9673$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	68
Obr. A.5: Obrázek <i>kodak03.bmp</i> pro $\text{bpp}=0,25$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,8393$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,8627$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	69
Obr. A.6: Obrázek <i>kodak03.bmp</i> pro $\text{bpp}=0,125$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,7767$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,7917$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	70
Obr. A.7: Obrázek <i>kodak07.bmp</i> pro $\text{bpp}=1$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,9768$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,9838$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	71
Obr. A.8: Obrázek <i>kodak07.bmp</i> pro $\text{bpp}=0,25$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,9082$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,9262$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	72
Obr. A.9: Obrázek <i>kodak07.bmp</i> pro $\text{bpp}=0,125$ při rovnoměrném rozdělení bitů (nahore, $\text{MSSIM}=0,8430$) a rozdělení podle kombinované metody pro $\alpha=0,7$ (dole, $\text{MSSIM}=0,8471$). Hodnoty MSSIM se vztahují k původnímu obrázku (obr. 5.6).....	73

SEZNAM TABULEK

Tab. 1.1: Naměřené hodnoty MSSIM obrázků při rozdílném počtu přidělených bitů a měnící se hodnotě váhy α u kombinované metody	58
------------------------------------------------------------------------------------------------------------------------------------------------------	----

ÚVOD

Digitální obraz, coby dvourozměrný signál, je při svém vzniku, následném zpracování a uložení na médium, případně během přenosu či kompresi, vystaven řadě různých druhů zkreslení, která mohou mít vliv na výslednou degradaci jeho vizuální kvality. V případech, kdy jsou obrazy prohlíženy a jejich obsah zkoumán lidmi (např. digitální fotografie, lékařské snímky apod.), je jedinou vhodnou metodou pro určení vizuální kvality obrazů subjektivní hodnocení. Tato metoda má však svá úskalí. Je závislá na pozorovateli, tudíž nemá přesně definovaný postup či matematický vzorec. Kvalitu obrazu posuzuje samotný pozorovatel podle parametrů, které mu nejvíce vyhovují. Aby subjektivní hodnocení mělo určitou vypovídající hodnotu, musí se výsledný obraz nechat posoudit statisticky dostatečným počtem lidí (řádově stovky až tisíce). To by však v praxi znamenalo drahou a časově náročnou operaci. Proto je snahou nalézt odpovídající řešení pro případ objektivního hodnocení kvality obrazu.

Dominantní roli v této oblasti zaujímá metrika MSE (Mean Squared Error). Její výpočet je jednoduchý a význam hodnoty zřejmý – udává, jaké množství chybového signálu je v obrázku obsaženo. V řadě aplikací pro zpracování obrazu (včetně komprese obrazových dat) však MSE dosahuje neuspokojivých výsledků. Cílem tedy bylo využít jinou metriku, která by více reflektovala skutečný vliv zkreslení na subjektivní kvalitu obrázku a umožnila tak toto zkreslení lépe potlačit. Jako vhodný nástroj pro tento účel byla představena metrika SSIM (Structural SIMilarity) [8]. Kvalita obrázku zde není hodnocena podle množství obsaženého chybového signálu, ale z hlediska vlivu chybového signálu na zkreslení struktury objektů v obraze. Tyto vlastnosti předurčují index SSIM pro využití zejména v oblasti ztrátové komprese, kde dochází k deformaci struktury objektů ve výsledném obrázku.

Diplomová práce se zabývá principem metody SSIM a porovnáním vlastností tohoto indexu s vlastnostmi metriky MSE. Cílem je analyzovat varianty použití SSIM indexu při kompresi obrazu. Rozebrány budou metody založené na vlnkové transformaci a kompresním algoritmu SPIHT (Set Partitioning In Hierarchical Trees), který představuje efektivní prostředek pro hledání optimální kvality výsledného obrázku při daném počtu přidělených bitů. U těchto metod se zaměříme na možné způsoby využití metriky SSIM, které by vedly k zlepšení kvality komprimovaného obrázku. Požadovaným řešením je navrhnout postup, na základě něhož můžeme optimalizovat algoritmus SPIHT podle kritéria SSIM.

1 METRIKY OBJEKTIVNÍHO HODNOCENÍ KVALITY OBRAZU

Metoda objektivního hodnocení kvality je užívána aplikacemi pro zpracování obrazu v mnoha funkcích. Její metrika se určuje výpočtem z dostupných obrazových dat. Metriky rozdělujeme podle toho, zda máme při výpočtu k dispozici originální (nezkreslený) obrázek v plném rozsahu, částečně anebo vůbec. My budeme dále uvažovat a zabývat se nejrozšířenější technikou nazývanou *full-reference*, kdy je pro porovnání se zkresleným obrazem dostupný originální obraz v plném rozsahu.

1.1 Mean Squared Error

Nejjednodušší a nejpoužívanější metrikou pro objektivní hodnocení kvality obrazu je střední kvadratická chyba – MSE (Mean Squared Error). Pro dvourozměrný obrazový signál ji definujeme takto[8]

$$\text{MSE} = \frac{1}{M \cdot N} \sum_{i=1}^M \sum_{j=1}^N (\mathbf{X}(i,j) - \mathbf{Y}(i,j))^2, \quad (1.1)$$

kde $\mathbf{X}(i,j)$ a $\mathbf{Y}(i,j)$ představují hodnoty pixelů dvou různých obrázků v rámci jednoho kanálu barvové informace (např. jasová složka) o rozměrech $M \times N$. Pokud \mathbf{X} budeme brát za původní (nezkreslený) obrázek a \mathbf{Y} za jeho zkreslenou verzi, u níž chceme stanovit vizuální kvalitu, potom MSE můžeme považovat za měřítko kvality obrazového signálu. Výraz $\mathbf{X}(i,j) - \mathbf{Y}(i,j)$ pak udává chybový signál, který vznikne mezi hodnotami původního a zkresleného obrazového signálu. Pro barevný obrázek (se složkami RGB) je nutno výpočet MSE podle rovnice (1.1) upravit přidáním sumy sčítající chyby jednotlivých barevných složek

$$\text{MSE} = \frac{1}{3 \cdot M \cdot N} \sum_{k=1}^3 \sum_{i=1}^M \sum_{j=1}^N (\mathbf{X}(i,j) - \mathbf{Y}(i,j))^2. \quad (1.2)$$

Z důvodu možnosti srovnání obrázků s rozdílnými dynamickými rozsahy se zavádí metrika PSNR (Peak Signal-to-Noise Ratio, špičkový odstup signál/šum), která vychází z výpočtu hodnoty MSE. Udává poměr mezi maximální možnou energií obrazového signálu a energií šumu v obraze. Rovnice má následující tvar [5]

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right), \quad (1.3)$$

kde MAX značí maximální hodnotu pixelu, což pro typickou hodnotu 8 bitů na kanál dává $MAX = 255$. PSNR se vyjadřuje v decibelech [dB]. V typických situacích se hodnoty pohybují v rozmezí 20 až 40 dB.

Důvodů, proč se MSE stalo nejpoužívanějším objektivním kritériem hodnocení kvality, je celá řada [7]:

- Především je to jednoduchost této metriky. Její výpočet je nenáročný a velice rychlý. Není třeba uvažovat žádné závislosti mezi jednotlivými pixely, jelikož kvadratická chyba je počítána pro každý pixel zvlášť, nezávisle na ostatních hodnotách pixelů.
- Metrika má jasný fyzikální význam – definuje energii chybového signálu. Tato energie je zachována i v případě lineárních transformací (jako např. Fourierova transformace). Vlastnost zachování energie zaručuje, že energie zkreslení v transformované oblasti bude stejná jako energie oblasti, pro níž jsou signály definovány.
- Výborné vlastnosti z hlediska možnosti optimalizace. Postup nalezení minimální hodnoty MSE je možné definovat analyticky – výpočet algoritmu je rychlý.
- Již od svého vzniku bylo MSE využíváno k porovnávání a optimalizaci v řadě různých aplikací pro zpracování signálů, zahrnující návrh filtrů, kompresi, třídění a restauraci signálů. Efektivita algoritmů byla hodnocena na základě hodnoty MSE (PSNR), tudíž i pro nově vzniklé algoritmy představovalo MSE (PSNR) nejběžnější společný způsob srovnávání. Takto došlo k značnému rozšíření této metriky.

Je zřejmé, že MSE má mnoho výhodných vlastností pro užití v různých aplikacích. Mohlo by se tedy zdát, že je ideálním prostředkem pro objektivní hodnocení kvality obrázku. V následující části kapitoly se přesvědčíme o tom, že toto tvrzení v řadě případů rozhodně neplatí.

1.1.1 Nedostatky kritéria MSE

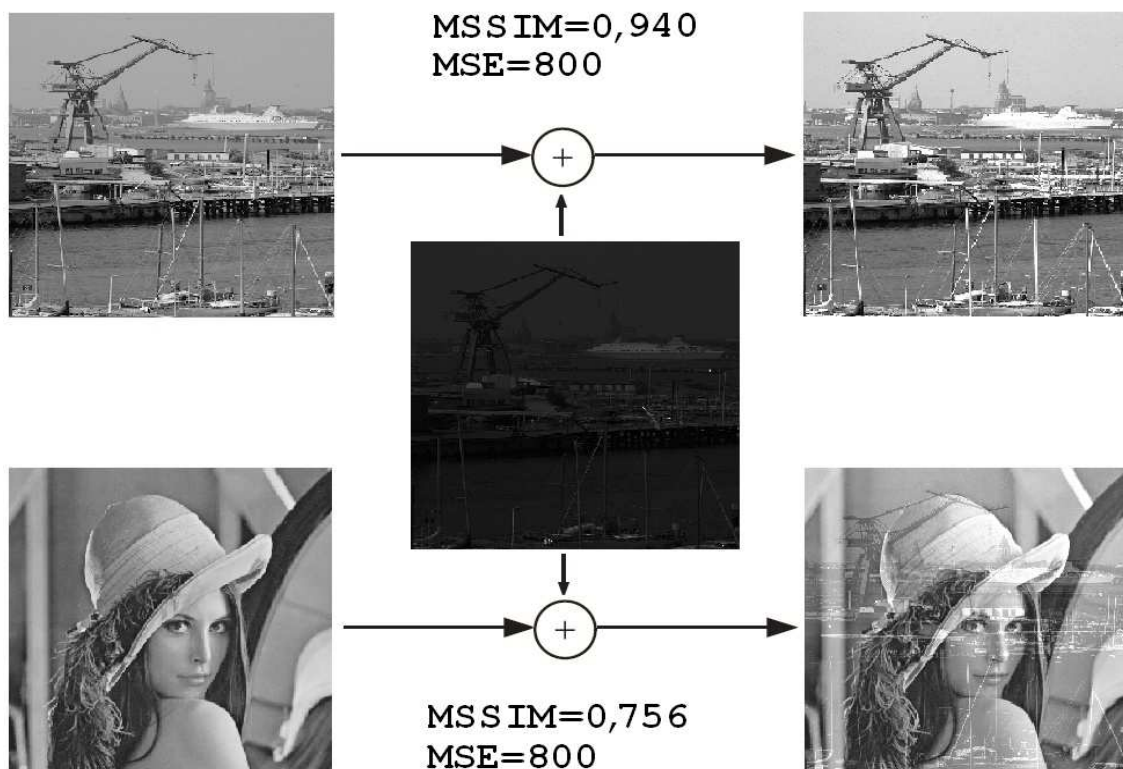
U metriky MSE, jakožto statistické veličiny, budeme předpokládat určité vlastnosti, které musíme vzít v úvahu, pokud hodnotíme věrnost obrazového signálu. Následně vyzkoušíme, jak se tyto vlastnosti projeví při měření objektivní kvality obrázků [7]. Pro srovnání je v popiscích **obr. 1.1**, **1.2**, **1.3** a **1.4** uvedena i metrika MSSIM, která rovněž může být použita pro hodnocení kvality obrázků, o ní bude v textu pojednáno později v kapitole 1.3.2.

Předpokládané vlastnosti jsou tyto [7]:

1. Výsledná kvalita obrazového signálu je nezávislá na vztahu mezi původním obrázkem a chybovým signálem. Pro daný chybový signál zůstává MSE nezměněno bez ohledu na to, ke kterému obrázku je chyba přičítána.

2. Výsledná kvalita obrazového signálu je nezávislá na vztazích mezi jednotlivými vzorky původního signálu v prostoru. Jinými slovy, pokud pozice pixelů původního a zkrasleného obrázku jsou přeuspořádány stejným způsobem, potom hodnota MSE u těchto obrázků zůstává nezměněna.
3. Výsledná kvalita obrazového signálu je nezávislá na znaménku vzorků chybového signálu.
4. Všechny vzorky obrazového signálu přispívají k výsledné kvalitě rovným dílem.

Na **obr 1.1** je znázorněna první předpokládaná vlastnost MSE. Stejný chybový signál (uprostřed) byl přidán ke dvěma rozdílným obrázkům. Signál byl vytvořen tak, že jeho vzorky zcela korelují se vzorky obrázku vlevo nahoře. Ačkoliv mají oba zkraslené obrázky (vpravo) stejnou hodnotu MSE vzhledem k jejich původnímu obrázku, je zřejmé, že vizuální zkraslení je mnohem patrnější u obrázku vpravo dole. Projevuje se zde významný nedostatek MSE. Korelace mezi výchozím obrázkem a chybovým signálem zásadně ovlivňuje vnímanou míru zkraslení v obrázku. Obrazový signál, jehož vzorky korelují s vzorky chybového signálu, vykazuje daleko vyšší pozorovatelnou kvalitu. Tento fakt MSE zcela opomíjí.

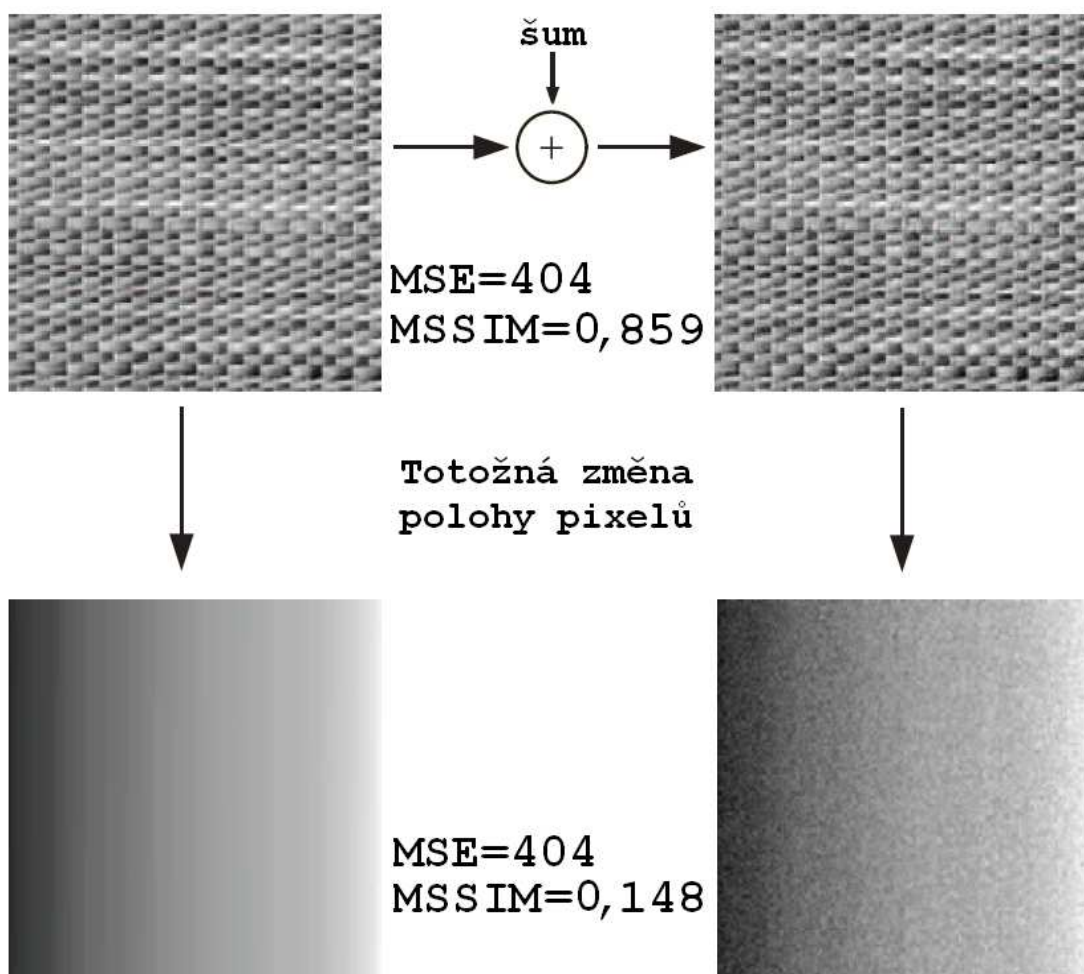


Obr. 1.1: Nezávislost MSE na vztahu mezi chybovým a obrazovým signálem (hodnoty metrik MSE a MSSIM se vztahují ke zkrasleným obrázkům)

Druhá předpokládaná vlastnost je zachycena na **obr. 1.2**. Obrázek vpravo nahoře byl vytvořen z původního obrázku přidáním chybového signálu v podobě Gaussova bílého šumu (signál s normálním rozdělením pravděpodobnosti výskytu velikosti

jeho vzorků a s rovnoměrnou spektrální výkonovou hustotou). Obrázek vlevo dole vznikl z původního přeuspořádáním polohy pixelů bez jakékoliv změny jejich hodnoty. Stejným způsobem byl získán ze zkrasleného obrázku i obrázek vpravo dole. Jelikož se nijak nezměnily hodnoty vzorků chybového signálu, je velikost MSE mezi dvěma horními a mezi dvěma spodními obrázky identická [7].

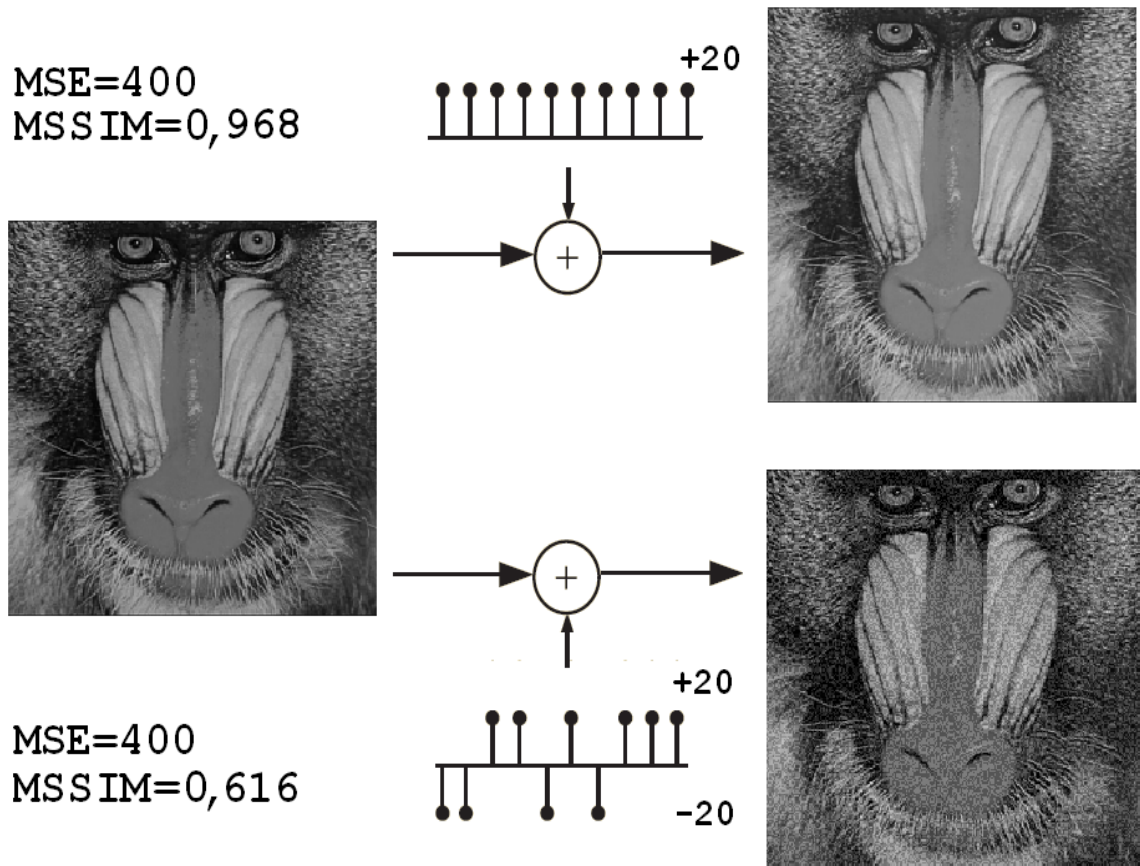
To však neodpovídá tomu, co ve skutečnosti vnímá naše oko. Zkraslení šumem se mnohem více projevuje u obrázku vpravo dole, takže i jeho výsledná vizuální kvalita se jeví daleko horší než u obrázku vpravo nahoře. Přirozené obrázky obsahují objekty s určitou strukturou (tvarem), která je dána rozmístěním a závislostmi mezi jednotlivými pixely v obraze. Použití MSE pro objektivní hodnocení kvality obrazu je v tomto případě nevhodné, jelikož tyto závislosti mezi jednotlivými vzorky obrazového signálu nebere vůbec v potaz.



Obr. 12: Nezávislost MSE na vztahu mezi jednotlivými pixely v obraze (hodnoty metrik MSE a MSSIM se vztahují ke zkrasleným obrázkům)

Vliv znaménka vzorků chybového signálu na hodnotu MSE a výslednou kvalitu obrázku je popsán na **obr. 13**. První zkraslený obrázek vznikl z původního obrázku přičtením konstantní kladné hodnoty ke všem jeho pixelům. Druhý zkraslený obrázek byl vytvořen stejným způsobem, avšak znaménko konstantní hodnoty se pro

jednotlivé pixely náhodně měnilo. Takto vypočítaná hodnota MSE mezi původním a dvěma zkreslenými obrázky je samozřejmě shodná, znovu ale můžeme pozorovat, že výsledná vizuální kvalita se pro oba případy zásadně liší.



Obr. 13: Vliv znaménka vzorků chybového signálu na hodnotu MSE (hodnoty metrik MSE a MSSIM se vztahují ke zkresleným obrázkům)

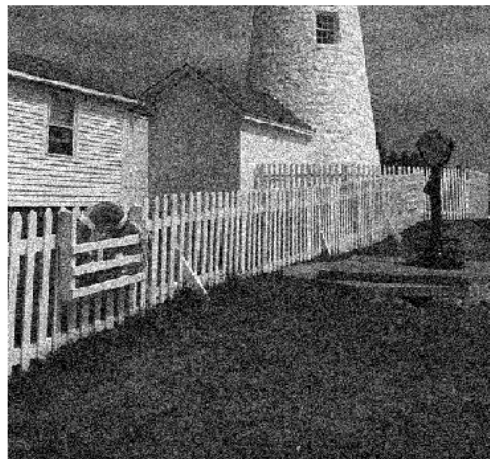
Obr. 14 znázorňuje současně předpoklady 2 a 4. Původní obrázek byl zkreslen přidáním Gaussova bílého šumu. Z pohledu na zašuměný obrázek je zřejmé, že stupeň zkreslení je pro různá místa v obraze značně odlišný. V oblastech s jemnými přechody (světlo - stín, obrysy objektů) působí přítomnost šumu mnohem rušivěji než v ostatních částech obrázku. Chybový signál, který získáme při výpočtu MSE (obrázek vpravo dole), má energii rozloženou rovnoměrně po celém obrázku. Vnímaná vizuální kvalita zkresleného obrázku se však mění v závislosti na tom, kterou část obrázku zrovna pozorujeme.

Tento fakt daleko lépe vystihuje obrázek vlevo dole (mapa hodnot SSIM reprezentující metriku MSSIM – bližší vysvětlení k tomuto obrázku bude v textu uvedeno později v kapitole 1.3.4). Jelikož MSE nijak nerozlišuje jednotlivé pixely v obraze, není možno s touto metrikou vyjádřit rozdílný vliv zkreslení na výslednou vizuální podobu obrázku.

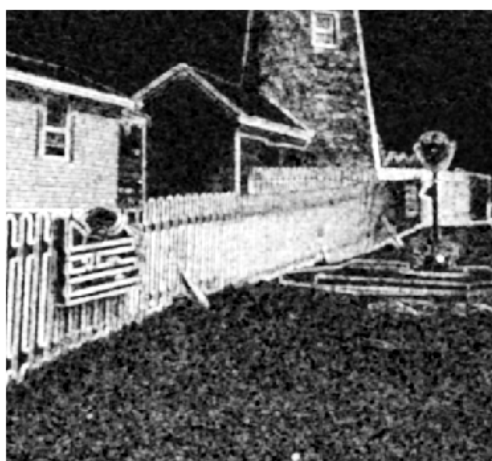
V předešlých příkladech jsme si ukázali, že MSE, i přes své široké využití v praxi, v některých případech naprosto selhává. Abychom objasnili příčiny těchto nedostatků, je potřeba vysvětlit, jakým způsobem vnímá člověk vizuální kvalitu obrazu.



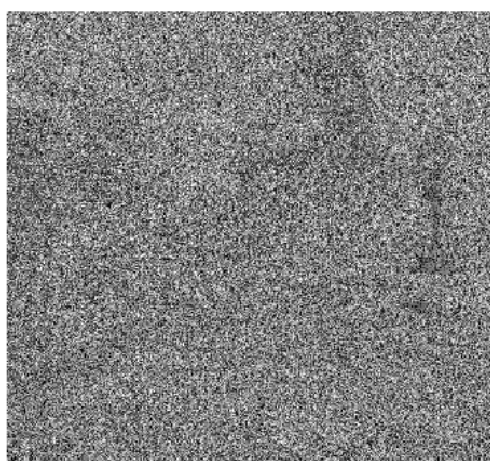
Původní obrázek



Obrázek zkreslený šumem



SSIM mapa



Rozdílový obrázek

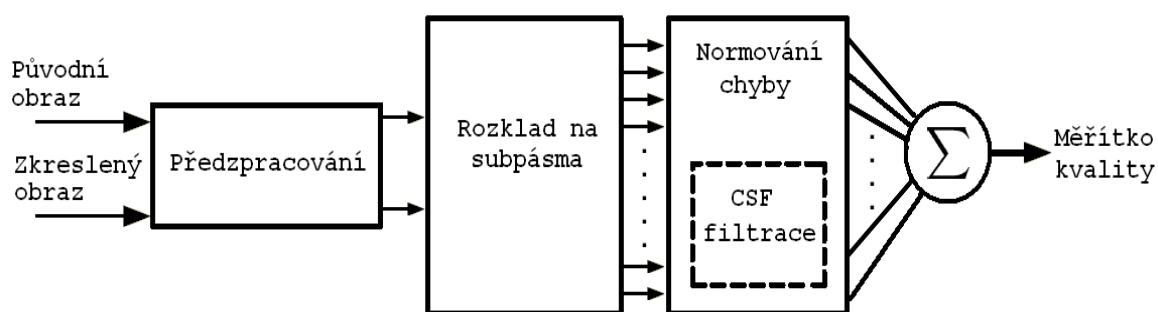
Obr. 1.4: Rozložení chybového signálu při výpočtu MSE (vpravo) a MSSIM (vlevo)

1.2 Model hodnocení kvality obrazu založený na lidském vnímání

Obrazový signál, jehož kvalitu hodnotíme, můžeme vyjádřit jako součet původního nezkresleného signálu a chybového signálu. Ztráta kvality může být poté chápána jako přímý nárůst viditelnosti chybového signálu v obraze. Z tohoto principu vychází rovněž metrika MSE, která určuje množství chybového signálu obsaženého v obrazovém signálu. Avšak dva zkreslené obrázky se stejnou hodnotou MSE, mohou

obsahovat radikálně odlišné druhy chyb, z nichž některé jsou mnohem více viditelné než ostatní.

K popisu vnímání kvality obrazu člověkem využíváme znalosti o modelu HVS (Human Visual System – lidský zrakový systém) [8]. Model vychází z poznatků o lidském vnímání, které byly zjištěny na základě neurofyziologických a psychologických měření. Způsob zpracování zrakového vjemu člověkem je velice složitý proces, stejně tak je složitá i samotná zraková soustava člověka. Model HVS je zjednodušený model, který se snaží vyjádřit jevy spojené s vnímáním zrakové informace pomocí matematických funkcí tak, aby bylo možné co nejlépe a zároveň co nejjednodušeji simulovat chování lidského zrakového systému. Blokový diagram modelu pro stanovení kvality obrazu, který vychází z modelu HVS, je znázorněn na **obr. 1.5** [8].



Obr. 1.5: Blokový diagram hodnocení kvality obrazu

Obrázek představuje zcela obecný model složený z několika bloků, které reprezentují určitou funkci při určování kvality obrazu, bližší popis nalezne čtenář v [8].

1.2.1 Předzpracování

V tomto bloku jsou nejprve původní a zkreslený obrazový signál upraveny do stejného rozsahu hodnot. Následně jsou signály převedeny do barevného modelu, který více odpovídá vlastnostem HVS (např. model YCbCr). Součástí bloku mohou být další úpravy signálů jako např. filtrace signálů dolní propustí, která má za úkol simulovat rozptylovací funkci oka.

1.2.2 Rozklad na subpásma

Obrazy jsou dále rozděleny na jednotlivá subpásma (kanály), která reprezentují signály o různých frekvencích, prostorovém umístění a orientaci. Z poznatků o HVS je známo, že neurony v primární zrakové kůře mozku jsou stimulovány zrakovými podněty specifickými právě těmito vlastnostmi. Transformace obrazů do subpásem tak vlastně simuluje chování mozku.

1.2.3 Normování chyby

V následující sekci je ze vzorků hodnot v subpásmech spočtena chyba mezi původním a zkresleným obrazem pro každý jednotlivý kanál. Normování chyby spočívá ve využití funkce rozlišitelnosti jednotlivých jasových hodnot CSF (Contrast Sensitivity Function) a maskovacího efektu. Tato funkce popisuje rozdílnou citlivost HVS na různé prostorové frekvence, které jsou obsaženy ve zřetelném podnětu. Jednotlivým subpásmům tak mohou být přiřazeny rozdílné váhové konstanty. Maskování pak bere v potaz fakt, že přítomnost určité složky v signálu snižuje viditelnost jiných složek signálu, které jsou nejbližšími sousedy této složky z hlediska frekvence, prostoru či orientace. Výsledná hodnota chyby, která je ukazatelem (měřítkem) kvality obrazového signálu, vznikne kombinací normovaných chyb z jednotlivých kanálů (subpásem).

1.3 Hodnocení kvality obrazu na základě strukturální podobnosti

Výše popsáný model však nebere v potaz důležitý fakt. Přirozené obrázky jsou vysoce strukturované – obsahují řadu objektů navzájem rozlišitelných svými obrysy a tvarem. Pixely těchto obrázků vykazují mezi sebou silné závislosti, zvláště pak, nacházejí-li se blízko sebe. Tyto závislosti v sobě nesou důležitou informaci o struktuře objektů uvnitř obrazu. To je nutné uvažovat i při výpočtu měřítka kvality obrazu, jelikož zraková soustava člověka je uzpůsobena tak, že objekty v obraze vnímá na základě jejich struktury. Změna struktury objektů původního obrazu se člověku jeví jako zhoršení jeho kvality.

Rozkladem obrazu do jednotlivých subpásem závislosti mezi pixely přetrvávají, takže pro správný výpočet chyby v daném subpásmu by bylo třeba zahrnout i závislé koeficienty ze sousedních subpásem. Tento postup by byl však zbytečně složitý a zdlouhavý. Místo toho bylo snahou nalézt jednodušší způsob jak porovnat strukturu původního a zkresleného obrázku. Řešením se stala metrika SSIM (z angl. Structural SIMilarity) [8].

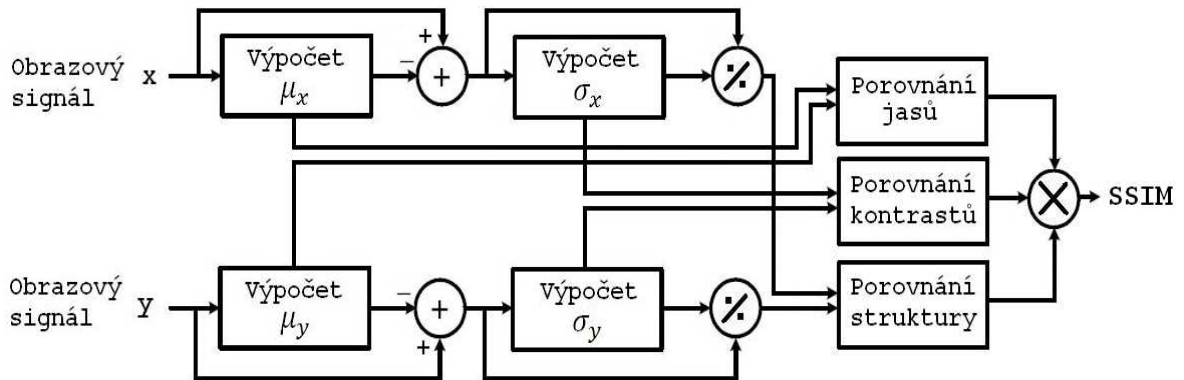
1.3.1 SSIM Index

Index strukturální podobnosti (SSIM) vychází při svém výpočtu z vlastností zjištěných o HVS. Kvalita obrazu není hodnocena podle množství obsaženého chybového signálu, ale z hlediska vlivu chybového signálu na zřetelnost struktury objektů v obraze.

Jas povrchu objektu, který pozorujeme, je výsledkem osvětlení povrchu a jeho odrazivosti. Struktura objektů je však na osvětlení nezávislá, takže abychom získali informaci o struktuře obrazu, je třeba vliv osvětlení nejprve odstranit. Strukturu obrazu poté definujeme jako atributy, které reprezentují strukturu objektů v obraze

nezávisle na průměrném jasu a kontrastu [8]. Blokový diagram výpočtu hodnoty SSIM je znázorněn na **obr. 1.6**. Veškeré rovnice uvedené v této kapitole jsou převzaty z [8].

Obrazové signály x a y představují bloky o stejných rozměrech, které byly vybrány ze dvou různých obrázků. Pokud budeme považovat jeden ze signálů za původní a druhý za jeho zkreslenou verzi, potom hodnota SSIM může sloužit jako měřítko kvality druhého obrazového signálu. Výpočet strukturální podobnosti je založen na porovnání tří parametrů: jasu, kontrastu a struktury.



Obr. 1.6: Blokový diagram výpočtu indexu SSIM

Nejprve je provedeno srovnání signálů z hlediska hodnot jejich jasů. U diskrétních obrazových signálů je střední hodnota jasu μ_x (pro signál x) vypočtena jako průměr přes všechny hodnoty pixelů x_i z celkového počtu N v daném bloku:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i. \quad (1.4)$$

Funkce srovnání jasů $l(x, y)$ následně porovnává střední hodnoty jasů μ_x a μ_y a je definována takto

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (1.5)$$

kde konstantu C_1 zavádíme z důvodu stability výpočtu pro případ, kdy se výraz $\mu_x^2 + \mu_y^2$ blíží nule. Její hodnotu definuje rovnice

$$C_1 = (K_1L)^2, \quad (1.6)$$

kde L značí maximální hodnotu pixelů (255 pro 8 bitů na kanál) a $K_1 \ll 1$ je malá konstantní hodnota [8].

Pro odhad kontrastu v obraze využijeme definici rovnice směrodatné odchylky σ_x (druhá odmocnina hodnoty rozptylu), která má pro obrazový signál \mathbf{x} tuto podobu

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2}. \quad (1.7)$$

Srovnání kontrastů dvou obrazových signálů provádí funkce $c(\mathbf{x}, \mathbf{y})$ na základě porovnání hodnot jejich směrodatných odchylek σ_x a σ_y následujícím způsobem [8]

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

kde $C_2 = (K_2L)^2$ a $K_2 \ll 1$ mají obdobný význam jako konstanty C_1 a K_1 v rovnici (1.6). Důležitým rysem této funkce je, že při stejném rozdílu kontrastů $\Delta\sigma = \sigma_y - \sigma_x$ je funkce méně citlivá na změny v případě vysokých hodnot kontrastů než v případě nízkých hodnot. To odpovídá vlastnostem HVS, které jsou v modelu popsány pomocí funkce CSF a efektu maskování.

Porovnání struktury obrazových signálů lze provést až po odečtení střední hodnoty jasu od hodnot jejich vzorků a normování vůči hodnotě jejich směrodatné odchylky, což můžeme zapsat pomocí výrazů $(\mathbf{x} - \mu_x)/\sigma_x$ a $(\mathbf{y} - \mu_y)/\sigma_y$. Korelace mezi těmito hodnotami odpovídá korelaci mezi vzorky signálů \mathbf{x} a \mathbf{y} a určuje míru strukturální podobnosti signálů. Samotnou funkci pro porovnání struktury $s(\mathbf{x}, \mathbf{y})$ tedy definujeme následovně [8]

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

Konstanta $C_3 = (K_3L)^2$ slouží podobně jako konstanty v předchozích případech k zajištění stability výpočtu. Symbol σ_{xy} reprezentuje zápis rovnice

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

Na rozdíl od předchozích funkcí může $s(\mathbf{x}, \mathbf{y})$ nabývat i záporných hodnot. Konečná podoba matematického vyjádření výpočtu indexu SSIM je kombinací dříve uvedených funkcí z rovnic (1.5), (1.8) a (1.9) a má tento obecný tvar

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha \cdot [c(\mathbf{x}, \mathbf{y})]^\beta \cdot [s(\mathbf{x}, \mathbf{y})]^\gamma, \quad (1.11)$$

kde $\alpha > 0$, $\beta > 0$ a $\gamma > 0$ jsou parametry, které lze využít k zvýraznění důležitosti některé z funkcí. Pro jednoduchost používáme $\alpha = \beta = \gamma = 1$.

Konkrétní podoba rovnice je poté následující

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \cdot \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \cdot \left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right). \quad (1.12)$$

Abychom výraz ještě o něco více zjednodušili, zavádíme $C_3 = C_2/2$. Rovnice (1.12) pak po drobných úpravách přechází ve finální tvar používaný při výpočtu indexu SSIM [8]

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \quad (1.13)$$

Metrika SSIM je symetrická: $\text{SSIM}(\mathbf{x}, \mathbf{y}) = \text{SSIM}(\mathbf{y}, \mathbf{x})$, což znamená, že při srovnání dvou obrázků vrací stejnou hodnotu bez ohledu na pořadí jejich srovnávání. Je rovněž omezená v rozsahu hodnot $-1 < \text{SSIM}(\mathbf{x}, \mathbf{y}) \leq 1$. Maximální hodnota $\text{SSIM}(\mathbf{x}, \mathbf{y}) = 1$ je dosažena tehdy a pouze tehdy, jsou-li dva obrazové signály \mathbf{x} a \mathbf{y} identické. O konkrétním použití a implementaci metriky SSIM při hodnocení kvality obrazů si povíme v následující kapitole.

1.3.2 Použití SSIM pro hodnocení kvality obrazu

Při hodnocení kvality obrazu je užitečné počítat hodnotu SSIM indexu pro menší oblasti s pevně definovanou velikostí spíše než pro celý obraz. Důvodů je hned několik. Zejména je to proměnlivost statistických vlastností obrazu v závislosti na pozici v obraze. Rovněž zkraslení obrazu, které nemusí nutně záviset na statistických vlastnostech určité oblasti, se mění s polohou v obraze. Další příčina vychází z vlastností HVS. Člověk nacházející se v určité vzdálenosti od obrazu je v daném časovém okamžiku schopen vnímat s vysokou rozlišovací schopností pouze určitou část obrazu [8]. Vypočítáme-li hodnotu SSIM pro více oblastí, získáme podrobnější informace o ztrátě kvality obrazu.

Hodnoty veličin μ_x (μ_y), σ_x (σ_y) a σ_{xy} , které jsou nezbytné pro stanovení indexu SSIM, tedy počítáme pro zvolený počet pixelů. Je možné použít čtvercové okno velikosti 8×8 pixelů, což je typická velikost bloku využívaná při zpracování obrazu po segmentech. Okno se v obraze posouvá s krokem jeden pixel a pro každý krok je spočtena hodnota SSIM uvnitř definovaného okna. Zmíněná metoda má svou nevýhodu. Jelikož jsou hodnoty SSIM počítány v rámci dané oblasti nezávisle na ostatních pixelech, projevují se ve výsledném obrazu sestaveném z těchto hodnot (SSIM mapě) nežádoucí „blokové“ artefakty známé zejména z kompresního algoritmu JPEG. Proto namísto předchozího okna byla zvolena oblast o velikosti 11×11 pixelů, uvnitř které jsou hodnoty pixelů váhovány Gaussovou funkcí $\mathbf{w} = \{w_i | i = 1, 2, \dots, N\}$ se směrodatnou odchylkou 1,5 vzorku [8]. Hodnoty vzorků Gaussovy funkce jsou normovány tak, aby jejich výsledný součet byl roven 1 a nezměnil se tak celkový jas uvnitř okna.

Výpočet hodnot μ_x , σ_x a σ_{xy} poté probíhá podle těchto rovnic [8]

$$\mu_x = \sum_{i=1}^N w_i x_i, \quad (1.14)$$

$$\sigma_x = \left(\sum_{i=1}^N w_i (x_i - \mu_x)^2 \right)^{1/2}, \quad (1.15)$$

$$\sigma_{xy} = \sum_{i=1}^N w_i (x_i - \mu_x)(y_i - \mu_y). \quad (1.16)$$

S použitím tohoto okna dochází k výraznému potlačení blokových artefaktů, což se příznivě projevuje i na celkovém vzhladu SSIM mapy.

Pro stabilizaci výpočtu SSIM indexu byly zvoleny hodnoty konstant $K_1 = 0,01$ a $K_2 = 0,03$. Tyto konstanty mohou být voleny takřka libovolně, jelikož nemají žádný výrazný vliv na algoritmus výpočtu SSIM, jedinou podmínkou je, aby jejich hodnota byla dostatečně nízká.

Abychom mohli porovnat SSIM s ostatními metrikami hodnocení kvality obrazu většinou vyžadujeme jedinou hodnotu indexu pro celý obraz. Tu získáme jako průměr všech hodnot SSIM vypočítaných pro jednotlivé oblasti. Metriku poté značíme jako MSSIM (Mean SSIM) a definujeme ji následovně [8]

$$\text{MSSIM}(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(\mathbf{x}_j, \mathbf{y}_j), \quad (1.17)$$

kde \mathbf{X} a \mathbf{Y} je původní resp. zkreslený obrázek, \mathbf{x}_j a \mathbf{y}_j jsou části obrázků vymezené j -tým posunem okna a M je celkový počet možných posunů okna v obrázku. Potřebujeme-li se zaměřit více na některou oblast v obraze, je možné přiřadit posunům okna různé váhovací konstanty a zvýraznit tak význam určité části obrazu pro výpočet MSSIM. My se však v dalším popisu budeme držet klasického přístupu.

1.3.3 Implementace výpočtu indexu SSIM v MATLABu

Za účelem srovnání metrik MSE a MSSIM při hodnocení kvality obrazu byla s využitím zdrojového kódu dostupného z [8] vytvořena jednoduchá aplikace v prostředí MATLAB. Základem této aplikace je funkce `ssim.m`, jejímž vstupem jsou dva porovnávané obrázky a výstupem pak mapa SSIM indexů a hodnota metriky MSSIM mezi těmito obrázky. Algoritmus výpočtu jednotlivých indexů SSIM je shodný

s postupem uvedeným v předešlé kapitole. Neuvažujeme úpravu rozlišení obrázků, kterou je možnou provést před samotným výpočtem. Hodnoty vzorků okna w velikosti 11×11 pixelů jsou nastaveny využitím předdefinované funkce MATLABu:

```
w = fspecial('gaussian', 11, 1.5).
```

Výpočet parametrů $\mu_x(\mu_y)$, $\sigma_x(\sigma_y)$ a σ_{xy} vychází z rovnic (1.14), (1.15) a (1.16). Okno w si můžeme představit jako konvoluční jádro, které aplikujeme na obraz, u něhož chceme potřebné parametry zjistit. Dvourozměrnou konvoluci jádra s obrazem lze v MATLABu realizovat funkcí `filter2`, tedy např. hodnoty $\mu_x(\mu_y)$ pro jednotlivé oblasti obrazu získáme takto

```
ux = filter2(w, img_x, 'valid'),
```

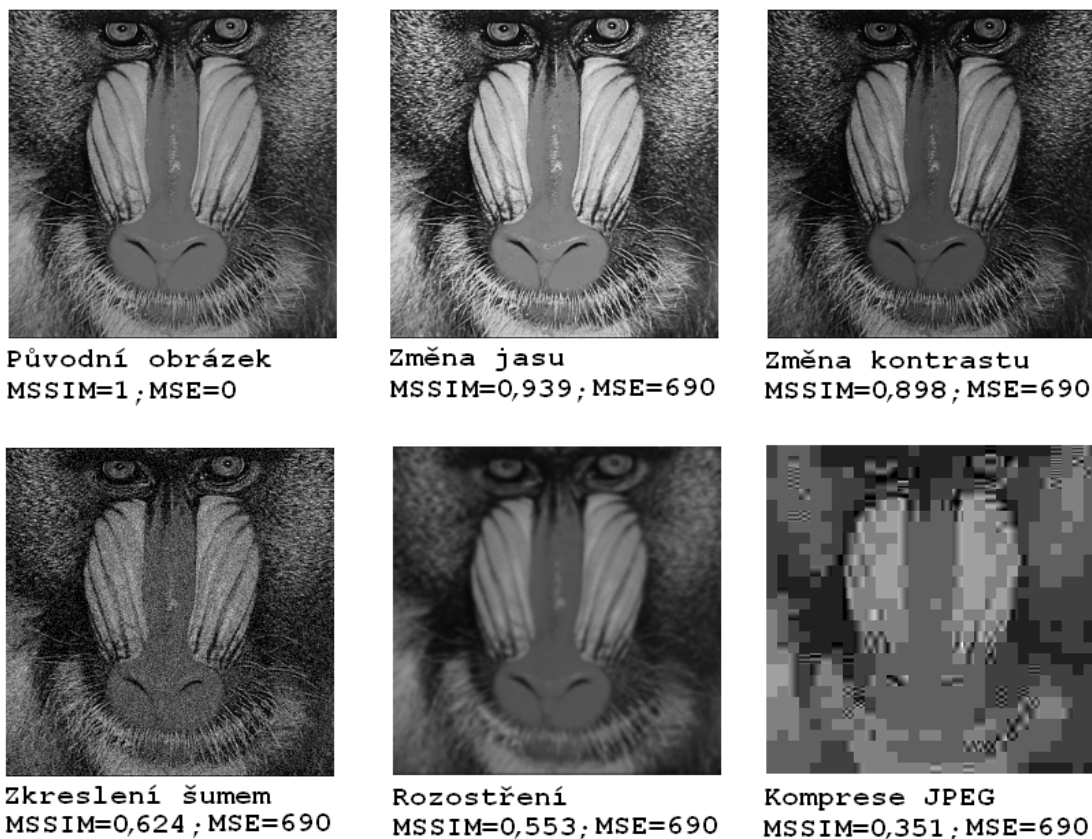
kde `img_x` představuje obraz, jehož parametry zjišťujeme. Obdobným způsobem se získají hodnoty i zbylých parametrů $\sigma_x(\sigma_y)$ a σ_{xy} . Mapa indexů SSIM vznikne kombinací hodnot SSIM vypočtených na základě rovnice (1.13). Výsledná metrika MSSIM se spočítá jako průměr indexů SSIM podle (1.17).

Aplikace umožňuje porovnat hodnoty MSE a MSSIM pro různé druhy zkreslení původního obrázku (šum, změna jasu, změna kontrastu, komprese, rozostření atd.) a zároveň zobrazit v podobě rozdílového obrázku a SSIM mapy, jak se u jednotlivých metrik projevuje chybový signál.

1.3.4 Porovnání MSE a MSSIM

Porovnání hodnot obou metrik lze pozorovat již v kapitole 1.1.1. U **obr. 1.1**, **1.2** i **1.3** je vždy patrné, že ačkoliv se subjektivní kvalita dvou zkreslených obrázků značně liší, hodnota MSE toto nijak nereflektuje. Oproti tomu index MSSIM daleko lépe vystihuje skutečný vliv chybového signálu na výslednou vnímanou podobu původních obrázků. Na **obr. 1.4** je vyobrazena SSIM mapa vytvořená sloučením hodnot SSIM indexů spočtených pro jednotlivé oblasti obrázku. Je zde možné dobře vidět, že v některých místech obrázku se zkreslení projevuje mnohem více, i když je hodnota přidaného šumu pro všechny pixely stejná (světlejší odstíny reprezentují vyšší kvalitu).

Následující sled obrázků (**obr. 1.7**) dokumentuje, proč je pro hodnocení kvality obrázků výhodnější použít metriku MSSIM spíše než MSE. Původní obrázek zde byl změněn různými druhy zkreslení tak, aby jeho hodnota MSE byla pro všechny zkreslené obrázky přibližně stejná. Je však zřejmé, že kvalita zkreslených obrázků je výrazně odlišná a její změna se odráží pouze v indexu MSSIM. Změna jasu a změna kontrastu, či obecně jakákoliv zkreslení, která nemají vliv na degradaci struktury objektů v obraze, vykazují vysoké hodnoty MSSIM. Naproti tomu rozostřený nebo silně zašuměný obrázek, u nichž dochází ke ztrátě informace o původní struktuře objektů, mají index MSSIM nižší. Nejvíce patrná je deformace struktury u obrázku podrobenému vysokému stupni komprese JPEG, a proto i hodnota MSSIM je v tomto případě nejnižší.



Obr. 1.7: Srovnání hodnot MSE a MSSIM pro různé druhy zkreslení

Ještě lépe lze efektivitu MSSIM vypočítat z **obr. 1.8**, kde je porovnávána SSIM mapa s rozdílovým obrázkem (chybový signál u metriky MSE). Původní obrázek zde byl zkomprimován algoritmem JPEG2000 na bitový poměr 0,8 bpp (0,8 bitů na pixel). V obou případech reprezentuje jas míru zkreslení v obraze, světlejší odstíny představují v dané metrice lepší kvalitu (za tímto účelem byly pro snazší srovnání s MSSIM upraveny hodnoty jasů chybového signálu u MSE). Oproti rozdílovému obrázku, který zobrazuje téměř rovnoměrné rozprostření chybového signálu, zvýrazňuje SSIM mapa místa, kde se v komprimovaném obrázku nejvíce projevuje zkreslení.

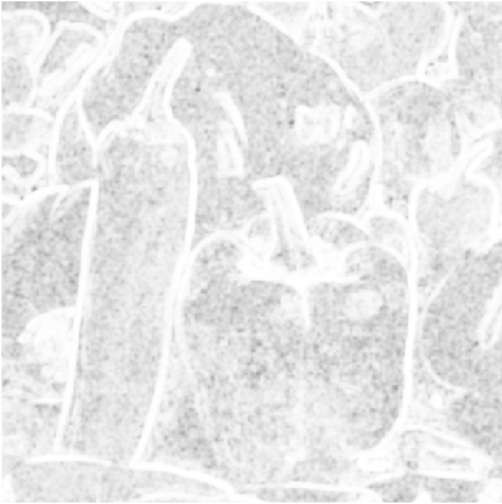
Z výše provedeného srovnání metrik MSSIM a MSE je zřejmé, že z hlediska hodnocení objektivní kvality obrázku se subjektivní kvalitě vnímané člověkem daleko více přibližuje metrika MSSIM. Vhodnou oblastí využití tohoto kritéria je zejména ztrátová komprese obrazových signálů, při které vždy dochází (ať už méně či více) k deformaci struktury objektů v obraze. Metriku MSSIM je zde užitečné použít nejen k hodnocení efektivity kompresních algoritmů, ale rovněž k jejich optimalizaci.



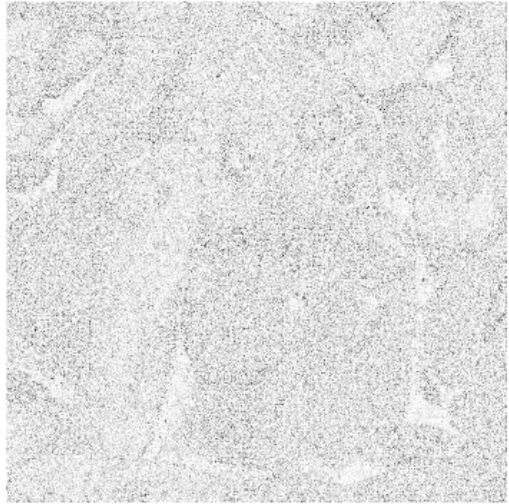
Původní obrázek



Kompresa JPEG2000 0,8 bpp



SSIM mapa



Rozdílový obrázek

Obr. 1.8: Srovnání SSIM mapy s rozdílovým obrázkem

2 OBJEKTIVNÍ HODNOCENÍ KVALITY U KOMPRESI OBRAZŮ

Oblast ztrátové komprese byla vždy odvětvím, kde bylo snahou dosáhnout co nejvyšší kvality obrazu v kombinaci s nejnižší možnou velikostí komprimovaných obrazových dat. Řešení spočívalo v optimalizaci kompresních algoritmů podle zvoleného kritéria. Toto kritérium nejčastěji představovala tradičně metrika MSE, kde úloha optimalizace řeší problém nalezení minimální hodnoty této metriky pro požadovaný kompresní poměr (přidělený počet bitů).

Z předchozích kapitol je však zřejmé, že z pohledu subjektivní kvality obrazu vnímané člověkem není výsledek této úlohy zcela ideálním řešením. V obrázku podrobenému kompresi sice dochází ke snížení celkového množství (energie) obsaženého chybového signálu, tato změna se však často neprojeví na vnímané kvalitě, jelikož MSE nijak nerozlišuje různý vliv chybového signálu na zkreslení jednotlivých oblastí obrázku.

Nabízí se proto jako kritérium optimalizace využít index SSIM, který daleko lépe vystihuje skutečný vliv komprese na pozorovatelnou kvalitu obrázku. Zvýšení celkové hodnoty MSSIM se rovněž příznivě odráží na výsledném vzhledu obrázku (viz např. **obr. 1.7**). Požadovaným řešením tedy bude nalézt takový způsob optimalizace, pro který by se zvolený kompresní algoritmus pokoušel při daném počtu přidělených bitů dosáhnout co nejvyšší hodnoty MSSIM mezi původním a komprimovaným obrázkem.

2.1 Použití indexu SSIM při ztrátové kompresi obrazu

Je třeba připomenout, že na rozdíl od MSE se index SSIM stanovuje nikoliv pro celý obraz, ale pouze pro určitý počet pixelů, který je dán velikostí výpočetního okna (typicky 11×11 pixelů [8]). Výsledná hodnota MSSIM je poté průměrem dílčích hodnot SSIM. Index SSIM se tedy vždy vztahuje k určité skupině pixelů umístěných v obraze blízko sebe. Kombinací indexů získáme mapu obrazu, která umožňuje odhalit místa s největším zkreslením v obraze.

Optimalizace kompresních algoritmů řeší tradiční problém: s dostupným počtem bitů (tzv. „bit budget“) zakódovat původní obrázek v co nejlepší kvalitě. V našem případě představuje lepší kvalitu vyšší hodnota SSIM. Zvýšit hodnotu SSIM pro určitou část obrazu znamená přidělit koeficientům příslušejícím pixelům dané části větší počet bitů. Koeficienty reprezentují pixely v transformované oblasti, do které převádíme obrazový signál z časové oblasti za účelem lepší interpretace jeho obsahu. V této oblasti pak dochází k rozdělení dostupného počtu bitů. Abychom mohli změnit hodnoty pixelů v časové oblasti, je nezbytné, aby transformace poskytovala informaci o tom, kterým koeficientům tyto pixely odpovídají v transformované oblasti. Takto jsme schopni upravit pixely v části obrázku, kde požadujeme zvýšení

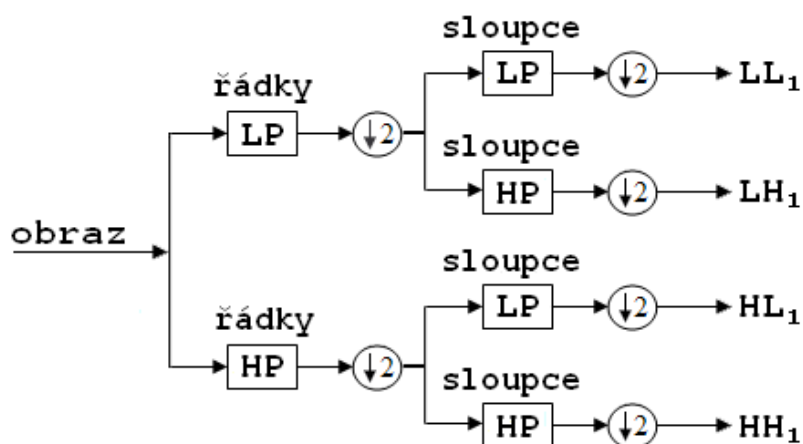
hodnoty SSIM. U kódování koeficientů je rovněž potřeba možnost jednoduchého přerozdělení bitů.

Těmto požadavkům nejlépe vyhovuje vlnková (waveletová) transformace společně v kombinaci s algoritmem SPIHT (Set Partitioning In Hierarchical Trees) [4]. Oba nástroje dohromady poskytují velmi efektivní způsob komprese obrazu. Pro naše účely je však výhodné zejména to, že postup algoritmu lze upravit takovým způsobem, že jednotlivé bity mohou být koeficientům přidělovány podle toho, ke které skupině pixelů v obraze tyto koeficienty náleží. Z tohoto důvodu jsme algoritmus SPIHT, který kóduje koeficienty vytvořené vlnkovou transformací, zvolili jako vhodný prostředek pro optimalizaci podle kritéria SSIM. Teorie k vlnkové transformaci i k samotnému SPIHT je značně obsáhlá, proto se v následující kapitole zmíníme pouze o základních principech, na nichž je algoritmus SPIHT postaven.

2.1.1 Vlnková transformace obrazu

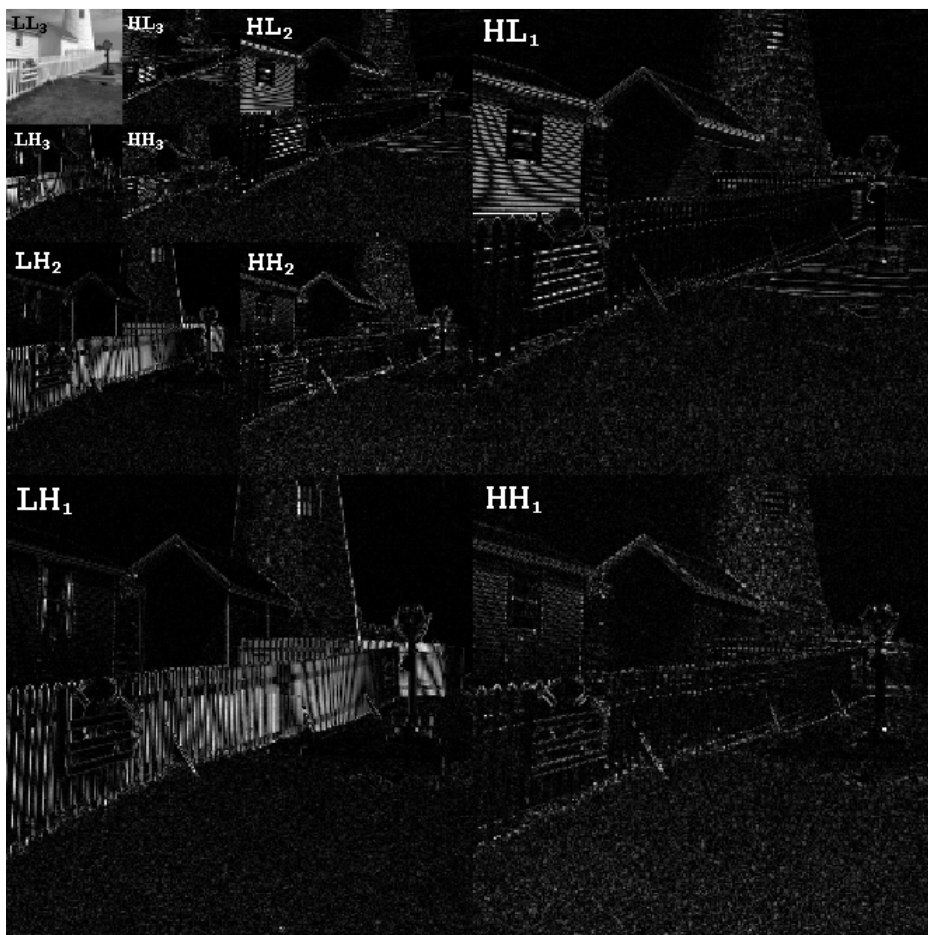
Jako první krok před samotnou aplikací algoritmu SPIHT je nejprve nutné provést transformaci obrazu s využitím dvourozměrné diskrétní vlnkové transformace v diskrétním čase (2D-DTWT). Výsledkem této operace je tzv. dekompoziční obrazec, který je složen z vlnkových koeficientů rozdělených v obrazci na čtyři různé části (subpásma). Tato subpásma označujeme LL_n , HL_n , LH_n a HH_n , kde L (low-pass) a H (high-pass) představují filtry typu dolní resp. horní propust a n značí tzv. hloubku dekompozice [4], [5]. Pořadí písmen udává filtraci obrazu v řádcích a následně ve sloupcích. Každé ze subpásem reprezentuje jinou obrazovou informaci.

Schéma algoritmu 2D-DTWT pro první úroveň dekompozice je znázorněno na **obr. 2.1**. Vstupní obraz je nejprve filtrován po řádcích dvojicí filtrů, dolní (LP) a horní (HP) propust. Následuje decimace dvěma (vypuštění každého sudého vzorku) a takto upravený výsledek filtrace je stejným způsobem zpracován po sloupcích. Výstupem algoritmu jsou čtyři skupiny koeficientů – aproximační LL_1 , horizontální LH_1 , vertikální HL_1 a diagonální HH_1 .



Obr. 2.1: Algoritmus 2D-DTWT pro první úroveň dekompozice

Koeficienty v subpásmu LL_n mohou být opakovaně podrobeny 2D-DTWT, čímž se zvyšuje hloubka dekompozice obrazu. Současně s každým dalším stupněm dekompozice rovněž roste i energie koeficientů v nejvyšším subpásmu (LL_n vlevo nahoře) [4]. Ukázka dekompozičního obrazce pro hloubku dekompozice $n = 3$ je na obr. 2.2.

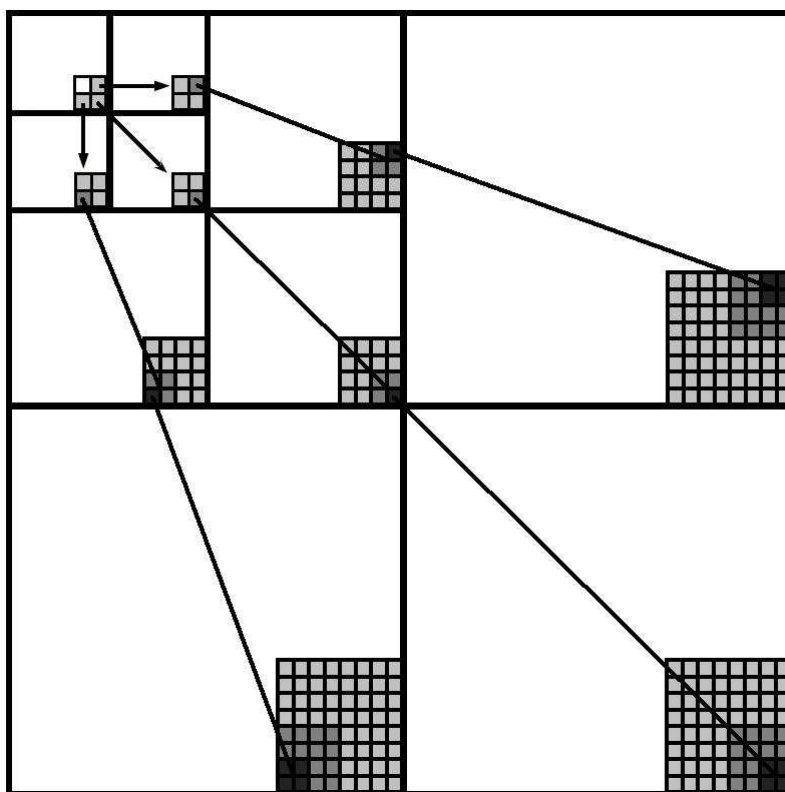


Obr. 2.2: Dekompoziční obrazec s hloubkou dekompozice $n = 3$

2.1.2 Algoritmus SPIHT

Algoritmus SPIHT následně zpracovává vlnkové koeficienty vytvořené 2D-DTWT. S dostupným počtem bitů kóduje jejich hodnoty, čímž přímo ovlivňuje výslednou kvalitu rekonstruovaného obrazu. Výhodou je, že postup kódování lze kdykoliv v jeho průběhu zastavit a na výstupu bude obraz v kvalitě odpovídající počtu použitých bitů. Myšlenka algoritmu SPIHT spočívá ve využití struktur, které nazýváme prostorově orientované stromy (SOT) [4]. Každé subpásmo obsahuje koeficienty, které souvisí se stejnou oblastí v původním obraze. Vazby mezi koeficienty vytváří v dekompozičním obraze stromové struktury. Místo s jednotlivými koeficienty se tedy pracuje s celými stromy.

Kořeny stromů se nacházejí v nejvyšším subpásmu jako součást bloku čtveřice koeficientů. V každé této čtveřici není kořenem žádného stromu koeficient z levé horní části bloku. Všechny kořeny stromů a následně i všechny jeho prvky mají vždy čtyři přímé potomky (s výjimkou prvků v nejnižších subpásmech). Souřadnice koeficientů umístěných v bloku čtveřice potomků vlevo nahoře jsou přesně dvojnásobkem hodnoty souřadnic jejich rodiče (neplatí pro nejvyšší subpásmo) [4]. Stromová struktura koeficientů je znázorněna na **obr. 2.3**.



Obr. 2.3: Stromová struktura používaná algoritmem SPIHT

Jelikož největší koeficienty se v dekompozičním obrazci nacházejí v nejvyšším subpásmu, je algoritmus iniciován právě zde. Nejprve se určí hodnota srovnávacího prahu jako nejvýznamnější bit koeficientu s největší hodnotou. Pro každý nový krok je vždy hodnota prahu snížena o jednu bitovou hladinu (tzn. na polovinu). Následuje tzv. *řadící průběh*, který v každém kroku rozděljuje prvky (koeficienty nebo množiny koeficientů v podobě stromů) podle významnosti vůči aktuální hodnotě prahu. Prvky s hodnotami většími nebo rovnými hodnotě prahu jsou v daném kroku označeny za významné (výstupní bit 1), prvky menší než práh za nevýznamné (výstupní bit 0) [4].

U významných koeficientů je kódován znaménkový bit, nevýznamné koeficienty nebo stromy (množiny koeficientů) se dále v daném kroku neuvažují. Významné stromy jsou rozděleny na podstromy, u nichž probíhá nové srovnání s prahem. Tento proces se neustále opakuje, pokud existuje nějaký významný podstrom nebo do doby, než se narazí na konec celého stromu (celý strom je rozložen až na jednotlivé pixely).

Za účelem uchování informace o významnosti koeficientů a množin koeficientů v jednotlivých krocích byly vytvořeny tři dynamické seznamy:

- LIP (list of insignificant pixels) – seznam nevýznamných koeficientů
- LSP (list of significant pixels) – seznam významných koeficientů
- LIS (list of insignificant sets) – seznam nevýznamných množin koeficientů

Prvky jsou v těchto seznamech reprezentovány svými souřadnicemi. V seznamech LIP a LSP představují souřadnice prvků přímo koeficienty, u seznamu LIS označují souřadnice rodiče daného stromu (množiny koeficientů). Před započítáním testování jsou do LIP uloženy všechny kořeny z nejvyššího subpásma a do LIS pouze ty, které jsou kořeny některého stromu. Seznam LSP je prázdný [4].

V *řadícím průběhu* jsou s aktuální hodnotou prahu srovnávány koeficienty v seznamu LIP (v předešlém kroku označeny za nevýznamné). Ty, které jsou významné, jsou přesunuty do LSP. Stejným způsobem jsou testovány množiny koeficientů v LIS. U významných množin (stromů) následuje jejich rozdělení na podmnožiny, původní množina je ze seznamu LIS smazána. Nevýznamné podmnožiny jsou uloženy na konec seznamu LIS. Pokud podmnožinu tvoří pouze samostatný koeficient, je přidána do LSP nebo do LIP v závislosti na významnosti koeficientu v příslušném kroku [4].

Na závěr každého kroku přichází *upřesňovací průběh*. Pro každý koeficient v seznamu LSP (kromě těch, které byly přidány v současném kroku) se na výstup posílá jeden bit na pozici odpovídající pozici nejvyššího bitu aktuálního prahu. Následně je práh redukován na poloviční hodnotu a pokračuje se dalším krokem. Algoritmus je zastaven, je-li vyčerpán veškerý počet přidělených bitů („bit budget“). V případě, kdy se algoritmus nechá proběhnout až do konce, představují výsledná data téměř bezztrátový obraz (chyba vzniká pouze v důsledku zaokrouhlování při výpočtech) [4].

Výhodou komprese obrazu pomocí SPIHT je, že vzniklý bitový tok lze progresivně dekódovat (s přesností na jediný bit), to znamená, že kvalita rekonstruovaného obrazu se s každým nově přichozícím bitem postupně zlepšuje. Na druhé straně je tato vlastnost příčinou zásadních nedostatků algoritmu. Výstupní bitový tok je vysoce citlivý na případné chyby, ty většinou vedou ke kompletnímu rozpadu informace o výsledném obrazu. Operace algoritmu jsou prováděny na celém obraze, tudíž je potřeba v dynamických seznamech (LIP, LIS, LSP) udržovat informace o všech koeficientech, v důsledku toho je realizace těchto seznamů hardwarově náročná. Výhodnější způsob pro nás představuje segmentace obrazu a následné kódování koeficientů v blocích. To umožňuje paralelní zpracování, menší velikost dynamických seznamů a rovněž nižší citlivost vůči chybám (v případě výskytu chyby je ovlivněn pouze blok, kde daná chyba nastala) [3], [4], [5], [10].

3 VYUŽITÍ KRITÉRIA SSIM U KOMPRESI OBRAZU ZALOŽENÉ NA SPIHT

Algoritmus SPIHT ve své základní podobě představuje efektivní kompresní nástroj, avšak výsledná vizuální kvalita komprimovaného obrázku nemusí být při daném počtu využitých bitů zcela ideální. Problém spočívá ve vhodném rozložení dostupných bitů v celém obrázku. SPIHT kóduje jednotlivé koeficienty podle jejich velikosti – nejprve jsou bity přidělovány koeficientům s nejvyšší hodnotou, avšak bez ohledu na to, ke které části obrázku se příslušné koeficienty vztahují. Jak již jsme však zmínili dříve, různé oblasti v obraze mohou být v závislosti na charakteru hodnot různě citlivé na případná zkreslení. Při shodném počtu přidělených bitů dvěma nezávislým oblastem se stejným počtem pixelů se může vnímaná kvalita těchto oblastí značně lišit. Více postiženy degradací kvality jsou zejména části obrázku obsahující nějaké detaily (např. ostré přechody apod.).

Snahou tedy bude nalézt určité optimální rozdělení bitů pro jednotlivé koeficienty, které se příznivě odrazí ve výsledné vizuální kvalitě obrázku. V předešlých kapitolách jsme ověřili, že vhodným indikátorem této kvality je index SSIM. Proto pro hledání ideálního rozložení bitů nám bude sloužit právě tato metrika.

Z důvodu prozkoumání různých variant využití indexu SSIM u kompresi obrazu založené na algoritmu SPIHT byly nastudovány literární prameny [3], [9], [10]. Metody [3] a [10] se pokouší o minimalizaci hodnoty MSE, avšak obecně lze v těchto případech zlepšit výslednou kvalitu komprimovaného obrázku i na základě použití kritéria SSIM. V následujících kapitolách si oblasti možného uplatnění indexu SSIM stručně popíšeme.

3.1 Zvyšování kvality části obrazu s minimální hodnotou indexu SSIM

Ústřední myšlenkou metody [9] je postupné rozdělování dostupných bitů v závislosti na hodnotě indexu SSIM v jednotlivých oblastech obrázku. Kódování obrázku algoritmem SPIHT probíhá iterativně. V každé iteraci dochází k přidělení bitů jednotlivým koeficientům podle hodnot SSIM mapy zjištěné mezi původním a komprimovaným obrázkem v předchozí iteraci. Cílem je zlepšit kvalitu té části obrazu, která vykazuje nejmenší hodnotu indexu SSIM a má tedy i nejnižší pozorovatelnou kvalitu ve výsledném obrázku.

Opodstatnění této metody vychází z vlastností HVS, kdy oblast s velmi nízkou vizuální kvalitou na sebe upíná pozornost pozorovatele na úkor zbytku obrázku a negativně tak ovlivňuje celkové hodnocení kvality obrázku.

Za účelem požadovaného rozdělení dostupných bitů byla definována tzv. „funkce ořezávání“ [9]. Před samotným kódováním pomocí algoritmu SPIHT jsou

koeficienty po transformaci upraveny touthle funkcí tak, že u nich dochází k „odříznutí“ části bitů pod stanovenou úrovní (bity pod touthle úrovní jsou vynulovány). Bitová hladina, na které je provedeno „odříznutí“ se určuje pro každý koeficient zvlášť v závislosti na jeho důležitosti. „Odříznutí“ vlastně odpovídá kvantování koeficientu na dané úrovni. Na takto upravené koeficienty je následně aplikováno kódování algoritmem SPIHT.

Celý postup této metody se skládá z následujících kroků:

1. Na počátku je nastavena „*funkce ořezávání*“ tak, že úroveň, od které jsou bity nulovány, je pro všechny koeficienty stejná. Tato úroveň musí být dostatečně vysoká, aby bylo možné pro daný počet přidělených bitů zakódovat všechny bity nad touto úrovní a zároveň určitá část bitů zůstala nevyužitá.
2. Obrázek s upravenými koeficienty je kódován a následně dekódován, čímž je získána výsledná podoba komprimovaného obrázku.
3. Mezi původním a komprimovaným obrázkem jsou vypočítány hodnoty indexů SSIM, které se vztahují k jednotlivým oblastem.
4. Z vypočtených indexů SSIM je vybrán ten, který má nejnižší hodnotu (značí se Min-SSIM). Současně je nalezena v obrázku oblast, ke které tento index přísluší. „*Funkce ořezávání*“ je upravena tím způsobem, že ke každému koeficientu, který se vztahuje k oblasti s minimální hodnotou SSIM, je přidán jeden bit. To představuje příslušné koeficienty ve všech subpásmech. U těchto koeficientů je tedy úroveň „*funkce ořezávání*“ snížena o jednu bitovou hladinu. Pokud se hodnota Min-SSIM po několik iterací nemění, je postup zastaven a „*funkce ořezávání*“ má svůj konečný tvar. V opačném případě se pokračuje další iterací od kroku 2.

K ukončení celého postupu dochází vždy, pokud je „*funkce ořezávání*“ upravena do takové míry, že při kódování koeficientů algoritmem SPIHT je vyčerpán veškerý počet dostupných bitů. Metoda vždy zaručeně konverguje a výsledkem je zřetelné zlepšení vizuální kvality v části obrázku s původně největším zkreslením. Podstatnou nevýhodou této metody je však vysoká výpočetní a časová náročnost, jelikož k nalezení potřebného řešení je nutný velký počet iterací [9].

3.2 Kódování koeficientů algoritmem SPIHT v nezávislých blocích

Metoda [10] byla navržena pro účely možného paralelního zpracování vlnkových koeficientů v blocích pomocí algoritmu SPIHT. Rovněž při postupném (sekvenčním) kódování bloků není třeba současně pracovat se všemi koeficienty obrazu, čímž klesají nároky na velikost potřebné (operační) paměti.

Koeficienty v dekompozičním obrazci jsou rozděleny do několika nezávislých bloků totožné velikosti. Blok je vždy složen ze všech koeficientů jednoho či více stromů. Počet stromů je volen tak, aby rozměry bloků měli požadovanou velikost, typicky 64×64 nebo 128×128 koeficientů [10]. Stromy jsou následně v blocích přeskupeny takovým způsobem, aby byly zachovány vztahy mezi jednotlivými koeficienty. Každý blok pak vlastně reprezentuje určitou oblast v původním obrázku.

Jednotlivé bloky jsou poté postupně zpracovány algoritmem SPIHT. Zde nastává problém, jelikož v době kódování bloku není známo, kolik z dostupného počtu bitů je možno koeficientům v tomto bloku přidělit. Z tohoto důvodu dochází k záměrnému zakódování bloku na vyšší počet bitů (typicky trojnásobný [10]), než je předpokládaný počet bitů pro výsledný blok v komprimovaném obraze. Každý blok je poté reprezentován nadměrným počtem bitů. Množství bitů pro jednotlivé bloky jsou upravena na jejich požadovanou velikost až následným přerozdělením bitů, kde dochází k odstranění nadbytečných bitů.

Operace přerozdělení bitů využívá informací získaných při kódování koeficientů v jednotlivých blocích. Účelem je rozdělit dostupný počet bitů mezi všechny bloky tak, aby byla minimalizována hodnota MSE mezi původním a komprimovaným obrázkem. Pro každý blok, v každém kroku (s daným srovnávacím prahem) a pro každý průběh (řadicí a upřesňovací) algoritmu SPIHT je spočten počet nově nalezených významných koeficientů, počet upřesněných koeficientů a počet výstupních bitů v řadicím a upřesňovacím průběhu. Hodnotám získaným pro vyšší srovnávací práh je přisuzována větší váha. Předpokládá se, že čím významnější bit většího koeficientu je v daném bloku dekódován, tím více se sníží zkreslení v odpovídající části obrázku. Pro každý blok je ze zjištěných hodnot sestavena funkce popisující pokles zkreslení při případném dekódování všech bitů v tomto bloku. Následná úloha přerozdělení bitů spočívá v postupném přidání bitů těm blokům, jejichž funkce má nejnižší pokles zkreslení a naopak odebrání bitů blokům s nejvyšším poklesem. Pro rozdělení může být využit pouze přidělený počet bitů („bit budget“).

V tomto případě je metoda [10] navržena pro přidání bitů blokům, jejichž koeficienty přísluší částem obrazu s nejvyšší hodnotou MSE. Obdobně by bylo možné provést přerozdělení bitů tak, aby došlo k zvýšení indexu SSIM pro bloky, jimž v obraze odpovídají oblasti s největším vizuálním zkreslením (nejnižším SSIM). Na straně kodéru je třeba provést dekódování bloků zakódovaných na vyšší bitový poměr a provést vždy srovnání příslušné části s původní částí obrazu na základě kritéria SSIM. Ze získaných hodnot SSIM je poté možné stanovit pro přidělený počet bitů nové rozdělení, jenž bude zvýhodňovat bloky s nejnižším indexem SSIM. Nevýhodou tohoto způsobu bitové alokace je vyšší časová a výpočetní náročnost, jelikož původní obrázek nelze zakódovat do jeho konečné podoby bez předchozího kódování bloků na vyšší bitový poměr.

3.3 Komprese barevného obrazu založená na algoritmu SPIHT s odděleným zpracováním kanálů

Jak je patrné již ze samotného názvu, tato metoda pracuje s barevnými obrazy. Původní algoritmus SPIHT byl navržen pouze pro zpracování jediného kanálu (šedotónové obrazy), tudíž bylo nutné nalézt požadované řešení i pro kódování obrazů s více kanály. Jednou z možností u barevných obrazů se složkami YCbCr (jasová a dvě chrominanční) je použití metody CSPIHT, která modifikuje původní SPIHT tak, že zavádí prostorové stromy i mezi subpásmy jasových a barevných složek. My se však dále zaměříme na jinou metodu, jejíž podstatou je oddělené kódování složek.

Metoda [3] rozděluje bity mezi roviny (složky) barevného obrazu na základě jejich důležitosti. Pro zakódování výsledného obrazu je přiděleno určité množství bitů („bit budget“), z kterého se kódují koeficienty příslušných složek. Snahou je přidělit bity jednotlivým kanálům v poměru, který by co nejvíce minimalizoval hodnotu MSE mezi původním a komprimovaným obrazem. K určení požadovaného rozdělení je využito vlastností mezi koeficienty v subpásmech zpracovávaných složek. Vychází se z poznatků, že efektivita komprese se přímo odvíjí od rozdílů hodnot mezi koeficienty, které mají být kódovány. Tyto rozdíly mohou být vyjádřeny pomocí rozptylu σ_x^2 , statistické veličiny popisující míru zkreslení. Pro hodnoty vzorků x_i z celkového počtu N ji vypočítáme takto

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)^2, \quad (3.1)$$

kde μ_x je střední hodnota vzorků, její definice je shodná s rovnicí (1.4). Vzorky v našem případě představují koeficienty libovolného subpásma. Jelikož množiny koeficientů se pro jednotlivá subpásma dosti liší, bylo by třeba počítat hodnotu rozptylu pro každé subpásma zvlášť a na závěr všechny hodnoty sečíst. Testováním v [3] však bylo zjištěno, že podstatný vliv na výsledný součet rozptylů má pouze hodnota rozptylu v nejvyšším subpásmu, značena jako σ_{LL}^2 . Hodnoty rozptylů ostatních subpásmech lze v tomto výpočtu opomenout.

Celkový počet dostupných bitů je poté rozdělen mezi složky podle hodnoty jejich rozptylu v nejvyšším subpásmu. Pro P složek je poměr přidělených bitů δ_p složce p určen následujícím způsobem [3]

$$\delta_p = \frac{\sigma_p^2}{\sum_{i=1}^P \sigma_i^2}, \quad 0 < \delta_p < 1. \quad (3.2)$$

Ze vzorce je zřejmé, že vyšší počet bitů je přidán složkám s většími rozdíly mezi jejich koeficienty. Tento způsob rozdělení však nebere v potaz rozdílné vlastnosti jednotlivých složek v modelu YCbCr, které vyplývají z vlastností HVS. Lidské oko je totiž citlivé na jasovou složku více než na složky barvonosné. Proto byla v rámci [3]

zvolena pro testování konstanta $1/2$, kterou se násobí hodnoty rozptylů barvonosných složek, což odpovídá podvzorkování složek v poměru 4:2:2, jenž se běžně používá při zpracování obrazu. Tuto konstantu je však možno nastavit libovolně, a to i pro každý kanál zvlášť.

Metoda umožňuje poměrně jednoduchým způsobem rozdělit dostupný počet bitů mezi složky původního barevného obrazu s ohledem na výslednou hodnotu MSE komprimovaného obrazu. Jelikož je kódování prováděno nezávisle, můžeme oddělené složky zpracovávat i paralelně. Výhodou této metody je nízká výpočetní náročnost při dosažení kvalitních výsledků. U některých testovaných obrázků však metoda selhávala, a proto nelze tvrdit, že toto řešení je všeobecně použitelné [3].

Pro naše účely by bylo výhodné nalézt odpovídající řešení pro metriku MSSIM, tedy provést takové rozdělení bitů mezi jednotlivé složky, které by se co nejpříznivěji projevilo na výsledné vnímané kvalitě komprimovaného obrazu. Jednou z variant je využít pro každou jednotlivou složku veškerý počet přidělených bitů („bit budget“). Následným dekódováním koeficientů složky a zpětným převodem do časové oblasti můžeme zjistit, jaký vliv měla komprese na hodnotu MSSIM u dané složky. Pro každou složku tedy vypočteme index MSSIM zvlášť a na základě těchto hodnot provedeme rozdělení původního počtu přidělených bitů. Poměr bitů přidělených složce p lze stanovit obdobně jako ve vztahu (2.2) takto

$$\delta_p = \frac{1 - \text{MSSIM}_p}{\sum_{i=1}^P (1 - \text{MSSIM}_i)}, \quad 0 < \delta_p < 1. \quad (3.3)$$

Složkám s původně nižší hodnotou MSSIM je přiděleno více bitů. Při rozdělování je opět třeba vzít v úvahu rozdílnou citlivost oka na jednotlivé složky (viz dříve). Nevýhodou tohoto postupu je nutnost provést kódování každé složky nejprve s vyšším počtem bitů a teprve poté je složka znovu zakódována do své finální velikosti podle hodnoty MSSIM, což představuje časově náročnou operaci.

Vhodnějším řešením je nalezení způsobu, jenž by umožňoval rozdělit přidělený počet bitů bez předchozího kódování složek. Problémem je určení kritéria, podle kterého se toto rozdělení provede. U metriky MSE, jakožto statistické veličiny, vychází rozdělení bitů z rozdílu hodnot mezi koeficienty, jež mají přímý vliv na výslednou hodnotu MSE. Obdobným způsobem by bylo možné postupovat i pro metriku MSSIM, kde k stanovení kritéria rozdělení bitů je třeba odhalit ty vlastnosti koeficientů, které se nějakým způsobem podílí na hodnotě MSSIM rekonstruovaného obrazu.

3.4 Shrnutí

Všechny výše uvedené metody představují různé možnosti využití metriky SSIM při kompresi obrazu pomocí algoritmu SPIHT. Metoda založená na postupném zvyšování kvality části obrazu s nejnižší hodnotou indexu SSIM je díky nutnosti provedení

velkého počtu iterací spojených s opakovaným kódováním a následným dekódováním koeficientů v praxi téměř nepoužitelná. Další variantou využití kritéria SSIM je optimalizace algoritmu SPIHT při kódování barevných obrazů. Cílem je rozdělit dostupný počet bitů mezi složky barevného obrazu takovým způsobem, aby rekonstruovaný obraz vykazoval co nejlepší kvalitu. V tomto případě hraje důležitou roli při přidělování bitů rozdílná citlivost oka na jasovou a barvonosné složky.

Z hlediska možností paralelního zpracování je nejperspektivnější metodou kódování vlnkových koeficientů v nezávislých blocích. Problémem je zde rozdělení bitů jednotlivým blokům, které jsou tvořeny koeficienty příslušející vždy určité části ve zpracovávaném obrazu. Bloky jsou kódovány nezávisle, což s sebou přináší výhody, ale i některé nevýhody. Výhodou je větší robustnost vůči chybám nebo ztrátě dat ve výstupním bitovém toku. Případná chyba má vliv pouze na koeficienty bloku, kde se daná chyba vyskytla (zanedbáme-li vliv přesahu mezi jednotlivými bloky způsobený délkou impulzních charakteristik filtrů při vlnkové transformaci obrazu). Nevýhodou blokového rozdělení je klesající efektivita komprese algoritmu SPIHT se snižující se velikostí zpracovávaných bloků. Je zřejmé, že důležitým parametrem při kódování obrázku touto metodou bude zvolená velikost zpracovávaných bloků (viz kap. 4.2.1).

Obdobným způsobem je možné provést rozdělení obrazu na bloky v časové oblasti (bloky jsou tvořeny z hodnot pixelů určité části obrazu). Oproti předchozí metodě má tento způsob členění na bloky značnou nevýhodu. Vlivem nezávislého zpracování pixelů v samostatných blocích pomocí vlnkové transformace se ve výsledném dekódovaném obrázku projevují nespojitosti mezi jednotlivými bloky (tzv. „blokové artefakty“ známé především z obrázků komprimovaných algoritmem JPEG).

Z výše uvedených důvodů byla pro účely praktické realizace vybrána metoda sdružující do jednoho bloku ty vlnkové koeficienty, které v dekompozičním obrazci vytváří společný prostorový strom. Metoda bude představovat ztrátové kompresní schéma, u něž se pokusíme nalézt vhodné rozložení bitů mezi jednotlivé bloky s využitím metriky SSIM. Popisem implementace této metody založené na algoritmu SPIHT se zabývá následující kapitola.

4 IMPLEMENTACE KOMPRESNÍHO SCHÉMATU

Obsahem této kapitoly je popis kompresního schématu použitého pro kódování vlnkových koeficientů sdružených do nezávislých bloků s využitím algoritmu SPIHT. Pro vlastní implementaci bylo využito prostředí programu MATLAB, který představuje ideální nástroj pro práci s obrazovými daty. Při realizaci jsme vycházeli z toho, že navržené kompresní schéma bude mít pouze demonstrační charakter. Z tohoto důvodu je schéma reprezentováno pouze jedinou funkcí, v rámci níž jsou volány pomocné funkce nutné pro dílčí výpočty (funkce realizující kódování a dekodování koeficientů, vlnkovou transformaci, výpočty metrik atd.).

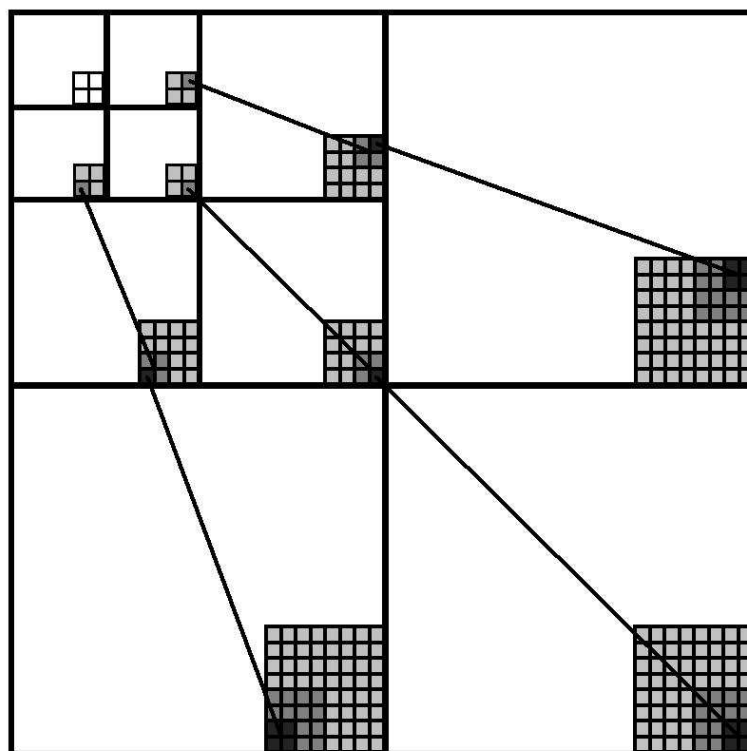
Vstupem kodéru jsou obrazová data, která chceme komprimovat (bitmapový obrázek ve stupních šedi) a zadaný počet bitů (kompresní poměr). Výstupem pak výsledná podoba obou obrázku (původního i komprimovaného), informace o kompresním poměru a hodnoty metrik PSNR a MSSIM. Lze tak snadno porovnat subjektivní kvalitu obrázku s objektivním hodnocením pomocí příslušných metrik.

4.1 Vlastnosti kodéru

Vstupní obrazová data jsou reprezentovaná obrázky ve stupních šedi (256 stupňů šedi) s rozměry, které jsou celočíselnými násobky mocniny dvou (snadnější dekompozice u 2D-DTWT). Pro účely testování byla nakonec vybrána velikost 512×512 pixelů, která představuje kompromis mezi dostatečným množstvím pozorovatelných detailů v obrázku a výpočetním časem komprese.

Výpočet vlnkové transformace je realizován pomocí algoritmu lifting schématu s použitím biortogonální vlnky CDF9/7 [5]. Hloubka dekompozice byla zvolena jako $n = 5$. Pro větší počet úrovní dekompozice již nedochází k výraznějšímu zvýšení efektivity komprese, pouze se zvyšuje výpočetní náročnost [5].

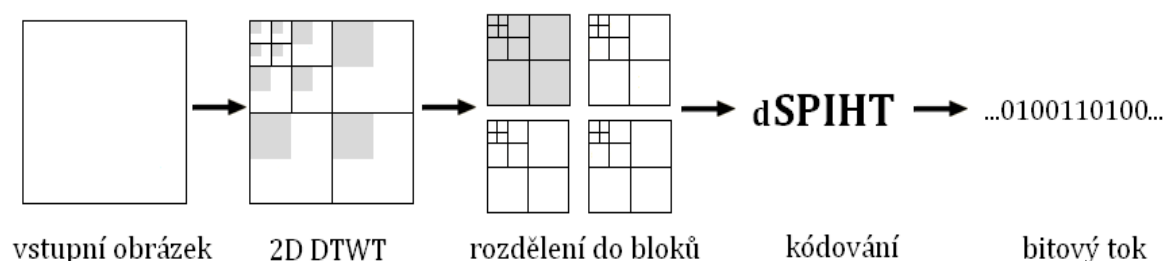
Pro kódování je použita zjednodušená varianta algoritmu SPIHT, kde ze struktury prostorově orientovaných stromů jsou vyjmuty koeficienty nejvyššího subpásma LL_n . Algoritmus má název dSPIHT („degraded tree“ SPIHT). Oproti „klasickému“ SPIHT se algoritmus dSPIHT odlišuje počátečním nastavením dynamických seznamů. Seznam LIP musí zahrnovat všechny koeficienty v nejvyšších subpásmech (včetně detailních koeficientů) a stejná množina bez koeficientů v nejvyšším subpásmu LL_n je přiřazena seznamu LIS. Toto zjednodušení má za následek snížení výpočetní náročnosti algoritmu SPIHT. Negativním důsledkem je nižší efektivita kódování, která je však potlačena pro vyšší úrovně dekompozice, kde se v nejvyšších subpásmech nachází pouze malá část z celkového počtu koeficientů. Na **obr. 4.1** je znázorněna stromová struktura algoritmu dSPIHT [3].



Obr. 4.1: Stromová struktura algoritmu dSPIHT

4.2 Kompresní schéma

Obr. 4.2 zachycuje proces komprese obrazových dat tak, jak je postupně prováděn příslušnými funkcemi v prostředí MATLAB. Obrázek ve stupních šedi je zpracován vlnkovou transformací s hloubkou dekompozice $n = 5$. Koeficienty v dekompozičním obrazci jsou dále rozděleny do samostatných bloků (podrobněji následující kapitola) a kódovány nezávisle na ostatních pomocí algoritmu dSPIHT. Výsledkem komprese je bitový tok reprezentující zakódované koeficienty daného bloku. Zpětná dekomprese bitového toku představuje inverzní proces, jehož výstupem je obrázek zkreslený kompresním algoritmem (v závislosti na zadaném kompresním poměru).



Obr. 4.2: Postup komprese obrazu

4.2.1 Rozdělení koeficientů do bloků

Základem rozdělení vlnkových koeficientů je zachování původní stromové struktury i v nově vzniklých samostatných blocích. Prostorové bloky jsou tvořeny množinou koeficientů z jednoho nebo více hierarchických stromů. Počet stromů je volen tak, aby vyhovoval požadované velikosti bloku. Stromy jsou v blocích přeskupeny takovým způsobem, aby byly ponechány původní vztahy mezi koeficienty daného stromu. Jednotlivé bloky jsou poté tvořeny koeficienty, které jsou nezbytné pro rekonstrukci určité souvislé části v původním obraze.

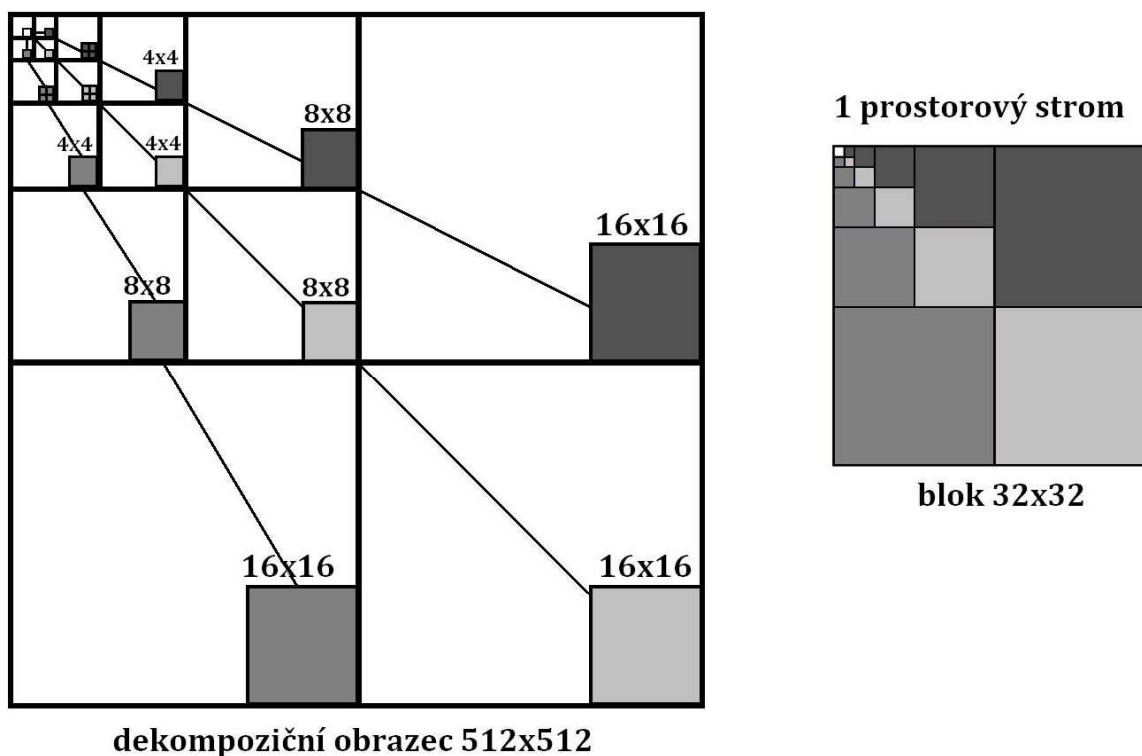
Důležitým aspektem ovlivňující kódování je volba velikosti zpracovávaných bloků. Větší velikost bloku znamená vyšší citlivost kódování na vzniklé chyby (případná chyba způsobí ztrátu informace o větší části původního obrazu). Jelikož velikost bloku v transformované oblasti odpovídá stejně velký blok v časové oblasti, je příliš velký blok rovněž nevhodný z hlediska popisu jeho vlastností pomocí metriky SSIM, která odráží spíše lokální charakter obrazových dat. Naopak pro menší velikosti bloků zase klesá efektivita komprese algoritmu SPIHT. Důvodem je nezávislé kódování koeficientů, které je určeno počátečním prahem stanoveným pro příslušný blok. Větší počet bloků má rovněž za následek zvýšenou režii spojenou s přenosem těchto bloků. Volba bloku tedy musí být vždy určitým kompromisem a je volena v závislosti na velikosti zpracovávaného obrázku.

V naší implementaci byla pro obrázky o rozměrech 512×512 pixelů vybrána velikost bloku 64×64 koeficientů, což představuje celkem 64 bloků tvořených příslušnými koeficienty dekompozičního obrazce. Pro menší velikosti bloků je snížena efektivita kódování, větší rozměry zase znamenají horší možnost popisu lokálních změn uvnitř bloku pomocí metriky SSIM.

Princip vytváření bloků z dekompozičního obrazce je naznačen na **obr. 4.3**. Pro zvolenou hloubku dekompozice $n = 5$ a velikost obrázku 512×512 pixelů se po vlnkové transformaci nachází v nejvyšším supásmu LL_5 celkem 256 koeficientů. Tyto prvky představují kořeny stromů tvořených prostorově závislými koeficienty, které se vztahují ke stejné oblasti v původním obraze. Stromy tak tvoří základní stavební jednotku bloku. Každý kořen stromu má tři přímé potomky v podobě detailních koeficientů na stejné úrovni dekompozice, stromová struktura se dále rozrůstá tak, že každý koeficient má vždy čtyři přímé potomky nacházející se v příslušném subpásmu o úroveň níže. Celkový počet koeficientů, které vytváří jeden strom, je tedy:

$$1 + 3 + 3 \cdot (2 \times 2) + 3 \cdot (4 \times 4) + 3 \cdot (8 \times 8) + 3 \cdot (16 \times 16) = 1024 \text{ koeficientů.}$$

To znamená, že minimální rozměr čtvercového bloku, který pojme všechny tyto koeficienty musí být 32×32. V případě námi zvolené velikosti 64×64 je blok složen ze čtyř sousedních stromů dekompozičního obrazce. Kořeny těchto stromů vytváří čtverec o rozměrech 2×2 a jejich vzájemná poloha musí být v bloku zachována.



Obr. 4.3: Způsob tvorby bloků ze stromové struktury dekompozičního obrazce

4.3 Výpočet metriky MSSIM

Obrázek, který prošel celým procesem komprese a následné zpětné dekomprese se na výstupu objevuje s určitou mírou zkreslení, která je způsobena zaokrouhlováním při výpočtech a především zvoleným kompresním poměrem. Aby toto zkreslení mohlo být při daném kompresním poměru co nejlépe redukováno, je potřeba nejprve výslednou kvalitu obrázku (míru zkreslení) ohodnotit pomocí některého objektivního kritéria. V předešlých kapitolách jsme ověřili, že vhodným kritériem pro tyto účely je metrika SSIM.

4.3.1 Výpočet MSSIM celého obrázu

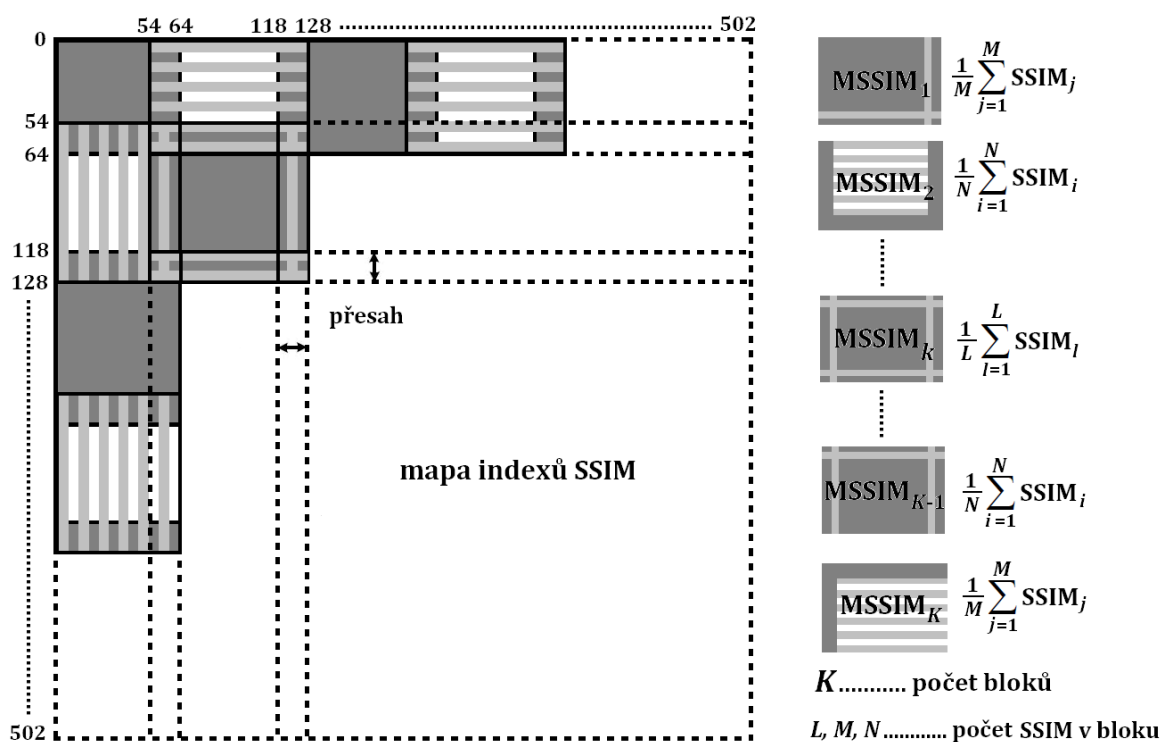
Výpočet hodnoty SSIM mezi původním obrázkem a obrázkem zkresleným kompresí probíhá podle postupu, jenž byl popsán v kap. 1.3.2 a kap. 1.3.3. Jsou srovnávány statistické vlastnosti odpovídajících si částí obou obrázků. Tyto části jsou vymezeny délkou okna 11×11 pixelů, které se v obraze posouvá s krokem jeden pixel. Po skončení výpočtu je k dispozici mapa SSIM indexů popisující kvalitu v různých částech komprimovaného obrázku. Jelikož při kompresi nedochází k inverzi hodnot pixelů, pohybují se hodnoty indexů SSIM prakticky pouze v rozmezí 0 až 1. Index MSSIM udávající „průměrnou“ kvalitu v celém obrázku se stanoví na základě rovnice (1.17).

4.3.2 Výpočet MSSIM bloku

Jelikož jsou při kompresi obrázku kódovány samostatné bloky koeficientů, bude nás zajímat hodnota MSSIM bloků, které těmto koeficientům odpovídají v časové oblasti, spíše než hodnota MSSIM celého obrazu. Metrika MSSIM nám umožní rozlišit rozdílnou míru zkreslení příslušných částí původního obrazu a zjistit tak, které bloky koeficientů vyžadují při kódování více bitů.

Prostorovému bloku o rozměrech 64×64 koeficientů odpovídá stejně velký blok pixelů v původním obrázku. Hodnoty koeficientů v bloku mají přímý vliv na hodnoty těchto pixelů. Délka impulsních charakteristik filtrů (vlnka CDF9/7) použitých pro vlnkovou transformaci pak určuje, do jaké míry budou tyto koeficienty ovlivňovat rovněž pixely v sousedních blocích. Z tohoto důvodu je potřeba při výpočtu indexu MSSIM uvažovat nejenom pixely odpovídajícího bloku obrázku, ale také část pixelů nacházejících se v sousedních blocích v blízkosti přechodu mezi těmito bloky.

Způsob určení hodnoty MSSIM jednotlivých bloků je znázorněn na **obr. 4.4**. Pro výpočet je využita již dříve vypočítaná mapa indexů SSIM. Ta má pro testované obrázky o rozměrech 512×512 pixelů a použité výpočetní okno 11×11 velikost 502×502 hodnot SSIM. V mapě jsou určeny oblasti, na jejichž hodnotách indexů SSIM se příslušný blok koeficientů podílí. Respektuje se při tom dříve zmíněný vliv přesahu do sousedních bloků. Rozměry výpočetního okna umožňují počítat s přesahem deset pixelů (to odpovídá přesahu deseti indexů SSIM). Z hodnot indexů SSIM náležejících této oblasti je poté obdobně jako ve vztahu (1.17) stanovena celková hodnota MSSIM charakterizující daný blok koeficientů.



Obr. 4.4: Princip výpočtu indexu MSSIM pro jednotlivé bloky

5 NÁVRH A TESTOVÁNÍ METOD PRO ROZDĚLENÍ BITŮ

V předchozí kapitole bylo popsáno kompresní schéma, pomocí něhož je prováděna ztrátová komprese zpracovávaných obrázků. Ztráta kvality v rekonstruovaném obrázku je způsobena omezeným počtem bitů při kódování vlnkových koeficientů algoritmem SPIHT. Míra tohoto zkreslení v různých částech obrázku je vyjádřena metrikou SSIM. Jelikož jsou koeficienty kódovány po blocích, nastává v okamžiku kódování problém, kolik příslušnému bloku přidělit bitů z celkového množství dostupných bitů, jejichž počet je určen zvoleným kompresním poměrem.

Otázce rozdělení bitů mezi jednotlivé bloky koeficientů je věnována následující kapitola. Cílem je dosáhnout toho, aby při daném kompresním poměru byla výsledná subjektivní kvalita komprimovaného obrázku co nejlepší. Převáděno do objektivního hodnocení kvality, aby hodnota metriky MSSIM mezi původním a rekonstruovaným obrázkem se co nejvíce blížila jedné (obrázek bez zkreslení). K určení optimálního rozdělení je proto potřeba využít kritéria SSIM jakožto indikátoru míry zkreslení v jednotlivých blocích.

5.1 Problém rozdělení bitů

Při hledání optimálního rozdělení bitů pro daný kompresní poměr budeme vycházet z parametrů kodéru a vstupních dat tak, jak byly definovány v kap. 4.1. Vstupní obrazová data jsou reprezentována obrázky ve stupních šedi ($8b/px = 256$ stupňů šedi) s rozměry 512×512 pixelů. Po provedení vlnkové transformace a rozčlenění vlnkových koeficientů získáme 64 bloků o velikosti 64×64 koeficientů. Každý z těchto bloků ovlivňuje hodnoty pixelů v určité části obrázku a je kódován nezávisle na ostatních blocích. To, do jaké míry bude příslušná oblast v obrázku zkreslena, určuje množství přidělených bitů odpovídajícímu bloku. Pro daný kompresní poměr (celkový počet dostupných bitů) tak volba poměru, v jakém budou rozděleny bity jednotlivým blokům, přímo ovlivňuje výslednou kvalitu rekonstruovaného obrázku. Vhodnou volbou lze dosáhnout při stejném kompresním poměru vylepšení kvality zkresleného obrázku.

5.1.1 Využití kritéria MSSIM

Abychom mohli určit optimální rozdělení, pro které bude mít zpracovávaný obrázek při daném kompresním poměru nejvyšší možnou kvalitu z hlediska metriky MSSIM, je potřeba popsat a rozlišit různý vliv kódování na různé části původního obrázku. Odlišné bloky koeficientů kódované stejným počtem bitů vykazují v komprimovaném obrázku různý stupeň zkreslení. Optimálním řešením by bylo nalézt způsob, který by

požadované rozdělení umožnil určit ještě před samotným kódováním bloků. Problém je v tom, že bez předchozího kódování bloku nejsme schopni určit, jaký vliv má komprese na zkreslení odpovídající části obrázku (tedy i na hodnoty SSIM). Z rovnice (1.13) je zřejmé, že index SSIM je na svých proměnných závislý nelineárně. Proto je obtížné nalézt přímou souvislost mezi zpracovávanými obrazovými daty a výslednou hodnotou MSSIM. Určitou tendenci lze vyzorovat ze statistických vlastností dané oblasti (např. nízký rozptyl hodnot pixelů značí oblast s malým počtem detailů, a tedy vliv komprese na zkreslení v této oblasti nebude tak výrazný), ale jedná se vždy pouze o hrubý odhad.

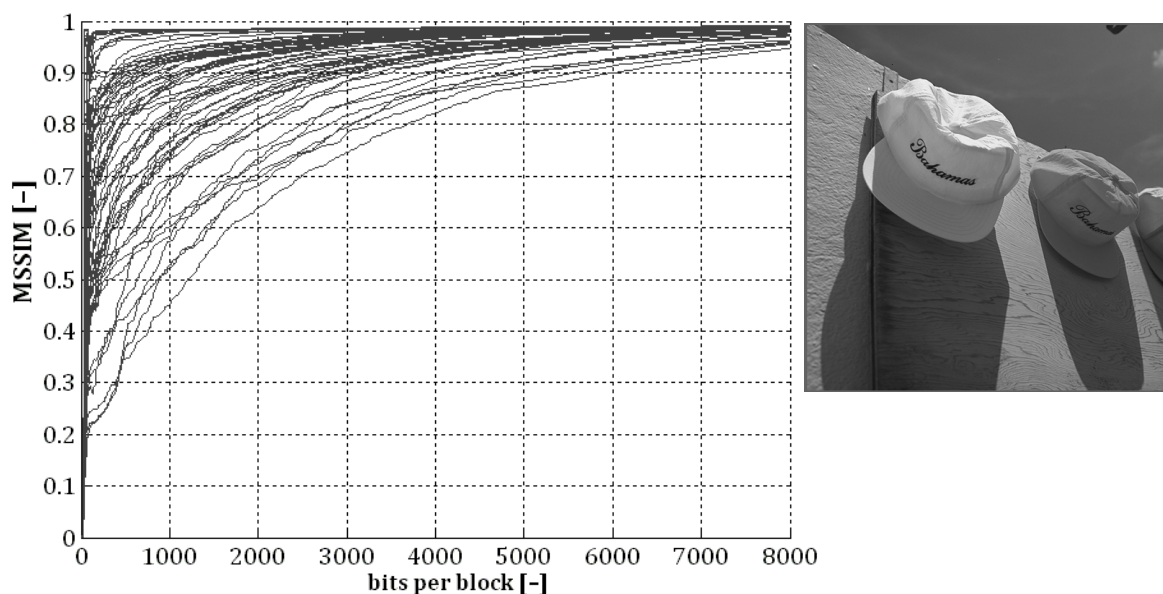
Provedeme tedy nejprve „testovací“ kompresi vstupních obrazových dat tím způsobem, že zakódujeme každý blok koeficientů stejným počtem bitů. Počet bitů je určen jednoduše jako poměrná část z celkového počtu dostupných bitů (pro K bloků to je $1/K$ z celkového množství bitů). Na výstupu kodéru dostaneme komprimovaný obrázek s různou degradací kvality jeho příslušných částí. Pro určení míry zkreslení využijeme hodnoty metrik MSSIM stanovených pro samostatné bloky. Způsob jejich výpočtu byl popsán v kap. 4.3.2. Získáme tak 64 indexů MSSIM, které udávají „průměrnou“ kvalitu dané části obrazu a jejichž hodnoty se liší v závislosti na rozdílném zkreslení příslušných částí. Na základě těchto hodnot můžeme určit bloky, které budou při rozdělování bitů preferovány. Bloky s nízkým indexem MSSIM indikují větší zkreslení způsobené kompresí. Abychom vliv tohoto zkreslení zmírnili, je potřeba těmto blokům přidělit více bitů. Naopak blokům s vysokou hodnotou MSSIM můžeme část bitů odebrat, aniž by se to nějak výrazně projevilo na jejich výsledné kvalitě.

Co však z jediné hodnoty MSSIM pro daný blok nelze určit, je množství bitů, které je bloku potřeba přidat nebo odebrat. Index SSIM je na rozdíl od metriky MSE, která udává pouze množství chybového signálu, silně závislý na charakteru vstupních obrazových dat a to tak, že nelineárně. Z tohoto důvodu je velmi obtížné určit analyticky nějaký vztah nebo rovnici, která by popisovala závislost indexu MSSIM (SSIM) na přiděleném počtu bitů pro daný blok. Tuto závislost se tedy pokusíme stanovit empiricky.

5.1.2 Závislost hodnoty MSSIM na počtu přidělených bitů

K určení závislosti MSSIM bloku na počtu přidělených bitů je nezbytné stanovit průběh křivky, která bude odrážet vztah mezi těmito veličinami. Algoritmus SPIHT umožňuje postupně vylepšovat kvalitu rekonstruované části obrázku s přesností na jediný bit. Takto nepatrná změna při kódování bloku se však na výsledné hodnotě MSSIM příslušné části obrazu prakticky neprojeví. Proto pro změření průběhu křivek jednotlivých bloků v dostatečném počtu bodů byl zvolen krok deset bitů. S každým krokem je každému bloku přidáno deset bitů a je provedena komprese a následná dekomprese. V rekonstruovaném obrázku se poté zjistí, jak se tato změna odrazila v hodnotě MSSIM odpovídající části obrazu.

Na **obr. 5.1** je zobrazen typický průběh křivek náležející jednotlivým blokům v obraze. Pro účely měření byl vybrán obrázek obsahující jak detailní části, tak části stejnorodé s nízkým výskytem detailů (*kodak03.bmp*). Závislost hodnot MSSIM pro 64 bloků byla změřena v rozsahu 0 až 8000 bitů, což pro celý obrázek (rozměry 512×512, hloubka 8 bitů/px) odpovídá postupnému snižování kompresního poměru od limitní hodnoty nekonečna až po poměr přibližně 1:4. Hodnoty mezi vypočítanými body křivky jsou doplněny lineární interpolací. Graf na **obr. 5.1** má pouze ilustrativní charakter, některé vybrané průběhy křivek budou rozebrány dále.

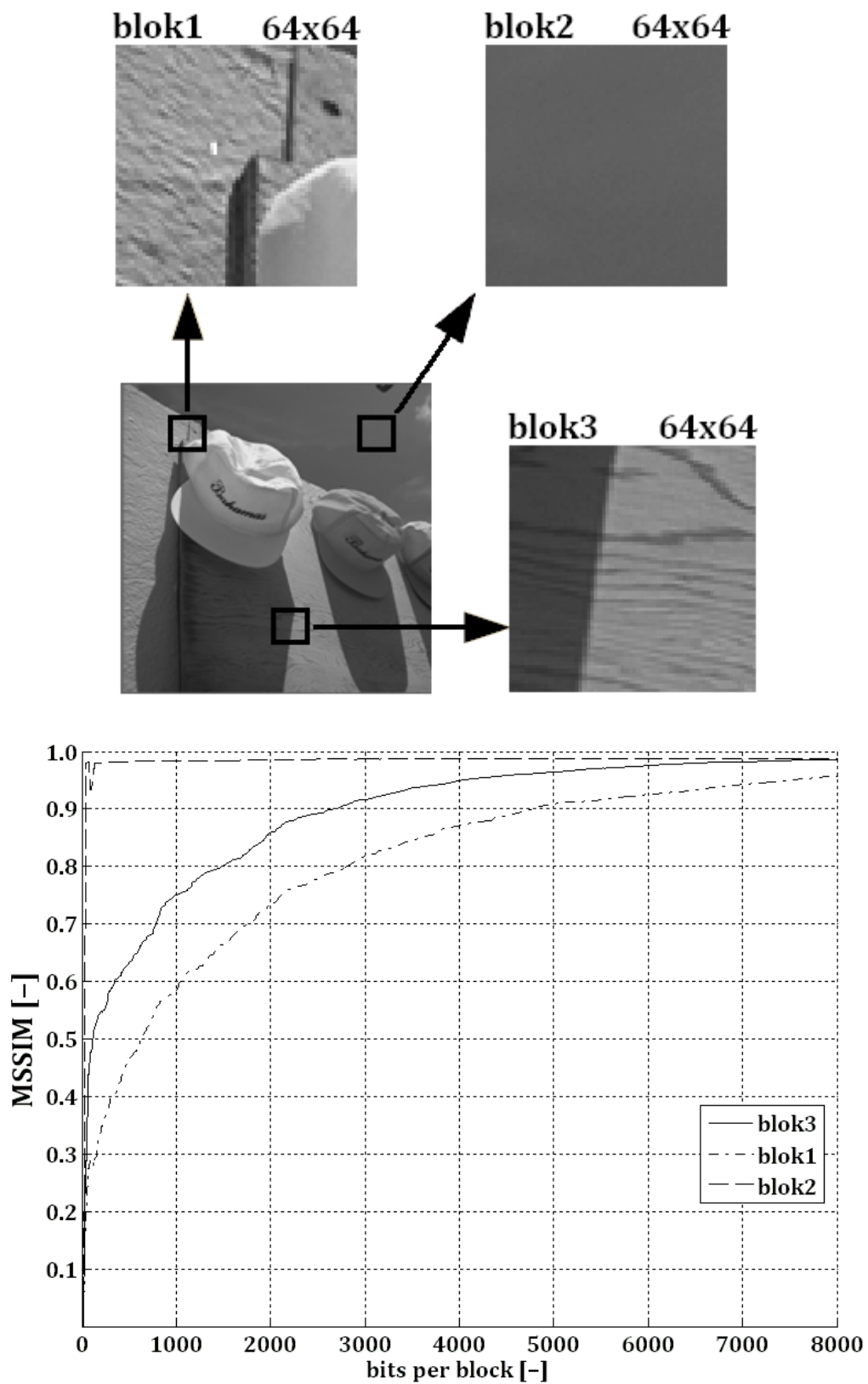


Obr. 5.1: Průběh závislosti hodnot MSSIM bloků obrázku na počtu přidělených bitů (64 bloků, rozměry obrázku 512×512 pixelů, graf má pouze ilustrativní charakter)

Podobné průběhy křivek bychom získali i pro jakýkoliv jiný obrázek stejných rozměrů s hodnotami pixelů ve stupních šedi. Vždy lze vypočítat určitou souvislost mezi průběhy křivek a hodnotami pixelů bloku, který popisují. Jak již bylo zmíněno dříve, hodnota indexu MSSIM výrazně závisí na charakteru obrazových dat. Jednotlivé křivky se proto od sebe liší podle toho, jaké množství detailů obsahuje blok, jemuž náleží. Bloky s vysokým výskytem detailů (s členitou strukturou) jsou náchylnější na zkreslení způsobená kompresí, což se odráží i v jejich průběhu hodnoty MSSIM.

Vezměme si konkrétní případ. Na **obr. 5.2** jsou znázorněny některé charakteristické průběhy křivek z **obr. 5.1** spolu s odpovídajícími oblastmi v obrázku. Je patrné, že blok, který je tvořen pouze stejnorodou plochou (*blok2*), má nejvyšší hodnotu MSSIM v celém měřeném průběhu. Tato hodnota se navíc od jistého minimálního počtu přidělených bitů prakticky nemění (konstantní průběh křivky). To znamená, že výsledná kvalita tohoto bloku je od určité hodnoty MSSIM téměř nezávislá na dalších přidělených bitech. Bloky s podobným průběhem křivky jsou vhodným „zdrojem“ bitů při rozdělování „bit budget“, jelikož pro jejich zakódování stačí pouze jistý minimální počet. Jiné je to u bloků, které obsahují nějaké detaily. Příkladem je *blok1*

na **obr. 5.2**. Při kompresi dochází k deformaci jeho členité struktury, a proto i index MSSIM nabývá v celém rozsahu měřených hodnot nejnižší velikosti. Křivka má logaritmický průběh, kdy hodnota MSSIM zpočátku rychle stoupá, dokud nedojde při

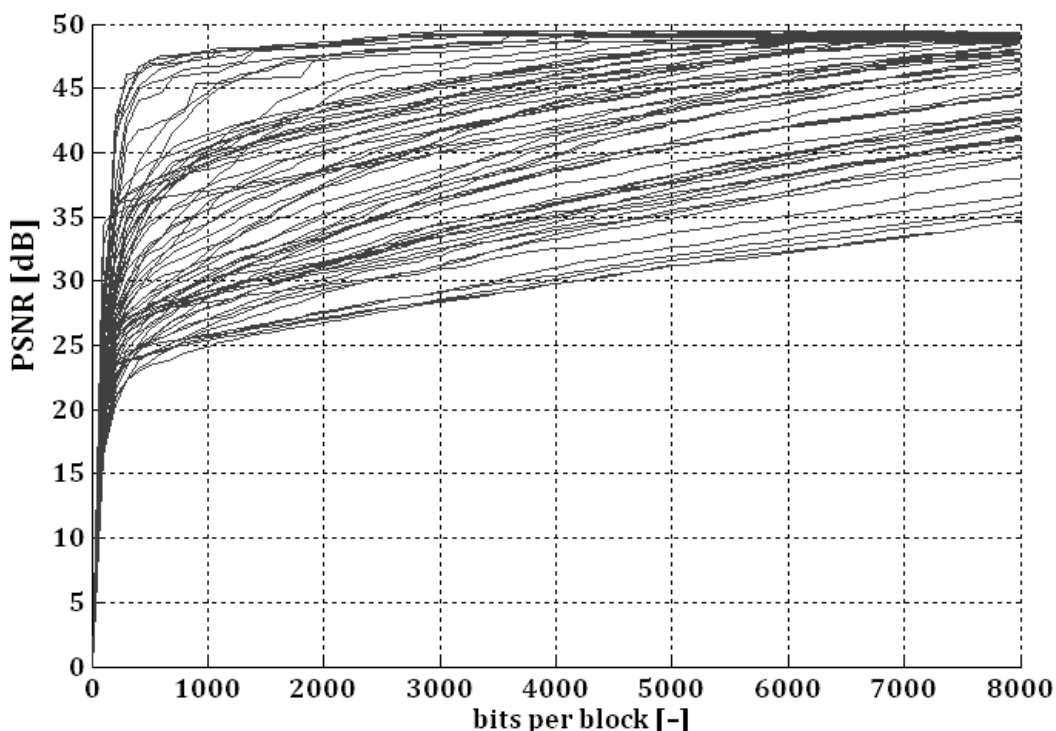


Obr. 5.2: Závislost MSSIM na počtu přidělených bitů u vybraných charakteristických bloků

určité hodnotě k jejímu „nasyčení“ a poté se s přibývajícím počtem bitů přibližuje k jedničce pozvolně. Tento typ bloků je naopak nutno při rozdělování bitů zvýhodnit nejvíce. Poslední křivka reprezentuje blok s nižším výskytem detailů (*blok3*). Její průběh je podobný průběhu křivky u *bloku2*, index MSSIM ale nabývá vyšších hodnot a k jeho ustálení dochází dříve.

Empiricky stanovený průběh závislosti MSSIM na počtu přidělených bitů potvrzuje, že nelze nalézt nějakou obecnou matematickou funkci, která by tento vztah dokázala vyjádřit. Funkce jsou odlišné pro každý jednotlivý blok v závislosti na obrazových datech v daném bloku. Další nevýhodou je nelineární průběh těchto funkcí. Požadovaná lineární závislost by nám umožnila zjednodušit problém stanovení optimálního rozdělení bitů.

Pro srovnání jsou na **obr. 5.3** uvedeny průběhy hodnot metriky PSNR v závislosti na rostoucím počtu bitů přidělených jednotlivým blokům. Testován byl stejný obrázek jako v případě indexu MSSIM (**obr. 5.1** vpravo). Metrika PSNR byla vypočítána z hodnot pixelů daného bloku podle vztahu (1.1) resp. (1.3). Z grafu je vidět, že od určitého počtu bitů roste hodnota PSNR lineárně. To přesně odpovídá interpretaci metriky MSE jakožto energie chybového signálu. Křivky jednotlivých bloků mají velmi podobné průběhy, které poskytují údaje o množství chyby obsažené v daném bloku, nelze z nich však vyvodit žádnou informaci o vlivu tohoto zkreslení na původní obrazová data.



Obr. 5.3: Průběh závislosti hodnot PSNR bloků obrázku na počtu přidělených bitů (64 bloků, rozměry obrázku 512×512 pixelů, graf má pouze ilustrativní charakter)

5.2 Návrh metod

V kapitole 5.1.2 jsme empiricky odvodili typické průběhy závislostí indexu MSSIM na počtu přidělených bitů. Z vlastností těchto křivek budeme dále vycházet i při hledání optimálního rozdělení bitů jednotlivým blokům. Prakticky je možno uvažovat dva různé přístupy pro vylepšení výsledné kvality komprimovaného obrázku. Záleží na tom, co konkrétně chceme zlepšit.

První přístup spočívá ve snižování zkreslení částí obrázku s nejnižšími hodnotami indexu SSIM. Myšlenka vychází z vlastností HVS, kdy oblast s velmi nízkou vizuální kvalitou na sebe upíná pozornost pozorovatele na úkor zbytku obrázku a negativně tak ovlivňuje celkové hodnocení kvality obrázku. Na tomto principu pracuje i algoritmus popsáný v kap. 3.1. V našem případě se jedná o vylepšení bloků, které mají nejnižší hodnoty MSSIM. Výsledkem je zvýšení kvality nejvíce zkreslených částí obrazu na úkor ostatních bloků, tento postup však obecně nevede k dosažení nejvyšší možné kvality z hlediska hodnoty MSSIM celého obrázku.

Tímto kritériem se řídí druhá varianta, která je založena právě na hledání maximální hodnoty MSSIM celého obrázku při daném kompresním poměru. Nalezení maxima nebo minima (obecně extrémů) funkcí podle určitého kritéria je cílem řady optimalizačních úloh. Při hledání optimálního rozdělení bitů mezi jednotlivé bloky podle metriky MSE je možné k nalezení řešení využít Lagrangeovu metodu [6]. Máme-li k dispozici jednotlivé průběhy závislostí MSE na počtu přidělených bitů (tzv. R-D charakteristiky, z angl. Rate-Distortion), lze optimální počet bitů pro daný blok u zvoleného kompresního poměru zjistit na základě hodnoty společné minimální směrnice λ (resp. $-\lambda$). V bodě, kde přímka s touto směrnicí (sklonem) protíná křivku závislosti MSE příslušného bloku, můžeme odečíst počet bitů přidělených danému bloku. Hodnota směrnice se iterativním způsobem upravuje tak, aby celkový počet bitů pro všechny bloky nepřekročil počet dostupných bitů a zároveň byl co nejvyšší. Využívá se toho, že křivky závislostí MSE jednotlivých bloků mají podobné průběhy (udávají velikost chybového signálu), tudíž je možné nalézt jejich společnou směrnicí [6]. V případě metriky MSSIM je toto problém, jelikož průběhy závislostí se u některých bloků liší dosti výrazně (viz **obr. 5.2**).

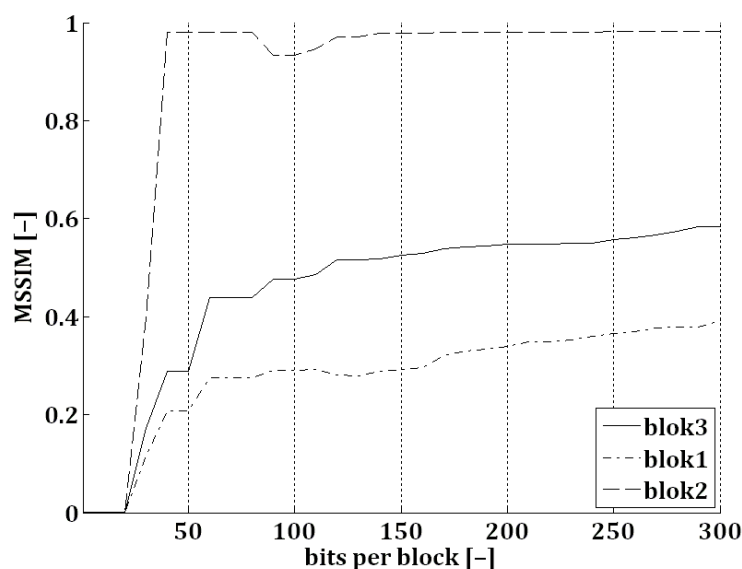
My se při hledání možných způsobu rozdělení bitů pokusíme zohlednit oba výše zmíněné přístupy. Při návrhu metod byly uvažovány obrázky o rozměrech 512×512 a bloky velikosti 64×64 koeficientů, to však žádným způsobem nebrání obecnému použití těchto metod.

5.2.1 Obecné poznatky

Během zkoumání vlastností průběhů závislostí MSSIM na počtu bitů přidělenému danému bloku, byly zjištěny některé poznatky, které je třeba brát v potaz před vlastním hledáním nejvhodnější metody.

Na **obr. 5.4** je zvětšen počátek průběhu křivek z **obr. 5.2**. Lze vyzorovat, že index MSSIM se u všech bloků ustálí na své minimální hodnotě až po určitém počtu přidělených bitů. Do té doby mají křivky spíše „skokovitý“ průběh a jejich hodnoty MSSIM se mění dosti nepředvídatě. Vyplývá to z vlastností kodéru SPIHT, který k zakódování alespoň nějaké informace o obrazových datech potřebuje určité minimální množství bitů. Pod touto minimální hranicí nemá přerozdělování bitů smysl. Rozdělení dostupných bitů tedy budeme provádět až od určitého minima bitů, které na počátku přidělíme všem blokům.

Tímto průběhy křivek prakticky na počátku „uřízneme“ a nastavíme jim novou počáteční hodnotu namísto nuly. Empirickým měřením křivek různých bloků bylo zjištěno, že pro obrázky rozměrů 512×512 pixelů a bloky velikosti 64×64 je tato minimální hodnota přibližně 100 bitů, což potvrzuje i graf na **obr. 5.4**. Tato úprava tak vlastně definuje i minimální velikost „bit budget“ (pro 64 bloků činí 6400 bitů).



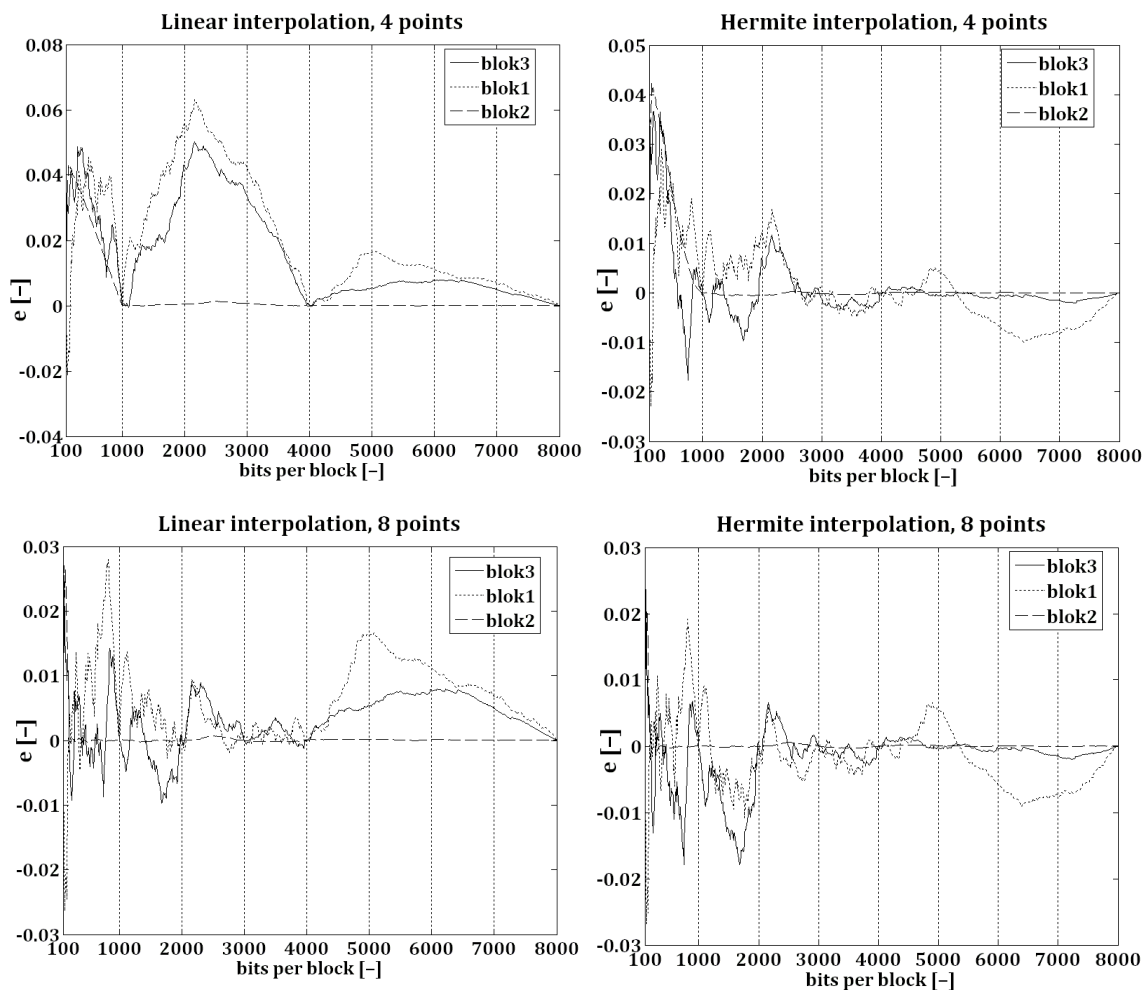
Obr. 5.4: Detail počátku průběhu křivek z obr. 5.2

K lepšímu studiu vlastností křivek je vhodnější popis křivek pomocí nějaké matematické funkce než pouze jejich bodová reprezentace. Z tohoto důvodu jsme se pokusili nahradit průběhy křivek rovnicemi funkcí, které by co nejlépe dokázali původní průběhy charakterizovat. Vycházíme-li z toho, že máme k dispozici přímo jednotlivé hodnoty křivky, je nejspolehlivějším řešením interpolace křivky z několika jejích bodů. Výhodou je, že takto můžeme získat průběhy křivek z omezeného počtu hodnot MSSIM (což také znamená omezený počet nutného opakovaného kódování a dekodování bloků).

K interpolaci křivek jsme využili předdefinované funkce MATLABu `interp1`, která umožňuje interpolaci pomocí různých metod. Funkce vždy interpoluje hodnoty mezi dvěma sousedními body podle zvoleného typu metody (jedná se tedy o interpolaci po částech). Pro srovnání jsme vybrali lineární interpolaci a interpolaci pomocí Hermitova po částech spojitého kubického interpolačního polynomu – dále

jen Hermitova kubická interpolace (v MATLABu funkce `linear` a `pchip`). Lineární interpolace nahrazuje průběh mezi dvěma sousedními body úsečkou (rovnicí přímky). Hermitova kubická interpolace prokládá sousední body kubikou (polynomem třetího stupně), u níž navíc platí, že v uzlech propojení s jinou kubikou má s touto kubikou shodnou nultou a první derivaci [1]. To zaručuje, že výsledná křivka bude hladká, i když je složena z více segmentů. Počet rovnic popisující průběh interpolované křivky je vždy o jeden nižší než celkový počet bodů použitých k interpolaci. V případě lineární interpolace jsou to rovnice přímek, u Hermitovy kubické interpolace rovnice kubik.

Při interpolaci jsme rovněž testovali různý počet použitých bodů původní křivky. Body byly voleny tak, aby co nejlépe pokryly celý měřený interval. Na **obr. 5.5** jsou uvedeny průběhy chybových funkcí e , které udávají absolutní rozdíl v hodnotách MSSIM mezi změřenými průběhy křivek z **obr. 5.2** a jejich interpolovanými verzemi. Jsou zde porovnány křivky, které byly interpolovány ze čtyř (hodnoty MSSIM pro 100, 1000, 4000 a 8000 bitů) a osmi (hodnoty MSSIM pro 100, 200, 400, 1000, 2000, 3000, 4000 a 8000 bitů) bodů původní křivky. K interpolaci byly použity metody popsané v předchozím odstavci.



Obr. 5.5: Průběhy chybových funkcí mezi původními (obr. 5.2) a interpolovanými hodnotami

Je vidět, že pro čtyři vybrané body je ještě interpolace dosti nepřesná, a to zejména v počátečním průběhu křivek detailních bloků. U lineární interpolace lze vyzorovat upínání chybové funkce v těchto bodech. Pro osm bodů je chyba již poměrně nízká, a to i při použití lineární interpolace (max. 3%). Rovnice kubik popisující jednotlivé průběhy křivek u Hermitovy kubické interpolace měly ve všech testovaných případech velmi nízké hodnoty koeficientů u kubického a kvadratického členu. To znamená, že při interpolaci křivek po částech si vystačíme při použití většího počtu vhodně zvolených bodů i s lineární interpolací.

Charakter průběhu závislosti MSSIM na počtu přidělených bitů tak v určitém omezeném intervalu křivky určují především hodnoty nulté (funkční hodnota) a první derivace (směrnice tečny). Hodnoty vyšších derivací jsou v tomto intervalu prakticky zanedbatelné.

5.2.2 Rozdělení bitů podle funkční hodnoty

Nejjednodušší možný způsob rozdělení dostupného počtu bitů spočívá ve využití funkční hodnoty (hodnoty nulté derivace) v jistém bodě křivky závislosti MSSIM na počtu přidělených bitů u daného bloku. Na základě zjištěných hodnot MSSIM při kódování bloků stejným počtem bitů se stanoví nové rozdělení bitů, které zvýhodní bloky s nižším indexem MSSIM. Poměr rozložení bitů mezi bloky je nastaven úměrně k velikostem hodnot MSSIM (čím nižší hodnota MSSIM tím více bitů je danému bloku přiděleno). Postup výpočtu této metody můžeme shrnout do následujících kroků:

1. Pro daný kompresní poměr (počet dostupných bitů – označíme jej BB) provedeme počáteční rozdělení, kdy každému bloku je přidělen stejně velký počet bitů $b = \text{round}(BB/K)$, kde K značí počet bloků a round operaci zaokrouhlení.
2. Bloky zakódujeme a dekodujeme, abychom získali podobu komprimovaných obrazových dat.
3. Pro každý blok vypočítáme hodnotu metriky MSSIM.
4. Nastavíme pro každý blok počáteční hodnotu přidělených bitů na 100 bitů.
5. Provedeme nové rozdělení bitů podle rovnice

$$b_k = \text{round} \left(\frac{1 - \text{MSSIM}_k}{\sum_{i=1}^K (1 - \text{MSSIM}_i)} \cdot (BB - K \cdot 100) \right), \quad (5.1)$$

kde b_k je počet bitů přidělený k -tému bloku s hodnotou indexu MSSIM_k .

6. Každý blok znovu zakódujeme, tentokrát s nově stanoveným počtem bitů.

Metoda vylepšuje bloky, které měly při daném kompresním poměru nejhorší kvalitu na úkor bloků s nejvyšší kvalitou. Počet bitů je přidělován úměrně velikosti zkreslení. Jedná se tedy o přístup, kdy se snažíme potlačit zkreslení v nejvíce zkreslených částech obrazu.

Variantou použití metody je způsob, kdy bloky zakódujeme na vyšší počet bitů, než kolik jich máme pro počáteční rozdělení k dispozici (typicky trojnásobek [10]). Dále provedeme kroky 2 až 5. Po určení nového rozdělení bitů ponecháme blokům pouze část zakódovaných bitů, jejichž počet je dán množstvím nově přidělených bitů, zbylé bity z datového toku odstraníme. Výhodou je, že takto provedeme operaci kódování, která je u algoritmu SPIHT poměrně výpočetně náročná, pouze jednou.

5.2.3 Rozdělení bitů podle velikosti směrnice

Myšlenka této metody vychází z následující úvahy: Jestliže je naší snahou při daném kompresním poměru dosažení maximální hodnoty MSSIM celého obrázku, musíme při rozdělování bitů preferovat takové bloky, které nejvíce přispívají k nárůstu celkové hodnoty MSSIM. Jinými slovy, hledáme takové bloky, u nichž přidělením dalšího počtu bitů dojde k největšímu zvýšení hodnoty MSSIM ze všech bloků, což znamená i největší nárůst MSSIM celého obrázku. Porovnáme-li průběhy závislosti hodnot MSSIM bloků na počtu přidělených bitů z **obr. 5.2**, je zřejmé, že budou-li mít všechny bloky přidělen stejný počet bitů, konkrétně např. hodnotu 500 bitů, tak přidáním dalšího stejně velkého množství bitů dojde k největšímu nárůstu hodnoty MSSIM u *bloku3*, ačkoliv tento blok původně neměl nejnižší hodnotu MSSIM.

Bude nás tedy zajímat tendence rychlosti růstu hodnot MSSIM jednotlivých bloků. Rychlost nárůstu nebo poklesu v určitém okolí bodu na křivce můžeme popsat pomocí velikosti směrnice s (hodnota první derivace křivky v daném bodě). Pro odlišení charakteru růstu křivek závislosti MSSIM na počtu přidělených bitů je třeba zvolit vhodný úsek křivky, kde budeme směrnicí určovat, jelikož při určitém počtu přidělených bitů již dojde k „nasycení“ hodnot MSSIM u většiny bloků a křivky poté mají podobné průběhy. Na **obr. 5.1** si lze všimnout, že toto „nasycení“ nastává přibližně od hodnoty 4000 bitů. Empirickým testováním bylo zjištěno, že vhodným úsekem pro popis vlastností křivek je okolí bodu 1000 bitů, kde jsou průběhy růstu křivek dosti odlišné.

Pro určení přesné velikosti směrnice (sklonu) v tomto bodě by bylo třeba vypočítat hodnotu první derivace křivky. K tomu bychom potřebovali znát přímo rovnici dané křivky, její zjišťování by však bylo zbytečně zdlouhavé a výpočetně náročné. V kap. 5.2.1 jsme testováním zjistili, že v určitém omezeném intervalu křivky je možno její průběh aproximovat pomocí rovnice přímky, aniž bychom se dopustili příliš velké odchylky od skutečného průběhu. Při určování velikosti směrnice (sklonu křivky) se tedy spokojíme s výpočtem hodnoty podle následujícího vztahu:

$$s = \frac{\text{MSSIM}_{1100} - \text{MSSIM}_{900}}{1100 - 900}, \quad (5.2)$$

kde MSSIM_{1100} a MSSIM_{900} jsou hodnoty MSSIM daného bloku pro přidělený počet 1100 resp. 900 bitů.

Jednotlivé kroky výpočtu požadovaného rozdělení bitů jsou obdobné jako v případě metody v kap. 5.2.2:

1. Provedeme počáteční kódování bloků rovnocenným počtem bitů $b = 1100$.
2. Dekódujeme všechny bity bloků a vypočítáme hodnoty indexů $MSSIM_{1100}$ pro jednotlivé bloky.
3. Dekódujeme pouze prvních 900 bitů každého bloku a určíme hodnoty $MSSIM_{900}$.
4. Podle vztahu (5.2) vypočítáme pro každý blok velikost směrnice s .
5. Nastavíme pro každý blok počáteční hodnotu přidělených bitů na 100 bitů.
6. Pro daný kompresní poměr provedeme rozdělení bitů podle této rovnice:

$$b_k = \text{round} \left(\frac{s_k}{\sum_{i=1}^K s_i} \cdot (BB - K \cdot 100) \right), \quad (5.3)$$

kde b_k je počet bitů přidělený k -tému bloku s velikostí směrnice s_k .

7. Bloky zakódujeme podle stanoveného rozdělení bitů.

První čtyři kroky výpočtu je nutné pro daný obrázek provést pouze jednou. Po určení velikosti směrnice pro jednotlivé bloky můžeme stanovit požadované rozdělení bitů pro jakýkoliv kompresní poměr (při respektování nutnosti minimálního „bit budget“). Metoda přiděluje bity jednotlivým blokům úměrně velikosti jejich směrnice. Preferovány jsou bloky s největším nárůstem hodnoty $MSSIM$ v určeném intervalu, což nutně nemusí být bloky s nejnižší hodnotou $MSSIM$. Snahou je vylepšit kvalitu obrázku z hlediska jeho celkové hodnoty $MSSIM$.

5.2.4 Kombinovaná metoda rozdělení bitů

Tento způsob rozdělení bitů využívá metody popsané v kap. 5.2.2 a 5.2.3. Při určování nejvhodnějšího rozložení bitů se tedy přihlíží jak k funkční hodnotě křivky (nulté derivaci), tak k velikosti její směrnice (první derivaci). Cílem je nastavit určitý správný poměr vlivu těchto veličin, aby při rozdělování bitů byly zohledněny oba přístupy (vylepšení nejvíce zkreslených bloků a co nejvyšší hodnota $MSSIM$ celého obrázku). Pro tyto účely byla zavedena váha α , jejíž hodnotou tento poměr určujeme. Výpočet rozdělení vychází z dříve popsaných metod:

1. Provedeme kroky výpočtu 1 až 5 uvedené v kap. 5.2.3.
2. Jako hodnotu nulté derivace ($MSSIM$) zvolíme u každého bloku $MSSIM_{1100}$.
3. Výsledný poměr rozdělení bitů určíme následujícím vztahem:

$$b_k = \text{round} \left(\left(\alpha \cdot \frac{1 - MSSIM_k}{\sum_{i=1}^K (1 - MSSIM_i)} + (1 - \alpha) \frac{s_k}{\sum_{i=1}^K s_i} \right) \cdot (BB - K \cdot 100) \right), \quad (5.4)$$

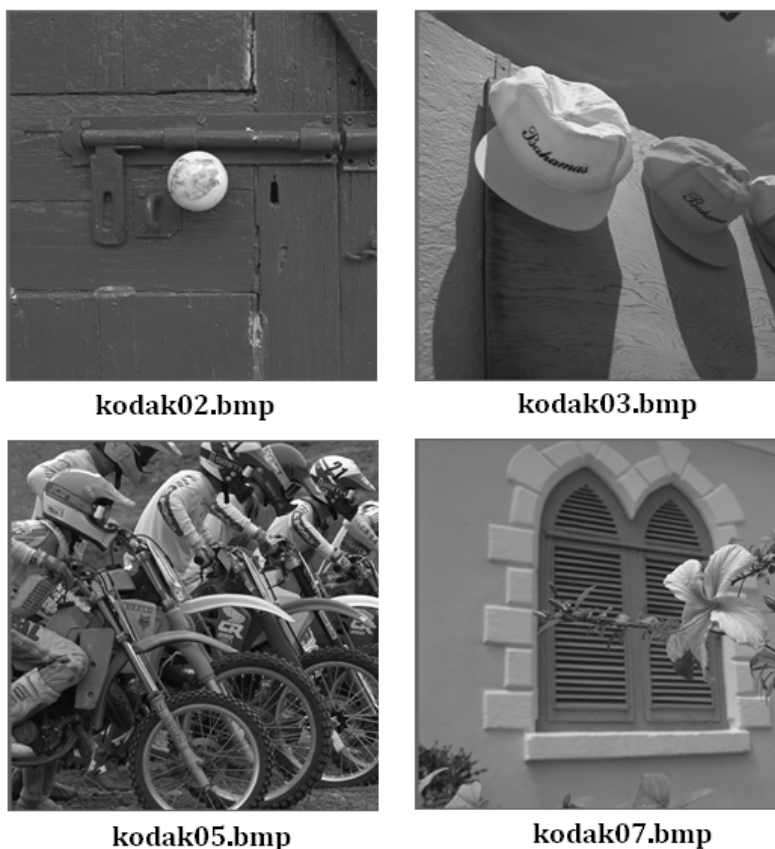
kde $0 \leq \alpha \leq 1$ a b_k je počet bitů přidělený k -tému bloku.

4. S novým rozložením bitů zakódujeme jednotlivé bloky.

Důležitým parametrem při výpočtu je hodnota váhy α , která rozhoduje o vlivu veličin na celkovou hodnotu b_k . Určením vhodné α se zabývá následující kapitola, v rámci níž byly jednotlivé metody testovány.

5.3 Testování navržených metod

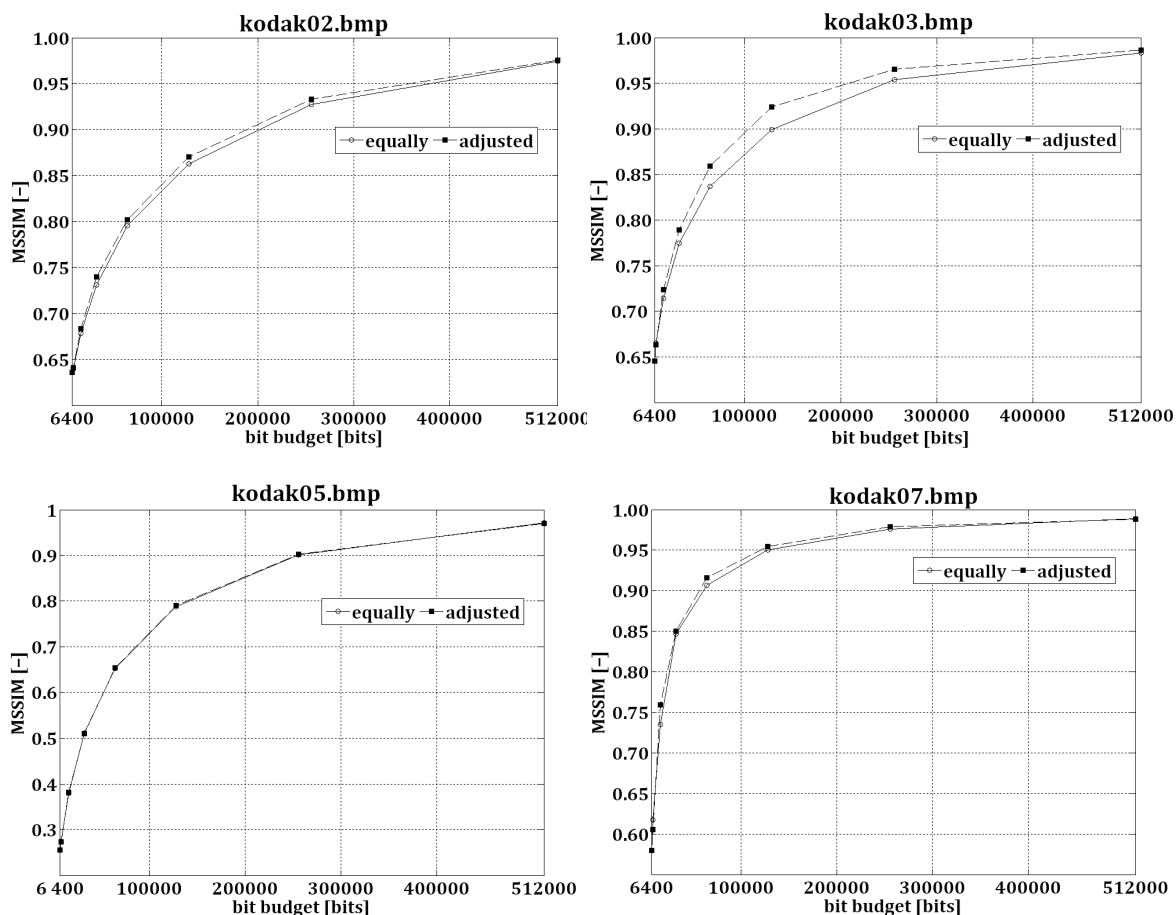
Pro testování účinnosti metod byly využity vybrané obrázky ze sady testovacích obrázků pro ztrátovou kompresi od společnosti Kodak [2]. Obrázky byly pro naše účely upraveny (převedeny do stupňů šedi a ořezány na rozměry 512×512 px). Podoba testovaných obrázků je uvedena na **obr. 5.6**. Obrázky *kodak02.bmp*, *kodak03.bmp* a *kodak07.bmp* představují běžné fotky, které obsahují jak detailní části, tak oblasti s nízkým výskytem detailů. Obrázek *kodak05.bmp* reprezentuje velmi detailní snímek. Při testování byly použity obrázky ve stupních šedi o rozměrech 512×512 px, hloubka dekompozice $n = 5$, vlnka CDF9/7, bloky 64×64 vlnkových koeficientů a kompresní algoritmus dSPIHT. Porovnávána byla vždy celková hodnota metriky MSSIM komprimovaného obrázku, a to při použití rovnoměrného rozložení bitů (každý blok obdrží stejný počet bitů) a rozdělení bitů podle příslušné metody.



Obr. 5.6: Testované obrázky

5.3.1 Rozdělení bitů podle funkční hodnoty

Na **obr. 5.7** je znázorněn průběh hodnot MSSIM komprimovaného obrázku pro zvyšující se počet přidělených bitů při rovnoměrném rozložení bitů mezi jednotlivé bloky a při rozdělení bitů podle hodnoty MSSIM zjištěné u tohoto rovnoměrného rozložení.

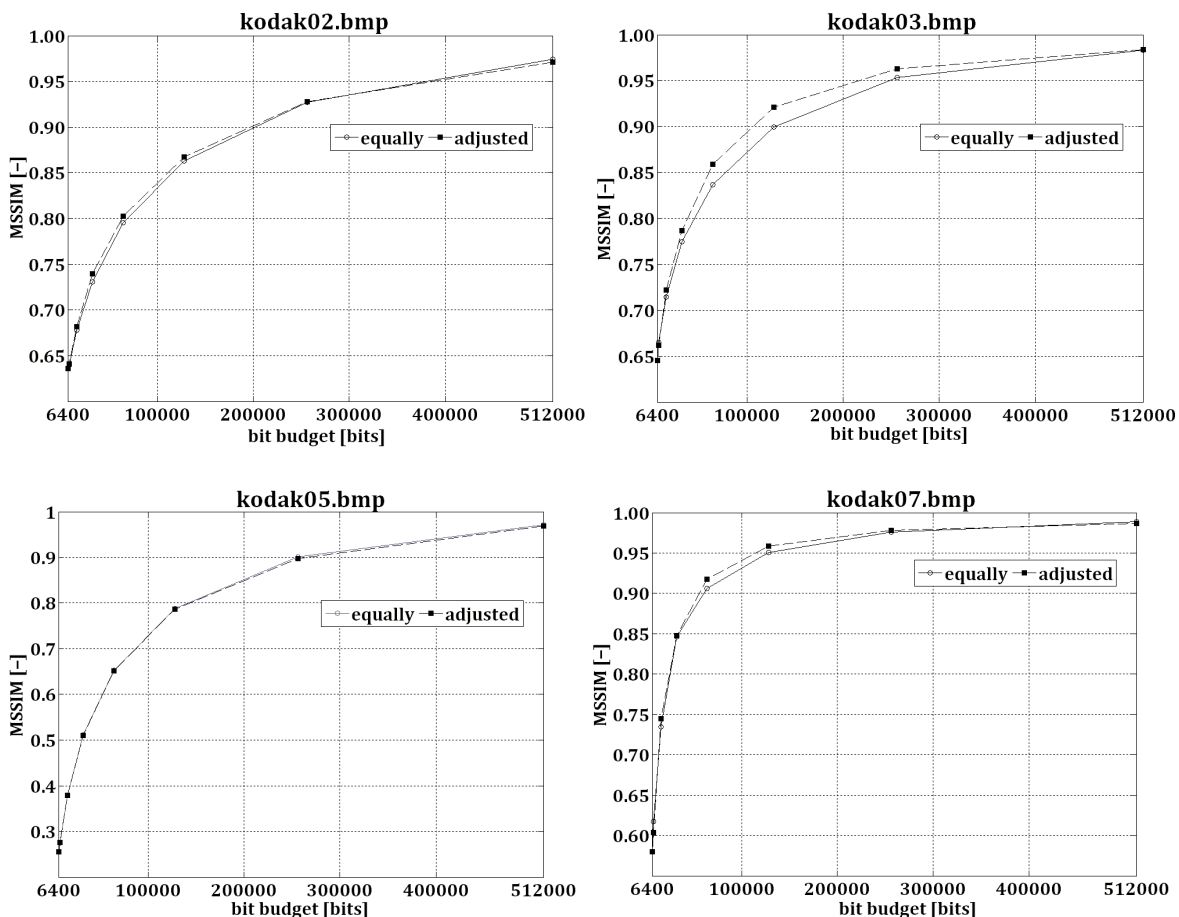


Obr. 5.7: Závislost hodnoty MSSIM testovaných obrázků na počtu přidělených bitů při rovnoměrném rozdělení (plná čára) a rozdělení podle funkční hodnoty (přerušovaná čára)

Z grafů lze vyčíst, že ve všech testovaných případech se hodnota MSSIM oproti rovnoměrnému rozdělení bitů zvýšila. Jedinou výjimkou byl obrázek *kodak05.bmp*, kde prakticky nedošlo k žádné změně (viz splývající křivky na **obr. 5.7**). Příčinou je zde velké množství detailů v celém obraze, jednotlivé bloky mají podobný charakter obrazových dat, tudíž je jim přidělen i přibližně stejný počet bitů. Největšího zlepšení kvality komprimovaného obrázku dle metriky MSSIM bylo dosaženo u obrázku *kodak03.bmp*. Důvodem je výskyt většího počtu bloků s homogenní strukturou. Tyto bloky je možné zakódovat na nižší počet bitů a ušetřit tak zbylé bity pro nejvíce zkraslené bloky obrázku.

5.3.2 Rozdělení bitů podle velikosti směrnic

Srovnání hodnot MSSIM obrázků s rovnoměrným rozdělením bitů a rozdělením bitů podle velikosti směrnic křivek jednotlivých bloků je uvedeno na **obr. 5.8**.



Obr. 5.8: Závislost hodnoty MSSIM testovaných obrázků na počtu přidělených bitů při rovnoměrném rozdělení (plná čára) a rozdělení podle velikosti směrnic (přerušovaná čára)

Je vidět, že průběh závislosti indexu MSSIM obrázků při rozdělení bitů podle velikosti směrnic je obdobný jako u předchozí metody. Nejlepších výsledků bylo dosaženo opět u obrázku *kodak03.bmp*, naopak u obrázku *kodak05.bmp* se výsledné hodnoty MSSIM téměř nezměnily. Naměřené hodnoty se lišily pouze nepatrně, o něco lepších výsledků dosahovala metoda rozdělení bitů podle funkční hodnoty.

5.3.3 Kombinovaná metoda rozdělení bitů

U této metody má na výsledné rozdělení bitů vliv váha α , proto před samotným použitím metody je nezbytné určit její hodnotu. Optimální hodnotu α jsme se pokusili stanovit empiricky, testováním metody na vybraných obrázcích. Výsledky tohoto měření jsou shrnuty v **tab. 1.1**.

Tab. 1.1: Naměřené hodnoty MSSIM obrázků při rozdílném počtu přidělených bitů a měnící se hodnotě váhy α u kombinované metody

α	<i>kodak02.bmp</i>					<i>kodak03.bmp</i>				
	bit budget [bits]					bit budget [bits]				
	8k	32k	64k	128k	512k	8k	32k	64k	128k	512k
0,0	0,6409	0,7395	0,8024	0,8673	0,9712	0,6615	0,7866	0,8589	0,9208	0,9841
0,2	0,6409	0,7397	0,8026	0,8689	0,9749	0,6624	0,7896	0,8604	0,9246	0,9870
0,5	0,6406	0,7382	0,8019	0,8686	0,9766	0,6652	0,7900	0,8612	0,9257	0,9887
0,7	0,6407	0,7395	0,8023	0,8709	0,9770	0,6652	0,7899	0,8672	0,9266	0,9893
1,0	0,6407	0,7397	0,8014	0,8698	0,9772	0,6623	0,7896	0,8606	0,9256	0,9887
α	<i>kodak05.bmp</i>					<i>kodak07.bmp</i>				
	bit budget [bits]					bit budget [bits]				
	8k	32k	64k	128k	512k	8k	32k	64k	128k	512k
0,0	0,2753	0,5104	0,6519	0,7863	0,9683	0,6038	0,8475	0,9174	0,9584	0,9868
0,2	0,2753	0,5113	0,6543	0,7884	0,9699	0,6034	0,8525	0,9216	0,9633	0,9895
0,5	0,2741	0,5115	0,6542	0,7905	0,9721	0,6177	0,8558	0,9245	0,9653	0,9911
0,7	0,2745	0,5108	0,6542	0,7895	0,9718	0,6178	0,8562	0,9265	0,9654	0,9916
1,0	0,2738	0,5108	0,6548	0,7907	0,9727	0,6177	0,8572	0,9191	0,9643	0,9907

Z naměřených hodnot je zřejmé, že index MSSIM se při změně váhy α pro různé počty přidělených bitů měnil nepatrně. To odpovídá dříve provedeným testům, kdy obě předchozí metody vykazovaly navzájem podobné výsledky. Většina naměřených indexů MSSIM dosahovala největších hodnot při velikosti váhy $\alpha = 0,7$. Proto hodnota této váhy byla vybrána pro danou metodu jako optimální.

Průběhy závislosti MSSIM na počtu přidělených bitů jsou při použití této metody u testovaných obrázků prakticky totožné s průběhy naměřenými u předchozích dvou metod (**obr. 5.7** a **obr. 5.8**). Pro zvolenou váhu $\alpha = 0,7$ dosahovala metoda pro většinu kompresních poměrů o něco málo vyšších hodnot MSSIM oproti zbylým dvěma testovaným metodám. Nutno však podotknout, že změny výsledné subjektivní kvality vůči subjektivní kvalitě obrázku při použití ostatních dvou metod byly minimální. Subjektivní kvalitu obrázků při rozdělení bitů mezi bloky pomocí této metody nebo při rovnoměrném rozdělení bitů mezi bloky je možné vizuálně porovnat v příloze A. Srovnání je provedeno pro obrázky *kodak02.bmp*, *kodak03.bmp* a *kodak07.bmp*. U obrázku *kodak05.bmp* se změny téměř neprojevily, proto srovnání není uvedeno. Je vidět, že navržené rozdělení bitů se příznivě odráží v kvalitě komprimovaných obrázků především při vyšším počtu dostupných bitů.

5.4 Shrnutí a interpretace dosažených výsledků

V rámci této kapitoly jsme provedli návrh a následné testování metod, které přerozdělují bity mezi jednotlivé bloky koeficientů za účelem dosažení co nejlepší subjektivní kvality komprimovaného obrázku. Metody pro výpočet požadovaného rozdělení využívají vlastností metriky MSSIM (SSIM), která ze všech objektivních metrik hodnocení kvality obrazu nejlépe koreluje s vnímáním obrazové informace u lidského zraku.

Je zřejmé, že žádná z navržených metod nepředstavuje optimální řešení z hlediska dosažení maximální hodnoty MSSIM obrázku při daném kompresním poměru (počtu dostupných bitů). Představují pouze určitý jednoduchý způsob rozdělení bitů mezi jednotlivé bloky, jelikož při výpočtu požadovaného rozdělení bitů vycházejí ze zjednodušených lineárních vztahů. Bity jsou vždy přidělovány úměrně podle velikosti zvoleného kritéria (funkční hodnota, velikost směrnice). Problém je v tom, že toto zjednodušení není přesné, protože hodnota metriky MSSIM je závislá na počtu bitů nelineárně (viz **obr. 5.2**). Tato závislost je silně ovlivněna statistickými parametry zpracovávaných obrazových dat. Proto je také velmi obtížné určit nějaký vztah nebo rovnici, která by popisovala závislost indexu MSSIM na přiděleném počtu bitů. Dalším důležitou vlastností metriky MSSIM (SSIM), kterou je nutné brát v potaz při hodnocení kvality obrázku, je to, že její velikost má vypovídací hodnotu pouze pro porovnávanou oblast obrazu. Jinými slovy, nelze porovnávat mezi sebou hodnoty SSIM vypočítané pro různé části obrazu, jelikož hodnota indexu SSIM v jedné části vyjadřuje jinou velikost zkreslení (kvalitu) než stejně velká hodnota SSIM v jiné části. Z tohoto důvodu není možné vylepšit kvalitu komprimovaného obrázku způsobem, jehož cílem by bylo dosažení přibližně stejné hodnoty MSSIM pro všechny bloky obrázku.

Zřejmě nejvhodnější způsob pro dosažení optimálního rozdělení bitů pro daný kompresní poměr u metriky MSSIM tak bude spočívat ve využití nějakého iteračního postupu, který optimální rozdělení nalezne (podobně jako např. Lagrangeova metoda u metriky MSE). Nevýhodou takového přístupu je vyšší časová a výpočetní složitost, jelikož k určení požadovaného řešení je nutno znát průběhy závislostí hodnot MSSIM na počtu přidělených bitů pro jednotlivé bloky obrázku. Pro přibližné určení těchto průběhů z omezeného počtu hodnot MSSIM je možno využít některé z interpolačních metod (viz kap. 5.2.1).

ZÁVĚR

Tato diplomová práce se zabývala principem objektivního hodnocení kvality obrazu pomocí indexu SSIM a jeho následným využitím při kompresi obrazu.

V úvodní části jsme nejprve zmínili, jaké jsou důvody hledání alternativních metod hodnocení kvality obrazu namísto nejpoužívanější metriky MSE. Následně byl uveden index SSIM a byl popsán způsob jeho výpočtu. Jako součást této kapitoly byla vytvořena aplikace v prostředí MATLAB, v rámci níž jsme provedli srovnání metrik MSE a MSSIM pro různé druhy zkreslení původního obrázku. Aplikace umožňuje vypočítat hodnotu obou metrik a zobrazit příslušné rozložení chybových signálů.

Kapitola 2 pokrývá teorii ke kompresi obrazu a významu objektivních kritérií pro její hodnocení. Především je rozebrán princip kódování vlnkových koeficientů obrazu pomocí algoritmu SPIHT, který představuje vhodný prostředek pro optimalizaci podle hodnoty kritéria SSIM.

V rámci kapitoly 3 byly uvedeny možnosti využití indexu SSIM při optimalizaci algoritmu SPIHT. Zmíněny byly tři různé oblasti použití. Pro své výhodné vlastnosti z hlediska možností paralelního zpracování byla pro optimalizaci vybrána metoda založená na kódování vlnkových koeficientů v nezávislých blocích.

Popisu implementace této metody v rámci realizovaného kompresního schématu je věnována navazující kapitola. Kompresní schéma umožňuje zpracovávat obrázky ve stupních šedi s rozměry násobků mocniny dvou. Postup při zpracování je naznačen na **obr. 4.2**. Obrázek je převeden do transformované oblasti pomocí vlnkové transformace, vlnkové koeficienty jsou sdruženy do bloků a poté jsou jednotlivé bloky kódovány zjednodušenou verzí algoritmu SPIHT, tzv. dSPIHT. Při kódování bloků je nezbytné určit počet bitů, který bude danému bloku přidělen. K tomuto účelu byla použita metrika MSSIM jakožto hodnotící kritérium zkreslení vzniklého v daném bloku. Způsob výpočtu MSSIM pro jednotlivé bloky je znázorněn na **obr. 4.4**. Jako optimální velikost bloku byly pro zpracovávané obrázky 512×512 pixelů zvoleny rozměry 64×64 koeficientů.

Poslední část práce se zabývá návrhem možných metod pro optimální rozdělení bitů mezi jednotlivé bloky koeficientů. Cílem je dosažení co nejvyšší hodnoty indexu MSSIM výsledného obrázku při daném kompresním poměru. Za tímto účelem byly nejprve prozkoumány průběhy závislosti indexu MSSIM na počtu přidělených bitů u jednotlivých bloků. Bylo ukázáno, že tyto průběhy jsou nelineární a značně závislé na charakteru zpracovávaných dat. Na základě zjištěných poznatků byly navrženy možné způsoby rozdělení bitů. Jako kritérium určující počet přidělených bitů danému bloku byla použita funkční hodnota a hodnota velikosti směrnice průběhů závislosti jednotlivých bloků. Dosažené výsledky při testování metod jsou znázorněny na **obr. 5.7** resp. **obr. 5.8**. Ukázalo se, že žádná z navržených metod neumožňuje dosáhnout optimálního rozložení bitů při daném kompresním poměru. Příčinou je použití příliš zjednodušených vztahů při výpočtu tohoto rozdělení. Nelineární závislost indexu MSSIM (SSIM) nelze takto jednoduše vyjádřit.

Úkolem dalšího zkoumání v této oblasti by bylo pokusit se nalézt požadované optimální řešení spíše pomocí nějakého iteračního postupu. Hodnota indexu SSIM je ovlivňována statistickými parametry zpracovávaných obrazových dat, proto nalezení nějakého obecného matematického vztahu popisující závislost na počtu přidělených bitů by bylo velmi obtížné.

SEZNAM POUŽITÉ LITERATURY

- [1] ČERMÁK, L.; HLAVIČKA, R. Numerické metody [online]. 1.11.2006 [cit. 20.5.2010]. Dostupné na internetu: <<http://mathonline.fme.vutbr.cz/UploadedFiles/246.pdf>>.
- [2] KODAK. Kodak losless true color image suite. Sada fotografií od společnosti Kodak [online]. 17.2.2010 [cit. 20.5.2010]. Dostupné na internetu: <<http://r0k.us/graphics/kodak/>>.
- [3] MALÝ, J.; VENARD, O.; PRŮŠA, Z. Color Image Compression Based on SPIHT with Separate Channel Processing. In *32nd International Conference on TELECOMMUNICATIONS AND SIGNAL PROCESSING*. Dunakiliti: ASSZISZTENCIA Congress Bureau, 2009. s. 1–5. ISBN: 978-963-06-7716-5.
- [4] PRŮŠA, Z.; MALÝ, J. Metoda SPIHT pro kompresi barevných obrazu a její implementace. *Elektrorevue – Internetový časopis* (<http://www.elektrorevue.cz>), 2009, roč. 2009, č. 17, s. 1–11. ISSN: 12131539.
- [5] PRŮŠA, Z. *Efektivní nástroj pro kompresi obrazu v jazyce Java*. Diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2008. 68 s. Vedoucí diplomové práce Ing. Jan Malý.
- [6] RACHMANDRAN, K.; ORTEGA, A. Rate-distortion methods for image and video compression. *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 23–50, 1998 [cit. 20.5.2010]. Available: <http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE523_files/Notes%20on%20Rate-Distortion%20Methods.pdf>
- [7] WANG Z.; BOVIK A.C. Mean Squared Error: Love It or Leave It?. *IEEE Signal Processing Magazine*, Vol. 26, No. 1, Jan 2009, pp. 98–117, ISSN: 1053-5888
- [8] WANG Z.; BOVIK A.C. SHEIKH H.R., and SIMONCELLI E.P., Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004 [online]. Available: <<http://www.cns.nyu.edu/~lcv/ssim/>>
- [9] WANG Z.; LI Q.; SHANG X. Perceptual image coding based on a maximum of minimal structural similarity criterion. *Proceedings of IEEE Int. Conf. Image Processing*, Sept. 2007, Vol 2., pp. 121–124.
- [10] WHEELER F. W. and PEARLMAN W. A. Low-memory packetized SPIHT image compression. In *Proceedings of the 33rd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 1999, vol. 2, pp. 1193–1197.

PŘEHLED POUŽITÝCH ZKRATEK

2D-DTWT (2D discrete time wavelet transform) – dvourozměrná diskrétní vlnková transformace v diskrétním čase

BB (Bit Budget) – počet dostupných bitů

bpp (bits per pixel) – počet bitů připadající na jeden pixel

bpb (bits per block) – počet bitů připadající na jeden blok

CDF9/7 – biortogonální Cohen Daubechies Feauveau 9/7 vlnka

CSF (Contrast Sensitivity Function) – funkce rozlišitelnosti jednotlivých jasových hodnot

dSPIHT („degraded tree“ SPIHT) – modifikace algoritmu SPIHT

HP – horní propust

HVS (Human Visual System) – lidský zrakový systém

LP – dolní propust

LIP (list of insignificant pixels) – seznam nevýznamných koeficientů

LIS (list of insignificant sets) – seznam nevýznamných množin koeficientů

LSP (list of significant pixels) – seznam významných koeficientů

MATLAB (MATrix LABoratory) – skriptovací programovací jazyk

MSE (Mean Squared Error) – střední kvadratická odchylka

MSSIM (Mean SSIM) – průměrná hodnota indexů SSIM

PSNR (Peak Signal to Noise Ratio) – špičkový odstup signál/šum

px (pixel, Picture Element) – obrazový bod

RGB (Red Green Blue) – aditivní barevný model

SOT (Spatial Oriented Tree) – prostorově orientovaný strom

SPIHT (Set Partitioning In Hierarchical Trees) – kódovací algoritmus vlnkových koeficientů

SSIM (Structural SIMilarity) – index strukturální podobnosti

YCbCr – barevný model s oddělenou jasovou a barevnými složkami