



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

SROVNÁNÍ METOD MĚŘICÍCH KVALITU OBRAZŮ

COMPARISON OF IMAGE QUALITY MEASURING METHODS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

ADAM ŠKURLA

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. DAVID BAŘINA, Ph.D.

BRNO 2018

Vysoké učení technické v Brně - Fakulta informačních technologií

Ústav počítačové grafiky a multimédií

Akademický rok 2017/2018

Zadání bakalářské práce

Řešitel: **Škurla Adam**

Obor: Informační technologie

Téma: **Srovnání metod měřících kvalitu obrazů**

Comparison of Image Quality Measuring Methods

Kategorie: Zpracování obrazu

Pokyny:

1. Seznamte se s metodami používanými k objektivnímu vyjádření podobnosti dvou obrazů.
2. Vyberte z nich ty významné, detailně je popište.
3. Tyto metody implementujte ve formě multiplatformní knihovny, vytvořte aplikaci demonstrující jejich funkčnost.
4. Na vhodných datech implementované metody srovnajte, zaměřte se měření kvality komprimovaných obrazů.
5. Diskutujte dosažené výsledky a vhodnost použití jednotlivých metod.

Literatura:

- Gayle D.; Mahlab H.; Ucar Y.; aj.: A Full-Reference Color Image Quality Measure In The DWT Domain.
- Wang, Z.; Bovik, A. C.; Sheikh, H. R.; aj.: Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing, ročník 13, č. 4, April 2004: s. 600-612.
- Zhou Wang, Alan C. Bovik: Mean Squared Error: Love It or Leave It? A new look at Signal Fidelity Measures. IEEE Signal Processing Magazine, vol. 26, issue 1, 2009

Pro udělení zápočtu za první semestr je požadováno:

- První dva body zadání.

Podrobné závazné pokyny pro vypracování bakalářské práce naleznete na adrese

<http://www.fit.vutbr.cz/info/szz/>

Technická zpráva bakalářské práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap (20 až 30% celkového rozsahu technické zprávy).

Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Bařina David, Ing., Ph.D., UPGM FIT VUT**

Datum zadání: 1. listopadu 2017

Datum odevzdání: 16. května 2018

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta informačních technologií
Ústav počítačové grafiky a multimédií
602 00 Brno, Božetěchova 2



doc. Dr. Ing. Jan Černocký
vedoucí ústavu

Abstrakt

Táto práca sa zaoberá metódami merajúcimi kvalitu obrazu a porovnaním ich výkonnosti v súvislosti s kompresnými štandardmi JPEG a JPEG 2000. V prvej časti približuje problematiku merania kvality obrazu, zmieňuje existujúce a používané metódy a detailnejšie popisuje vybrané z nich. V druhej časti popisuje použitý postup pri získaní subjektívneho a objektívneho hodnotenia kvality obrázkov. Ďalej porovnáva získané subjektívne výsledky a výsledky určené objektívnymi metódami. Záver poskytuje vyhodnotenie získaných výsledkov a určenie najvhodnejších metód na meranie kvality obrázkov vytvorených pomocou vyššie uvedených kompresných štandardov.

Abstract

This thesis is aimed at image quality assessment methods and comparison of their performance regarding compression standards JPEG and JPEG 2000. First part provides information about image quality assessment basics, mentioning existing and used methods. The second part describes approach used to acquire subjective and objective rating of image quality. Further it compares obtained subjective ratings and ratings determined by objective methods. Conclusion provides an evaluation of obtained results and specification of the most suitable methods for measuring quality of images created by compression standards stated above.

Klíčové slová

obraz, kvalita, stratová kompresia, JPEG, JPEG 2000, PSNR, PSNR-B, SSIM, UQI, MAD, Butteraugli, porovnanie výkonnosti, hodnotenie kvality obrazu

Keywords

image, quality, lossy compression, JPEG, JPEG 2000, PSNR, PSNR-B, SSIM, UQI, MAD, Butteraugli, performance comparison, image quality assessment

Citácia

ŠKURLA, Adam. *Srovnání metod měřících kvalitu obrazů*. Brno, 2018. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. David Bařina, Ph.D.

Srovnání metod měřících kvalitu obrazů

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána doktora Davida Bařiny. Ďalšie informácie mi poskytol pán docent Martin Āadík. Uviedol som všetky literárne pramene a publikácie, z ktorých som Āerpal.

.....
Adam Škurla
6. mája 2018

Podakovanie

Rád by som poĀakoval vedúcemu práce pánovi Ing. Davidovi Bařinovi, Ph.D. za poskytnuté rady a odbornú pomoc pri vytváraní tejto bakalárskej práce. Tiež by som chcel poĀakovať mojim rodiĀom za ich poskytnutú pomoc pri realizácii subjektívneho hodnotenia kvality obrázkov a za ich podporu.

Obsah

1	Úvod	2
2	Digitálny obraz a kompresia	3
2.1	Formáty obrázkových súborov	3
2.2	Stratová kompresia	4
2.3	JPEG	5
2.4	JPEG 2000	10
3	Kvalita obrazu a metódy hodnotenia kvality	14
3.1	Subjektívne hodnotenie kvality	14
3.2	Objektívne hodnotenie kvality	15
3.3	Peak Signal-to-Noise Ratio	16
3.4	Peak Signal-to-Noise Ratio – Blocking	16
3.5	Structural Similarity	18
3.6	Universal Quality Index	19
3.7	Most Apparent Distortion	20
3.8	Butteraugli	23
4	Testovanie výkonnosti metód	24
4.1	Dataset	24
4.2	Subjektívne hodnotenie datasetu	25
4.3	Objektívne hodnotenie datasetu	31
4.4	Meranie výkonnosti	40
4.5	Výpočet koeficientov a odchýlok	42
4.6	Výsledky testovania	43
4.7	Testovanie na vyššom bitrate	47
5	Záver	53
	Literatúra	54

Kapitola 1

Úvod

V súčasnosti je používanie obrázkov v digitálnej forme nenahraditeľnou súčasťou veľkého množstva oblastí a s rozvojom technológií sa neustále rozširuje. S digitálnymi obrázkami prichádzame do styku na dennej báze, či už ide o používanie internetu, kde ich nájdeme na väčšine stránok, pri fotografovaní, či úprave fotografií, pri práci s mobilným telefónom, tabletom alebo smart hodinkami. Každá z týchto skupín má svoje vlastné nároky na kvalitu obrazu, jeho veľkosť a formát, ktorý sa môže odvíjať od podporovanosti jednotlivými zariadeniami. Medzi najpoužívanejšie z dostupných obrázkových formátov dnes patrí štandard JPEG, ktorý vznikol ešte v roku 1992. Jeho novší nastupca JPEG 2000 zatiaľ nemá tak širokú podporu napriek ponúkaným vylepšeniam.

Pri používaní obrázkov je podstatné vedieť posúdiť ich kvalitu, ktorá závisí od spôsobu použitia. Ak je podstatné rozlíšenie a nezáleží na kapacite, ktorú daný obrázok zaberie, volíme minimálnu alebo žiadnu úroveň kompresie a tým vysokú kvalitu. Na druhej strane, pokiaľ chceme mať webové stránky s vysokou rýchlosťou načítania aj pri pomalšom internetovom pripojení, je podstatná menšia veľkosť obrázkov a volíme vyššiu úroveň kompresie a teda nižšiu kvalitu. Pri kompresii dochádza k vynechaniu niektorých detailov obrazu, pričom o ktoré detaily sa jedná určuje použitý štandard. Po odstránení týchto detailov môžu v obraze vzniknúť pre ľudské oko na prvý pohľad viditeľné artefakty. Ich množstvo a umiestnenie vrámci obrazu je kľúčové pre posúdenie kvality komprimovaného obrázku, ktorá sa môže každému hodnotiacemu javiť inak. Ak chceme objektívne posúdiť kvalitu či porovnať originál s komprimovaným obrazom, môžeme využiť metódy vytvorené za týmto účelom. Ich účinnosť sa môže líšiť, pričom tiež záleží na použitom formáte obrázku, nakoľko druh vzniknutých artefaktov sa môže viazať na konkrétnu kompresnú metódu.

Cielom tejto práce je vybrať vhodné metódy merania kvality obrazu a určiť, ktorá z týchto metód dosahuje najlepšiu koreláciu objektívneho hodnotenia kvality obrázkov so subjektívnym hodnotením kvality ľuďmi so zameraním na kompresné štandardy JPEG a JPEG 2000.

Práca je rozdelená do niekoľkých logických častí. V ďalšej kapitole čitateľa oboznámi s pojmom digitálny obraz, existujúcimi súborovými formátmi obrázkov, s procesom stratovej kompresie a štandardmi JPEG a JPEG 2000. V tretej kapitole priblíži problematiku merania kvality obrázkov, tiež viaceré metódy používané pre objektívne hodnotenie kvality a spôsob, ktorým kvalitu vyhodnocujú. V predposlednej kapitole je popísaný použitý postup pri vytvorení datasetu a získavaní subjektívneho a objektívneho hodnotenia. Ďalej sa kapitola venuje samotnému porovnávaniu výkonnosti metód a vyhodnoteniu získaných výsledkov. Záverečná kapitola obsahuje celkové zhrnutie a vyhodnotenie, ktorá z vybraných metód sa javí ako najvhodnejšia pre jednotlivé kompresné štandardy.

Kapitola 2

Digitálny obraz a kompresia

Pojem digitálny obraz reprezentuje dvojrozmerný obrázok uložený vo forme núl a jednotiek na pamäťovom médiu. Zvyčajne sa jedná o bitmapu alebo rastrový obrázok, ktorý má konečný počet digitálnych hodnôt nazývaných pixely. Obsahuje pevný počet riadkov a stĺpcov pixelov. Pixely sú najmenšou jednotkou obrazu a reprezentujú farbu a jej jas v danom mieste. V súvislosti s obrazom je možné hovoriť napríklad o jeho rozlíšení, farebnej hĺbke, formáte obrázkového súboru a taktiež o kvalite obrázku, ktorá môže byť subjektívna.

2.1 Formáty obrázkových súborov

Tieto formáty predstavujú mnoho štandardizovaných spôsobov uchovávaní digitálnych obrázkov na pamäťovom médiu. Obrázkové súbory pozostávajú z digitálneho obsahu v konkrétnom formáte, ktorý zvyčajne značí koncovka súboru predstavujúca skratku z celého pomenovania formátu.

Digitálne dáta reprezentujúce vizuálne informácie sú pri procese zobrazenia obrázku na výstup rasterizované. Výstupom je myslená napríklad obrazovka počítača alebo tiež tlačiareň. Rasterizácia predstavuje prevedenie digitálnych dát na hodnoty reprezentujúce nastavenie jednotlivých pixelov zariadenia, na ktorých bude obsah obrázku zobrazený.

Základné druhy formátov

Digitálne dáta obrázku môžu byť uchovávané buď v rastrovom, či vektorovom formáte. Rastrové formáty sa ďalej delia na komprimované a nekomprimované.

Pokiaľ je obrázok uchovávaný v komprimovanom formáte, pôvodné množstvo údajov reprezentujúcich obsah obrázku bolo zredukované na menšie množstvo, čím je dosiahnutá nižšia veľkosť výsledného súboru. Tento proces môže byť buď zvratný, kedy sa dá z komprimovaného stavu obnoviť pôvodný stav s pôvodnými vlastnosťami, napríklad rozlíšenie a farby obrázku. Alebo proces nezvratný, kedy sú niektoré údaje pri procese komprimácie nenávratne stratené, a nie je možné obnoviť pôvodný stav pred komprimáciou. Obrázok uchovávaný v nekomprimovanom formáte obsahuje pôvodné informácie vzniknuté pri vytvorení obrázku bez zásahov. Primárne sa jedná o obrázky v *raw*, čiže surovom formáte, ktoré sú produkované senzorom napríklad digitálnej kamery alebo skenera. Obrázky uchovávané vo vektorovom formáte sa skladajú zo základných útvarov, medzi ktoré patria body, krivky, priamky a mnohoúhelníky. Každý z bodov má definovanú polohu na osi x a osi y a určuje dráhu, ktorá môže mať viacero vlastností, ako napríklad farbu a hrúbku.

Veľkosť súboru

V prípade, že sa jedná o rastrový obrázok je veľkosť súboru zvyčajne podmienená rozlíšením obrázku a farebnou hĺbkou, pričom so stúpajúcim rozlíšením alebo farebnou hĺbkou stúpa veľkosť súboru. Medzi najznámejšie rastrové formáty patrí najmä JPEG, ktorý bude ďalej podrobne popísaný spolu s jeho novším nástupcom JPEG 2000. Ďalej PNG (Portable Network Graphics), využívaný najmä na webe pre dobrú veľkosť súboru so zachovaním dostatočnej kvality, potom napríklad formát GIF (Graphics Interchange Format) s obmedzením na 8-bitovú farebnú hĺbkou alebo tiež formát BMP, ktorého súbory sú nekomprimované, tým pádom veľké no bezstratové.

Na druhej strane v prípade vektorových obrázkov nie je veľkosť súboru podmienená rozmermi, ale stúpa pridaním ďalších útvarov. Tu medzi najznámejšie patrí najmä SVG (Scalable Vector Graphics), ktorý je otvoreným štandardom.

2.2 Stratová kompresia

Stratová, alebo tiež *lossy* kompresia obrazu je proces, pri ktorom vzniká nepresná aproximácia pôvodného obrazu, pričom sa čiastočne zahadzujú údaje popisujúce jeho obsah. Hlavným cieľom je redukcia veľkosti obrázkového súboru a ide o nezvratný proces. Obrázky 2.1 a 2.2 ukazujú dve situácie, kedy v prvom prípade ide o vysokú kvalitu a nízku úroveň kompresie, v druhom o nízku kvalitu a vysokú úroveň kompresie.



Obr. 2.1: Nízka úroveň kompresie.



Obr. 2.2: Vysoká úroveň kompresie.

Dobre navrhnutá metóda stratovej kompresie môže značne znížiť veľkosť súboru a užívateľ si vôbec nemusí všimnúť pokles kvality. Častokrát je však požadovaná ešte intenzívnejšia úroveň kompresie (menšia veľkosť obrazu) kvôli zníženiu množstva prenášaných údajov, zníženiu potrieb na úložisko a podobne. V niektorých situáciách býva obrázok stlačený aj v pomere 10:1. V mnohých prípadoch obraz obsahuje viac informácií, ako je pre daný účel potrebné, napríklad môže obsahovať viac detailov, než dokáže ľudské oko rozlíšiť.

Niektoré formáty stratovej kompresie trpia generačnou stratou, kedy opakovaná kompresia a dekompresia spôsobuje postupné znižovanie kvality, no existuje aj bezstratová kompresia, pri ktorej k tomuto javu nedochádza. Základnou schémou kompresie je *transformačné* alebo *prediktívne* kódovanie. Transformačné kódovanie je používané zvyčajne pre stratové kodeky. Obraz sa rozdelí na malé segmenty – bloky, ktoré sú transformované do nového základného priestoru a *kvantizované*. Počas kvantizácie sa prevádza množina vstupných hodnôt na výstupnú množinu obsahujúcu menší počet hodnôt. Výsledné kvantizované hod-

noty sú potom kódované do bitového toku pomocou entropického kódovania. Prediktívne kódovanie je používané najčastejšie pri bezstratových kodekoch. Využívajú sa predchádzajúce a/alebo nasledujúce dáta na predpoveď aktuálnych dát nasledované porovnaním predpovede so skutočnou hodnotou. Rozdiel medzi predpovedanými a skutočnými dátami je prevedený entropickým kódovaním do bitového toku.

Obrázok vzniknutý kompresiou môže byť značne odlišný od originálu na bitovej úrovni, pričom pre ľudské oko môže ísť o nerozoznatelný rozdiel. Viacero kompresných metód sa sústreďuje na ľudský vizuálny systém, teda berie ohľad na zložky, ktoré ľudské oko vidí a vypúšťajú tie málo viditeľné.

Artefakty

Obrazové artefakty sú pre pozorovateľa obrázku viditeľné skreslenia jeho obsahu. Môžu vzniknúť napríklad pri zachycovaní obrazu digitálnym snímačom či pri stratovej kompresii. Pokiaľ algoritmus vykonávajúci kompresiu nebol schopný reprodukovať dostatok dát vo výslednom obrázku, obraz má nižšiu pozorovateľnú kvalitu a je tu priestor pre vznik artefaktov. Kompresný algoritmus tiež nemusí brať ohľad na niektoré pre ľudské vnímanie podstatné detaily. Podľa charakteru, množstva a umiestnenia artefaktov je ovplyvnená užívateľom subjektívne vnímaná kvalita.

Pri stratovej kompresii s použitím JPEG sa objavujú najmä artefakty nazývané *ringing* a *blocking*. Pre JPEG 2000 je typický takisto *ringing* a tiež *blurring*. Všetky tri artefakty budú ďalej podrobne opísané vždy pri príslušnom kompresnom štandarde, u ktorého sa najčastejšie vyskytujú.

2.3 JPEG

Skratka JPEG je odvodená od Joint Photographic Experts Group, čo je výbor stojaci za vytvorením štandardu. Ide o v súčasnosti široko používaný štandard stratovej kompresie digitálneho obrazu. Je možné nastaviť úroveň kompresie, čo umožňuje výber správneho pomeru medzi kvalitou a veľkosťou obrazu. Zvyčajne sa dosahuje kompresie s pomerom 10:1 pri len miernom znížení pozorovateľnej kvality.

JPEG je vhodný pre fotografie reálnych scén obsahujúcich rôzne farby a tóny. Nie veľmi sa hodí pre digitálne kresby obsahujúce ostré hrany alebo línie, ďalej ikony a obrázky obsahujúce text. Kontrast na hranách medzi susednými pixelmi tu môže vytvárať viditeľné artefakty. JPEG by nemal byť používaný v situáciách, kedy je vyžadovaná presná reprodukcia údajov, ako napríklad v medicínskych a iných vedeckých aplikáciách. Taktiež nie je vhodné takéto obrázky viacnásobne upravovať, pretože sa zvyčajne s každou ďalšou úpravou stráca kvalita.

Základný súborový formát pre tento štandard s názvom JIF (JPEG Interchange Format) sa často nepoužíva, hlavne z dôvodu náročnosti vytvorenia kodéru a dekodéru, ktoré by plne implementovali všetky aspekty štandardu. Namiesto tohoto základného boli vyvinuté viaceré dodatočné formáty. Väčšina zo zariadení zachytávajúcích obraz, ktoré produkujú JPEG vytvárajú súbory v EXIF (Exchangeable Image File Format) formáte, ktorý sa stal štandardom v kamerovom priemysle. Väčšina softvérov na úpravu obrazu zas ukladá JPEG v JFIF (JPEG File Interchange Format). Súbory, ktoré obsahujú JPEG kompresiu sa často nazývajú len JPEG súbory. Oba formáty, EXIF aj JFIF používajú bytové rozloženie JIF, ale každý z nich používa iba jeden z rozširujúcich bodov normy JIF nazývaný aplikačná

značka. Pre JFIF je to APP0 a pre EXIF APP1. Najčastejšie používané koncovky súborov používajúcich JPEG sú .jpeg a .jpg, niekedy je to tiež .jpe, .jif alebo .jif.

Blocking a ringing artefakty

Blocking je artefakt typický pre tento štandard. Vyskytuje sa pri použití vyššej úrovne kompresie a je viditeľný vo forme blokov, pričom sú zreteľné hranice týchto blokov, na ktorých prebehla diskretná kosínusová transformácia (DCT). Koeficienty každého bloku sú totiž po transformácii kvantizované. Čím vyššia úroveň kompresie, tým viac koeficientov je kvantizovaných na nulu. Všeobecne majú obrázky viac nízkofrekvenčného, ako vysokofrekvenčného obsahu a po kvantizácii zostáva primárne nízkofrekvenčný obsah. V krajných situáciách len DC koeficient z DCT, ktorý reprezentuje priemernú farbu bloku, čiže vznikne jednofarebný blok. Porovnanie pôvodného obrázku s komprimovaným je možné vidieť na obrázkoch 2.3 a 2.4, kedy druhý obrázok, ktorý je výsledkom vysokej úrovne kompresie obsahuje viditeľné bloky.



Obr. 2.3: Pôvodný obrázok.



Obr. 2.4: Obrázok obsahujúci blocking.

Kvantizácia je aplikovaná na každý blok zvlášť, takže koeficienty susediacich blokov sú kvantizované odlišným spôsobom.

Ringing artefakty sa objavujú ako rušivé signály v blízkosti ostrých prechodov viditeľné ako prstence, či pásy v blízkosti okrajov. Medzi hlavné príčiny patrí prechod signálu cez dolnopriepustný filter, čím sú odstránené vysoké frekvencie.

Ringing je možné vidieť na obrázku 2.6, kde maják oproti tomu na obrázku 2.5 obklopuje obrys, ktorý je okom rozlíšiteľný od okolitej oblohy.



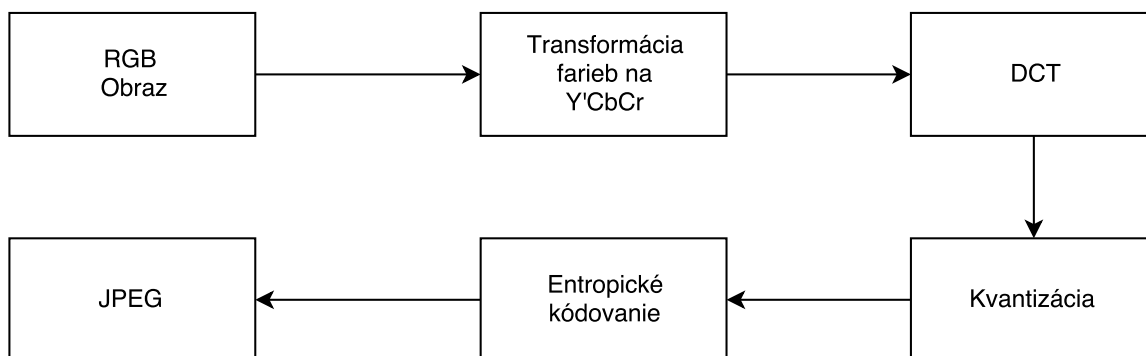
Obr. 2.5: Pôvodný obrázok.



Obr. 2.6: Obrázok obsahujúci ringing.

Kódovanie

Počas procesu zakódovania obrázku pri JPEG kompresii sú jeho farby prevedené z farebného modelu RGB (Red, Green, Blue) do $Y'CbCr$, kde Y' je jeden *luma* komponent predstavujúci jas, a dva *chroma* komponenty Cb a Cr predstavujúce farbu. Obrázok je rozdelený do blokov o veľkosti 8×8 pixelov a pre každý blok každého z Y' , Cb a Cr komponent podstúpi DCT. Následne sú amplitúdy frekvenčných zložiek kvantizované. Ľudský zrak je citlivejší na malé rozdiely vo farbe alebo jase na veľkých plochách ako na rozdiely vysokofrekvenčných odchýlok jasu, kvôli čomu je rozsah vysokofrekvenčných zložiek uložený s menšou presnosťou oproti nízkofrekvenčným. Výsledné dáta všetkých blokov s rozmermi 8×8 sú potom skomprimované bezstratovým algoritmom založenom na Huffmanovom kódovaní.



Obr. 2.7: Schéma kompresie JPEG.

Obrázok 2.7 zobrazuje proces kompresie obrázku pomocou štandardu JPEG. Opačný proces, proces dekódovania postupuje opačným smerom, pričom vynechá kvantizáciu nakoľko ide o nezvratný proces.

Redukcia farieb a delenie na bloky

Rozlíšenie chroma komponent môže byť zredukované, typicky faktorom 2 alebo 3. Tento krok predpokladá, že oko je menej citlivé na farebné detaily, než na detaily jasu. Po transformácii z RGB do Y'CbCr farebného modelu môže nasledovať redukcia zložiek Cb a Cr. Využíva sa viacero prístupov, zvyčajne je to 4:4:4 – bez redukcie, 4:2:2 – redukcia faktorom 2 v horizontálnom smere alebo najčastejšie 4:2:0 – redukcia faktorom 2 v horizontálnom aj vertikálnom smere. Podľa zvolenej redukcie chroma zložiek farebného modelu je každý kanál rozdelený na bloky. Bez redukcie sú to bloky o veľkosti 8×8 , pri redukcii 4:2:2 veľkosť 16×8 a najčastejšie pri redukcii 4:2:0 bloky 16×16 .

Keď nie je dostatok pixelov na vyplnenie bloku – napríklad na hrane obrazu, kódér musí vyplniť prázdne miesto blokov nejakými dátami. Ak sú chýbajúce dáta vygenerované nesprávnym spôsobom, môžu na hranici obrázku vznikáť artefakty.

Diskrétna kosínusová transformácia

Najprv sú hodnoty každého bloku posunuté z kladného rozsahu do rozsahu so stredom v nule. Ak uvažujeme obraz s 8-bitovou farebnou hĺbkou, každá hodnota pixelu môže byť v rozmedzí $[0, 255]$. Stredná hodnota je v tomto prípade 128, tá sa odčíta od každej hodnoty, čím vzniká nové rozmedzie hodnôt $[-128, 127]$. Týmto krokom je zredukovaný dynamický rozsah potrebný v DCT transformácii.

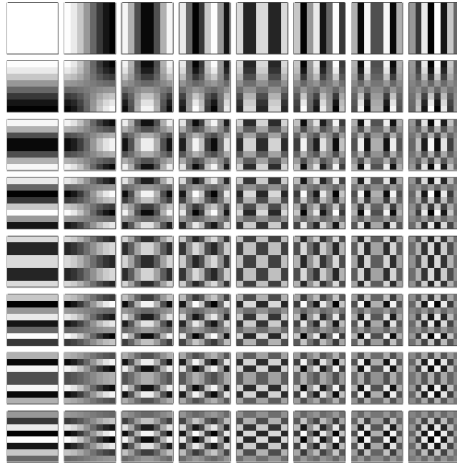
Cieľom DCT transformácie je dekorovať (odstrániť závislosti v signáloch) dáta obrazu [4]. Po dekorácii môže byť každý koeficient kódovaný nezávisle bez straty kompresnej účinnosti. DCT transformáciu je možné si predstaviť, ako súčet viacerých kosínusoid s rôznymi frekvenciami. Keďže ide o diskretnú transformáciu, aplikuje sa na konečný počet bodov. Jednodimenzionálnu transformáciu predstavuje

$$C(x) = \alpha(x) \sum_{n=0}^{N-1} f(n) \cos \left[\frac{\pi(2n+1)x}{2N} \right]. \quad (2.1)$$

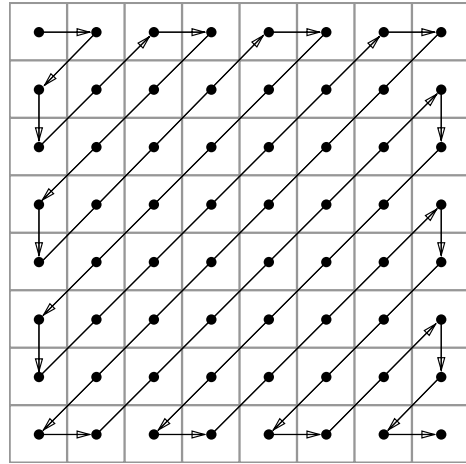
Obraz je reprezentovaný dvomi dimenziami, čiže je aplikovaná dvojdimenzionálna kosínusová transformácia

$$C(x, y) = \alpha(x)\alpha(y) \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} f(n, k) \cos \left[\frac{\pi(2n+1)x}{2N} \right] \cos \left[\frac{\pi(2k+1)y}{2N} \right]. \quad (2.2)$$

DCT transformácia je aplikovaná na maticu reprezentujúcu blok s hodnotami posunutými na stred v nule. Vzniká nová matica s najväčšími hodnotami sústredenými najmä okolo ľavého horného rohu. Hodnota ľavého horného koeficientu sa nazýva DC-koeficient a definuje základný odtieň bloku, zvyšné koeficienty sa nazývajú AC-koeficienty. DCT transformácia môže dočasne zvýšiť bitovú farebnú hĺbkou, napr. koeficienty 8-bitového obrazu môžu po transformácii zabrať aj viac ako 10 bitov. Kódér môže tým pádom dočasne zdvojnásobiť používané bity na ukladanie hodnôt. V nasledujúcom kroku kvantizácie sú hodnoty zredukované naspäť na pôvodnú bitovú hĺbkou.



Obr. 2.8: Bázové funkcie DCT.¹



Obr. 2.9: Prechádzanie cik-cak.

DCT transformuje blok vstupných hodnôt na lineárnu kombináciu vzorov ukázaných na obrázku 2.8. Tieto vzory sa nazývajú *bázové funkcie* dvojdimenzionálnej DCT.

Kvantizácia a kódovanie

Po uskutočnení DCT transformácie nasleduje krok nazývaný kvantizácia. V tomto kroku sú zredukované informácie, najmä vo vysoko frekvenčných komponentoch. Každý komponent je vydelený konštantou a následne zaokrúhlený na najbližšie celé číslo. Mnoho komponentov s vysokou frekvenciou je zaokrúhlených na nulu. Počas zaokrúhľovania sú stratené informácie, čím sa dosahuje zníženie potrebnej kapacity pamäte na uloženie obrazu (okrem redukcie chroma komponent).

Entropické kódovanie, ktoré nasleduje po kvantizácii spracováva jednorozmerný signál. Matica s koeficientami vzniknutými v predchádzajúcom kroku je prechádzaná spôsobom *cik-cak* zobrazenom na obrázku 2.9. Postupnosť koeficientov získaná z cik-cak prechodu maticou je nakoniec zakódovaná použitím Huffmanovho kódovania. Štandard JPEG poskytuje všeobecné Huffmanove tabuľky, kódér však môže vygenerovať optimalizovanú verziu tabuľky pre frekvencie dané konkrétnym obrazom. Algoritmus vytvorí kódovú tabuľku s kódovaním pre každý zdrojový symbol a toto kódovanie má premennú dĺžku. Tabuľka je odvodená na základe početnosti výskytu (váhy) každej zo zdrojových hodnôt. Častejšie sa vyskytujúce hodnoty sú reprezentované menším počtom bitov ako menej časté hodnoty, čím dochádza k zmenšeniu miesta potrebného na uloženie zakódovaných hodnôt.

Výpočetná náročnosť a kompresný pomer

JPEG nepatrí medzi výpočetne náročné štandardy. Známu implementáciou jeho kodeku je nástroj `libjpeg`, ktorého autorom je Independent JPEG Group. V marci 2017 vydala spoločnosť Google open source projekt s názvom `Guetzli`, ktorý má značne dlhší čas kódovania v porovnaní napríklad s `libjpeg`, no dosahuje vyššiu kvalitu obrázkov v kombinácii s menšou veľkosťou výstupného súboru [1]. Štandard JPEG dosahuje v lepších prípadoch kompresný pomer 10:1 z hľadiska veľkosti súboru pri zachovaní veľmi podobnej vizuálnej kvality výstupného obrázku.

¹<https://upload.wikimedia.org/wikipedia/commons/2/24/DCT-8x8.png>

2.4 JPEG 2000

JPEG 2000 je štandard kompresie obrazu vytvorený v roku 2000, od čoho je aj odvodený jeho názov, pričom mal za cieľ nahradiť pôvodný JPEG z roku 1992. Pri komprimácii využíva na rozdiel od štandardu JPEG a jeho DCT diskretnú vlnkovú transformáciu, ktorá mu dodáva lepšie vlastnosti. Oproti predchodcovi ponúka tento štandard viacero výhod, medzi ktorými je napríklad lepší kompresný pomer – pri nižšom dátovom toku sú artefakty menej viditeľné a nevyskytujú sa takmer žiadne blocking artefakty. Ďalej štandard poskytuje progresívne dekódovanie a škálovateľnosť pomeru signálu k šumu. Poskytuje účinnú organizáciu toku kódu, napríklad ak užívateľ prijme zo servera menšiu časť celku, najprv sa mu zobrazí obrázok s nižšou kvalitou a následne sa kvalita postupne zlepšuje počas sťahovania ďalších dát. Takisto poskytuje výber medzi stratovou a bezstratovou kompresiou v jednej architektúre. JPEG 2000 ďalej rozkladá obraz v priebehu procesu kompresie do viacerých rozlíšení, ponúka podporu ľubovolnej farebnej bitovej hĺbky, ľubovolného farebného priestoru a tiež podporu priehľadnosti. Ako pri štandarde JPEG, tak aj tu môžu vzniknúť viaceré druhy artefaktov. Vyskytujú sa najmä v obrázkoch s nižším bitrate a vyšším kompresným pomerom a dominuje najmä spomínaný blurring a ringing.

Blurring

Blurring artefakty predstavujú vyhladenie, ktoré je výraznejšie na skomprimovanom obrázku, než na pôvodnom. Pre ľudské oko sa jedná o jednoducho spozorovateľný rozdiel. Sú zachované údaje o tvare no stráca sa textúra, pričom tento druh artefaktu postihuje najmä tento štandard na rozdiel od klasického JPEG.



Obr. 2.10: Pôvodný obrázok.



Obr. 2.11: Obrázok obsahujúci blurring.

Príklad ukazujú obrázky 2.10 a 2.11, kedy na druhom obrázku vzniknutom kompresiou s využitím tohoto štandardu je možné pozorovať rozmazanie ostrých hrán, napríklad na hornom konci vertikálne umiestneného dreveného trámu.

Redukcia farieb a delenie na dlaždice

Pri procese komprimácie je potrebné rovnako ako pri štandarde JPEG transformovať obraz z RGB farebného priestoru na iný priestor, aby došlo k zredukovaniu množstva dát potrebných na reprezentáciu obsahu obrázku. K tomuto účelu sa používajú dva prístupy: zvrtná transformácia farieb, ktorá používa farebný priestor typu YUV – jeden luma (Y) a dva chroma komponenty (UV). A druhý prístup predstavujúci nezvratnú transformáciu farieb do farebného priestoru Y'CbCr používaného aj pri JPEG kompresii.

Po transformácii farieb je obraz rozdelený do štvorcových dlaždíc nazývaných *tiles*, ktoré sú transformované a kódované oddelene. Majú voliteľnú veľkosť, pričom môže byť celý obraz ako jedna dlaždica. Rozdelenie obrazu na dlaždice poskytuje výhodu v podobe menšej náročnosti na pamäť počas dekódovania. Pri použití príliš veľkého množstva dlaždíc hrozí vznik artefaktov typu blocking.

Diskrétna vlnková transformácia

Dlaždice vzniknuté rozdelením v predchádzajúcom kroku sú následne podrobené diskrétny vlnkovej transformácii (DWT). Používajú sa dve rozdielne vlnkové transformácie: zvrtná CDF 5/3 (Cohen-Daubechies-Feauveau) vlnková transformácia využívajúca len celočíselné koeficienty, u ktorej výstup nevyžaduje kvantizáciu a nezavádza žiadny kvantizačný šum. Druhá je nezvratná CDF 9/7 vlnková transformácia, ktorá zavádza kvantizačný šum závisiaci na presnosti dekodéru.

DWT transformácia [2] vytvára signál z *vlniek*, vstupný signál transformuje pomocou filtrov dolná priepusť a horná priepusť. Filter dolnej priepusti prepúšťa signály nízkej frekvencie a filter hornej priepusti signály vysokej frekvencie. Transformáciou vzniknú subpásma zoskupené do štyroch zón. Podoba takýchto zón pri reálnej fotografii je ukázaná na obrázku 2.12. Tieto zóny sú kombinácie filtrov dolnej (L) a hornej (H) priepusti. Proces transformácie sa aplikuje vo viacerých iteráciách, pre kompresiu obrazu zvyčajne 4 až 5 iterácií. Vstupom každej ďalšej iterácie je potom výstup dolnej priepusti, čiže LL kvadrant, čo znázorňuje obrázok 2.13.



Obr. 2.12: Kvadranty reálnej fotografie.

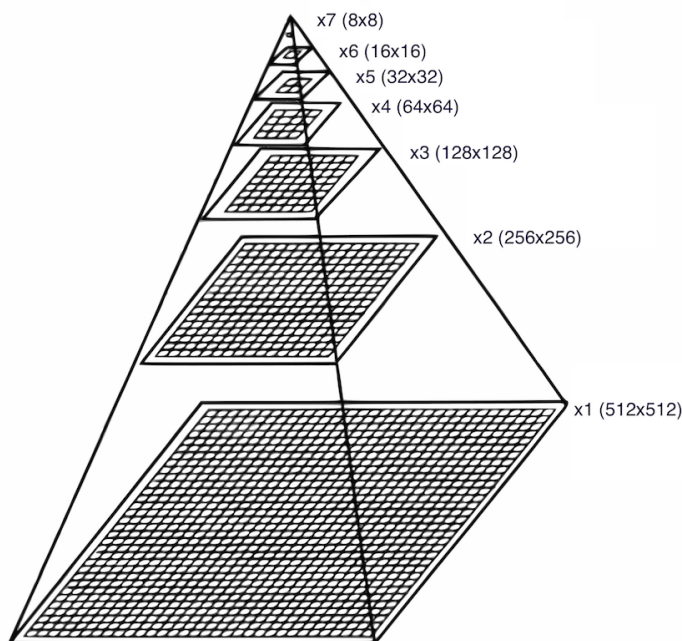
LL ₂	HL ₂	HL ₁
LH ₂	HH ₂	
LH ₁		HH ₁

Obr. 2.13: Kvadranty po dvoch iteráciách.

Kvantizácia a kódovanie

Podobne ako pri JPEG, aj po vlnkovej transformácii sú na úkor kvality koeficienty kvantizované za účelom redukcie počtu bitov potrebných na ich reprezentáciu. Výstupom je súbor celých čísel, ktoré musia byť zakódované. Je možné zmeniť kvantizačný krok: čím väčší krok, tým nižšia kvalita a väčšia úroveň kompresie. Výsledkom kvantizácie je kolekcia *subpásiem*, ktoré predstavujú niekoľko odlišných mierok. Subpásmo je súbor koeficientov, reálnych čísel reprezentujúcich aspekty obrazu určitej priestorovej oblasti a frekvenčného rozsahu. Subpásma sú následne rozdelené do *častí* vo vlnkovej doméne, ktoré sa ďalej rozdelia na kódové *bloky*. Bloky sa nachádzajú v jednom subpásme a až na tie na okrajoch obrazu majú rovnaké veľkosti. Úlohou kodéru je zakódovať bity všetkých koeficientov kódového bloku od najvýznamnejších po menej významné.

Využíva sa kódovanie so schémou EBCOT [10], každý bit bloku je zakódovaný počas troch prechodov a bity vybrané týmito prechodmi sú ďalej zakódované binárnym aritmetickým kodérom. Výsledná postupnosť bitov je rozdelená na *pakety*, pričom každý paket predstavuje všetky kódové bloky z jednej časti a zoskupuje ich do jedinej jednotky. Pakety všetkých subpásiem sú zhromaždené vo *vrstvách*. Kodek zvyčajne vytvára vrstvy tak, že s každou ďalšou vrstvou kvalita postupne stúpa a skreslenie sa znižuje. Tieto vrstvy predstavujú pyramidovú reprezentáciu, do ktorej tento štandard vstupný obrázok ukladá. Príklad takejto pyramídy znázorňuje obrázok 2.14, kde pri každej vrstve je uvedené rozlíšenie v pixeloch.



Obr. 2.14: Ilustrácia pyramídy.²

Výpočetná náročnosť a kompresný pomer

Oproti pôvodnému kompresnému štandardu je JPEG 2000 omnoho viac zložitý v prípade výpočetnej náročnosti. Vyššie spomínané kroky ako delenie na dlaždice, transformácia fa-

²<https://pinterest.com/pin/833588212249788129/>

rebných komponentov a DWT transformácia nie sú náročné, no entropické kódovanie a kódovanie EBCOT zaberú v procese kompresie najviac času. Štandard v porovnaní s jeho predchodcom ponúka zmenšenie výstupného súboru približne o 20%, pričom táto hodnota sa môže meniť v závislosti na obsahu obrázku.

Kapitola 3

Kvalita obrazu a metódy hodnotenia kvality

Táto vlastnosť obrazu meria vnímanie degradácie obrazu zvyčajne v porovnaní s ideálnym či dokonalým obrazom. Vzťahuje sa na subjektívne vnímanie obrazu človekom pozeraajúcim sa na obraz a je ovplyvňovaná viacerými faktormi. Jedným z nich je napríklad *šum*, ktorý môže vzniknúť počas stratovej kompresie. Softvér na redukciu šumu zvyčajne môže znížiť jeho viditeľnosť vyhladzovaním, no v niektorých prípadoch je nevýhodou strata jemných detailov v oblastiach s nízkym kontrastom. Ďalej napríklad *skreslenie* viditeľné vo forme nepriamych, pôvodne priamych línií, či *ostrosť* určujúca množstvo detailov dostupných v obraze. Má na ňu vplyv technika použitá pri vytváraní obrazu – šošovka a senzor, chvenie kamery, optická a softvérová stabilizácia a presnosť zaostrenia senzoru. Na kvalitu majú takisto vplyv artefakty rôzneho druhu.

3.1 Subjektívne hodnotenie kvality

Subjektívna kvalita obrazu je kvalita vnímaná ľuďmi. Pri získavaní subjektívneho hodnotenia kvality obrázkov ide o psychofyzikálne experimenty, pri ktorých divák hodnotí vybranú sadu obrázkov nazývanú *dataset*. Tieto testy bývajú náročné na prípravu a prevádzku. Pre vyhodnotenie subjektívnej kvality obrázku sa zvyčajne podnikajú tieto kroky: na testovanie sa vyberie pôvodný obrázok bez skreslenia. Vyberú sa nastavenia, ktoré sa použijú na vytvorenie upravených obrázkov a aplikujú sa na originálny obrázok, čím vzniknú testovacie obrázky. Vyberie sa spôsob, akým sa tieto obrázky prezentujú divákovi, ktorých názor sa bude zaznamenávať. Následne sú tieto sekvencie obrázkov hodnotené divákmi, pričom na hodnotenie sa zvyčajne využíva stupnica MOS (Mean Opinion Score). Po dokončení hodnotenia je vypočítaný výsledok pre každý obrázok. Testovanie by malo obsahovať dostatočný počet rôznych obrázkov s odlišným obsahom a zdrojové obrázky by mali byť čisté, bez akýchkoľvek viditeľných artefaktov znižujúcich ich kvalitu.

Aby bolo získané štatisticky významné hodnotenie, je potrebný určitý počet hodnotiacich. Podľa [15] je minimálny potrebný počet 10 ohodnotení každého obrázku a akýkoľvek počet vyšší než 40 ohodnotení neprináša významnú pridanú hodnotu. Diváci by zároveň nemali pôsobiť v oblasti spracovania obrazu, či podobných sférach. Výsledné skóre konkrétneho obrázku v stupnici MOS sa po ukončení hodnotenia počíta, ako súčet všetkých hodnotení tohoto obrázku vydelený počtom ľudí, ktorí obrázok ohodnotili. Toto skóre je vyjadrené ako jedno racionálne číslo v zvolenom rozsahu. Dolná hranica rozsahu je najhoršie

vnímaná kvalita a horná hranica je najlepšie vnímaná kvalita. Zvyčajne sa používa stupnica s hodnotami 1 až 5, kde 5 predstavuje najvyššiu kvalitu.

3.2 Objektívne hodnotenie kvality

Metódy merajúce kvalitu obrazu majú za cieľ predvídať kvalitu vnímanú ľudským pozorovateľom. Záujem o objektívne posúdenie kvality sa za posledné roky zväčšuje, preto vývoj nových a presnejších metód pokračuje. Metódy využívané na objektívne posúdenie kvality obrázkov sa delia do niekoľkých skupín podľa viacerých vlastností. Z hľadiska dostupnosti pôvodného obrázku pri posudzovaní kvality sú metódy rozdelené na *full-reference* metódy, ktoré priamo porovnávajú testovaný obrázok s originálom v ideálnej kvalite, *reduced-reference* metódy, ktoré porovnávajú vlastnosti extrahované z oboch obrázkov a *no-reference* metódy určujúce kvalitu testovaného obrázku bez akejkoľvek referencie na originálny obrázok. Metódy sú tiež rozdelené na základe spôsobu vyhodnocovania kvality a to na metódy založené na modeloch *zdroja* obrazu, ktorými sú často štatistické modely zachytávajúce štatistické vlastnosti nízkej úrovne obrázkov a dokážu vyhodnotiť informácie obrazu efektívnym spôsobom, napr. metóda Mean Squared Error. Ďalej na metódy založené na zachytávaní *skreslení*, ktoré sú najužitočnejšie, pokiaľ je dostupná informácia o druhoch skreslení vyskytujúcich sa v obraze, napr. metóda Peak Signal-to-Noise Ratio-Blocking zameraná na blokové artefakty. A nakoniec na metódy založené na modeloch *ľudského vizuálneho systému* (HVS). V tomto prípade sa zapájajú psychologické a psychofyzické štúdie, napr. metóda Most Apparent Distortion.

Kvôli zisteniu, ktorá metóda je najvhodnejšia na testovanie kvality obrázkov vytvorených pomocou kompresných štandardov JPEG a JPEG 2000 som vykonal celkovo 2 testovania výkonnosti metód. Tieto testovania mali za cieľ zistiť koreláciu výsledkov metód so subjektívnym hodnotením kvality ľuďmi. Pre tieto účely som vybral viacero metód s predpokladom, že dané metódy by mohli dosahovať dobré výsledky a sú vhodné na bežné použitie bez nutnosti rozsiahlych príprav testovania. Všetky vybrané metódy patrili do skupiny full-reference metód. Prvé testovanie sa zameriavalo na meranie kvality obrázkov so širokým rozpätím bitrate a do tohoto testovania boli zaradené metódy: Peak Signal-to-Noise Ratio (PSNR), Peak Signal-to-Noise Ratio - Blocking (PSNR-B), Structural Similarity (SSIM), Universal Quality Index (UQI) a Most Apparent Distortion (MAD). Druhé testovanie sa zameriavalo na meranie výkonnosti metód len na obrázkoch s vyšším bitrate, čo v tomto prípade predstavovalo 0,6 bitov/pixel a vyššie. Do tohoto testovania bola zaradená okrem už spomínaných piatich metód tiež metóda Butteraugli, ktorá sa sústreďuje na testovanie práve takýchto obrázkov a nie na určovanie kvality širokého rozpätia bitrate.

Nasledujúca časť popisuje viaceré existujúce metódy a princípy, podľa ktorých posudzujú kvalitu. Takisto popisuje dôvody, prečo boli vybrané do testovania práve spomínané metódy.

Mean Squared Error

Cieľom merania Mean Squared Error (MSE) je porovnať dva signály a poskytnúť kvantitatívne skóre, ktoré popisuje stupeň chyby/skreslenia medzi nimi. Využíva sa strednej kvadratickej odchýlky, pričom jeden zo signálov je originál a druhý je skreslený, alebo obsahujúci chyby. Výsledkom je vždy kladné číslo a hodnoty bližšie k nule znamenajú väčšiu podobnosť signálov, teda lepšiu kvalitu. Ide o jednu z prvých metód využívaných na meranie kvality obrazu. Nech $x = x_i | i = 1, 2, \dots, N$ a $y = y_i | i = 1, 2, \dots, N$ sú dva diskkrétne

signály konečnej dĺžky – v tomto prípade testovaný a referenčný obrázok, kde N je počet pixelov a x_i a y_i sú hodnoty i -teho pixelu, potom

$$\text{MSE}(x, y) = \frac{1}{N} \sum_{i=0}^N (x_i - y_i)^2. \quad (3.1)$$

Metódu som nezahrnul do testovania, pretože už je v testovaní metóda PSNR, ktorá je na MSE založená. Zároveň samotné MSE neposkytuje [13] hodnotenie, ktoré by sa dostatočne podobalo subjektívnemu hodnoteniu v porovnaní s inými často používanými metódami.

3.3 Peak Signal-to-Noise Ratio

Jednou z najčastejšie používaných metód je metóda PSNR, ktorá je daná pomerom maximálnej možnej sily signálu a šumu. Je vyjadrená logaritmickou decibelovou stupnicou. Pri porovnávaní dvoch vzorkov, kde jeden je originál a druhý zase testovaný obrázok sa PSNR snaží zohľadniť ľudské vnímanie kvality obrazu [5]. Pre výpočet využíva MSE a je definovaná, ako

$$\text{PSNR}(x, y) = 10 \log_{10} \left(\frac{2^{B-1}}{\text{MSE}(x, y)} \right), \quad (3.2)$$

kde B je počet bitov na pixel. Pri 8-bitovej farebnej hĺbke sa typické hodnoty výsledkov pohybujú v rozmedzí 30 až 50 dB a pri 16-bitovej hĺbke sú typicky medzi 60 až 80 dB. Platí, že čím viac dB, tým je obraz kvalitnejší (podobnejší originálu).

Metódu považujem za vhodnú na zahrnutie do testovania, pretože ide o jednu z najznámejších a široko používaných metód pri zisťovaní kvality obrazu [9].

3.4 Peak Signal-to-Noise Ratio – Blocking

Metóda PSNR-B je založená na PSNR so zahrnutím faktora blokového efektu [16]. PSNR-B štandardne predpokladá veľkosť bloku 8×8 , čo predstavuje typickú veľkosť pri JPEG štandarde, pričom tento parameter je možné upraviť. Pre zjednodušenie predpokladajme celočíselný počet blokov tvoriaci obrázok a hranice blokov sa vyskytujú vo vertikálnom alebo horizontálnom smere. Nech N_H a N_V sú horizontálny a vertikálny rozmer obrázku I . Nech \mathcal{H} je množina horizontálnych susediacich párov pixelov. Nech $\mathcal{H}_B \subset \mathcal{H}$ je množina horizontálnych susediacich párov pixelov, ktoré ležia na hranici bloku. Nech \mathcal{H}_B^C je množina horizontálnych susediacich párov pixelov neležiacich na hraniciach blokov, takže platí $\mathcal{H}_B^C = \mathcal{H} - \mathcal{H}_B$. Podobne nech je \mathcal{V} množina vertikálnych susediacich párov pixelov a \mathcal{V}_B množina vertikálnych susediacich párov pixelov ležiacich na hraniciach blokov. A nech \mathcal{V}_B^C je množina vertikálnych susediacich párov pixelov neležiacich na hraniciach blokov, takže platí $\mathcal{V}_B^C = \mathcal{V} - \mathcal{V}_B$.

Ďalej nech N_{H_B} , $N_{H_B^C}$, N_{V_B} a $N_{V_B^C}$ sú počty párov pixelov v \mathcal{H}_B , \mathcal{H}_B^C , \mathcal{V}_B a \mathcal{V}_B^C v uvedenom poradí. Keď B je veľkosť bloku, potom

$$N_{H_B} = N_V \left(\frac{N_H}{B} \right) - 1, \quad (3.3)$$

$$N_{H_B^C} = N_V(N_H - 1) - N_{H_B}, \quad (3.4)$$

$$N_{V_B} = N_H \left(\frac{N_V}{B} \right) - 1, \quad (3.5)$$

$$N_{V_B^C} = N_H(N_V - 1) - N_{V_B}. \quad (3.6)$$

Na obrázku 3.1 je ukázaný príklad, ako vyzerajú bloky pixelov v obrázku s rozmermi $N_H = 8$, $N_V = 8$ a $B = 4$. Hrubé čiary reprezentujú hranice blokov.

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}
y_{17}	y_{18}	y_{19}	y_{20}	y_{21}	y_{22}	y_{23}	y_{24}
y_{25}	y_{26}	y_{27}	y_{28}	y_{29}	y_{30}	y_{31}	y_{32}
y_{33}	y_{34}	y_{35}	y_{36}	y_{37}	y_{38}	y_{39}	y_{40}
y_{41}	y_{42}	y_{43}	y_{44}	y_{45}	y_{46}	y_{47}	y_{48}
y_{49}	y_{50}	y_{51}	y_{52}	y_{53}	y_{54}	y_{55}	y_{56}
y_{57}	y_{58}	y_{59}	y_{60}	y_{61}	y_{62}	y_{63}	y_{64}

Obr. 3.1: Ilustrácia blokov obrázku [16].

Pre tento príklad ďalej platí, že $N_{H_B} = 8$, $N_{H_B^C} = 48$, $N_{V_B} = 8$ a $N_{V_B^C} = 48$. Množiny majú v tomto prípade tvar

$$\mathcal{H}_B = \{(y_4, y_5), (y_{12}, y_{13}), \dots, (y_{60}, y_{61})\}, \quad (3.7)$$

$$\mathcal{H}_B^C = \{(y_1, y_2), (y_2, y_3), (y_3, y_4), (y_5, y_6), \dots, (y_{63}, y_{64})\}, \quad (3.8)$$

$$\mathcal{V}_B = \{(y_{25}, y_{33}), (y_{26}, y_{34}), \dots, (y_{32}, y_{40})\}, \quad (3.9)$$

$$\mathcal{V}_B^C = \{(y_1, y_9), (y_9, y_{17}), (y_{17}, y_{25}), (y_{33}, y_{41}), \dots, (y_{56}, y_{64})\}. \quad (3.10)$$

Ďalej je definovaný stredný kvadratický rozdiel hraničných pixelov D_B a stredný kvadratický rozdiel nehraničných pixelov D_B^C pre obrázok \mathbf{y}

$$D_B(\mathbf{y}) = \frac{\sum_{(y_i, y_j) \in \mathcal{H}_B} (y_i - y_j)^2 + \sum_{(y_i, y_j) \in \mathcal{V}_B} (y_i - y_j)^2}{N_{H_B} + N_{V_B}}, \quad (3.11)$$

$$D_B^C(\mathbf{y}) = \frac{\sum_{(y_i, y_j) \in \mathcal{H}_B^C} (y_i - y_j)^2 + \sum_{(y_i, y_j) \in \mathcal{V}_B^C} (y_i - y_j)^2}{N_{H_B^C} + N_{V_B^C}}. \quad (3.12)$$

Na týchto dvoch kvadratických rozdieloch je založený výpočet blokového faktoru

$$\text{BEF}(\mathbf{y}) = \eta \cdot [D_B(\mathbf{y}) - D_B^C(\mathbf{y})]. \quad (3.13)$$

Tu sa predpokladá, že viditeľnosť blokového efektu rastie s veľkosťou bloku. Z blokového efektu je následne odvodené

$$\text{MSE-B}(x, y) = \text{MSE}(x, y) + \text{BEF}(\mathbf{y}), \quad (3.14)$$

ktoré zohľadňuje na rozdiel od klasického MSE aj blokový efekt. Z tohoto MSE-B je v závere vypočítaná finálna hodnota kvality testovaného obrázku

$$\text{PSNR-B}(x, y) = 10 \log_{10} \frac{255^2}{\text{MSE-B}(x, y)}. \quad (3.15)$$

Metódu som vybral k porovnaniu, pretože sa zameriava na detekciu blokových artefaktov, ktoré môžu vzniknúť pri použití oboch kompresných štandardov JPEG aj JPEG 2000 a tiež kvôli predpokladu, že by mohla ukázať lepšie výsledky v porovnaní s PSNR.

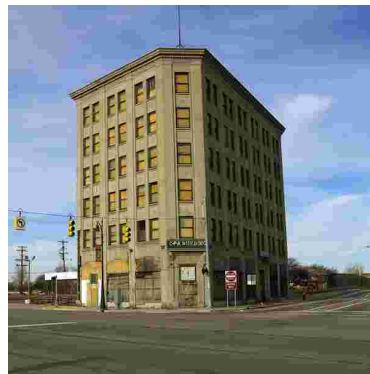
3.5 Structural Similarity

Metóda SSIM je založená na pozorovaní, že pre ľudské vnímanie degradácia obrazu znamená zmenu v jeho štruktúrálnych informáciach [13]. Zahŕňa tiež vnemové javy ako maskovanie jasom a kontrastom. Maskovaním jasom je myslený jav, pri ktorom sú obrazové artefakty menej viditeľné v oblastiach s vysokým jasom a maskovaním kontrastom fakt, že deformácia je menej viditeľná na miestach obrazu s výraznou aktivitou – textúrou. Výsledok sa počíta buď pre každú farebnú zložku obrazu, alebo len pre jasovú zložku.

Výsledkom SSIM môže byť buď číslo popisujúce kvalitu testovaného obrázku, alebo tiež indexová mapa, na ktorej je možné vidieť miesta s rozdielmi medzi dvomi obrázkami. Obrázok 3.2 ukazuje fotografiu, ktorá je podrobená JPEG kompresii. Výsledok tejto kompresie je zobrazený na obrázku 3.3, na ktorom je možné vidieť blokové artefakty, najmä v miestach s oblohou. Posledný obrázok 3.4 zobrazuje indexovú mapu vytvorenú metódou SSIM, ktorá zachytáva rozdiely medzi originálnym a testovaným obrázkom.



Obr. 3.2: Pred kompresiou.



Obr. 3.3: Po kompresii.



Obr. 3.4: Indexová mapa.

Metóda predpokladá, že medzi pixelmi vo vzájomnej blízkosti existuje silná závislosť. Táto závislosť obsahuje dôležité informácie o štruktúre a objektoch scény. SSIM vzorec je založený na troch porovnávacích meraniach vzorkov x a y . Meria jas, kontrast a štruktúru

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}, \quad (3.16)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \quad (3.17)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}, \quad (3.18)$$

kde $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$, $c_3 = c_2/2$. Predvolené hodnoty pre k sú $k_1 = 0.01$ a $k_2 = 0.03$. L je dynamický rozsah hodnoty pixelu, μ_x je stredná intenzita a σ_x je štandardná odchýlka vzorku x . Celkový výsledok metódy zložený z váženej kombinácie týchto meraní je daný

$$\text{SSIM}(x, y) = [l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma]. \quad (3.19)$$

Metódu som vybral na porovnávanie, pretože ukázala dobré výsledky [13] pri hodnotení kvality obrazov poškodených stratovou kompresiou a dokáže hodnotiť kvalitu farebných obrázkov, teda bez nutnosti ich prevodu na odtiene šedej.

Multiscale Structural Similarity

Upravená forma SSIM, ktorá poskytuje väčšiu flexibilitu ako metódy s jednou mierkou, zahŕňa tiež rozlíšenie obrazu a rôzne varianty podmienok pri zobrazení obrazu [14]. Túto metódu som na testovanie nevybral, lebo je už vybraná metóda SSIM.

3.6 Universal Quality Index

Objektívny ukazateľ kvality UQI, ktorý modeluje každé skreslenie obrazu ako kombináciu troch faktorov: strata korelácie, skreslenie jasu a skreslenie kontrastu [12]. Má silnú schopnosť merať stratu informácií počas degradačného procesu obrazu, ako je napríklad kompresia. Nech $x = x_i | i = 1, 2, \dots, N$ a $y = y_i | i = 1, 2, \dots, N$ sú signály originálu a testovaného obrazu v uvedenom poradí, potom

$$\text{UQI} = \frac{\sigma_{xy}}{\sigma_x\sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2}, \quad (3.20)$$

kde σ_x^2 je odchýlka pre x , σ_y^2 odchýlka pre y a σ_{xy} súčiniteľ vzájomnej korelácie x a y . Prvý komponent je korelačný koeficient, ktorý meria stupeň lineárnej korelácie medzi x a y . Aj keby sú x a y lineárne príbuzné, stále medzi nimi môžu existovať relatívne skreslenia, ktoré sú vyhodnotené druhým a tretím komponentom. Druhý komponent meria podobnosť jasu obrazov x a y a tretí komponent podobnosť kontrastu obrazov.

Dynamický rozsah metódy UQI je $[-1, 1]$. Najlepšia hodnota 1 je dosiahnutá len v prípade, že sa signály rovnajú. Najnižšia hodnota -1 je dosiahnutá, pokiaľ $y_i = 2\bar{x} - x_i$ pre všetky $i = 1, 2, \dots, N$. Metóda je aplikovaná na lokálne oblasti obrazu s použitím prístupu posuvného okna. Začína sa v ľavom hornom rohu a toto posuvné okno o veľkosti $B \times B$ sa posúva pixel po pixeli horizontálne a vertikálne cez všetky riadky a stĺpce obrazu, až kým je dosiahnutý pravý dolný roh. Pri i -tom kroku je vypočítaný index kvality UQI_i v rámci posuvného okna. Ak je dokopy M krokov, potom je celkový index kvality

$$\text{UQI} = \frac{1}{M} \sum_{i=1}^M \text{UQI}_i. \quad (3.21)$$

Metódu považujem za vhodnú na porovnanie, pretože nie je výpočetne náročná, ďalej podľa autorov ukázali ich testy konzistentnú koreláciu so subjektívnym hodnotením kvality obrázkov človekom.

3.7 Most Apparent Distortion

Metóda MAD predpokladá, že HVS zapája pri posudzovaní kvality obrázkov dve rôzne stratégie [6]. Pokiaľ HVS vidí obrázky obsahujúce skreslenia na hrane pozorovateľnosti, snaží sa pozeráť sa cez predmet obrázku a nájsť skreslenia. Táto stratégia sa nazýva *detection-based* stratégia a pre odhad skreslení používa maskovanie jasom a kontrastom. Keď sa HVS snaží vidieť obrázky obsahujúce jasne viditeľné skreslenia, pokúsi sa pozrieť cez skreslenia, hľadajúc predmet obrázku. Táto druhá stratégia nesie názov *appearance-based* a využíva rozdiely v lokálnych štatistikách priestorových frekvenčných komponentov.

Detection-based stratégia

Stratégia využíva poznatok, že keď HVS vidí vysoko kvalitný obraz, snaží sa pozeráť cez predmet obrázku hľadajúc skreslenia. Skladá sa z dvoch fáz: zistenie umiestnení viditeľných skreslení a výpočet vnímaného skreslenia. Aby MAD algoritmus popísal nelineárny vzťah medzi hodnotami pixelov a fyzickým jasom displeja zariadenia, transformuje pixely referenčného a testovaného obrázku na hodnoty jasú

$$L = (\beta + \alpha I)^\gamma, \quad (3.22)$$

kde L je jas obrázku, I je referenčný alebo testovaný obrázok, β , α a γ sú konštanty špecifické pre zariadenie. Používajú sa konštanty 0, 0,02874 a 2,2 v uvedenom poradí, ktoré boli vypočítané s použitím 8 bitových hodnôt pixelov a sRGB displeja. Pomocou (3.22) sú vypočítané hodnoty L_{ref} a L_{test} pri aplikovaní I_{ref} a I_{test} . Získané hodnoty L sú následne umocnené na $\frac{1}{3}$, čo zohľadňuje nelineárnu odpoveď HVS na jas, čím vzniknú nové hodnoty \hat{L}_{ref} a \hat{L}_{test} nazývané vnímaný jas. Následne je vypočítaný chybový obrázok

$$\hat{L}_{err} = \hat{L}_{ref} - \hat{L}_{test}. \quad (3.23)$$

Pre popis variácií v citlivosti autori metódy vytvorili funkciu CSF (contrast sensitivity function), ktorá je aplikovaná na referenčný a chybový obrázok, čím sú získané I'_{ref} a I'_{err} . Keďže obsah obrázku môže ovplyvniť detekciu skreslení, používa sa ďalej meranie priestorovej domény maskovania kontrastu. Je vypočítaný kontrast $C_{ref}(b)$ a $C_{err}(b)$ pre každý blok b o veľkosti 16×16 obrázkov I'_{ref} a I'_{err} . Po výpočte kontrastov pre referenčný a chybový obrázok je vypočítaná mapa lokálnych viditeľných skreslení $\xi(b)$ z týchto kontrastov. Z tejto mapy je vypočítaná hlavná hodnota z detection-based stratégie

$$d_{detect} = \left(\frac{1}{B} \sum_b [\xi(b) \times \text{MSE}(b)]^2 \right)^{\frac{1}{2}}, \quad (3.24)$$

kde B je celkový počet blokov.

Appearance-based stratégia

Táto stratégia predpokladá, že keď HVS vidí obraz s nízkou kvalitou, snaží sa prejsť cez skreslenia a hľadá obsah obrázku. Táto stratégia pozostáva z dvoch fáz: výpočet lokálnej štatistickej rozdielovej mapy a dekompozícia referenčného a testovaného obrázku pomocou log-Gabor filtra

$$\mathbf{G}_{s,o}(\omega, \theta) = \exp \left[-\frac{\log \left(\frac{\omega}{\omega_s} \right)}{\sqrt{2}\sigma_r} \right] \times \exp \left[-\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2} \right], \quad (3.25)$$

kde s je priestorová škála a o orientácia. Parametre ω a θ sú normalizovaná radiálna frekvencia a orientácia v uvedenom poradí. Ďalej ω_s je normalizovaná stredná frekvencia, σ_r kontroluje šírku pásma filtra, θ_0 a σ_θ sú stredná orientácia a uhlový rozptyl filtra. Keď sa pre log-Gabor dekompozíciu napríklad použije 5 škál a 4 orientácie, výsledkom je 20 subpásiev na obrázok. V ďalšom kroku je vygenerovaná lokálna štatistická mapa rozdielov $\eta(b)$. Je získaná porovnaním lokálnych subpásmových štatistík referenčného a testovaného obrázku. Pre každý blok b o veľkosti 16×16 je

$$\eta(b) = \sum_{s=1}^5 \sum_{o=1}^4 \ell[|\sigma_{s,o}^{ref}(b) - \sigma_{s,o}^{test}(b)| + 2|\zeta_{s,o}^{ref}(b) - \zeta_{s,o}^{test}(b)| + |\kappa_{s,o}^{ref}(b) - \kappa_{s,o}^{test}(b)|], \quad (3.26)$$

kde $\sigma_{s,o}(b)$, $\zeta_{s,o}(b)$ a $\kappa_{s,o}(b)$ sú štandardná odchýlka, šikmosť a špicatosť subpásma 16×16 so škálou s , orientáciou o a blokom b . Po výpočte $\eta(b)$, je vypočítaná konečná skalárna hodnota vnímaného skreslenia

$$d_{appear} = \left(\frac{1}{B} \sum_b \eta^2(b) \right)^{\frac{1}{2}}. \quad (3.27)$$

Výsledok sa nachádza v intervale $[0, \infty)$ a je rovný 0 v prípade, že v testovanom obrázku nie je žiadne vnímané skreslenie. So stúpajúcou hodnotou d_{appear} stúpa tiež vnímané skreslenie a znižuje sa kvalita obrázku. Na záver je vypočítaná celková hodnota MAD s využitím výsledkov oboch stratégií

$$\text{MAD} = (d_{detect})^\alpha (d_{appear})^{1-\alpha}, \quad (3.28)$$

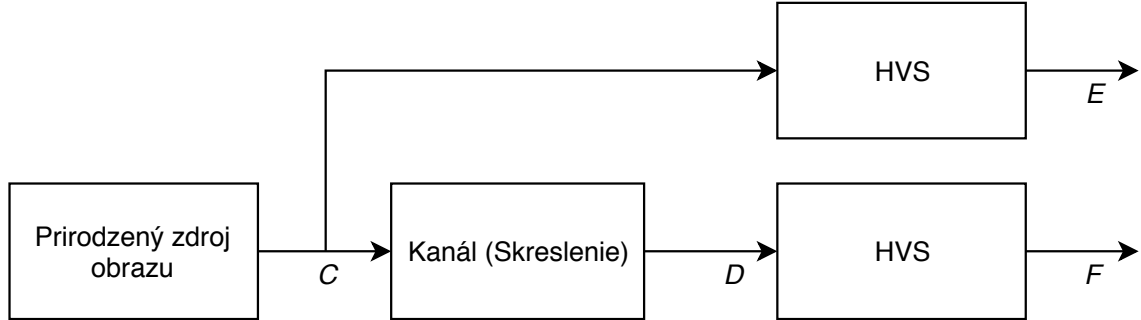
kde α je konštanta vyjadrujúca váhu a indikuje relatívnu dôležitosť každej zložky.

Metódu som vybral na porovnanie, pretože dokáže vyhodnotiť kvalitu farebných obrázkov, podľa autorov ukázala dobrú koreláciu so subjektívnym hodnotením, nie je potrebný konkrétny rozmer testovaných obrázkov (napríklad rovnaká šírka a výška) a metóda výrazne berie ohľad na HVS.

Visual Information Fidelity

Metóda Visual Information Fidelity (VIF) posudzuje kvalitu poškodeného obrázku jeho porovnaním s pôvodnou podobou. Využíva štatistiky prirodzených scén (NSS) a zohľadňuje spôsob extrahovania obrazových informácií ľudským vnímaním [9]. Prirodzenými scénami sú myslené obrazy viditeľného prostredia zachytené použitím vysoko kvalitných zariadení, operujúcich v človekom viditeľnom spektre. Tieto scény tvoria v priestore všetkých možných signálov malý podpriestor. Väčšina reálne sa vyskytujúcich skreslujúcich procesov narúša

štatistiky scén a spôsobuje, že obrazové signály pôsobia neprirodzené. Prirodzené obrázky bezchybnej kvality môžu byť modelované ako výstup stochastického zdroja. Pokiaľ chýbajú akékoľvek skreslenia, signál prejde cez HVS kanál pred vstupom do mozgu pozorovateľa, čo je znázornené na obrázku 3.5.



Obr. 3.5: Spoločné informácie medzi C a E kvantifikujú informácie, ktoré by mohol mozog extrahovať v ideálnom prípade z referenčného obrázku. Spoločné informácie medzi C a F kvantifikujú informácie, ktoré môžu byť extrahované z testovaného obrázku.

Metóda VIF je odvodená z kvantifikácie dvoch informačných množstiev – spoločná informácia medzi vstupom a výstupom kanála HVS bez prítomnosti skresľujúceho kanálu a spoločná informácia medzi vstupom skresľujúceho kanálu a výstupom HVS kanálu pre testovaný obrázok. Prirodzené obrázky sú modelované vo vlnkovej doméne s použitím *Gaussian mixture* modelu (GSM) [11]. Na ich modelovanie sa ukázala byť užitočnou vlnková analýza (alebo tiež *scale-spaced-orientation* analýza). Každé subpásmo vlnkovej dekompozície sa modeluje ako náhodné pole GSM. Koefficienty každého subpásma sú rozdelené do neprelínajúcich sa blokov a každý blok i sa modeluje ako vektor \vec{C}_i .

$I(\vec{C}^N; \vec{E}^N | s^N)$ predstavuje informácie o referenčnom obrázku, pričom vizuálna kvalita sa vzťahuje na množstvo týchto informácií, ktoré môže mozog extrahovať z referenčného obrázku. Pokiaľ je možné extrahovať z testovaného obrázku napríklad 2 bity/pixel a z príslušného referenčného obrázku je to 2,1 bitov/pixel, môže mozog obnoviť väčšinu obsahu referenčného obrázku z toho testovaného. Na druhej strane, pokiaľ by bolo možné z referenčného obrázku extrahovať až 5 bitov/pixel, potom by boli stratené 3 bity/pixel v skresľujúcom kanáli a vizuálna kvalita testovaného obrázku by bola podradná. Ďalej $I(\vec{C}^N; \vec{F}^N | s^N)$ predstavuje informácie, ktoré by mohli byť v ideálnom prípade mozgom extrahované z testovaného obrázku. Metóda VIF je definovaná ako pomer týchto informácií pre jednotlivé subpásma vlnkovej dekompozície

$$\text{VIF} = \frac{\sum_{j \in \text{subbands}} I(\vec{C}^{N,j}; \vec{F}^{N,j} | s^{N,j})}{\sum_{j \in \text{subbands}} I(\vec{C}^{N,j}; \vec{E}^{N,j} | s^{N,j})}. \quad (3.29)$$

Metódu som nevybral na porovnanie z toho dôvodu, že pokiaľ sa v testovanom obrázku nachádzajú iné skreslenia než biely šum a blurring, napríklad artefakty typické pre JPEG, metóda nedokáže dostatočne reprodukovat túto skutočnosť [9]. Ďalej je potrebné, aby bol testovaný obrázok v odtieňoch šedej, čo zbytočne zvyšuje množstvo krokov potrebných pri vykonávaní testovania.

3.8 Butteraugli

Metóda Butteraugli od spoločnosti Google je založená na odhade psychovizuálnej podobnosti dvoch obrazov [3]. Je využívaná pri JPEG kodeku Guertzli, ktorý dosahuje 29% až 45% redukciu vo veľkosti súboru pre danú vzdialenosť vnímania [1]. Produkuje vizuálne nerozlíšiteľné obrázky v porovnaní s pôvodným obrázkom s nižším bitrate oproti ostatným kodekom. Táto metóda hodnotí, aká kvalita výsledného komprimovaného obrázku je ešte prijateľná a ktorá už nie. Výsledkom môže byť skalárna hodnota alebo priesorová mapa úrovni rozdielov medzi obrázkami. Najlepšiu účinnosť metóda údajne dosahuje pri posudzovaní obrázkov s nastaveným parametrom kvality v rozsahu 90 až 95 pri štandarde JPEG. Jej výkonnosť v iných situáciách nie je udávaná.

Metódu som vybral až do druhého testovania, ktoré bude uskutočnené len na vyšších hodnotách bitrate, aby bolo možné určiť výkonnosť vybraných metód len na vyššom bitrate a zároveň zahrnúť do porovnávania aj túto metódu, pretože sa zameriava len na určovanie kvality obrázkov s vyšším bitrate.

Kapitola 4

Testovanie výkonnosti metód

4.1 Dataset

Pre účely testovania s cieľom zistiť, výsledky ktorej metódy budú najviac odpovedať subjektívnemu hodnoteniu kvality obrázkov som vytvoril dataset pozostávajúci z 10 rôznych fotografií reálnych scén. Náhlady týchto fotografií sú v obrázku 4.1. Obrázky boli vybrané z celkového počtu 30 z LIVE datasetu [8]. Bitrate týchto fotografií v plnej kvalite sa pohyboval približne na úrovni 8 bitov/pixel.



Obr. 4.1: Obrázky datasetu.

Bitrate obrázkov

Z každého obrázku som vytvoril viacero obrázkov s rôznymi úrovňami kvality pomocou oboch kompresných štandardov – 10 úrovní kvality pre každý obrázok s použitím JPEG a 10 úrovní s použitím JPEG 2000, čo je dohromady 200 obrázkov. Keďže pri vyšších bitrate sa pozorovateľná kvalita obrazu menila len mierne, zvolil som logaritmické rozloženie bitrate obrázkov. Ako dolnú hranicu som zvolil približne 0,15 bitov/pixel a ako hornú hranicu 3,5 bitov/pixel. Horná hranica bola zvolená nižšie ako spomínaných 8 možných bitov/pixel, pretože voľným okom sú obrázky s takýmto vysokým bitrate nerozlišiteľné. K vypočítaniu logaritmického rozloženia bitrate v konkrétnom rozsahu bol využitý nástroj Wolfram Mathematica¹ a príkaz `10^Array[# &, 10, Log10@x,y]` vždy s príslušnou dolnou a hornou hodnotou pre konkrétny obrázok na mieste x a y , ktoré boli získané po komprimácii s využitím nástroja `cjpeg` na všetky úrovne kvality, čo bude detailne popísané v nasledujúcej podsekcii.

¹<https://www.wolfram.com/mathematica/>

Pre jednotlivé obrázky bol použitý rôzny bitrate, ktorý je podrobne uvedený v tabuľke 4.1.

Obr.	Bitrate									
1.	0,152	0,222	0,307	0,427	0,601	0,874	1,223	1,744	2,467	3,476
2.	0,159	0,222	0,307	0,427	0,601	0,908	1,255	1,778	2,481	3,529
3.	0,152	0,203	0,293	0,438	0,616	0,876	1,260	1,780	2,509	3,609
4.	0,145	0,203	0,291	0,410	0,603	0,847	1,193	1,719	2,476	3,440
5.	0,166	0,247	0,340	0,452	0,635	0,901	1,244	1,754	2,466	3,416
6.	0,151	0,213	0,306	0,429	0,602	0,844	1,170	1,703	2,411	3,348
7.	0,147	0,213	0,299	0,430	0,607	0,855	1,221	1,740	2,500	3,552
8.	0,145	0,202	0,290	0,407	0,573	0,802	1,135	1,554	2,135	3,172
9.	0,145	0,198	0,300	0,401	0,580	0,827	1,152	1,635	2,289	3,275
10.	0,147	0,207	0,290	0,418	0,614	0,867	1,221	1,732	2,455	3,498

Tabuľka 4.1: Bitrate obrázkov datasetu.

Obrázky JPEG a JPEG 2000

Pri vytváraní obrázkov s pomocou kompresného štandardu JPEG nie je možné vygenerovať obrázok s nastavením konkrétneho bitrate pomocou štandardných nástrojov, z tohoto dôvodu boli vygenerované všetky obrázky s nastavením parametru `-quality` na celočíselné hodnoty od 1 do 100 a s využitím nástroja `cjpeg`. Následne skript napísaný v jazyku Bash vypočítal bitrate pre každý z týchto obrázkov. Pre dolnú hranicu bol zvolený obrázok, ktorého bitrate sa najviac približoval stanovenej dolnej hranici 0,15 bitov/pixel a podobným spôsobom obrázok pre hornú hranicu stanovenú na 3,5 bitov/pixel. Pre každý z 10 originálnych obrázkov datasetu tak boli použité iné hodnoty pre výpočet logaritmickeho rozloženia. Po vypočítaní rozloženia som hľadal zvyšných 8 obrázkov s najbližšou hodnotou bitrate, opakovane pre každý z 10 pôvodných obrázkov. Štandard JPEG 2000 podporuje nastavenie vlastného bitrate. Za týmto účelom bol využitý nástroj `kakadu`² a potrebný výstupný bitrate bol zabezpečený nastavením parametru `-rate`. Tento parameter bol nastavený pre jednotlivé obrázky na rovnaký bitrate, aký vznikol pri komprimácii pomocou `cjpeg` a ktorý sa najviac blížil tomu pôvodnému v logaritmickeho rozsahu približne 0,15 až 3,5 bitov/pixel.

Po vytvorení všetkých 200 obrázkov pre testovanie ich skript napísaný v jazyku Bash prekonvertoval na bezstratový formát BMP. Všetky obrázky mali tak jednotný formát, v ktorom boli ďalej použité na testovanie objektívnej kvality a mohli byť tiež jednoducho prekonvertované na formát PNG pre účely webovej aplikácie uvedenej v nasledujúcej sekcii.

4.2 Subjektívne hodnotenie datasetu

Subjektívne hodnotenie kvality všetkých obrázkov datasetu vytvorených v predchádzajúcom kroku bolo od ľudí získané s využitím webovej aplikácie. Užívatelia hodnotili kvalitu jednotlivých obrázkov v stupnici MOS.

²<http://kakadusoftware.com/software/>

Webová aplikácia

Vytvoril som webovú aplikáciu s využitím jazykov HTML, JavaScript, PHP a SQL. Hodnotiacemu sa po príchode zobrazila úvodná správa zobrazená na obrázku 4.2, ktorá vysvetľovala ako bude hodnotenie prebiehať a tiež aké možnosti môže pri hodnotení využiť.

Hodnotenie kvality obrazu

Pre hodnotenie slúžia tlačidlá:

1 2 3 4 5

Ohodnotiť môžete Vami zvolený počet obrázkov.

Výsledky hodnotenia sa ukladajú priebežne, hodnotenie môžete prerušiť v ľubovoľnej chvíli.

Testované obrázky tvorí 10 fotografií rôznej kvality, dokopy 200 obrázkov.

Kvalitu zobrazeného obrázku ohodnoťte podľa Vášho subjektívneho dojmu.

Vyberte jednu z piatich možností, ktorá reprezentuje kvalitu obrázku na úrovni:

- 5: vynikajúca
- 4: dobrá
- 3: stredná
- 2: slabá
- 1: zlá

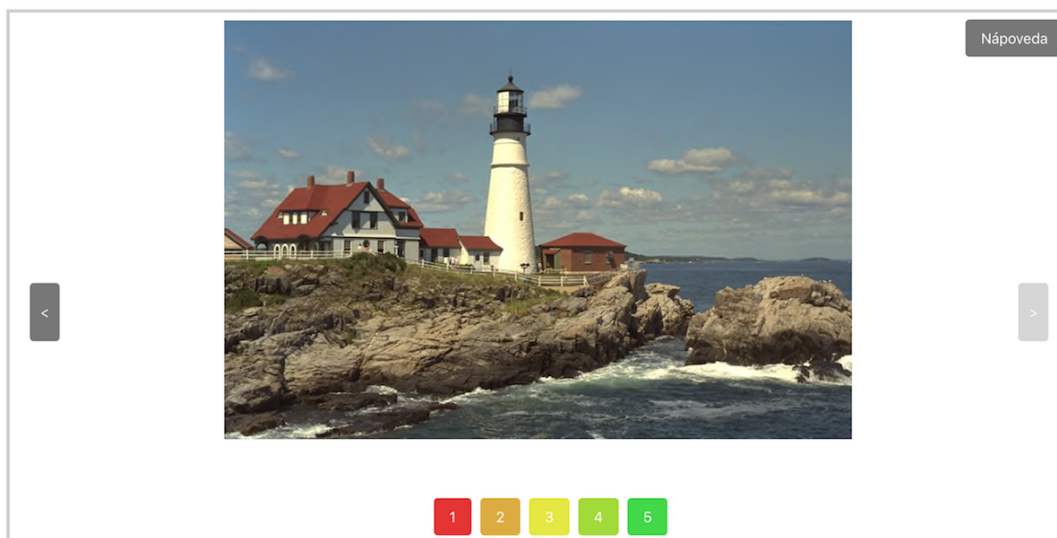
Prejsť k hodnoteniu

Obr. 4.2: Úvodná správa.

Pred reálnym hodnotením sa hodnotiacemu zobrazili 3 testovacie obrázky, na ktorých si mohol vyskúšať proces hodnotenia a nadobudnúť predstavu o kvalite testovaných obrázkov. Hodnotiaci nebol upozornený, že sa jedná o testovacie obrázky.

Aplikácia podporovala vrátenie sa na už ohodnotenú obrázku, čo bolo možné využiť napríklad, pokiaľ si hodnotiaci rozmyslel svoje ohodnotenie predošlého obrázku po zhliadnutí toho nasledujúceho. Táto navigácia bola umožnená s využitím navigačných tlačidiel umiestnených na bokoch stránky ale takisto s využitím šípok na klávesnici vľavo a vpravo. Nebolo umožnené prejsť na ešte neohodnotenú obrázku. Hodnotiaci si mohol zobrazit nároveň v prípade, že si potreboval pripomenúť možnosti, ktoré proces hodnotenia ponúkal alebo význam hodnotiacej stupnice MOS.

Na obrázku 4.3 je vzhľad aplikácie, kde je možné vidieť rozmiestnenie týchto prvkov. Navigačné tlačidlo pre prechod na nasledujúci obrázok je v tomto prípade zakázané.



Obr. 4.3: Webová aplikácia.

Hodnotiaci mohol hodnotenie prerušiť v ľubovoľnej chvíli, pretože výsledky boli priebežne ukladané do databázy. Obrázkov bolo veľké množstvo, takže sa nepredpokladalo, že každý bude hodnotiť všetky obrázky. Ak mal niekto záujem ohodnotiť všetky obrázky, no bola by to pre dotyčného časovo alebo psychicky náročná úloha, mohol stránku zavrieť a vrátiť sa k hodnoteniu neskôr. V tomto prípade webová aplikácia po opätovnom otvorení nastavila hodnotenie na obrázok nasledujúci za posledným ohodnoteným. Hodnoty, ktoré hodnotiaci udelil jednotlivým obrázkom a index posledného ohodnoteného obrázku boli ukladané do lokálneho úložiska (local storage) prehliadača a boli použité pri opätovnom otvorení webovej aplikácie, kedy hodnotiaci pokračoval tam, kde naposledy skončil.

Obrázky sa hodnotiacemu zobrazovali v pomiešanom poradí. Keďže bolo potrebné získať približne rovnaký počet ohodnotení každého obrázku, obyčajné náhodné poradie obrázkov by pre účely tohoto testovania nestačilo, nakoľko sa počítalo s variantou, že hodnotiaci neohodnotí všetkých 200 obrázkov. Z tohoto dôvodu boli pred vygenerovaním postupnosti obrázkov stiahnuté dáta z databázy, ktoré obsahovali počet ohodnotení jednotlivých obrázkov. Na základe týchto údajov bolo 200 indexov rozdelených do intervalov podľa počtu ohodnotení. Napríklad 120 obrázkov malo 15 ohodnotení a zvyšných 80 obrázkov malo už 16 ohodnotení, tým pádom vznikli 2 intervaly. Tieto boli zvlášť náhodne pomiešané a následne spojené do jednej postupnosti indexov, pričom poradie intervalov bolo od najmenšieho počtu ohodnotení po najväčší.

imageName	compressionType	bitrate	q1	q2	q3	q4	q5
image1	jpe	1	20	1	0	0	0
image1	jpe	2	15	4	0	2	0
image1	jpe	3	6	12	1	1	1
image1	jpe	4	1	11	8	1	0
image1	jpe	5	1	4	11	3	2
image1	jpe	6	1	3	7	9	1
image1	jpe	7	1	2	2	12	4
image1	jpe	8	2	1	2	11	5
image1	jpe	9	0	3	2	10	6
image1	jpe	10	0	1	3	3	14
image1	jp2	1	6	10	4	1	0
image1	jp2	2	1	11	9	0	0

Obr. 4.4: Záznamy v databáze hodnotení.

Ohodnotenia boli ukladané do databázy vytvorenej s využitím jazyka SQL. Táto databáza obsahovala jednu tabuľku `imageRecord` s položkami: meno obrázku, napríklad `image4` pre obrázok 4, ďalej typ kompresného štandardu – buď `jpe` pre JPEG alebo `jp2` pre JPEG 2000. Potom bitrate, ktorý nadobúdal pre zjednodušenie celočíselné hodnoty od 1 do 10 a predstavoval logaritmické rozloženie bitrate pre jednotlivé pôvodné obrázky datasetu. A nakoniec pri každom zázname bolo 5 položiek s názvom `q1` až `q5` pre 5 stupňov kvality podľa MOS stupnice. Záznamy v databázi sú zobrazené na obrázku 4.4.

Pri hodnotení vo webovej aplikácii sa po každom kliku na jedno z 5 tlačidiel odoslala požiadavka na server pre uloženie hlasu. Telo požiadavky malo tvar napríklad – `jpe image1 1 4 0`, kde ako oddelovač slúžil znak medzery. Prvý reťazec značí typ kompresného štandardu, druhý reťazec značí poradie obrázku v datasete, prvé z troch čísel predstavuje bitrate z celkového počtu 10 rôznych, druhé číslo predstavuje udelené hodnotenie a posledné číslo predstavuje predchádzajúce udelené hodnotenie. Ako je možné vidieť v tomto prípade, posledné číslo je rovné 0 čo značí, že obrázok predtým nebol ohodnotený. Ak by obrázok už bol týmto užívateľom ohodnotený a užívateľ si rozmyslel svoje ohodnotenie a klikol na inú hodnotu, server by na základe tohoto údaju vedel, z ktorého hodnotenia ubrať jeden hlas pred pričítaním toho nového.

Na serveri boli uložené dva skripty napísané v jazyku PHP. Tento jazyk umožňuje spúšťať SQL dotazy, čo je potrebné pre prácu s databázou na strane serveru. Prvý skript spracovával požiadavky na uloženie hodnotení. Po príchode požiadavky príchodziu správu rozdelil na miestach so znakmi medzery. Následne vytvoril SQL dotaz s príkazom `UPDATE` s hodnotami získanými v predchádzajúcom kroku. Ak nešlo o prvé ohodnotenie, pred samotnou inkrementáciou hodnoty na mieste nového hodnotenia bola dekrementovaná hodnota na mieste predchádzajúceho hodnotenia. Druhý PHP skript bol používaný pri načítavaní štatistík, ktoré budú popísané v ďalšej podsekcii. Skript načítal všetky dáta z databázy s SQL príkazom `SELECT * FROM imageRecord`. Tieto dáta následne rozparsoval a vytvoril z nich objekt vo formáte JSON a uložil ho do textového súboru, z ktorého si webová aplikácia so štatistikami načítala údaje pre grafy.

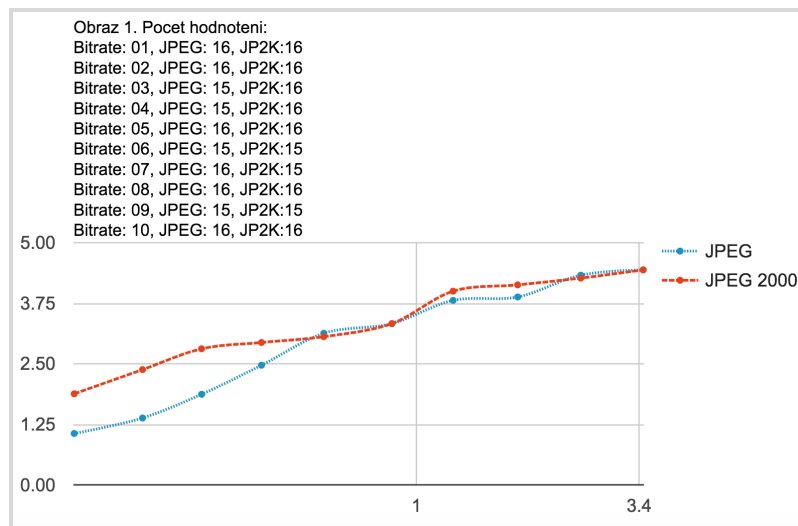
Webová aplikácia zároveň ošetrovala prípad, kedy by hodnotiaci otvoril aplikáciu v malom okne, kde by nemohol zreteľne vidieť zobrazený obrázok so všetkými detailmi a

správne ohodnotiť jeho kvalitu. Túto hranicu som stanovil na 850 pixelov pre šírku a 450 pixelov pre výšku okna prehliadača. Zároveň, pokiaľ užívateľ pristúpil k aplikácii cez mobilné zariadenie, ktoré by malo s vysokou pravdepodobnosťou displej s menším rozmerom než minimálnym potrebným, bolo zobrazené upozornenie vysvetľujúce dopad použitia malého displeja a výzva k použitiu prehliadača na počítači a nebolo umožnené hodnotiť.

Navigácia vrámci obrázkov dopredu bola zabezpečená automaticky po kliknutí na niektoré z 5 tlačidiel stupnice MOS. Obrázky sa automaticky v predstihu jeden obrázok načítavali zo serveru do cache prehliadača, aby sa predišlo čakaniu na nový obrázok po ohodnotení. Všetky obrázky datasetu boli vrámci webovej aplikácie používané vo formáte PNG, ktorý je vrámci internetu široko využívaný. Dôvodom je najmä nízka veľkosť výsledných súborov a ide o bezstratový formát, takže všetky artefakty vzniknuté stratovou kompresiou zostali zachované bez zmien. Rozlíšenie väčšiny obrázkov datasetu sa pohybovalo okolo hodnôt 768x512 pixelov, čomu bola prispôbená veľkosť obrázku vrámci stránky. Išlo o dosiahnutie vhodného zobrazenia na štandardnom monitore s rozlíšením Full-HD tak, aby boli zreteľne viditeľné detaily obrázkov. Po ohodnotení posledného obrázku sa zobrazila záverečná správa, pričom hodnotiaci už nemal možnosť sa k hodnoteniu vrátiť.

Aktuálny prehľad ohodnotení

Kvôli jednoduchému zisteniu aktuálneho počtu ohodnotení jednotlivých obrázkov a zobrazeniu výsledkov ohodnotenia som vytvoril stránku zobrazujúcu aktuálny stav výsledkov. Stránka využívala Google Charts³, pričom aktuálne ohodnotenia v stupnici MOS boli získané z databázy zo servera v dobe načítania stránky. Pre každý z pôvodných 10 obrázkov datasetu bol zobrazený graf s poradovým číslom obrázku, ďalej počet ohodnotení jednotlivých obrázkov vytvorených z pôvodného obrázku zvlášť pre oba kompresné štandardy JPEG a JPEG 2000 a pre jednotlivé bitrate hodnoty. Tieto hodnoty tu majú číselné označenie od 1 do 10, no predstavujú logaritmické rozloženie z tabuľky 4.1. Časť zo štatistík je zobrazená na obrázku 4.5, týkajúca sa prvého z 10 pôvodných obrázkov datasetu.

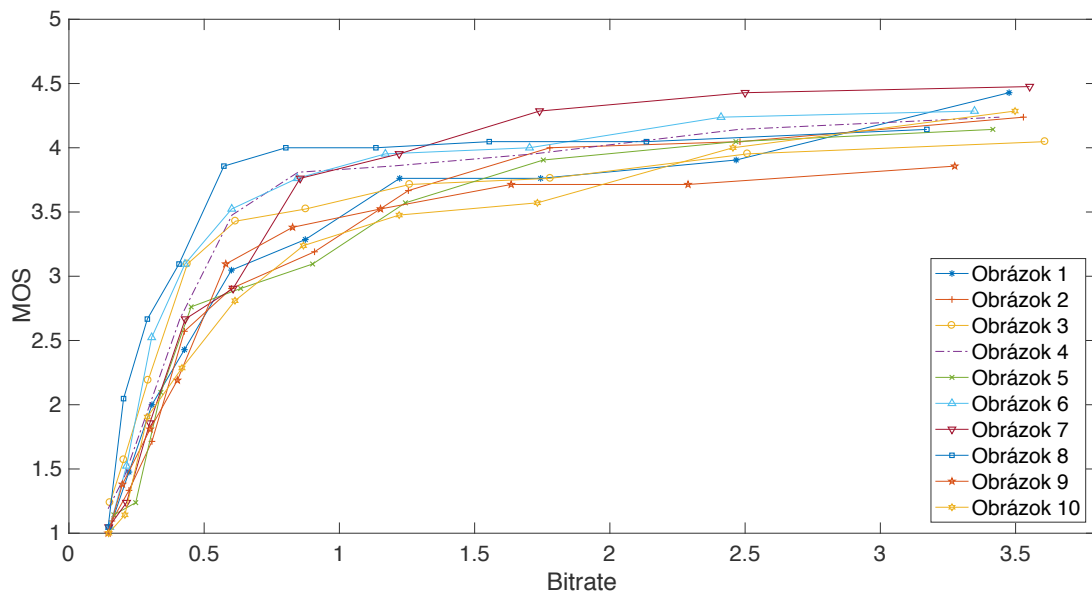


Obr. 4.5: Priebežný prehľad ohodnotení obrázku č. 1.

³<https://developers.google.com/chart/>

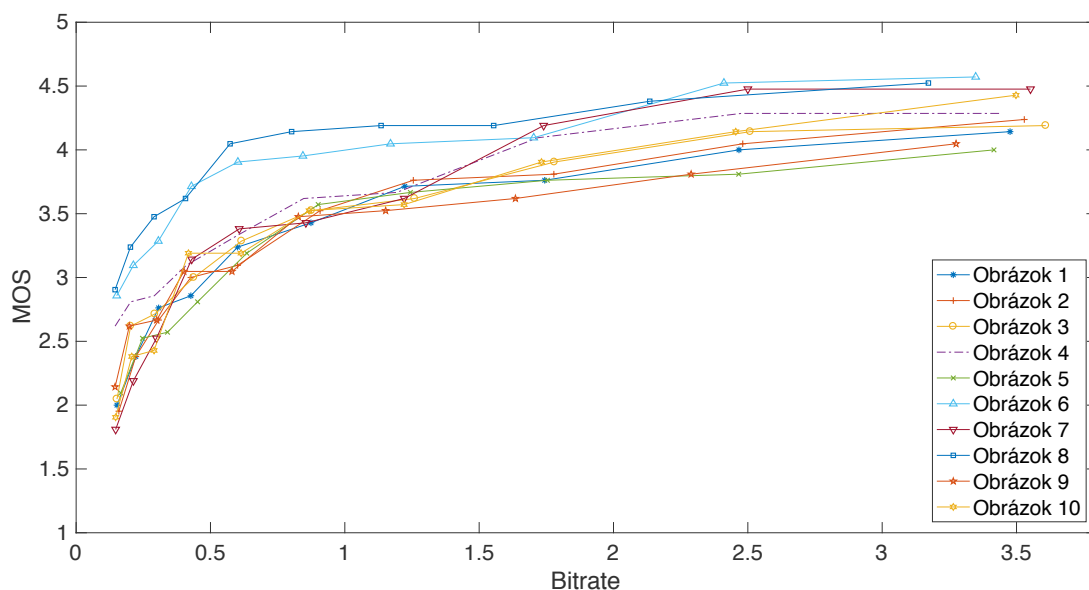
A nakoniec bol pri každom obrázku graf zobrazujúci priemery z ohodnotení. Boli rozlíšené funkcie pre JPEG a JPEG 2000. Os x je s logaritmickým rozložením za účelom zvýšenia prehľadnosti. Po ukončení subjektívneho hodnotenia kvality mal každý z celkového počtu 200 obrázkov presne 21 ohodnotení v stupnici MOS. Toto množstvo dostatočne prevyšovalo minimálnu odporúčanú hranicu 10 ohodnotení každého obrázku.

Obrázok 4.6 zobrazuje priemerné získané MOS ohodnotenie na osi y každého z 10 pôvodných obrázkov datasetu pre jednotlivé úrovne bitrate, ktorý je zobrazený na osi x . Z obrázku je možné vidieť, že obrázky s najnižším bitrate dosahovali hodnotenie od hodnoty približne 1, čo v tejto stupnici znamená najhoršiu pozorovateľnú kvalitu. Obrázok zobrazuje výsledky len pre kompresný štandard JPEG.



Obr. 4.6: Bitrate a získané MOS hodnotenie pre JPEG.

Obrázok 4.7 zobrazuje rovnakú situáciu ako v predchádzajúcom prípade, no tentoraz pre JPEG 2000. Je možné vidieť, že hodnotenia obrázkov s najnižším bitrate začínajú vyššie oproti pôvodnému štandardu JPEG. JPEG 2000 dokáže vytvoriť obrázky s malým bitrate a poznateľne vyššou kvalitou oproti JPEG, čo sa premietlo aj do hodnotenia.



Obr. 4.7: Bitrate a získané MOS hodnotenie pre JPEG 2000.

4.3 Objektívne hodnotenie datasetu

Táto kapitola popisuje postup použitý pri vypočítaní objektívneho hodnotenia kvality obrázkov datasetu vybranými metódami. Pre všetky výpočty uvádzané v tejto kapitole bola použitá platforma MATLAB⁴ alebo jazyk Bash. MATLAB je programovacia platforma navrhnutá pre technikov a vedcov a využíva programovací jazyk s rovnakým názvom. Pre skripty v tomto jazyku sa používajú súbory s koncovkou `.m` a ide o jazyk na báze matic, ktorý dovoľuje vyjadrenie výpočetnej matematiky.

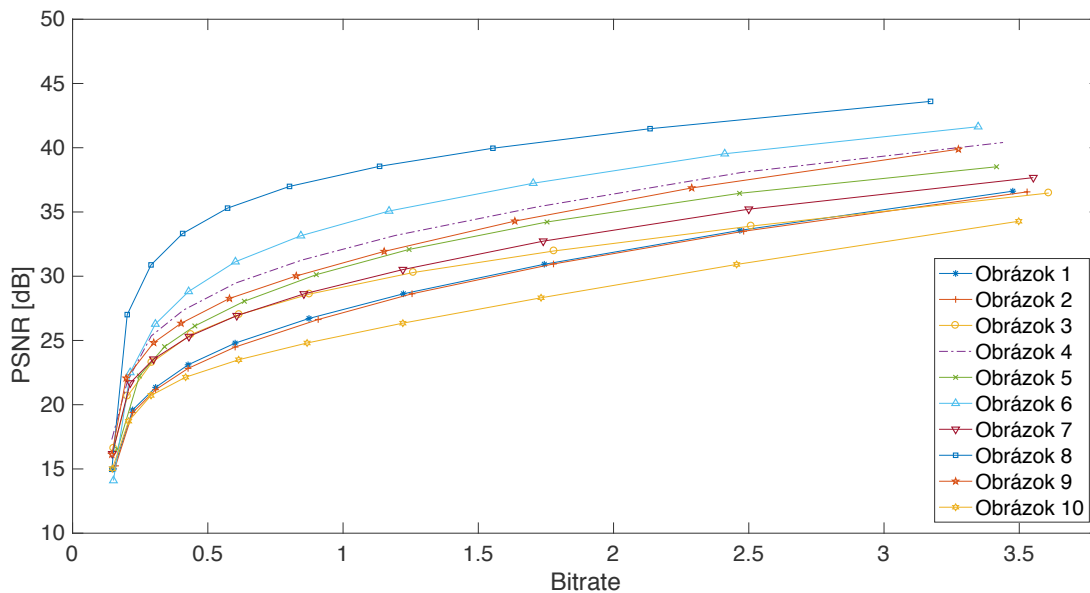
Pre výpočet hodnotení metód jednotlivých obrázkov bolo využitých viacero skriptov v programovacom jazyku MATLAB. Najprv skript pre každú z metód, ktorý prechádzal všetky priechy s obrázkami rozdelenými podľa kompresného štandardu a podľa mena. Načítal referenčný obrázok, ktorý neprešiel stratovou kompresiou a vo `for` cykloch v rozsahoch od 1 do 10, čo značilo bitrate úrovně, volal príslušnú metódu hodnotenia kvality, ktorej predal tento referenčný obrázok spolu s obrázkom, na ktorý ukazoval index cyklu. Výsledok metódy bol uložený do matice s odpovedajúcim menom. Po vyhodnotení kvalít všetkých obrázkov jednotlivými metódami ďalší skript zjednotil výsledky do jedného celku a vyexportoval hodnoty vo forme `.csv` súboru.

PSNR

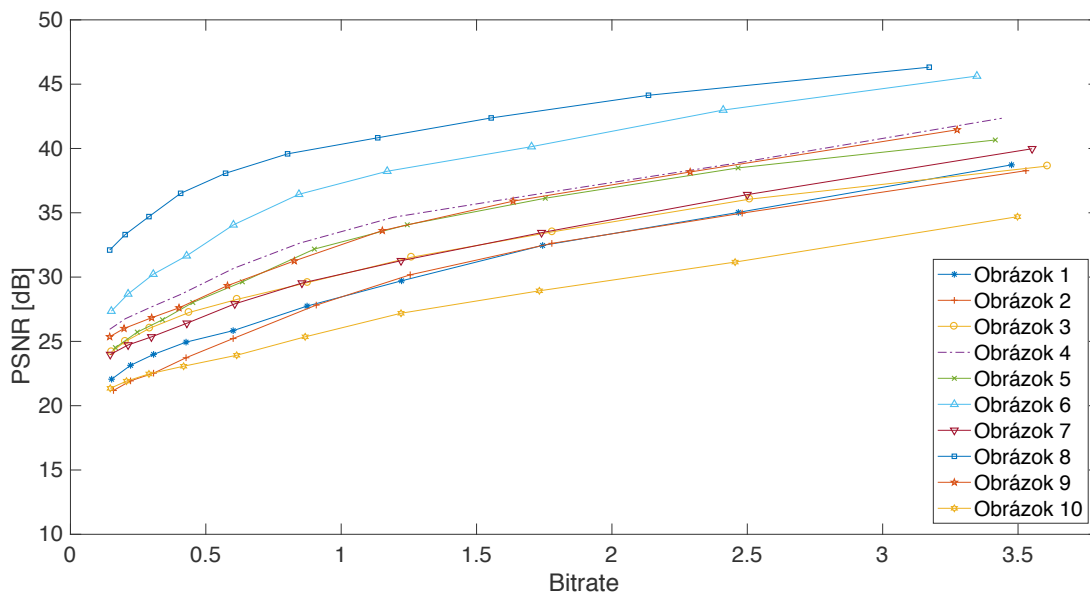
Výpočet výsledkov PSNR bolo možné uskutočniť bez využitia doplnkových knižníc pre platformu MATLAB, pretože natívne obsahuje funkciu `psnr(A, ref)`, kde `A` je testovaný obrázok a `ref` referenčný obrázok. Výsledky hodnotenia touto metódou boli v dB a čím vyššia je výsledná hodnota, tým vyššia by mala byť kvalita testovaného obrázku.

⁴<https://www.mathworks.com/discovery/what-is-matlab.html>

Obrázok 4.8 zobrazuje výsledky PSNR pre jednotlivé obrázky v závislosti na bitrate pre JPEG. Výsledky sa pohybovali v intervale od hodnoty 14,0951 do 43,6078. Výsledky počas celého priebehu stúpajú, čo zodpovedá zvyšujúcemu sa bitrate.



Obr. 4.8: Bitrate a výsledky PSNR pre JPEG.



Obr. 4.9: Bitrate a výsledky PSNR pre JPEG 2000.

Obrázok 4.9 ukazuje hodnoty výsledkov pre jednotlivé obrázky v závislosti na bitrate pre JPEG 2000. V prípade JPEG 2000 dolná hranica začínala na hodnote 21,1832 a horná

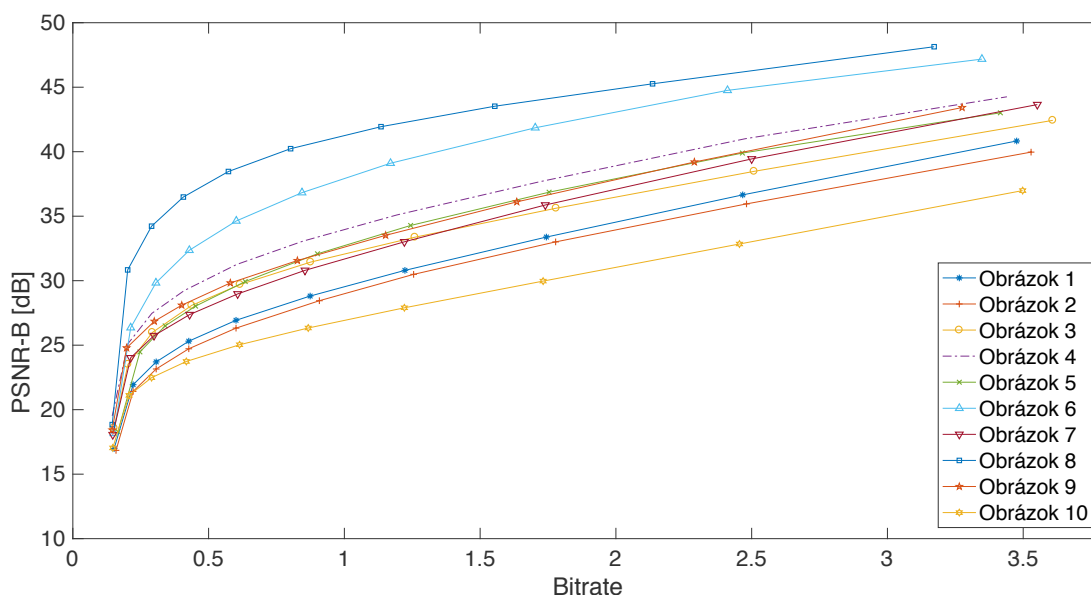
hranica bola 46,3199, čo je mierne vyššia dolná hranica oproti JPEG. Výsledky takisto stúpajú pre všetky obrázky počas celého priebehu.

PSNR-B

Pre získanie výsledkov metódy PSNR-B bola použitá funkcia `compute_psnrb(im_gnd, im)`⁵ pre MATLAB, kde `im_gnd` je referenčný a `im` testovaný obrázok.

Veľkosť bloku B bola pre JPEG nastavená na hodnotu 8. Pri testovaní obrázkov vytvorených pomocou JPEG 2000 bola veľkosť bloku B nastavená na 64, čo je východzia hodnota parametru `Cblk` kompresného nástroja Kakadu, s ktorým boli obrázky vytvorené. Výsledná hodnota udáva kvalitu obrázku v porovnaní s referenčným. Čím je hodnota vyššia, tým vyššia by mala byť aj kvalita a naopak.

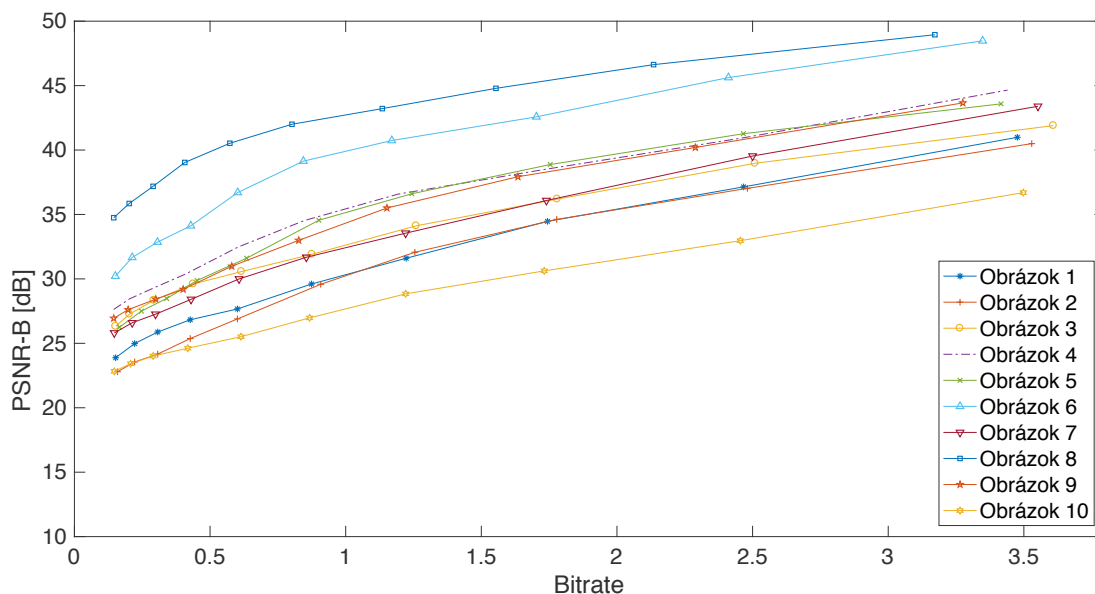
Obrázok 4.10 zobrazuje priebeh výsledkov metódy PSNR-B v súvislosti s bitrate obrázkov pre JPEG. Metóda poskytovala pre JPEG hodnoty v rozmedzí 16,8441 až 48,1381 a hodnoty stúpajú so zvyšujúcim sa bitrate.



Obr. 4.10: Bitrate a výsledky PSNR-B pre JPEG.

⁵<https://drive.google.com/file/d/0B7tU5Pj1dfCMYWRFMGw2cGg3b1U/view>

Obrázok 4.11 ukazuje priebeh výsledkov pre JPEG 2000. Dolné hodnoty sú v porovnaní s výsledkami pre JPEG vyššie a na začiatku nie je strmé stúpanie. Pre JPEG 2000 boli výsledky v rozmedzí 22,7979 až 48,9486 a hodnoty takisto, ako v predchádzajúcom prípade len stúpajú.

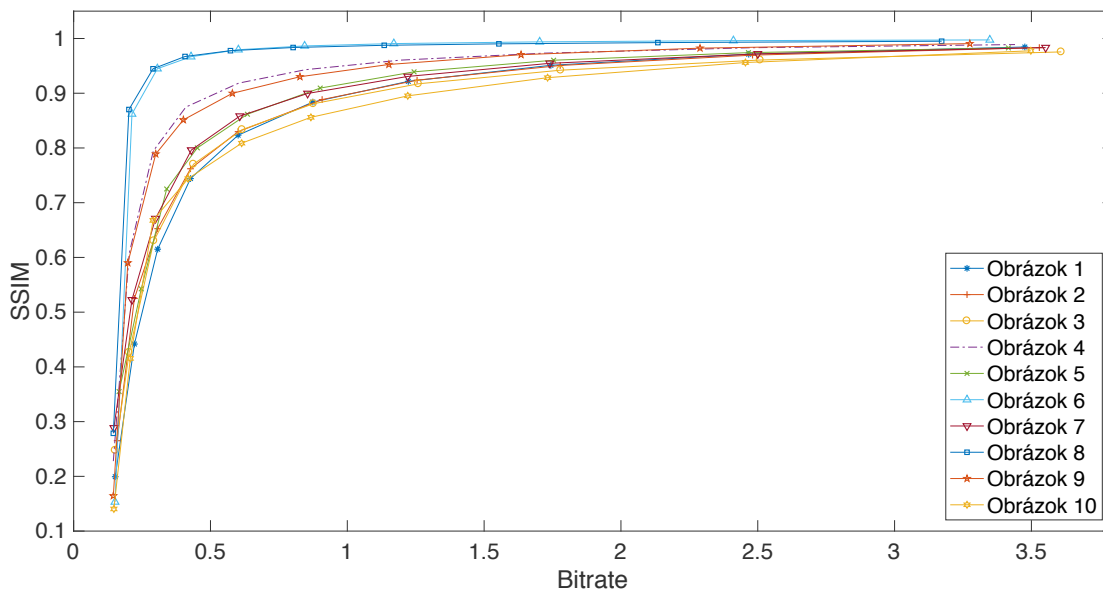


Obr. 4.11: Bitrate a výsledky PSNR-B pre JPEG 2000.

SSIM

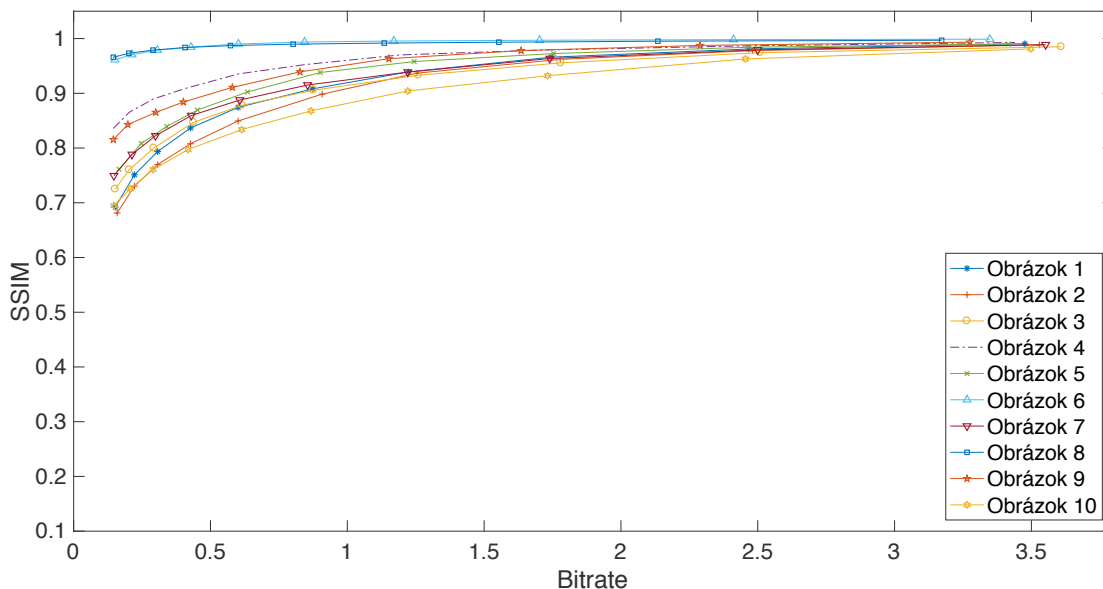
MATLAB obsahuje funkciu pre výpočet SSIM, ktorá má tvar `ssim(A, ref)`, kde význam parametrov je rovnaký ako v prípade funkcie `psnr`. Jedná sa o pôvodnú a neupravenú podobu metódy SSIM. Takisto ako v predchádzajúcich prípadoch, skript ohodnotil všetky obrázky datasetu a výsledky ukladal do matice. Výsledky metódy SSIM sa môžu pohybovať v intervale -1 až 1 , pričom pri výsledku 1 ide o identické obrázky.

Obrázok 4.12 zobrazuje priebeh hodnotení pri JPEG obrázkoch. Výsledky testovania sa pohybovali v intervale od $0,14042$ do $0,99741$. Pri spodných hodnotách bitrate hodnotenie takmer kolmo stúpa, a ďalej sa rýchlo dostáva do skoro rovnobežného smeru s osou x . Pri najvyšších úrovniach bitrate sú rozdiely vo výsledkoch veľmi malé, čo je pre túto metódu typické. Hodnoty pre každý z 10 pôvodných obrázkov datasetu stúpajú počas celého priebehu, aj keď toto stúpanie nezodpovedá použitému bitrate v takom rozsahu, ako napríklad pri PSNR.



Obr. 4.12: Bitrate a výsledky SSIM pre JPEG.

Na obrázku 4.13 sú zobrazené hodnotenia pre JPEG 2000, kde je priebeh opäť na začiatku menej strmý. Výsledné hodnoty metódy pre tento štandard sa pohybujú v intervale od 0,68135 do 0,99886 a celý čas stúpajú pre všetky testované obrázky.

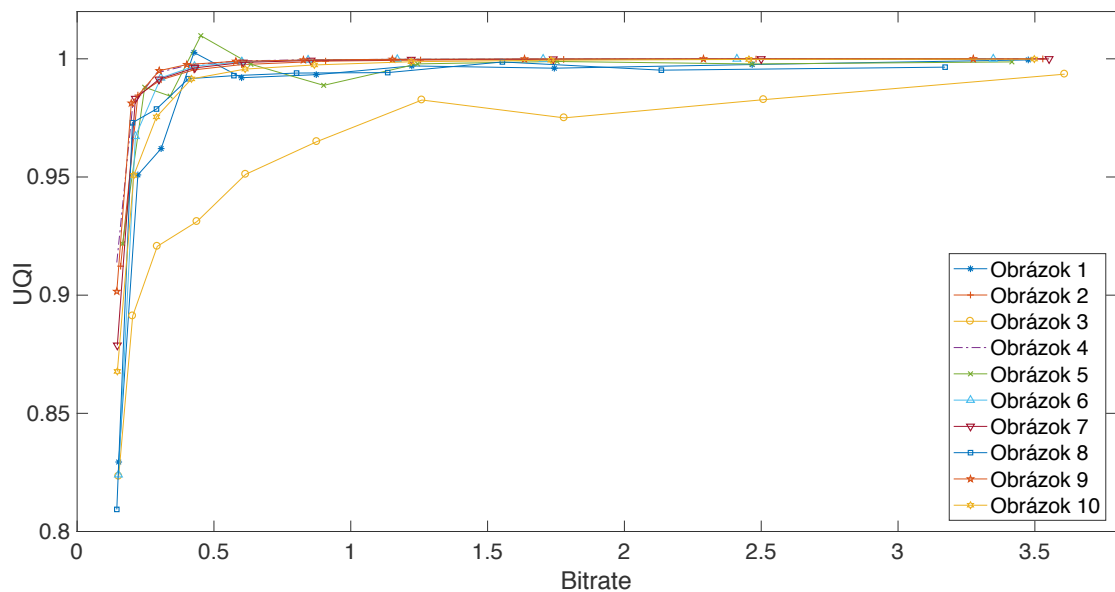


Obr. 4.13: Bitrate a výsledky SSIM pre JPEG 2000.

UQI

Funkcia pre výpočet hodnotenia UQI nie je obsiahnutá v základnom vybavení platformy MATLAB. Pre výpočet som preto použil funkciu `imageQualityIndex(A, ref)`⁶. Hodnota výsledkov pri UQI môže byť rovnako ako pri SSIM v intervale -1 až 1 , pričom ich význam je rovnaký.

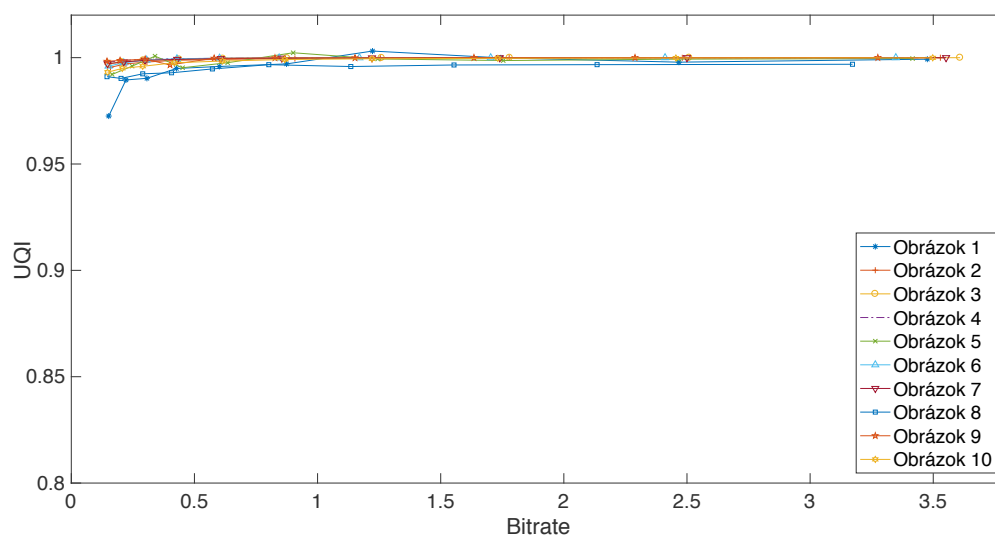
Na obrázku 4.14 sú výsledky pre JPEG. Hodnoty výsledkov nestúpajú v celom rozsahu, ako to bolo v prípade všetkých predchádzajúcich metód. V niektorých prípadoch naopak klesajú nasledované opätovným stúpaním. Výsledky sa pohybovali v intervale 0,80935 až 1,0098.



Obr. 4.14: Bitrate a výsledky UQI pre JPEG.

Obrázok 4.15 ukazuje výsledky pre JPEG 2000. Výsledky opäť v mnohých prípadoch striedavo stúpajú a klesajú. Hodnoty boli v intervale od 0,96235 do 1,01.

⁶<https://www.mathworks.com/matlabcentral/fileexchange/29500-image-error-measurements>



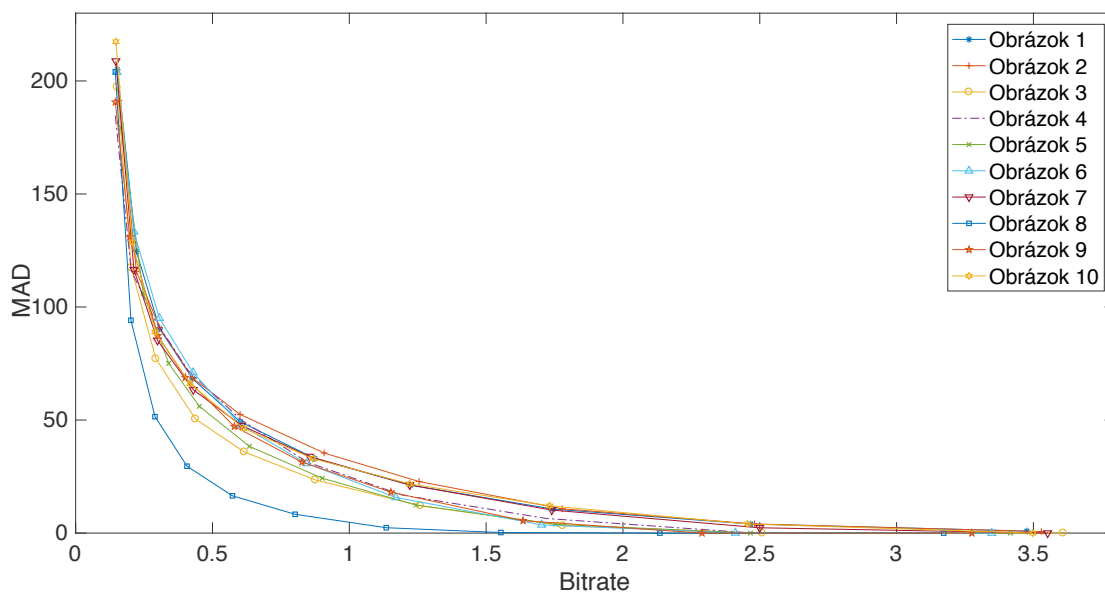
Obr. 4.15: Bitrate a výsledky UQI pre JPEG 2000.

MAD

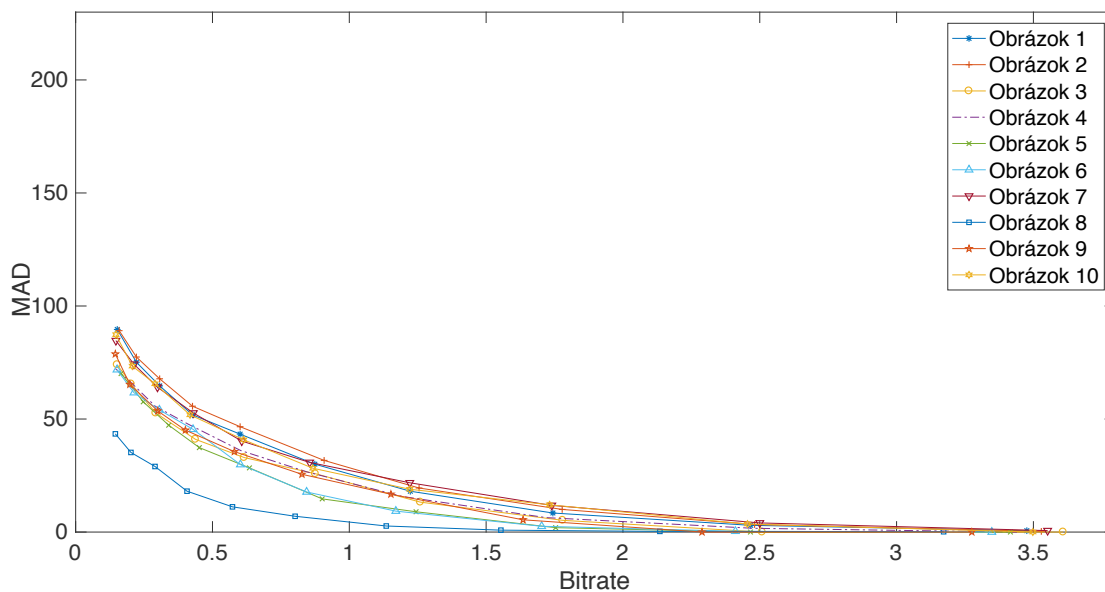
Na výpočet hodnotení bola použitá funkcia $MAD_index(ref, A)$ ⁷ vytvorená autormi metódy. Čím nižšia je hodnota výsledku, tým lepšia by mala byť kvalita testovaného obrázku. V najlepšom prípade by sa mala výsledná hodnota rovnať 0.

Na obrázku 4.16 je ukázaný priebeh výsledkov pre JPEG. Hodnoty pre všetky obrázky počas celého priebehu klesajú, pričom v úvode je toto klesanie takmer kolmé, čo by mohlo zodpovedať použitému bitrate a danej kvalite. Testované obrázky dosahovali hodnotenie kvality v rozsahu 0 až 217,3674.

⁷http://sse.tongji.edu.cn/linzhang/IQA/Evaluation_MAD/eva-MAD.htm



Obr. 4.16: Bitrate a výsledky MAD pre JPEG.



Obr. 4.17: Bitrate a výsledky MAD pre JPEG 2000.

Obrázok 4.17 zobrazuje výsledky na obrázkoch JPEG 2000. Hodnoty pre všetky obrázky počas celého priebehu klesajú rovnako ako pri JPEG a sú v rozsahu 0 až 88,6164.

Po vypočítaní všetkých objektívnych ohodnotení pre daný kompresný štandard ďalší skript zlúčil všetky výsledky do jednej veľkej matice a tú vyexportoval do .csv súboru. Nakoniec tak vznikli dva .csv súbory s výsledkami, jeden pre JPEG a druhý pre JPEG

2000. Pre každú metódu boli dva stĺpce, v prvom bol celý názov obrázku a v druhom výsledok ohodnotenia tohoto obrázku danou metódou.

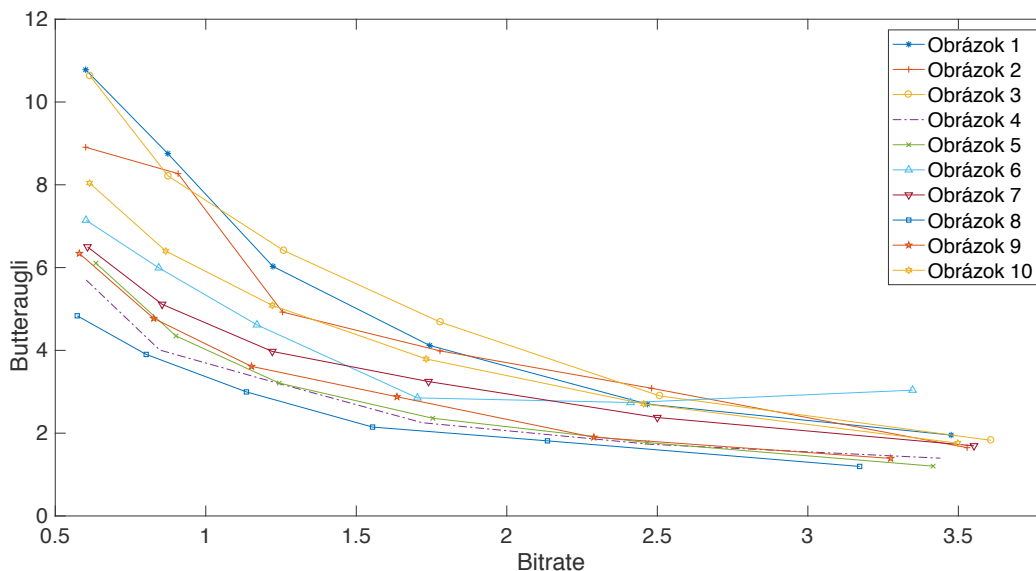
Butteraugli

Metóda Butteraugli bola zahrnutá len do druhého testovania výkonnosti metód a to na vyššom bitrate oproti prvému testovaniu. Toto druhé testovanie sa uskutočňovalo len na horných 6 obrázkoch s najvyšším bitrate z celkového rozsahu 10 pre každý z pôvodných 10 obrázkov datasetu. Pre získanie ohodnotení kvality obrázkov datasetu touto metódou bola využitá implementácia⁸ od Google. Zdrojové súbory boli skompilované s využitím poskytnutých inštrukcií a nástroja Bazel na spustiteľný binárny súbor.

Ďalej bol použitý skript v jazyku Bash, ktorý opakovane spúšťal skompilovaný porovnávací nástroj spolu so všetkými obrázkami datasetu. Porovnávací nástroj bol spúšťaný v tejto podobe: `./butteraugli reference${n}.png image${n}q${i}.png`, pričom skript pracoval na základe 3 for cyklov. Vonkajší cyklus bežal v rozsahu od 1 do 10 s priradenou premennou `n`, čo predstavovalo 10 pôvodných obrázkov datasetu. Vnútri hlavného cyklu boli dva vnorené cykly pre JPEG a JPEG 2000 s rozsahom od 6 do 10 s premennou `i`, čo predstavovalo horných 6 obrázkov. Tieto dva vnorené cykly vždy vchádzali do priechov s obrázkami, prvý cyklus pre JPEG a druhý pre JPEG 2000.

Pre porovnávanie kvality touto metódou musel byť použitý formát obrázkov PNG, pretože porovnávací nástroj podporuje len formáty PNG a PNM. Skript ukladal výsledky do textového súboru oddelené medzerou a vždy na nový riadok pre každý obrázok. Čím nižia bola hodnota výsledku, tým by mala byť kvalita testovaného obrázku vyššia.

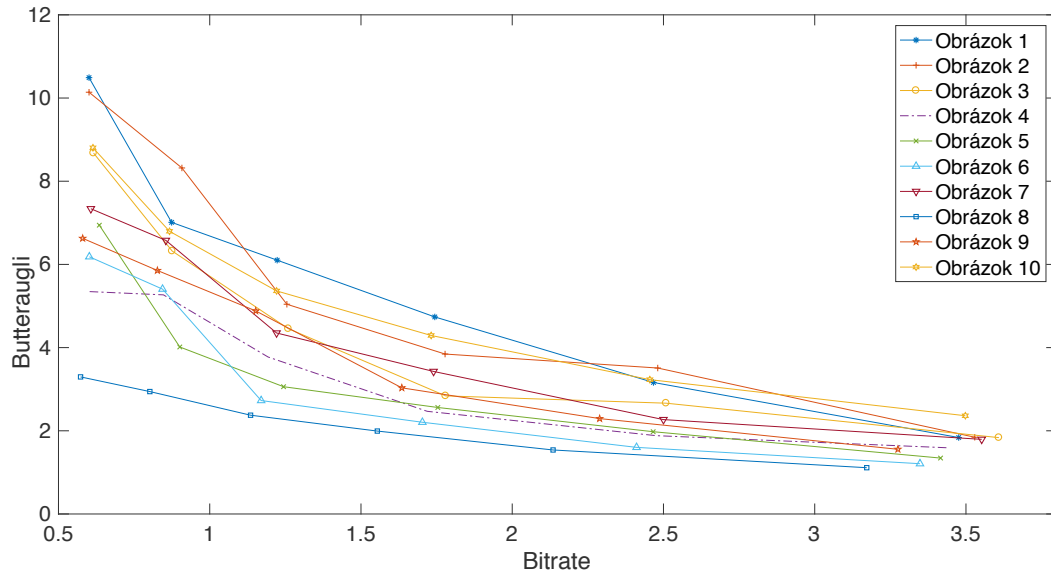
Na obrázku 4.18 sú ukázané výsledky testovania na štandarde JPEG. Bitrate zobrazený na osi x sa tentoraz od predchádzajúcich grafov odlišuje, pretože zahŕňa len bitrate posledných 6 obrázkov z celkových 10 úrovní. Získané výsledky len klesajú a dosahovali hodnoty v rozsahu 1,2030 až 10,7798.



Obr. 4.18: Bitrate a výsledky Butteraugli pre JPEG.

⁸<https://github.com/google/butteraugli>

Obrázok 4.19 zobrazuje výsledky pre JPEG 2000. Hodnoty klesajú v celom priebehu a sú v rozsahu 1,2088 až 10,4903.



Obr. 4.19: Bitrate a výsledky Butteraugli pre JPEG 2000.

4.4 Meranie výkonnosti

Nasledujúca časť vychádza z viacerých už citovaných zdrojov a tiež z [7]. Pred samotným vyhodnotením výsledkov je potrebné zohľadniť možnú nelinearitu subjektívnych hodnotení, ktorá mohla počas subjektívneho testovania vzniknúť. Za týmto účelom sa vykonáva nelineárne mapovanie objektívneho skóre poskytnutého metódami pred samotným meraním korelácie medzi subjektívnymi a objektívnymi hodnotami. Pre získanie lineárneho vzťahu medzi objektívnymi hodnotami poskytnutými metódou a príslušným subjektívnym skóre je každá hodnota metriky x namapovaná na hodnotu

$$q(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(x - \beta_3))} \right) + \beta_4 x + \beta_5. \quad (4.1)$$

Tieto nové namapované hodnoty objektívnych metód sú následne používané pri meraní výkonnosti metód a s pôvodnými hodnotami sa po získaní týchto namapovaných už nepracuje. Počiatočné hodnoty parametrov $\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5\}$ môžu byť zvolené viacerými spôsobmi. Zvolil som riešenie, kedy sa parametre β rovnajú hodnotám $\{\max(mos), \min(mos), \text{mean}(objectiveValues), 0.1, 40\}$. Toto riešenie bolo použité v [17] pre LIVE dataset, s ktorým som pracoval v mojom testovaní.

Pre meranie korelácie medzi subjektívnym a namapovaným objektívnym skóre metódy sa najčastejšie používajú korelačné koeficienty – Spearmanov, Kendallov a Pearsonov koeficient a tiež 2 typy odchýlok – stredná kvadratická odchýlka a stredná absolútna chyba.

Spearmanov korelačný koeficient

Spearmanov korelačný koeficient (SRCC) predstavuje korelačný koeficient medzi predikovaným objektívnym hodnotením a subjektívnym hodnotením v stupnici MOS. Koeficient meria monotónnosť metódy, napríklad limit, na ktorý sa zhoduje hodnotenie metódy s relatívnou veľkosťou subjektívneho hodnotenia

$$\text{SRCC} = 1 - \frac{6 \sum_{i=1}^{M_d} d_i^2}{M_d(M_d^2 - 1)}, \quad (4.2)$$

kde d_i je rozdiel medzi objektívnym a subjektívnym hodnotením i -teho obrázku. Koeficient je nezávislý na nelineárnom mapovaní objektívnych hodnotení.

Kendallov korelačný koeficient

Kendallov korelačný koeficient (KRCC) je miera používaná na meranie korelácie medzi dvoma nameranými množstvami. Ide o neparametrický test hypotéz na štatistickú závislosť

$$\text{KRCC} = \frac{M_c - M_{dc}}{\frac{1}{2}M_d(M_d - 1)}, \quad (4.3)$$

kde M_c je počet konkordantných a M_{dc} počet diskordantných párov v datasete. Takisto ako predchádzajúci SRCC, je KRCC mierou monotónnosti, a patrí medzi najčastejšie používané koeficienty pri porovnávaní výkonnosti metód merajúcich kvalitu obrázkov.

Pearsonov lineárny korelačný koeficient

Pearsonov lineárny korelačný koeficient (PLCC) je lineárny koeficient medzi predpovedaným skóre metódou a subjektívnym skóre v stupnici MOS. Koeficient je miera presnosti predpovede danej metódy merajúcej kvalitu obrazu, v tomto prípade schopnosť metódy predpovedať subjektívne skóre s malou chybou. Koeficient

$$\text{PLCC} = \frac{\sum_{i=1}^{M_d} (q_i - \bar{q})(s_i - \bar{s})}{\left(\sum_{i=1}^{M_d} (q_i - \bar{q})^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{M_d} (s_i - \bar{s})^2\right)^{\frac{1}{2}}}, \quad (4.4)$$

kde s_i je subjektívne, a q_i namapované objektívne ohodnotenie i -teho obrázku datasetu o veľkosti M_d . Ďalej \bar{s} a \bar{q} sú priemery subjektívnych a namapovaných objektívnych ohodnotení.

Stredná kvadratická odchýlka

Stredná kvadratická odchýlka (RMSE) je miera rozdielov medzi hodnotami predpovedanými modelom a reálne pozorovanými hodnotami

$$\text{RMSE} = \left(\frac{1}{M_d} \sum_{i=1}^{M_d} (q_i - s_i)^2\right)^{\frac{1}{2}}. \quad (4.5)$$

Takisto ako PLCC, je RMSE miera presnosti predpovede.

Stredná absolútna chyba

Stredná absolútna chyba (MAE) tiež porovnáva predpovedané a pozorované hodnoty a určuje odchýlku

$$\text{MAE} = \frac{1}{M_d} \sum_{i=1}^{M_d} |q_i - s_i|. \quad (4.6)$$

Ide o mieru presnosti predpovede, podobne ako u PLCC a RMSE.

Pri posudzovaní, ktorá z testovaných metód je najpresnejšia je potrebné zistiť, ktorá z nich má vyššie hodnoty koeficientov SRCC, KRCC a PLCC a zároveň čo najmenšie odchýlky RMSE a MAE.

4.5 Výpočet koeficientov a odchýlok

Pre výpočet korelačných koeficientov a odchýlok bol vytvorený skript pre platformu MATLAB. Jeho výstupom bola matica so spomínanými 3 koeficientami a 2 odchýlkami pre príslušnú metódu a pre daný pôvodný obrázok datasetu.

Pred spustením bolo potrebné inicializovať premenné s názvami PSNR, PSNRB, SSIM, UQI, MAD a MOS, ktoré boli inicializované vždy zvlášť pre každý z 10 pôvodných obrázkov datasetu. Každá z týchto premenných vždy obsahovala 10 hodnôt, ktoré predstavovali výsledky metód, v prípade premennej MOS priemery subjektívnych ohodnotení každého z 10 obrázkov s rôznymi bitrate úrovňami.

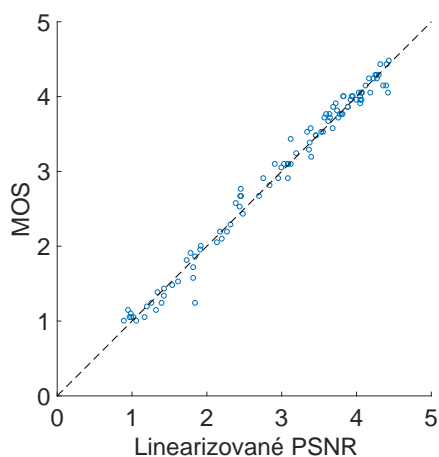
Po inicializácii premenných bol spustený skript, ktorý postupne pre každú metódu spúšťal iný skript vždy s premennou MOS a výsledkami danej metódy. V opakovane spúšťanom skripte sa na začiatku nastavili počiatočné hodnoty β parametrov spôsobom uvedeným na začiatku predchádzajúcej sekcie. Následne bola zavolaná funkcia `nlinfit` využívaná na nelineárnu regresiu. Jej vstupnými parametrami boli objektívne výsledky metódy, subjektívne MOS skóre, modelová funkcia (4.1) uvedená na začiatku predchádzajúcej sekcie a nakoniec β parametre s počiatočnými hodnotami. Výstupom boli nové hodnoty β parametrov, rezíduá a tiež Jacobiho matica.

Následne bola zavolaná funkcia `nlpredci`, ktorej vstupom bola modelová funkcia (4.1), objektívne hodnoty metódy, β parametre vypočítané v predchádzajúcom kroku, rezíduá a Jacobiho matica. Výsledkom funkcie boli namapované objektívne hodnoty. Tieto hodnoty boli následne použité na výpočty jednotlivých koeficientov a odchýlok potrebných pre porovnanie výkonnosti metód. Na výpočet každého koeficientu a odchýlky bola zavolaná príslušná funkcia a výsledky uložené do jednej matice pre každú z testovaných metód a pre každý z 10 pôvodných obrázkov datasetu.

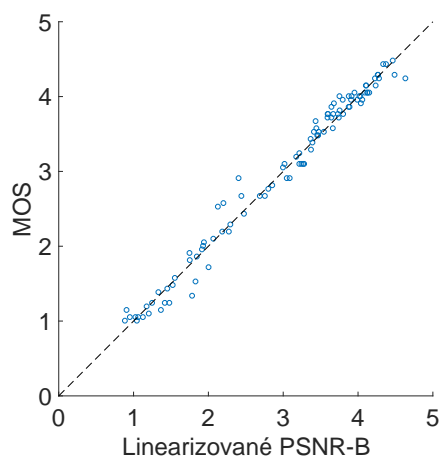
Po inicializácii premenných PSNR, PSNRB, SSIM, UQI, MAD a MOS a po spustení úvodného skriptu tak bolo získaných 5 matíc – jedna matica pre každú testovanú metódu a táto matica obsahovala všetky 3 koeficienty v poradí KRCC, SRCC a PLCC nasledované odchýlkami MAE a RMSE. Obsah tejto matice ukazoval výkonnosť konkrétnej metódy pre daný pôvodný obrázok datasetu a pre konkrétny kompresný štandard. Vykonaním tohoto postupu pre všetky testované obrázky a pre oba kompresné štandardy boli získané čiastkové výsledky – 5 matíc s korelačnými koeficientami a odchýlkami pre každý z pôvodných 10 obrázkov datasetu pre JPEG a rovnako aj pre JPEG 2000.

4.6 Výsledky testovania

Táto sekcia obsahuje výsledky získané v prvom realizovanom porovnávaní vybraných metód. Obrázky 4.20 až 4.24 zobrazujú bodové grafy znázorňujúce namapované linearizované hodnoty výsledkov jednotlivých metód pre kompresný štandard JPEG v závislosti na príslušnom subjektívnom hodnotení v stupnici MOS.

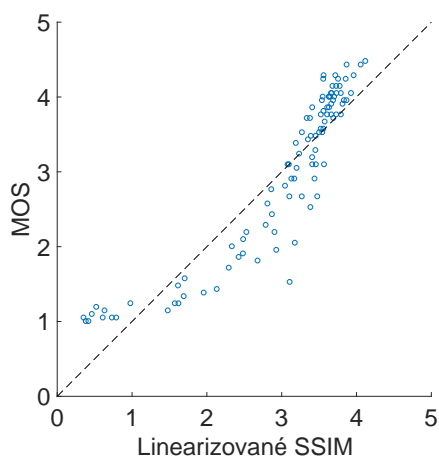


Obr. 4.20: PSNR výsledky pre JPEG.

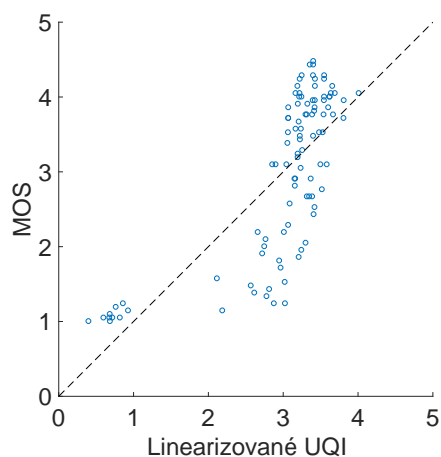


Obr. 4.21: PSNR-B výsledky pre JPEG.

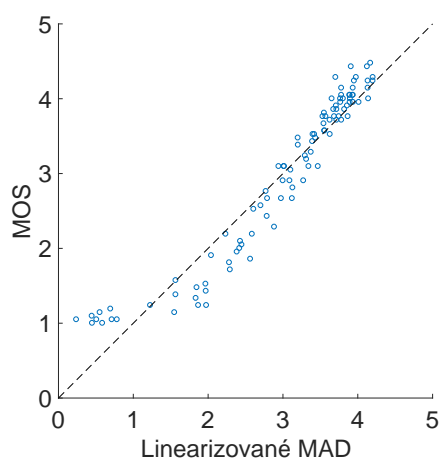
Z obrázkov sa dá pozorovať, že metódy PSNR a PSNR-B predpovedali kvalitu konzistentne so subjektívnym hodnotením, pretože body sú umiestnené relatívne v blízkosti stredu, ktorý predstavuje funkcia $y = x$. Najhoršie hodnotenie podľa obrázkov dosiahla metóda UQI, ktorej výsledky sa nezhodovali so subjektívnym hodnotením najmä pri stredných bitrate hodnotách.



Obr. 4.22: SSIM výsledky pre JPEG.

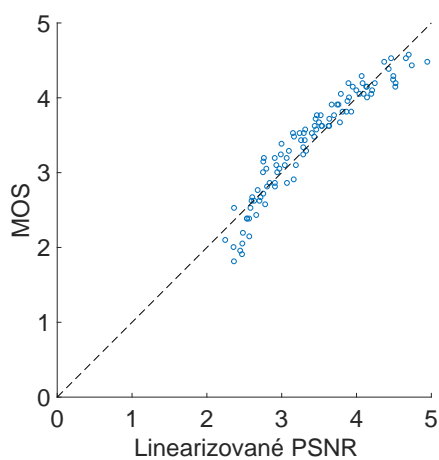


Obr. 4.23: UQI výsledky pre JPEG.

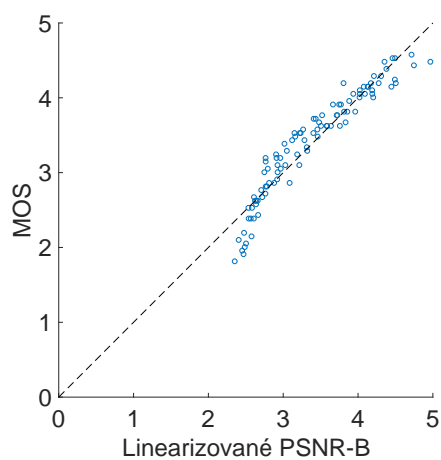


Obr. 4.24: MAD výsledky pre JPEG.

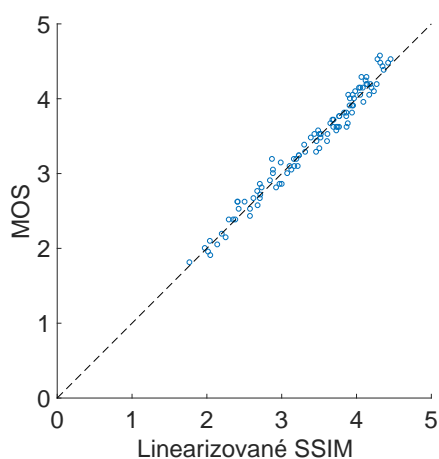
Ďalej obrázky 4.25 až 4.29 ukazujú rovnakú situáciu pre JPEG 2000. Z obrázkov je možné pozorovať, že PSNR a PSNR-B už neukazuje takú koreláciu so subjektívnym hodnotením ako v prípade JPEG obrázkov. Ďalej je možné vidieť, že SSIM a MAD ukazujú najlepšiu presnosť. Metóda UQI dosiahla rovnako ako pri JPEG najväčšie odchylenie od subjektívneho hodnotenia zo všetkých testovaných metód.



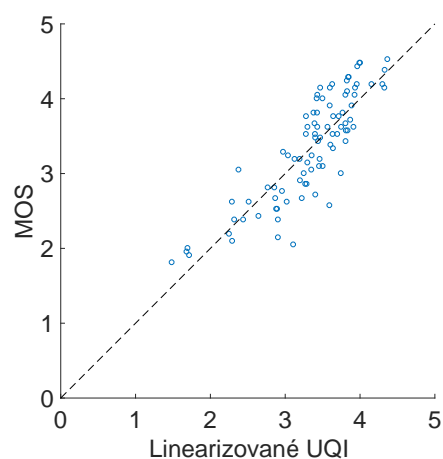
Obr. 4.25: PSNR výsledky pre JP2.



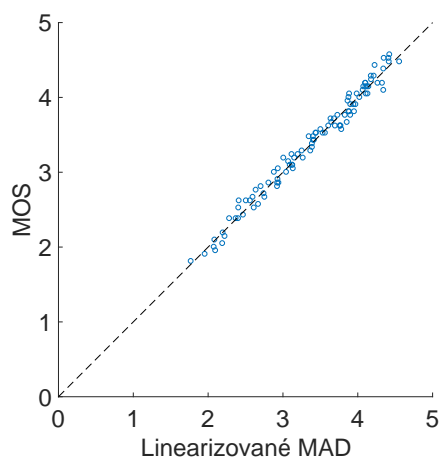
Obr. 4.26: PSNR-B výsledky pre JP2.



Obr. 4.27: SSIM výsledky pre JP2.



Obr. 4.28: UQI výsledky pre JP2.



Obr. 4.29: MAD výsledky pre JP2.

Koeficienty

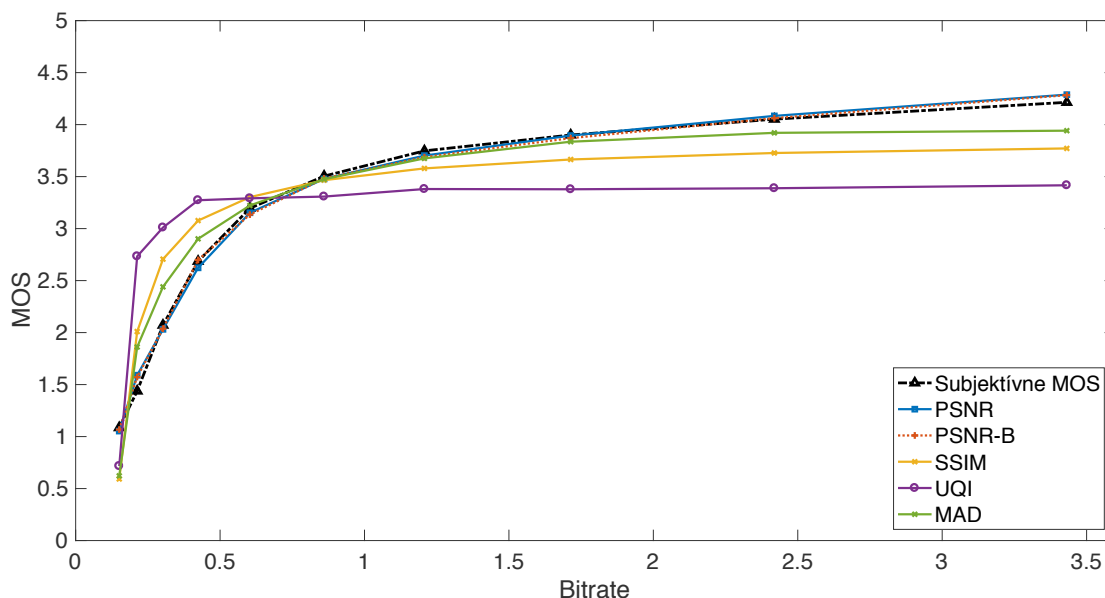
Výsledné koeficienty a odchýlky boli vypočítané súčtom čiastkových výsledkov pre jednotlivé obrázky, zvlášť pre oba kompresné štandardy a následným vydelením počtom pôvodných obrázkov datasetu, čiže 10. Získané hodnoty sú uvedené v tabuľkách 4.2 a 4.3. Výsledky sú zaokrúhlené na 4 desatinné miesta. Najväčšia hodnota v stĺpcoch s koeficientami je zvýraznená, takisto najmenšia hodnota v stĺpcoch s chybami. V prípade, že stĺpec neobsahoval jedinú najväčšiu hodnotu pre koeficienty alebo jedinú najmenšiu hodnotu pre chyby sú zvýraznené všetky výskyty.

Metóda	KRCC	SRCC	PLCC	MAE	RMSE
PSNR	0,9911	0,9988	0,9918	0,0955	0,1156
PSNR-B	0,9867	0,9970	0,9909	0,1043	0,1247
SSIM	0,9911	0,9988	0,9064	0,3636	0,4246
UQI	0,9067	0,9364	0,7468	0,6017	0,7000
MAD	0,9800	0,9967	0,9581	0,2458	0,2975

Tabuľka 4.2: Korelačné koeficienty a odchýlky pre JPEG.

Pre JPEG sa podľa získaných výsledkov v tabuľke 4.2 javí ako najvhodnejšia metóda PSNR. Táto metóda dosiahla v testovaní rovnaké hodnoty korelačných koeficientov KRCC a SRCC ako v prípade SSIM, no v ostatných kategóriách dosiahla lepšie výsledky. Koeficient PLCC dosiahol u nej najvyššiu hodnotu zo všetkých metód a zároveň dosiahlo PSNR najmenšie chyby MAE a RMSE zo všetkých testovaných metód. Metóda PSNR teda vychádza ako najvhodnejšia metóda pre zistenie objektívneho hodnotenia kvality obrázkov vytvorených s využitím kompresného štandardu JPEG, pokiaľ je cieľom merať kvalitu okrem obrázkov s vysokým bitrate aj obrázky s bitrate nižším, kedy je dolná hranica približne 0,15 bitov/pixel. Veľmi tesne za PSNR sa však umiestnila metóda PSNR-B s len mierne horšími výsledkami vo všetkých kategóriách.

Obrázok 4.30 zobrazuje linearizované výsledky všetkých metód zahrnutých v tomto testovaní nad JPEG obrázkami spolu s priemerom subjektívnych hodnotení v závislosti na bitrate obrázkov. Os x zobrazuje v jednotlivých bodoch priemery 10 podobných bitrate každého z pôvodných 10 obrázkov datasetu skomprimovaných na 10 rôznych úrovni. Os y zobrazuje linearizované výsledky jednotlivých metód obsiahnutých v testovaní spolu so subjektívnym hodnotením v stupnici MOS. Metódy PSNR a PSNR-B majú podobný priebeh, a je možné vidieť ich koreláciu so subjektívnym hodnotením.



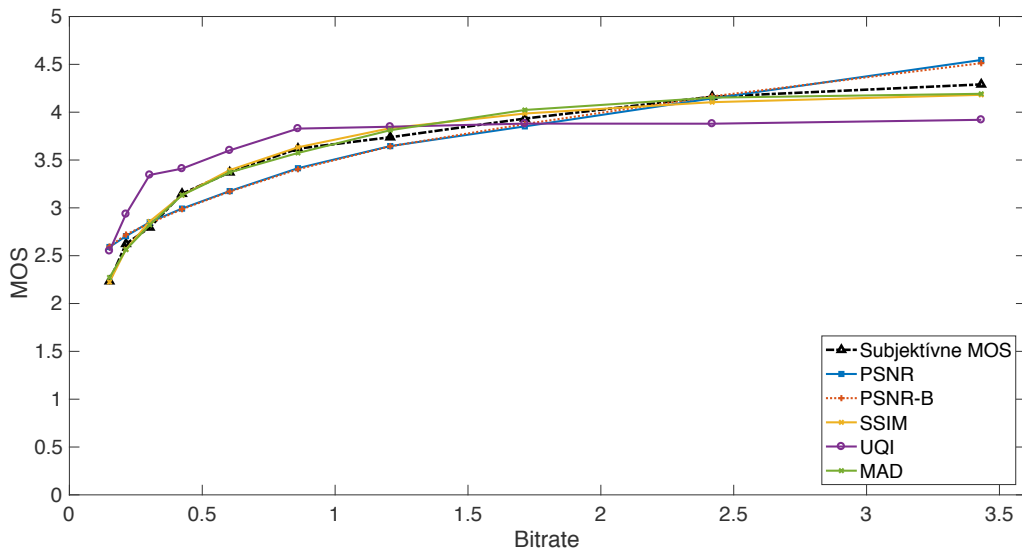
Obr. 4.30: Výsledky hodnotení v stupnici MOS v závislosti na bitrate pre JPEG.

Metóda	KRCC	SRCC	PLCC	MAE	RMSE
PSNR	0,9844	0,9973	0,9495	0,1712	0,2060
PSNR-B	0,9889	0,9985	0,9486	0,1662	0,2041
SSIM	0,9889	0,9985	0,9841	0,0901	0,1067
UQI	0,8178	0,8782	0,8127	0,3947	0,4494
MAD	0,9778	0,9964	0,9881	0,0792	0,0940

Tabuľka 4.3: Korelačné koeficienty a odchýlky pre JPEG 2000.

V prípade štandardu JPEG 2000 podľa výsledkov v tabuľke 4.3 dosiahla metóda SSIM spolu s PSNR-B najvyššie hodnoty korelačných koeficientov KRCC a SRCC, no najvyšší koeficient PLCC spolu s najnižšími chybami MAE a RMSE dosiahla metóda MAD. Testovaním v tomto prípade nie je jednoznačne určená jediná najvhodnejšia metóda, ako tomu bolo v prípade JPEG. Podľa dosiahnutých výsledkov sa javí ako najvhodnejšia metóda buď SSIM alebo MAD, ktoré dosiahli najvyššie hodnoty odlišných koeficientov a metóda MAD dosiahla najmenších odchýlok zo všetkých testovaných metód.

Obrázok 4.31 ukazuje linearizované výsledky všetkých metód zahrnutých v tomto testovaní nad JPEG 2000 obrázkami spolu s priemerom subjektívnych hodnotení v závislosti na bitrate obrázkov. Výsledky metód PSNR a PSNR-B sa opäť mierne podobajú, no už v takej miere nekorelujú so subjektívnym hodnotením a takisto sa podobajú aj výsledky metód MAD a SSIM, ktorých koreláciu je naopak možné vidieť.



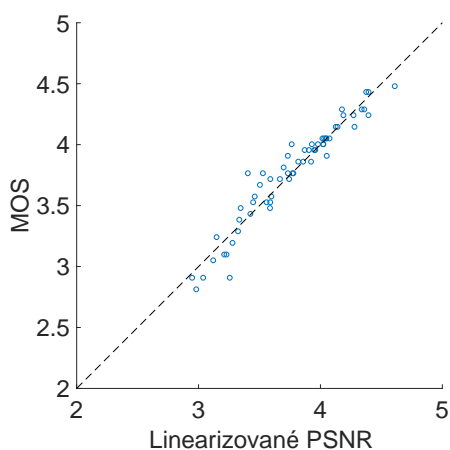
Obr. 4.31: Výsledky hodnotení v stupnici MOS v závislosti na bitrate pre JPEG 2000.

4.7 Testovanie na vyššom bitrate

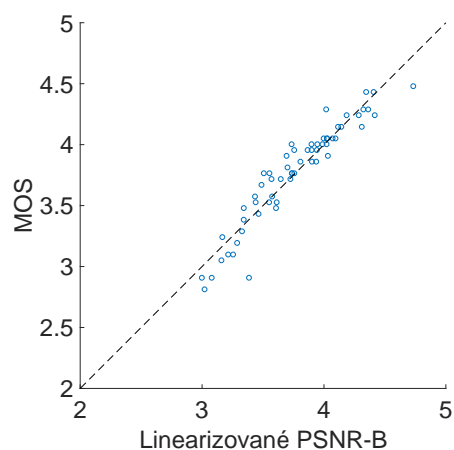
Druhé testovanie výkonnosti metód sa týkalo zistenia výkonnosti vybraných metód na vyšších úrovniach bitrate, ktorá by sa mohla líšiť od výkonnosti ukázanej na širokom rozsahu bitrate datasetu s hodnotami od 0,15 bitov/pixel do 3,5 bitov/pixel. Ďalším dôvodom bolo otestovanie metódy Butteraugli a porovnanie jej výkonnosti s ostatnými vybranými metódami. Táto metóda sa sústreďuje na obrázky s vyšším bitrate a jej chovanie pri nižších

hodnotách nie je udávané ako smerodajné. Do tohoto testovania boli zapojené všetky predchádzajúce vybrané metódy a navyš spomínaná metóda Butteraugli používaná pri JPEG kodeku Guertzli na hodnotenie kvality. Dolná hranica bitrate testovaných obrázkov sa v tomto prípade zmenila z 0,15 bitov/pixel na približne 0,6 bitov/pixel. Druhé testovanie sa teda týkalo horných 6 z celkového počtu 10 obrázkov s rôznym bitrate, na ktoré bol každý z 10 pôvodných obrázkov datasetu skomprimovaný.

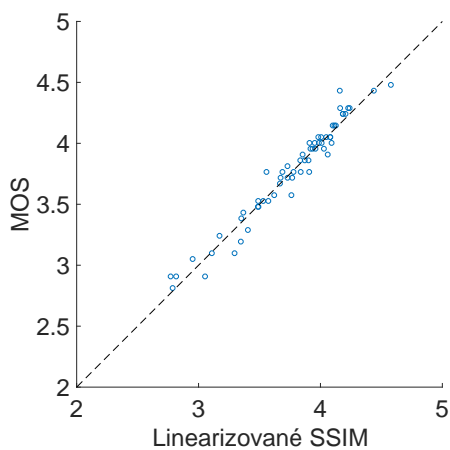
Obrázky 4.32 až 4.37 zobrazujú bodové grafy ukazujúce linearizované hodnoty výsledkov zahrnutých metód pre kompresný štandard JPEG v závislosti na subjektívnych hodnoteniach daných obrázkov.



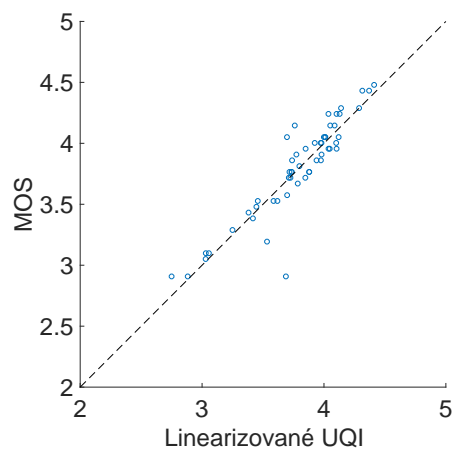
Obr. 4.32: PSNR výsledky pre JPEG.



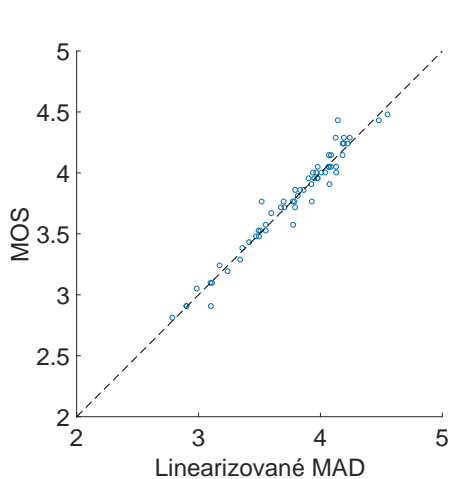
Obr. 4.33: PSNR-B výsledky pre JPEG.



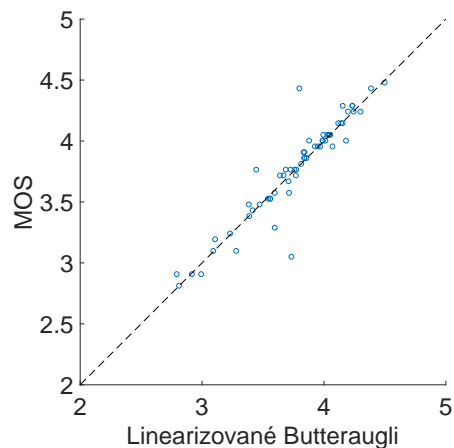
Obr. 4.34: SSIM výsledky pre JPEG.



Obr. 4.35: UQI výsledky pre JPEG.

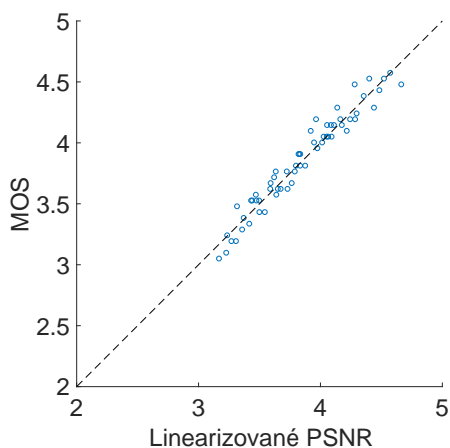


Obr. 4.36: MAD výsledky pre JPEG.

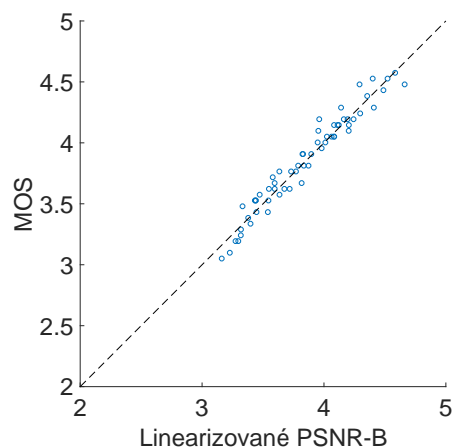


Obr. 4.37: Butteraugli výsledky pre JPEG.

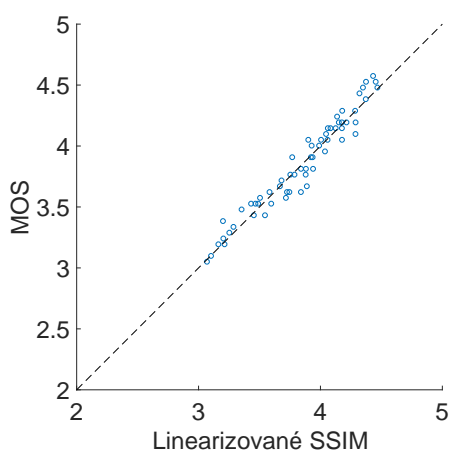
Obrázky 4.38 až 4.43 zobrazujú rovnakú situáciu pre štandard JPEG 2000. Podľa vizuálneho porovnania dosahuje najmenšie odchylenie od stredu metóda Butteraugli spoločne s metódou SSIM a MAD. Podrobné porovnanie bude v tabuľke 4.5 obsahujúcej hodnoty koeficientov a odchýlok.



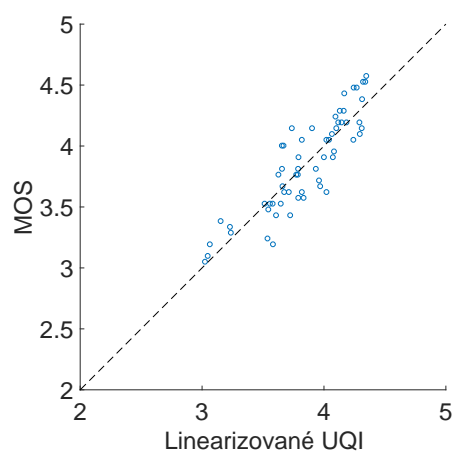
Obr. 4.38: PSNR výsledky pre JP2.



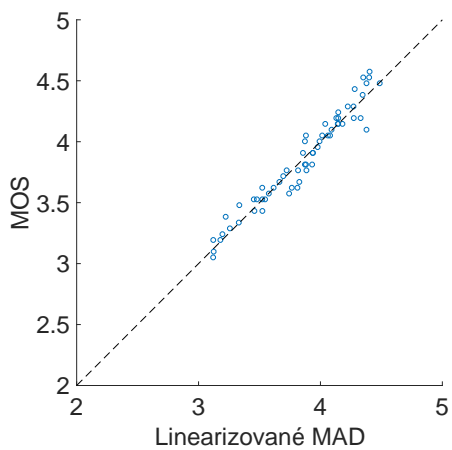
Obr. 4.39: PSNR-B výsledky pre JP2.



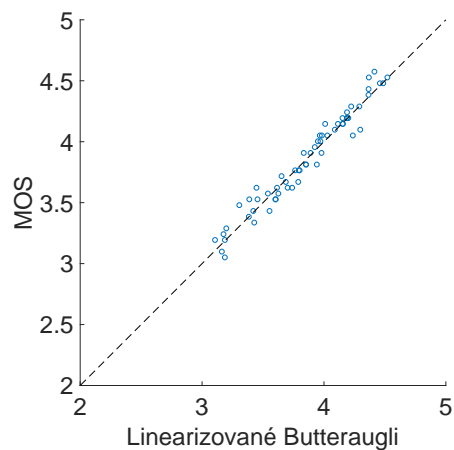
Obr. 4.40: SSIM výsledky pre JP2.



Obr. 4.41: UQI výsledky pre JP2.



Obr. 4.42: MAD výsledky pre JP2.



Obr. 4.43: Butteraugli výsledky pre JP2.

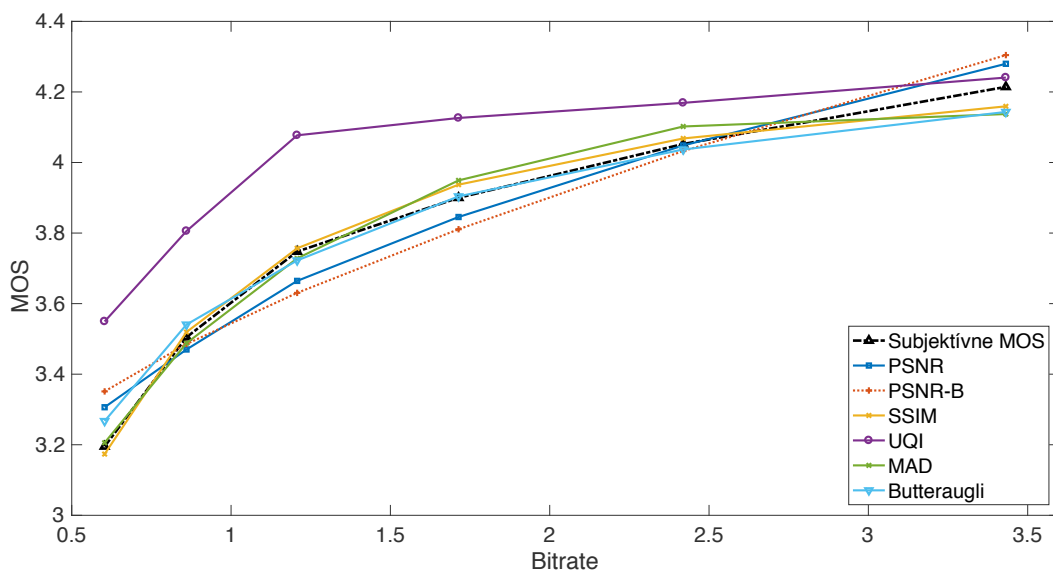
V tabuľkách 4.4 a 4.5 na ďalších stranách sú uvedené výsledky korelačných koeficientov a odchýlok získaných v tomto testovaní. Najvyššia hodnota v stĺpcoch s koeficientami je zvýraznená, takisto najnižšia hodnota v stĺpcoch s odchýlkami. V prípade, že bolo viacero metód s rovnakým výsledkom v jednom stĺpci, nie je zvýraznená žiadna hodnota.

Metóda	KRCC	SRCC	PLCC	MAE	RMSE
PSNR	0,9733	0,9943	0,9630	0,0805	0,0959
PSNR-B	0,9733	0,9943	0,9430	0,0992	0,1142
SSIM	0,9733	0,9943	0,9720	0,0646	0,0774
UQI	0,8933	0,9457	0,8821	0,3186	0,3433
MAD	0,9400	0,9843	0,9688	0,0571	0,0708
Butteraugli	0,9467	0,9771	0,9811	0,0491	0,0618

Tabuľka 4.4: Korelačné koeficienty a odchýlky pre JPEG.

Z hodnôt v tabuľke 4.4 je možné pozorovať, že v prípade obrázkov vytvorených s využitím štandardu JPEG dosiahla pridaná metóda Butteraugli najnižšie odchýlky zo všetkých metód. Zároveň u nej koeficient PLCC presahuje hodnoty tohoto koeficientu pri ostatných metódach. Nedosiahla však najvyššie hodnoty KRCC a SRCC, no hodnoty týchto koeficientov u nej nie sú výrazne nižšie oproti najlepším, napríklad v porovnaní s UQI. Metóda Butteraugli sa podľa tohoto pozorovania javí ako najvhodnejšia pre meranie kvality obrázkov vytvorených štandardom JPEG pri zameraní sa na vyšší bitrate.

Na obrázku 4.44 sú zobrazené linearizované výsledky metód zahrnutých v tomto testovaní pre JPEG obrázky spolu s priemerami subjektívnych hodnotení kvality pre jednotlivé úrovne bitrate. Priebeh výsledkov metód má mierne podobný charakter, až na výsledky metódy UQI.



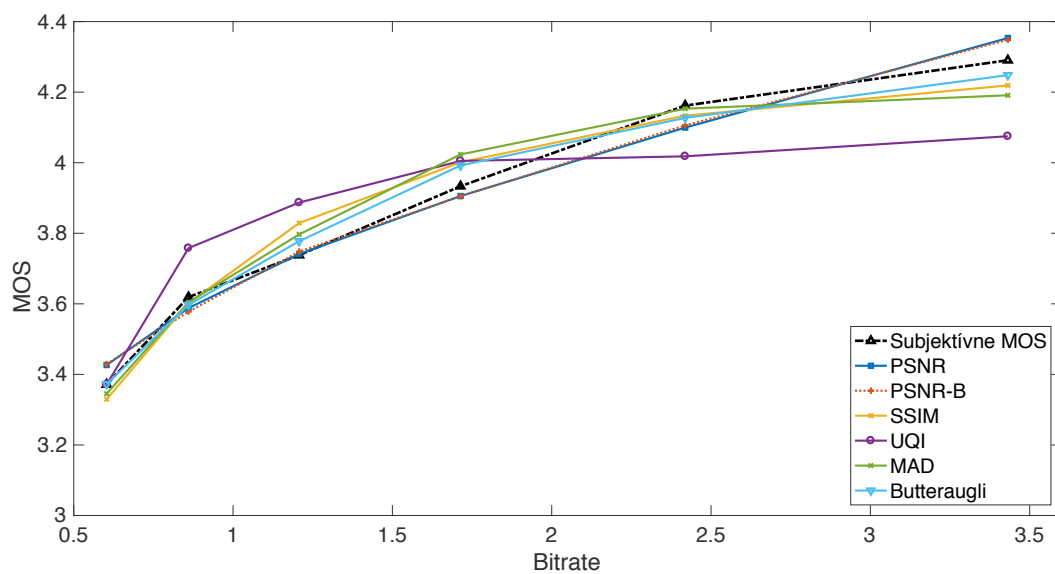
Obr. 4.44: Výsledky hodnotení v stupnici MOS v závislosti na bitrate pre JPEG.

Metóda	KRCC	SRCC	PLCC	MAE	RMSE
PSNR	0,9800	0,9957	0,9644	0,0739	0,0842
PSNR-B	0,9800	0,9957	0,9670	0,0708	0,0809
SSIM	0,9800	0,9957	0,9549	0,0722	0,0863
UQI	0,7600	0,8314	0,7562	0,1502	0,1746
MAD	0,9600	0,9914	0,9382	0,0712	0,0889
Butteraugli	0,9800	0,9957	0,9614	0,0625	0,0711

Tabuľka 4.5: Korelačné koeficienty a odchýlky pre JPEG 2000.

V prípade kompresného štandardu JPEG 2000 najvyššiu hodnotu PLCC dosiahla metóda PSNR-B, aj keď metódy PSNR a Butteraugli mali hodnotu tohoto koeficientu veľmi podobnú. Najmenšie hodnoty odchýlok dosiahla opäť metóda Butteraugli. Najvyššie hodnoty KRCC a SRCC dosiahlo viacero metód, medzi nimi aj Butteraugli. Táto metóda sa podľa nameraných výsledkov javí ako najvhodnejšia aj v tomto prípade.

Obrázok 4.45 ukazuje linearizované výsledky metód spolu s priemerom subjektívnych hodnotení v závislosti na bitrate. Výsledky metód PSNR a PSNR-B majú mierne podobný priebeh a tiež výsledky metód MAD a SSIM a Butteraugli.



Obr. 4.45: Výsledky hodnotení v stupnici MOS v závislosti na bitrate pre JPEG 2000.

Kapitola 5

Záver

Neustále sa zväčšujúci dopyt po technológiach využívajúcich digitálne obrázky vytvára potrebu presných metód objektívne merajúcich kvalitu obrazu. Túto kvalitu ovplyvňuje mnoho faktorov, pričom medzi hlavné patria najmä artefakty spôsobené stratovou kompresiou, ktorá je s použitím štandardu JPEG veľmi rozšírená. Jeho nástupca, novší štandard JPEG 2000 síce poskytuje omnoho väčšie možnosti pri procese komprimácie a ovela viac nastavitelných parametrov, no svoju popularitu si doposiaľ nezískal, za čo môže hlavne vyššia výpočetná náročnosť a problémy s licenciami.

Porovnanie výkonnosti metód malo za cieľ zistiť, ktorá z vybraných metód sa ukáže ako najvhodnejšia na meranie kvality obrázkov vytvorených pomocou týchto kompresných štandardov. Takáto metóda mala ukázať čo najlepšiu koreláciu so subjektívnym hodnotením kvality obrázkov a zároveň dosahovať čo najmenšie odchýlky.

Ako ukázali výsledky v predchádzajúcej kapitole, v prípade že je cieľom zisťovať kvalitu obrázkov so širokým rozpätím bitrate, pre kompresný štandard JPEG sa javí ako najvhodnejšia metóda PSNR. Táto metóda dosiahla najlepšie hodnoty v prípade koeficientu PLCC a odchýlok, pričom hodnoty koeficientov SRCC a KRCC vyšli takisto najvyššie, ale rovnaké v porovnaní s metódou SSIM. Pre testovanie obrázkov vytvorených kompresným štandardom JPEG 2000 a so širokým rozpätím bitrate dosiahli najlepšie výsledky metódy SSIM a MAD. SSIM spolu s PSNR-B dosiahli najvyššie hodnoty KRCC a SRCC, SSIM ale ukázala lepšie výsledky v ostatných hodnotách oproti PSNR-B. MAD na druhej strane dosiahla najvyššiu hodnotu PLCC a najmenšie odchýlky zo všetkých testovaných metód. Je teda na dotyčnom vykonávanom testovaní, či si zvolí metódu SSIM alebo MAD.

Druhé realizované testovanie bolo zamerané na meranie výkonnosti len na vyššom bitrate a bola zahrnutá tiež metóda Butteraugli. Táto metóda dosiahla najvyššiu hodnotu koeficientu PLCC a tiež najmenšie odchýlky zo všetkých metód v prípade JPEG obrázkov. Pre JPEG 2000 dosiahla metóda Butteraugli takisto najmenšie odchýlky. Najvyššiu hodnotu koeficientu PLCC dosiahla metóda PSNR-B, pričom hodnoty pre PSNR a Butteraugli sa k nej veľmi blížili. Butteraugli sa podľa výsledkov druhého testovania javí ako najvhodnejšia metóda pre oba kompresné štandardy, pokiaľ je cieľom zistiť kvalitu obrázkov s bitrate vyšším ako 0,6 bitov/pixel.

Existuje viacero príležitostí pre budúci rozvoj práce, napríklad otestovať aj iné metódy, ktoré neboli zahrnuté do testovania, nakoľko táto práca všetky existujúce nepokrýva. Takisto by bolo možné uskutočniť testovanie aj na ďalších datasetoch, prípadne upravovať nastavenia nástroja na vytváranie obrázkov pomocou kompresného štandardu JPEG 2000.

Literatúra

- [1] Alakuijala, J.; Obryk, R.; Stoliarchuk, O.; aj.: Guetzli: Perceptually Guided JPEG Encoder. [Online; navštíveno 14.4.2018].
URL <https://arxiv.org/abs/1703.04421>
- [2] Balan, V.; Condea, C.: *Wavelets and Image Compression*. Telecommunication Standardization Sector of ITU, Január 2003.
- [3] Google: Google Butteraugli. [Online; navštíveno 14.4.2018].
URL <https://github.com/google/butteraugli>
- [4] Khayam, S. A.: *The Discrete Cosine Transform (DCT): Theory and Application*. Department of Electrical and Computer Engineering, Michigan State University, 2003.
- [5] Korhonen, J.; You, J.: Peak signal-to-noise ratio revisited: Is simple beautiful? In *2012 Fourth International Workshop on Quality of Multimedia Experience*, Júl 2012, s. 37–38.
- [6] Larson, E. C.; Chandler, D. M.: Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging*, ročník 19, č. 1, 2010: s. 011006:1–011006:21.
- [7] Mohammadi, P.; Ebrahimi-Moghadam, A.; Shirani, S.: Subjective and Objective Quality Assessment of Image: A Survey. *Majlesi Journal of Electrical Engineering*, ročník 9, č. 1, Marec 2015: s. 55–83.
- [8] Sheikh, H.; Z.Wang; Cormack, L.; aj.: LIVE Image Quality Assessment Database Release 2. [Online; navštíveno 05.11.2017].
URL <http://live.ece.utexas.edu/research/quality>
- [9] Sheikh, H. R.; Bovik, A. C.: Image information and visual quality. *IEEE Transactions on Image Processing*, ročník 15, č. 2, Február 2006: s. 430–444, ISSN 1057-7149.
- [10] Taubman, D.: High performance scalable image compression with EBCOT. *IEEE Transactions on Image Processing*, ročník 9, č. 7, Júl 2000: s. 1158–1170, ISSN 1057-7149.
- [11] Wainwright, M. J.; Simoncelli, E. P.; Wilsky, A. S.: Random Cascades on Wavelet Trees and Their Use in Analyzing and Modeling Natural Images. *Applied and Computational Harmonic Analysis*, ročník 11, č. 1, November 2001: s. 89–123, ISSN 1063-5203.

- [12] Wang, Z.; Bovik, A. C.: A universal image quality index. *IEEE Signal Processing Letters*, ročník 9, č. 3, Marec 2002: s. 81–84, ISSN 1070-9908.
- [13] Wang, Z.; Bovik, A. C.; Sheikh, H. R.; aj.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, ročník 13, č. 4, Apríl 2004: s. 600–612, ISSN 1057-7149.
- [14] Wang, Z.; Simoncelli, E. P.; Bovik, A. C.: Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, ročník 2, November 2003, s. 1398–1402.
- [15] Winkler, S.: *On the properties of subjective ratings in video quality experiments*. IEEE, Quality of Multimedia Experience, 2009.
- [16] Yim, C.; Bovik, A. C.: Quality Assessment of Deblocked Images. *IEEE Transactions on Image Processing*, ročník 20, č. 1, Január 2011: s. 88–98, ISSN 1057-7149.
- [17] Zhang, L.: Evaluation of VIF. [Online; navštíveno 22.2.2018].
URL http://sse.tongji.edu.cn/linzhang/IQA/Evaluation_VIF/eva-VIF.htm