



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

VYUŽITÍ VLNKOVÉ TRANSFORMACE PŘI ZPRACOVÁNÍ OBRAZU

WAVELET TRANSFORM IN IMAGE PROCESSING

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. MARTIN DOSTÁL

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. DAVID BAŘINA

BRNO 2015

Abstrakt

Vlnková transformace je využívána již několik desetiletí a je stále předmětem výzkumu – především její modifikace používající tzv. vlnky druhé generace. Oproti ostatním integrálním transformacím má několik výhod, těmi nejdůležitějšími jsou schopnost lokalizace zároveň v čase a frekvenci a schopnost dobře dekorelovat některé reálné nestacionární signály – např. obraz. Z těchto důvodů se vlnková transformace stala často využívaným nástrojem při řešení mnohých úkolů ve zpracování obrazu, jako je např. jeho komprese, detekce hran nebo vylepšení kontrastu. V této práci je nejprve vysvětlen teoretický základ samotné vlnkové transformace a její implementace pro použití s dvourozměrnými diskrétními signály. V další části je představeno a popsáno několik aplikací vlnkové transformace – detekce hran, odstranění šumu, vylepšení kontrastu, zobrazování HDR a několik dalších. Provedené testy potvrdily vhodnost využití vlnkové transformace ve vybraných aplikacích zpracování obrazu.

Abstract

The wavelet transform has been used for several decades and it is still an object of research – especially its recent modifications which are using the so-called second generation wavelets. It has several advantages over other integral transformations. The most important of them are the ability to localize both in time and frequency and an ability to decorrelate some real non-stationary signals such as images. For this reasons, the wavelet transform became an often used tool in many image processing tasks, for example in image compression, edge detection or contrast enhancement. In this thesis, the wavelet transform is explained, including the theoretical foundation and implementation for use with two-dimensional discrete signals. Some of the applications of the wavelet transform are presented and described. The wavelet transform showed to be suitable tool for edge detection, noise reduction, contrast enhancement and HDR compression.

Klíčová slova

vlnková transformace, zpracování obrazu, detekce hran, odstranění šumu, vylepšení kontrastu, zobrazování HDR

Keywords

wavelet transform, image processing, edge detection, noise reduction, contrast enhancement, HDR compression

Citace

Martin Dostál: Využití vlnkové transformace při zpracování obrazu, diplomová práce, Brno, FIT VUT v Brně, 2015

Využití vlnkové transformace při zpracování obrazu

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením Ing. Davida Bařiny.

.....
Martin Dostál
27. května 2015

Poděkování

Rád bych poděkoval mému vedoucímu diplomové práce, Ing. Davidu Bařinovi, za jeho odborné vedení a cenné rady.

© Martin Dostál, 2015.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

1	Úvod	2
2	Vlnková transformace	3
2.1	Diskrétní vlnková transformace	5
2.2	Stacionární vlnková transformace	21
3	Aplikace vlnkové transformace	23
3.1	Detekce hran	23
3.2	Odstranění šumu	28
3.3	Vylepšení kontrastu	32
3.4	Zobrazování HDR	35
3.5	Ostatní aplikace	37
4	Testování a výsledky	42
4.1	Detekce hran	42
4.2	Odstranění šumu	45
4.3	Vylepšení kontrastu	49
4.4	Zobrazování HDR	52
5	Závěr	55
A	Obsah CD	56

Kapitola 1

Úvod

Vlnková transformace je jednou z nejpoužívanějších integrálních transformací. Její využití najdeme například ve zpracování signálů, statistice, ale i lékařství, ekonomice a dalších odvětvích. V této práci se zaměřím na využití vlnkové transformace ve zpracování obrazu. V počítačové grafice je obraz chápán jako konečný dvojrozměrný signál. Pro provedení jeho frekvenční analýzy obvykle využijeme jedné z integrálních transformací, jako je Fourierova, kosinová, nebo vlnková transformace. Poslední jmenovaná má výhodu spočívající ve schopnosti s určitou přesností časově lokalizovat výskyt určité frekvence (platí zde Heisenbergův princip neurčitosti), oproti pouhé globální informaci o výskytu frekvencí v obraze u transformací předchozích. Jedná se tedy o časově-frekvenční analýzu. Této vlastnosti se dá využít v mnohých aplikacích, například při kompresi obrazu, detekci hran, odstranění šumu, vylepšení kontrastu, kolorizaci nebo tzv. inpaintingu.

V této práci je nejprve popsána samotná vlnková transformace (kapitola 2), včetně její diskrétní a stacionární verze (kapitoly 2.1 a 2.2), které jsou nejčastěji používané v praktických aplikacích. Také jsou zde popsány vlastnosti vlnek a vybrané vlnky jsou podrobněji představeny. Nakonec jsou představeny a podrobně popsány vybrané metody pro detekci hran (kap. 3.1), odstranění šumu (kap. 3.2), vylepšení kontrastu (kap. 3.3) a zobrazování HDR¹ obrazu (kap. 3.4). V kapitole 3.5 je představeno několik dalších aplikací, které už ale nebudou podrobnějším předmětem výzkumu této práce. Kapitola 4 je zaměřena především na testování různých parametrů vybraných metod a na srovnání využití vlnkové transformace v daných aplikacích s ostatními používanými přístupy.

¹High Dynamic Range, obraz s vysokým dynamickým rozsahem

Kapitola 2

Vlnková transformace

V oblasti zpracování signálů po dlouhou dobu dominovala (a je stále často používaná) Fourierova transformace (FT). S její pomocí získáme informaci o frekvencích obsažených v analyzovaném signálu, nicméně nedokážeme zjistit čas, ve kterém se tyto frekvence vyskytují. To je dáno povahou bázové funkce, kterou je v FT komplexní exponenciála, tedy periodická funkce s energií rozloženou po celém jejím průběhu. Kvůli této vlastnosti nelze získat odezvu od konkrétního místa (času) v signálu. Poměrně jednoduchým řešením je neprovádět analýzu celého signálu, ale postupně po menších částech, definovaných plovoucím oknem. Tento přístup je znám jako krátkodobá Fourierova transformace (short-time Fourier transform, STFT). Zde ale nastává problém s volbou velikosti plovoucího okna. Čím je okno menší, tím je lepší časové rozlišení analýzy. Se zmenšujícím se oknem ale také ubývá množství frekvencí, které lze ve zkoumaném signálu identifikovat. Toto platí samozřejmě i naopak, tedy se zvětšujícím se oknem se zvyšuje frekvenční rozlišení a zmenšuje rozlišení časové. Tento problém je úzce svázan s tzv. Heisenbergovým principem neurčitosti. [36]

Tento nedostatek STFT řeší vlnková transformace (wavelet transform, WT) volbou jiné bázové funkce – vlnky s většinou energie umístěné v malém konečném intervalu. Rozdíl mezi charakteristikami jednotlivých transformací – FT, STFT a WT je znázorněn na obrázku 2.1

Spojité vlnkové transformace rozkládá signál pomocí dilatovaných a posunutých vlnek. Vlnka ψ je funkce z prostoru funkcí integrovatelných s druhou mocninou ($\psi \in L^2(\mathbb{R})$) s následujícími vlastnostmi [26]:

- má nulový průměr (tedy osciluje):

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2.1)$$

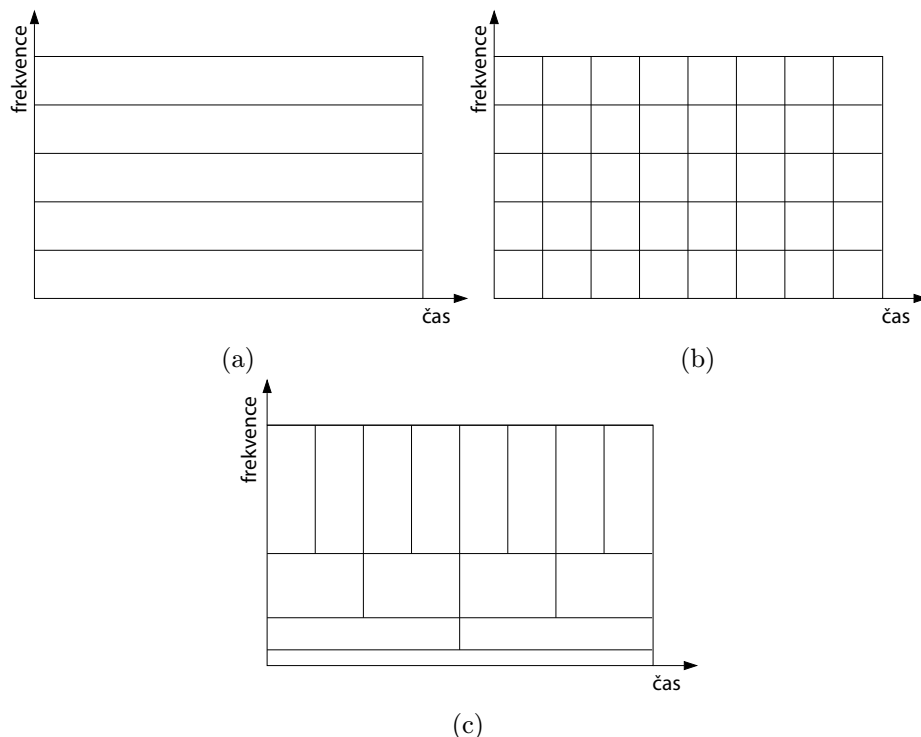
- je normalizovaná:

$$\|\psi\| = 1 \quad (2.2)$$

- je centrovaná v čase $t = 0$

Další, obecnější vlastnosti definoval ve své práci Wim Sweldens [33]. Podle něj musí funkce splňovat následující tři podmínky, aby mohla být nazývána vlnkovým systémem:

- Vlnky jsou základními stavebními bloky pro obecné funkce. Tedy funkce může být reprezentována ve vlnkovém prostoru jako nekonečná řada vlnek.



Obrázek 2.1: Porovnání časového a frekvenčního rozlišení jednotlivých transformací: (a) FT, (b) STFT, (c) WT

- Vlnky mají schopnost časově-frekvenční lokalizace. Většina jejich energie je koncentrována v konečném intervalu a transformace obsahuje pouze frekvence z určitého frekvenčního pásma.
- Vlnky podporují rychlé a účinné transformační algoritmy. Vlnkové transformace často potřebují pouze $O(n)$ operací, což znamená, že počet násobení a sčítání roste lineárně s délkou signálu.

Množinu \mathcal{D} bázových funkcí – vlněk posunutých o u a dilatovaných faktorem s lze popsat rovnicí 2.3.

$$\mathcal{D} = \left\{ \psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right) \right\}_{u \in \mathbb{R}, s \in \mathbb{R}^+} \quad (2.3)$$

Její dosazením do rovnice obecné integrální transformace $F(n)$ (2.4) potom získáme rovnici spojité vlnkové transformace (CWT) funkce $f \in L^2(\mathbb{R})$ (2.5):

$$F(n) = \int f(t) \psi(t, n) dt \quad (2.4)$$

$$\text{W} f(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt \quad (2.5)$$

Analogicky lze definovat rovnici pro inverzní vlnkovou transformaci:

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{s}} \text{W} f(u, s) \psi \left(\frac{t-u}{s} \right) \frac{du ds}{s^2}, \quad (2.6)$$

kde ψ je duální funkcí k ψ^* . Rekonstrukce je možná pouze pokud je definován faktor C_ψ , který je nazýván konstanta přípustnosti.

Spojité vlnkové transformace v podobě dle rovnice 2.5 je ale vysoce redundantní, díky čemuž její přímá implementace nemá velký význam.

2.1 Diskrétní vlnková transformace

Jak již bylo zmíněno, CWT je kvůli své redundanci nevhodná pro přímou implementaci. Pro praktické použití byla odvozena její varianta pro diskrétní vlnky – diskrétní vlnková transformace (DWT). Pro její popis je nejprve potřeba definovat koncept multirozkladu a s ním spojené škálovací a vlnkové funkce.

2.1.1 Multirozklad

Koncept multirozkladu (neboli víceúrovňové analýzy) byl poprvé zaveden Stéphane Mallattem v roce 1989 [28]. Pomocí vnořených podprostorů a škálovací funkce definuje vztah mezi vlnkovými funkcemi a bankami filtrů, které se používají v diskrétní vlnkové transformaci. [36]

Vnořené podprostory

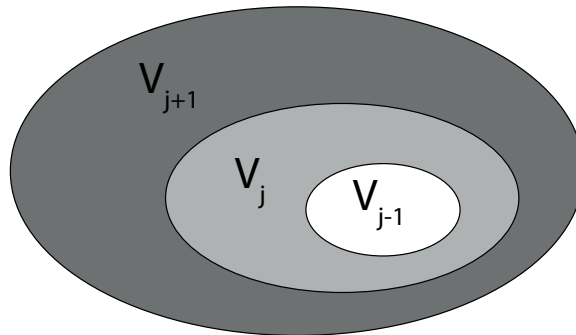
Základní podmínkou pro multirozklad je existence množiny podprostorů $L^2(\mathbb{R})$ s různou úrovní rozlišení:

$$V_{-\infty} \cdots \subset V_{-1} \subset V_0 \subset \dots \subset V_\infty = L^2(\mathbb{R}) \quad (2.7)$$

Tento vztah je pro tři po sobě jdoucí podprostory znázorněn na obr. 2.2.

Pokud $f(t) \in V_j$, pak $f(2t) \in V_{j+1}$, tedy podprostor obsahující vyšší rozlišení obsahuje ty s nižším rozlišením. Obecně platí

$$f(t) \in V_0 \implies f(2^k t) \in V_k \quad (2.8)$$



Obrázek 2.2: Zanořené podprostory

Škálovací a vlnková funkce

Škálovací funkce je nutná k využití výhod víceúrovňové analýzy. Je definovaná rovnicí

$$\varphi_k(t) = \varphi(t - k) \quad k \in \mathbb{Z} \quad \varphi \in L^2(\mathbb{R}) \quad (2.9)$$

Se svými posunutými verzemi tvoří ortonormální bázi prostoru V_0 . To znamená, že kterákoliv funkce z tohoto prostoru může být vyjádřena lineární kombinací škálovacích koeficientů se škálovací funkcí:

$$f(t) = \sum_k c_k \varphi_k(t) = \sum_k c(k) \varphi(t - k) \quad (2.10)$$

Podobně jako škálovací funkce je definována tzv. vlnková funkce:

$$\psi_k(t) = \psi(t - k) \quad k \in \mathbb{Z} \quad \psi \in L^2(\mathbb{R}) \quad (2.11)$$

Vlnková funkce se svými posunutými verzemi tvoří ortonormální bázi prostoru W_0 , který je doplňkem $V_0 \subset V_1$:

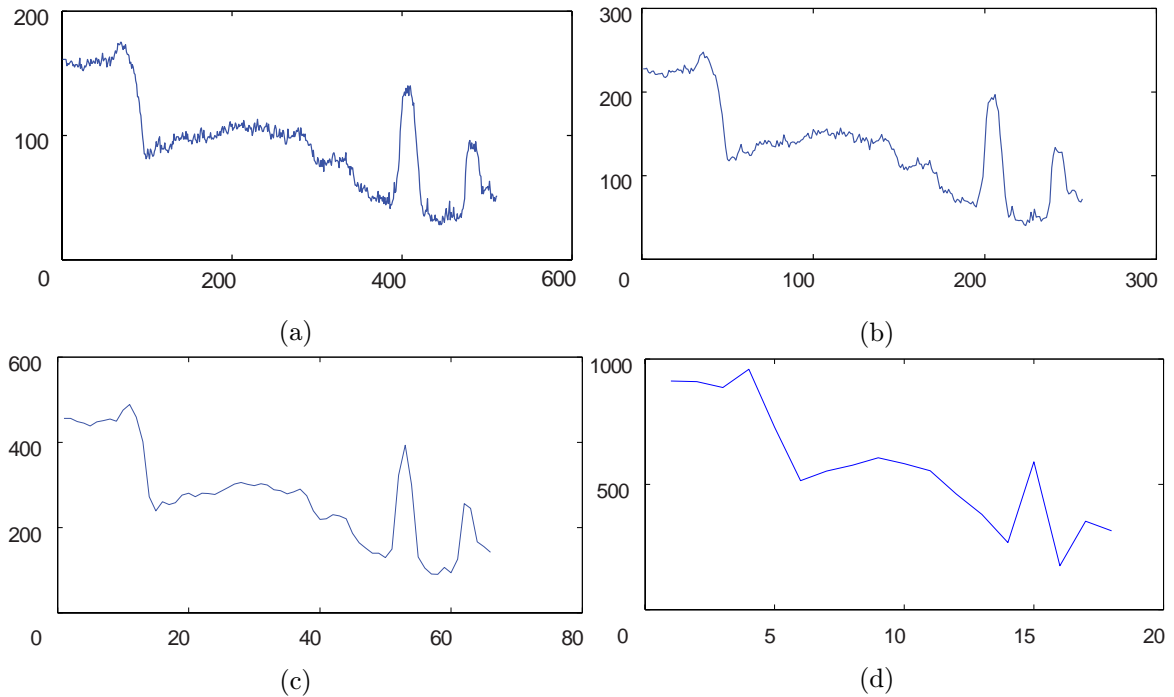
$$V_1 = V_0 \oplus W_0 \quad (2.12)$$

Funkce $\psi(t)$ se někdy nazývá také mateřská vlnka a podobně jako u škálovací funkce platí:

$$f(t) = \sum_k d_k \psi_k(t) = \sum_k d(k) \psi(t - k), \quad (2.13)$$

kde $d(k)$ jsou tzv. vlnkové koeficienty.

Na obr. 2.3 je několik úrovní rozkladu reálného 1D signálu. Zobrazeny jsou aproximační koeficienty získané škálovací funkcí Daubechies (viz obr. 2.13).



Obrázek 2.3: Pohled na reálný 1D signál v různých měřících. (a) Původní signál. (b) - (d) 1., 3. a 5. úroveň rozkladu škálovací funkcí Daubechies. Platí zde vztah z rovnice 2.8, kde původní signál odpovídá f_0 a $k = \{1, 3, 5\}$.

2.1.2 Definice DWT

Spojitou vlnkovou transformaci lze částečně diskretizovat omezením koeficientů s a u na celá čísla následujícím způsobem:

$$s = s_0^m, u = s_0^m b_0, \quad (2.14)$$

kde $a_0 > 1, b_0 > 0$ a $m \in \mathbb{Z}, n \in \mathbb{Z}$. Nejčastěji se volí hodnoty $s = 2^m, u = n2^m$, výsledná transformace se potom nazývá dyadická vlnková transformace a její rovnice má tvar

$$W f(m, n) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^m}} \psi^*(2^{-m}t - n) dt, \quad (2.15)$$

kde hodnoty m a n se nazývají kmitočtové resp. časové měřítko. Tato transformace ale stále pracuje se spojitým časem, počet dilatací tedy není omezen.

Jakoukoliv funkci z $L^2(\mathbb{R})$ lze rozložit pomocí škálovací funkce a vlnkových funkcí definovaných v kapitole 2.1.1 (viz rovnice 2.16, více v [36]). Úpravou vztahu pro dyadickou transformaci (2.15) potom můžeme konečně získat rovnici 2.17 pro dyadickou vlnkovou transformaci s diskretním časem (často označovanou pouze jako diskretní vlnková transformace).

$$f(t) = \sum_k c_j(k) \varphi_{j,k}(t) + \sum_{n=j}^{+\infty} \sum_{k=-\infty}^{+\infty} d_n(k) \psi_{n,k}(t) \quad (2.16)$$

$$y_m[n] = \sum_{k=-\infty}^{\infty} f[k] h_m[2^m n - k] \quad (2.17)$$

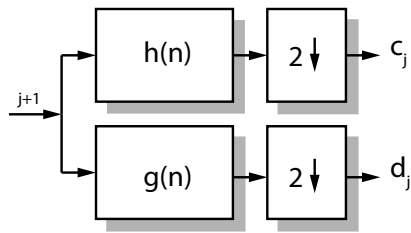
Index j , který lze zvolit, představuje „nejhrubější“ úroveň detailů. Koeficienty $c_j(k)$ a $d_n(k)$ jsou výstupem diskretní vlnkové transformace funkce $f(t)$.

h_m představuje impulzní charakteristiku filtru, který odpovídá vlnce ψ_m^* pro danou úroveň m .

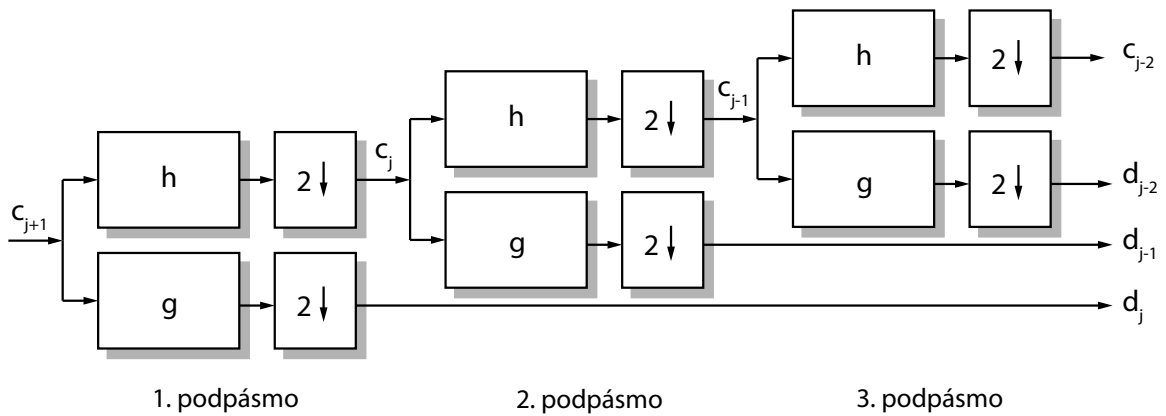
2.1.3 Implementace bankou filtrů

V praktických aplikacích využívajících vlnkovou transformaci je potřeba pracovat s diskretními koeficienty namísto škálovací a vlnkových funkcí. K tomu jsou potřeba tzv. banky filtrů, které rozkládají signál na jeho přibližnou část a část s detaily. Banka filtrů je tvořena dvojicí zrcadlových FIR¹ filtrů – z dolní propusti ($h(n)$, odpovídá škálovací funkci φ) a horní propusti ($g(n)$, odpovídá vlnkové funkci ψ), následovaných decimátory, které provádějí podvzorkování faktorem 2. Dekompozice signálu tak sestává ze dvou operací – filtrování a podvzorkování. Banka filtrů je znázorněna na obr. 2.4. Pro provedení více úrovní rozkladu se banky filtrů řadí kaskádovitě za sebe, jak znázorňuje obr. 2.5. Toto schéma je známé jako kódování podpásem (subband coding).

¹Finite impulse response – filtr s konečnou impulsní odezvou

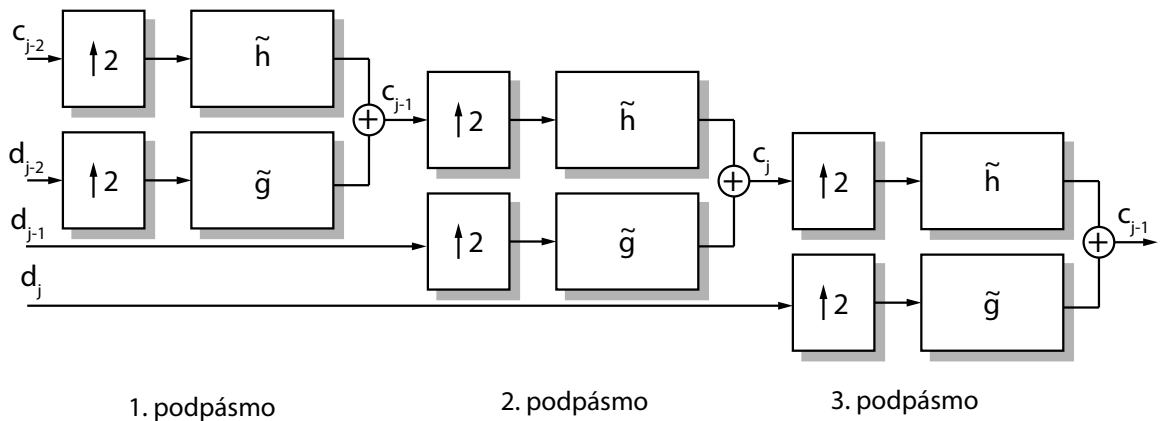


Obrázek 2.4: Banka rozkladových filtrů



Obrázek 2.5: Tříúrovňová kaskáda filtrů pro rozklad signálu

Pro sestavení původního signálu ze škálovacích a detailních koeficientů existuje obrácená verze banky filtrů. Ta provádí nejprve nadzorkování faktorem 2 a následně filtrování příslušnými rekonstrukčními filtry ($\tilde{g}(n)$ a $\tilde{h}(n)$). Tříúrovňová kaskáda filtrů pro znovusestavení signálu je zobrazena na obr. 2.6.



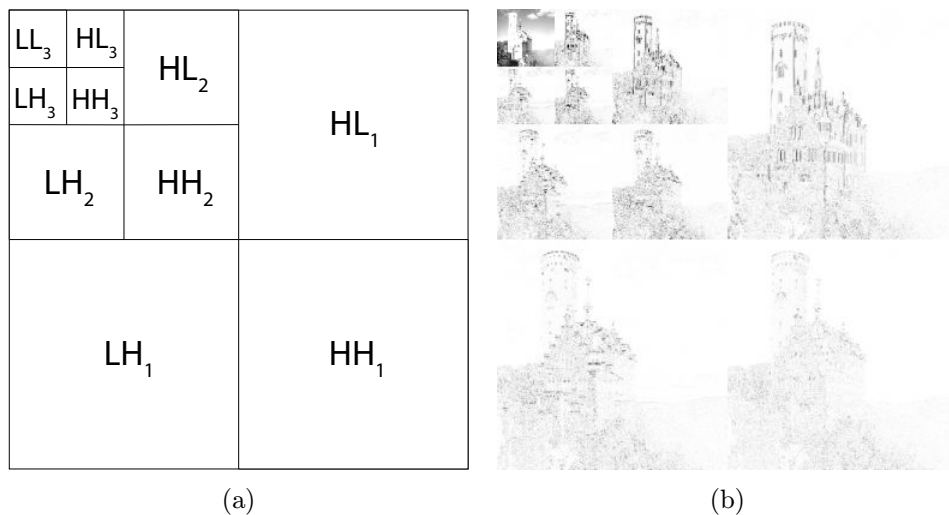
Obrázek 2.6: Tříúrovňová kaskáda filtrů pro zpětnou rekonstrukci signálu

2.1.4 2D diskretní vlnková transformace

Pro účely zpracování obrazu je vhodnější použít dvourozměrnou verzi diskretní vlnkové transformace. Existuje několik implementací 2D DWT, jedna z nejpoužívanějších je tzv. Mallatův pyramidový algoritmus. [26] V něm se 2D signál pro každou úroveň rozkladu rozdělí do 4 podpásem a na každé z nich se použije jiný filtr:

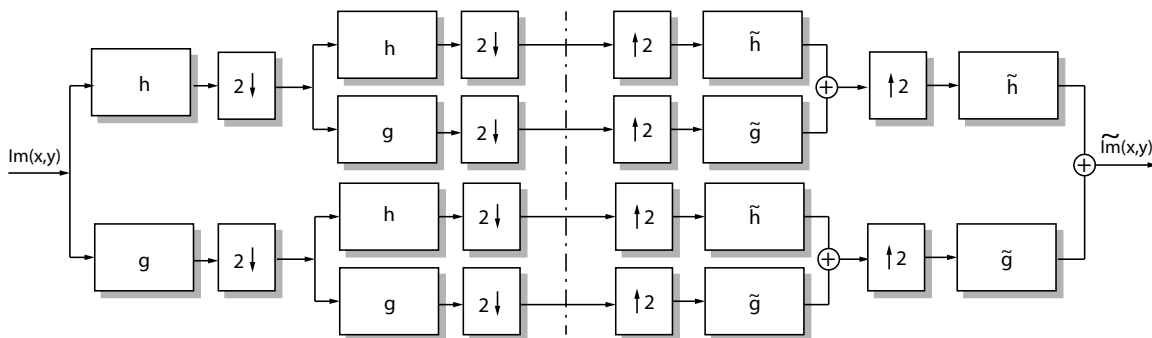
$$\begin{aligned}
 \varphi(x, y) &= \varphi(x)\varphi(y) \\
 \psi^H(x, y) &= \psi(x)\varphi(y) \\
 \psi^V(x, y) &= \varphi(x)\psi(y) \\
 \psi^D(x, y) &= \psi(x)\psi(y)
 \end{aligned}
 \tag{2.18}$$

Postup je tedy podobný jako při 1D transformaci, pouze se signál zpracovává pro každý směr (x a y) zvlášť. Výsledkem je potom 1 podobraz s aproximačními koeficienty a $3n$ podobrazů s detailními koeficienty pro n úrovní rozkladu (viz obr. 2.7).



Obrázek 2.7: Dělení 2D signálu do podpásem podle Mallatova algoritmu pro 3 úrovně rozkladu. (a) schéma (b) reálný signál (fotografie Lichtenštejnského hradu)

Na obrázku 2.8 jsou znázorněny kaskády filtrů implementujících Mallatův algoritmus pro rozklad a zpětnou rekonstrukci obrazu.



Obrázek 2.8: Kaskády filtrů implementující Mallatův pyramidový algoritmus.

2.1.5 Lifting

Lifting je výpočetní schéma DWT. Představil jej v r. 1995 Wim Sweldens [32]. Používá se díky rychlosti výpočtu transformace (až dvojnásobně oproti klasické implementaci bankou filtrů) a také díky možnosti konstrukce tzv. vlnek druhé generace.

Schéma liftingu se skládá ze tří fází – rozdělení, predikce a aktualizace. V první fázi jsou vstupní data $a^0[n]$, kde index 0 značí nejvyšší úroveň, **rozdělena** do dvou podmnožin C a F (hodnoty signálu odpovídající těmto množinám jsou potom značeny $a_C^0[n]$ a $a_F^0[n]$), kde C značí aproximační a F detailní data. Způsob dělení ani výsledné velikosti podmnožin nejsou dány, nicméně nejčastěji se signál dělí na sudé a liché vzorky, čímž se zajistí vysoká korelace mezi množinami C a F . V další fázi jsou hodnoty a_C^0 použity k **predikci** detailních hodnot a_F^0 . Je třeba nalézt operátor $\mathcal{P} : C \mapsto F$, využívající korelace mezi oběma množinami k predikci. Chyba predikce

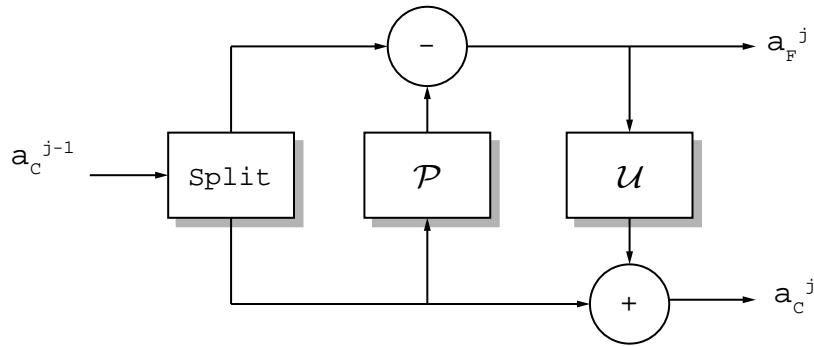
$$a_F^1[n] = a_F^0[n] - \mathcal{P}(a_C^0)[n] \quad (2.19)$$

je potom použita jako detailní koeficient pro další úroveň rozkladu. Protože často je potřeba, aby koeficienty a_C^j pro všechny úrovně zachovávaly určitou vlastnost původního signálu (v případě obrazu např. stejnou průměrnou intenzitu), zavádí se poslední krok, **aktualizace**. Tu provádí další operátor, $\mathcal{U} : F \mapsto C$, který průměruje aproximační koeficient pomocí okolních detailních koeficientů:

$$a_C^1[n] = a_C^0[n] + \mathcal{U}(a_F^1)[n] \quad (2.20)$$

Hodnoty $a_C^1[n]$ jsou potom aproximační koeficienty další úrovně rozkladu. Nyní je transformace kompletní. Výše popsané kroky mohou být podobně jako u klasické implementace DWT rekurzivně prováděny pro více úrovní rozkladu.

K provedení inverzní transformace stačí obrátit pořadí operací a zaměnit odčítání za sčítání a naopak. Schéma liftingu je znázorněno na obr. 2.9.



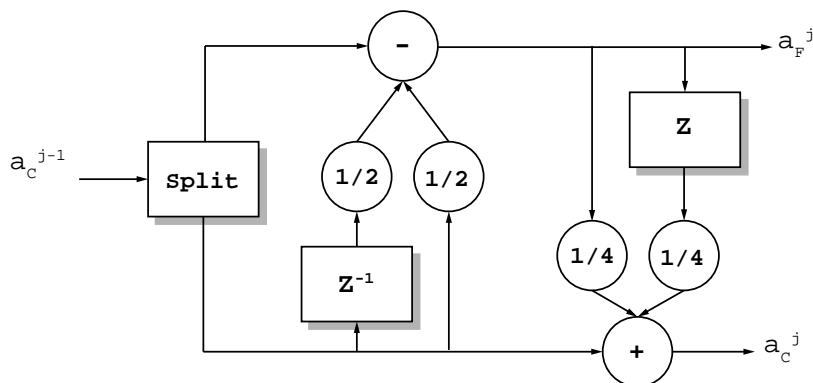
Obrázek 2.9: Schéma liftingu – tři fáze: rozdělení (split), predikce a aktualizace.

Příklad pro vlnku CDF 5/3

V práci Geerta Uytterhoevena et al. [38] je popsán postup konstrukce biortogonálních vlnek pomocí liftingu. Signál je nejprve rozdělen na sudé (a_C^0) a liché (a_F^0) vzorky. Pro vlnku CDF 5/3 použitou například ve standardu JPEG 2000 (nědy také označovanou jako biortogonální vlnka 2,2), jsou potom provedeny kroky liftingu (predikce a aktualizace) s operátory:

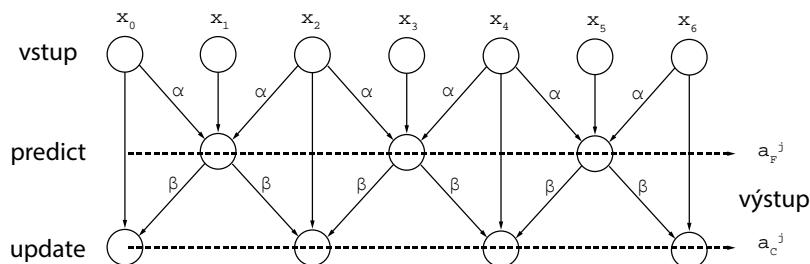
$$\begin{aligned} \mathcal{P}(a_C^j)[n] &= \frac{1}{2} (a_C^j[n-1] + a_C^j[n+1]) \\ \mathcal{U}(a_F^{j+1})[n] &= \frac{1}{4} (a_F^{j+1}[n-1] + a_F^{j+1}[n+1]) \end{aligned} \quad (2.21)$$

Pro tuto konkrétní vlnku potom vypadá schéma liftingu následovně:



Obrázek 2.10: Schéma liftingu pro CDF 5/3.

Schéma lze pro lepší představu znázornit pomocí toku dat v jednotlivých krocích liftingu:



Obrázek 2.11: Tok dat při liftingu vlnkou CDF 5/3. Řádky představují kroky liftingu, sloupce jednotlivé vzorky signálu.

x_i jsou vzorky signálu předchozí úrovně rozkladu (tedy a_C^{j-1}), hodnoty α a β závisí na zvolené vlnce. Pro CDF 5/3 je $\alpha = -\frac{1}{2}$, $\beta = \frac{1}{4}$.

2.1.6 Druhy vlnek

Základní vlastnosti vlnek byly popsány v kapitole 2. Vlnkových funkcí s různými vlastnostmi již bylo vyvinuto nepřeberné množství. Při návrhu aplikace využívající vlnkovou transformaci je důležité zvolit vlnku vhodnou právě pro danou konkrétní aplikaci. Všímáme si především těchto vlastností: [1]

- Existence kompaktního nosiče

Vlnka s kompaktním nosičem je omezená v časové doméně a teoreticky neomezená v doméně frekvenční. V praxi to znamená, že nenulové koeficienty se vyskytují pouze v krátkém časovém úseku, což snižuje výpočetní náročnost transformace (za cenu sníženého frekvenčního rozlišení). Vlnky s kompaktním nosičem lze implementovat FIR filtry. Opakem jsou úzkopásmové vlnky, které jsou omezené ve frekvenční, nikoli časové doméně a implementují se IIR filtry.

- Symetrie

V některých aplikacích (např. komprese obrazu, redukce šumu) je výhodné použít symetrické nebo antisymetrické vlnky, díky své vlastnosti zachování lineární fáze. Konstruuji se pomocí FIR filtrů s lineární fází.

- Počet nulových momentů

m -tý moment vlnky je definován jako $\int t^m \psi(t) dt$. Pokud prvních M momentů je nulových, všechny signály ve tvaru

$$x(t) = \sum_{0 < m < M} c_m t^m \quad (2.22)$$

budou mít (téměř) nulové detailní koeficienty. Toho se dá využít např. při kompresi obrazu, kdy při volbě vlnky s velkým počtem nulových momentů lze dosáhnout vysokého kompresního poměru.

- Hladkost

Hladkost neboli regularita je úzce spojená s počtem nulových momentů – čím více nulových momentů vlnka obsahuje, tím je hladší. Hladkost zaručuje numerickou stabilitu a lepší rekonstrukční vlastnosti. Je důležitá například pro kompresi obrazu nebo řešení diferenciálních rovnic.

V následujícím textu budou uvedeny některé z nejpoužívanějších diskretních vlnkových funkcí pro různé aplikace.

Haar

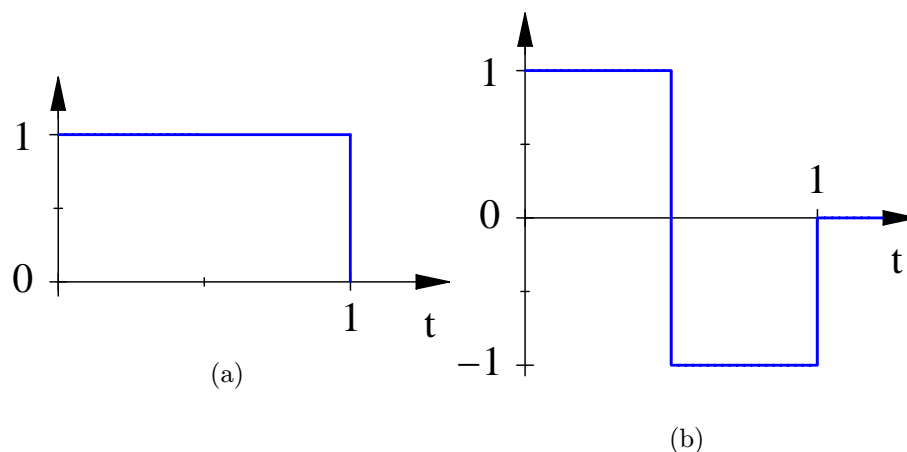
Haarova funkce (Alfred Haar, 1909) je nejjednodušší možnou vlnkovou funkcí:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2, \\ -1 & 1/2 \leq t < 1, \\ 0 & \text{jinde.} \end{cases} \quad (2.23)$$

Její škálovací funkce je:

$$\varphi(t) = \begin{cases} 1 & 0 \leq t < 1, \\ 0 & \text{jinde.} \end{cases} \quad (2.24)$$

Přesto, že byla vyvinuta dávno před samotnou vlnkovou transformací, lze ji pro její jednoduchost a efektivitu použít v aplikacích jako je komprese signálu. Nicméně pro lepší výsledky je vhodné zvolit některou ze složitějších, později vyvinutých vlnek. Funkce jsou znázorněny na obr. 2.12.



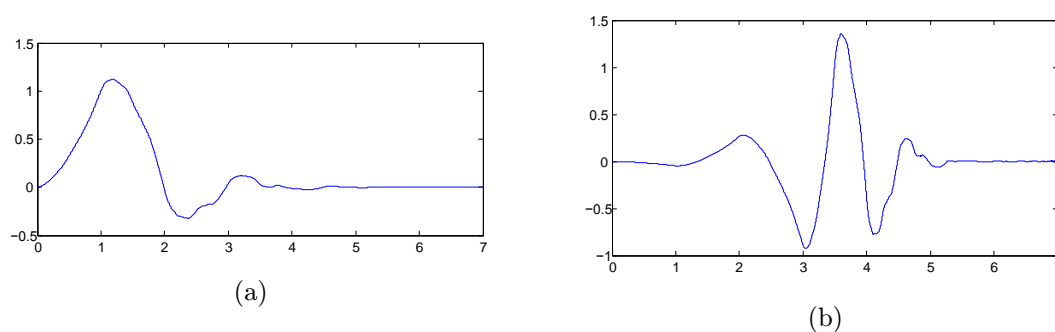
Obrázek 2.12: Haarova škálovací (a) a vlnková (b) funkce

Daubechies

Vlnky Daubechies jsou pojmenovány podle jejich autorky, belgické matematicky Ingrid Daubechies. Jedná se o rodinu ortogonálních vlnek s maximálním počtem nulových momentů pro danou velikost nosiče. Používají se v diskretní vlnkové transformaci. Nejsou explicitně definovány, koeficienty filtrů se musí spočítat. Ke každé vlnkové funkci existuje škálovací funkce, se kterou dohromady definují multirozklad signálu. Vlnky Daubechies řádu p jsou:

- asymetrické (kromě $p = 1$)
- ortogonální
- délka filtru je $2p$
- mají p nulových momentů

Vlnka Daubechies řádu 1 je Haarova funkce (2.1.6). Příklad Daubechies vlnkové a škálovací funkce pro $p = 4$ je uveden na obr. 2.13.



Obrázek 2.13: Škálovací (a) a vlnková (b) funkce Daubechies řádu 4.

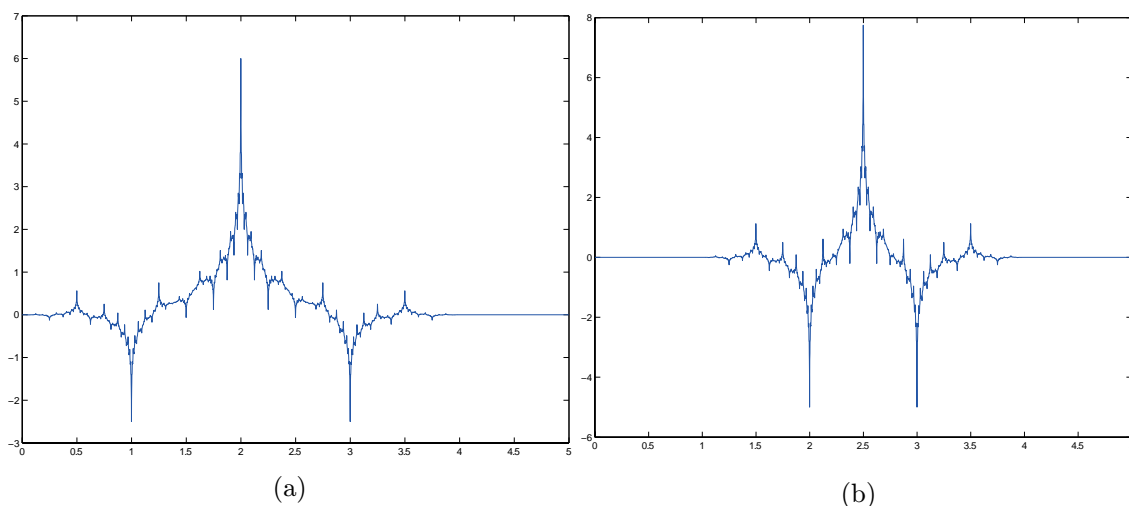
CDF

CDF (Cohen-Daubechies-Feauveau, 1992) je první rodina biortogonálních vlnek. Ačkoli postup při tvorbě vlnek CDF je podobný jako u vlnek Daubechies, nejsou si příliš podobné

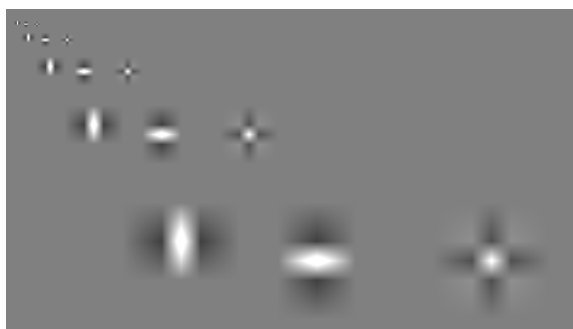
ve tvaru ani vlastnostech. Jsou generovány B-spliny, proto se někdy nazývají také biortogonální spline vlnky. Vlnky CDF jsou známé především díky jejich použití ve standardu pro kompresi obrazu JPEG 2000 [22] (konkrétně jejich verze CDF 5/3 pro bezztrátovou a CDF 9/7 pro ztrátovou kompresi) a standardu pro kódování otisků prstů FBI Fingerprint Image Compression Standard [2]. Tyto vlnky jsou:

- symetrické (kromě $p = 1$)
- biortogonální
- s různou délkou filtrů
- s různou délkou kompaktního nosiče
- mají nulové momenty

Na obr. 2.14 jsou vlnkové a škálovací funkce CDF 5/3 použité ve standardu JPEG 2000. Diskrétní 2D variantu vlnkové funkce znázorňuje obrázek 2.15.



Obrázek 2.14: Škálovací (a) a vlnková (b) funkce CDF 5/3 pro dekompozici obrazu.



Obrázek 2.15: Diskrétní 2D varianta vlnky CDF 5/3 v různých dyadických měřítcích. Pro každé měřítko zobrazena (zleva doprava) vlnka zachycující horizontální, vertikální a diagonální detaily.

Při implementaci bankou filtrů jsou koeficienty filtrů následující:

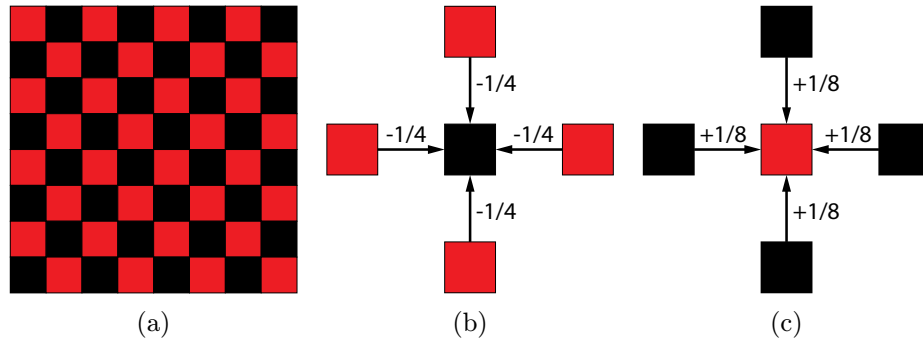
$$\begin{aligned}
 h &: \frac{\sqrt{2}}{8} \cdot (-1, 2, 6, 2, -1) \\
 g &: \frac{\sqrt{2}}{4} \cdot (1, -2, 1) \\
 \tilde{h} &: \frac{\sqrt{2}}{4} \cdot (1, 2, 1) \\
 \tilde{g} &: \frac{\sqrt{2}}{8} \cdot (-1, -2, 6, -2, -1)
 \end{aligned} \tag{2.25}$$

Implementace liftingem byla popsána v kapitole 2.1.5.

Red-Black wavelets

Transformace pomocí všech dříve zmíněných vlnek probíhá konvolucí signálu s různě posunutými a dilatovanými verzemi jedné mateřské vlnky. Takové vlnky se potom nazývají vlnkami první generace. Jako vlnky druhé generace označujeme vlnky složitější, kde nepracujeme s různými verzemi jedné vlnky, ale tvoříme je úplně odlišným způsobem, tzv. liftingem, viz 2.1.5). Jsou to např. vlnky na ohraničených intervalech, vlnky na křivkách a površích, vážené vlnky nebo vlnky přizpůsobené nepravidelně vzorkovaným datům.

Jedním z příkladů vlnek druhé generace jsou tzv. Red-Black wavelets, neboli černo-černé vlnky. Poprvé je představil ve své práci Geert Uytterhoeven [39]. V ní pomocí lifting schematu definoval neseparabilní dvourozměrnou verzi vlnky CDF 5/3. Název Red-Black odvodil od šachovnicového rozdělení pixelů v prvním kroku liftingu. Rozdělení dat a kroky predikce a aktualizace lze nejnázorněji popsat obrázkem 2.16.



Obrázek 2.16: Schéma první části liftingu pro Red-Black vlnky. (a) rozdělení dat, (b) predikce, (c) aktualizace

Matematicky lze predikci a aktualizaci definovat následovně:

$$x_{i,j} = x_{i,j} - \frac{1}{4}(x_{i-1,j} + x_{i,j-1} + x_{i,j+1} + x_{i+1,j}), \tag{2.26}$$

kde $i \bmod 2 \neq j \bmod 2$ (predikce).

$$x_{i,j} = x_{i,j} + \frac{1}{8}(x_{i-1,j} + x_{i,j-1} + x_{i,j+1} + x_{i+1,j}), \tag{2.27}$$

kde $i \bmod 2 = j \bmod 2$ (aktualizace).

Dále je třeba všechny kroky liftingu opakovat pro pixely podél diagonál: červené pixely rozdělíme na žluté a modré a provedeme predikci a aktualizaci podobně jako v předchozím kroce. Vše je znázorněno na obr. 2.17. Rovnice jsou potom následující:

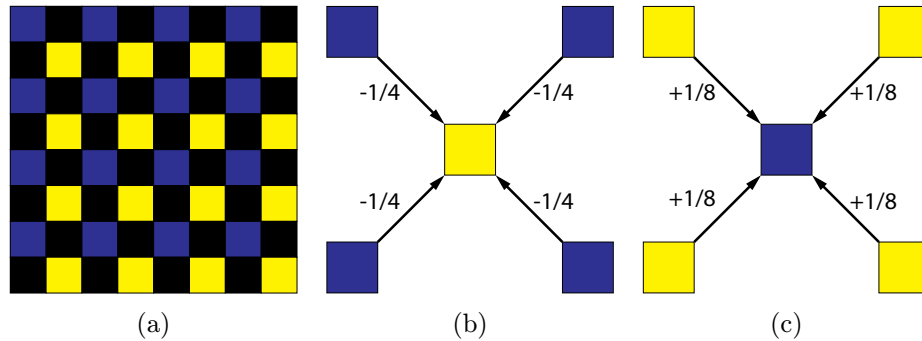
$$x_{i,j} = x_{i,j} - \frac{1}{4}(x_{i-1,j-1} + x_{i-1,j+1} + x_{i+1,j-1} + x_{i+1,j+1}), \quad (2.28)$$

kde $i \bmod 2 = 1 \wedge j \bmod 2 = 1$ (predikce).

$$x_{i,j} = x_{i,j} + \frac{1}{8}(x_{i-1,j-1} + x_{i-1,j+1} + x_{i+1,j-1} + x_{i+1,j+1}), \quad (2.29)$$

kde $i \bmod 2 = 0 \wedge j \bmod 2 = 0$ (aktualizace).

Modré pixely představují aproximační koeficienty transformace, černé se žlutými představují koeficienty detailní. V další úrovni rozkladu se tedy pokračuje dalším rozdělením modrých pixelů.



Obrázek 2.17: Schéma druhé části liftingu pro Red-Black vlnky. (a) rozdělení dat, (b) predikce, (c) aktualizace

Podle autorů tato transformace vykazuje nižší anizotropii než separabilní transformace vlnkou CDF. Diskrétní 2D Red-Black vlnka je zobrazena na obr. 2.18. Díky neseparabilitě jsou vlnky zachycující horizontální a vertikální detaily stejné.



Obrázek 2.18: Diskrétní 2D Red-Black vlnka v různých dyadických měřítcích. Pro každé měřítko zobrazena (zleva doprava) vlnka zachycující horizontální, vertikální a diagonální detaily.

Edge avoiding wavelets

Vlnky druhé generace, jejichž tvar je závislý na vstupních datech, tzv. Edge avoiding wavelets, představil ve své práci Raanan Fattal [18]. Implementace probíhá pomocí liftingu, kdy

je definován operátor predikce vážící pixely na základě podobnosti s predikovaným pixelem. Váhy jsou definovány pomocí tzv. edge-stopping funkce [25], viz rovnice 2.30. Tato funkce je použita v predikčním i v aktualizacím operátoru.

$$w_n^j[m] = (|a^j[n] - a^j[m]|^\alpha + \epsilon)^{-1} \quad (2.30)$$

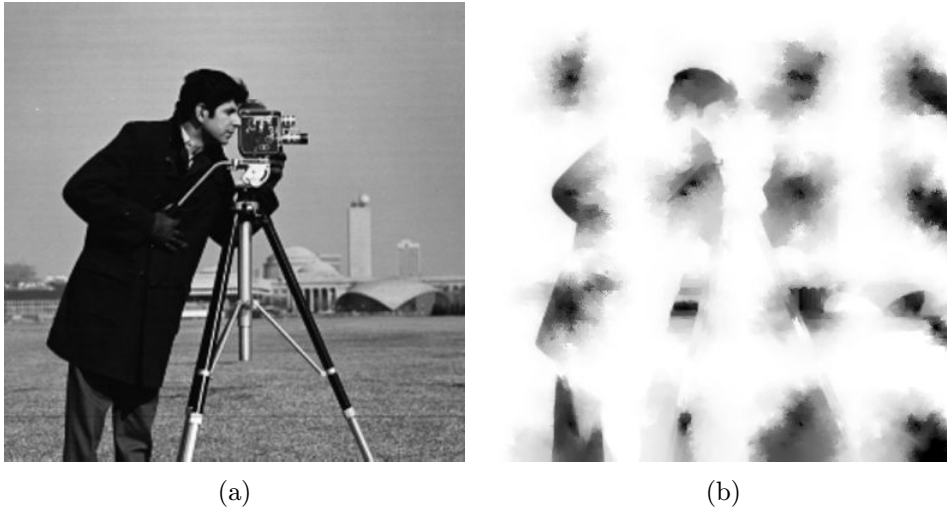
$w_n^j[m]$ značí váhu m -tého pixelu vzhledem k n -tému, na úrovni rozkladu j . a^j jsou koeficienty vlnkové transformace, α konstanta mezi 0,8 a 1,2 a $\epsilon = 10^{-5}$.

Váhová funkce není vázaná na žádnou konkrétní vlnku. Rovnice 2.31 ukazuje její použití v operátorech predikce a aktualizace při transformaci liftingem vlnkou CDF 5/3.

$$\begin{aligned} \mathcal{P}(a_C^j[x]) &= \frac{w_x^j[x-1]a_C^j[x-1] + w_x^j[x+1]a_C^j[x+1]}{w_x^j[x-1] + w_x^j[x+1]} \\ \mathcal{U}(a_F^{j+1}[x]) &= \frac{w_x^j[x-1]a_F^{j+1}[x-1] + w_x^j[x+1]a_F^{j+1}[x+1]}{2(w_x^j[x-1] + w_x^j[x+1])} \end{aligned} \quad (2.31)$$

Analogicky lze definovat operátory predikce a aktualizace pro Red-Black vlnkovou transformaci. Výsledné vlnky se potom nazývají vážené CDF nebo vážené Red-Black vlnky (WCDF, WRB).

Díky použití váhové funkce je dosaženo vlastnosti, kdy vlnková funkce je hladká v oblastech s nízkými frekvencemi ale poblíž silných hran je náhle ukončena (v závislosti na α z 2.30). Oblast podpory vlnky je tedy definována pouze mezi hranami, odtud název edge avoiding wavelets. Tuto vlastnost ilustruje za použití vlnek WRB obr. 2.19. Výsledkem je snížená odezva od hran a s tím spojená korelace mezi jednotlivými úrovněmi rozkladu, která je mj. hlavním zdrojem nežádoucích halo efektů². Tyto vlnky je proto výhodné využít v různých oblastech zpracování obrazu jako je komprese dynamického rozsahu, vyhlazení zachovávající hrany, zvýraznění detailů nebo kolorizace.



Obrázek 2.19: WRB transformace. (a) vstupní obraz, (b) 16 škálovacích funkcí rovnoměrně rozmístěných po obraze.

²Jev, kdy dojde k přílišnému zvýšení kontrastu kolem hran. Kolem tmavých hran vznikají světlé obrysy a naopak.

Curvelety

2D vlnkovou transformací první generace lze efektivně zakódovat bodovou nespojitost – je jí ovlivněno pouze malé množství koeficientů. Nicméně nespojitost podél křivky ovlivní všechny koeficienty podél této křivky, taková transformace již není efektivní. S rostoucím množstvím koeficientů je obtížnější kvalitně provést např. redukci šumu v obraze, nastává známý problém s nalezením kompromisu mezi šetrným vyhlazením zachovávajícím detaily a přílišným vyhlazením, které dobře redukuje šum, ale má za následek přílišnou ztrátu detailů.

Emmanuel Candes a David Donoho [6] v r. 1999 představili novou metodu – tzv. curvelet transformaci, která se s výše popsáním nedostatkem WT vypořádává – dokáže popsat objekt s hranami minimálním množstvím koeficientů. Curvelet transformace se v praxi realizuje jako sekvence následujících kroků [31]:

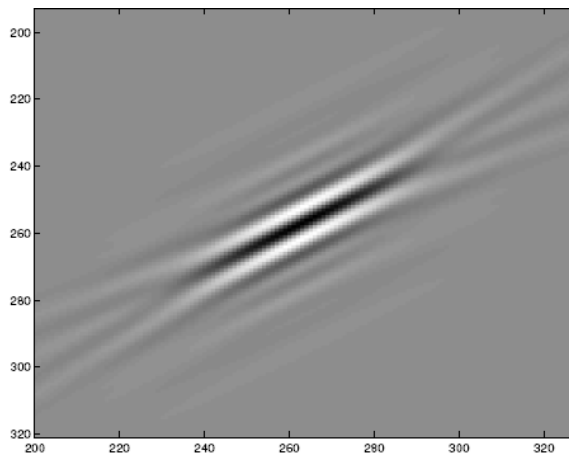
- Dekompozice do podpásem. Zatímco v klasické WT se obraz většinou dekomponuje do dyadických podpásem $[2^j, 2^j + 1]$, v curvelet transformaci jsou použity pod pásma $[2^{2^j}, 2^{2^{j+2}}]$. Lze využít stejný algoritmus jako v SWT, který obraz I rozloží na aproximační koeficienty c_J a detailní koeficienty w_j :

$$I(x, y) = c_J(x, y) + \sum_{j=1}^J w_j(x, y) \quad (2.32)$$

- Rozdělení do oken. Každé podpásmo je rozděleno do čtvercových oken s délkou strany 2^{-j}
- Ridgelet analýza. Každý čtverec je analyzován pomocí diskrétní ridgelet transformace. Transformovány jsou pouze měřítka 1 až J . Tzv. ridgelet je základní stavební prvek curveletové transformace. Lze jej popsat rovnicí

$$\psi_{a,b,\theta}(x_1, x_2) = a^{-\frac{1}{2}} \psi((x_1 \cos \theta + x_2 \sin \theta - b)/a), \quad (2.33)$$

kde $a > 0$ je parametr udávající měřítko, $b \in \mathbb{R}$ pozici a $\theta \in [0, 2\pi)$ orientaci. Ve 2D prostoru ridgelet vypadá jako podlouhlá vlnka (angl. ridge = hřeben, hřbet), která má navíc svou orientaci v prostoru θ . Příklad ridgeletu je na obr. 2.20. Ridgelet transformací (podrobněji např. v [31]) získáme koeficienty udávající míru podobnosti dané části obrazu s úsečkou o určité orientaci.

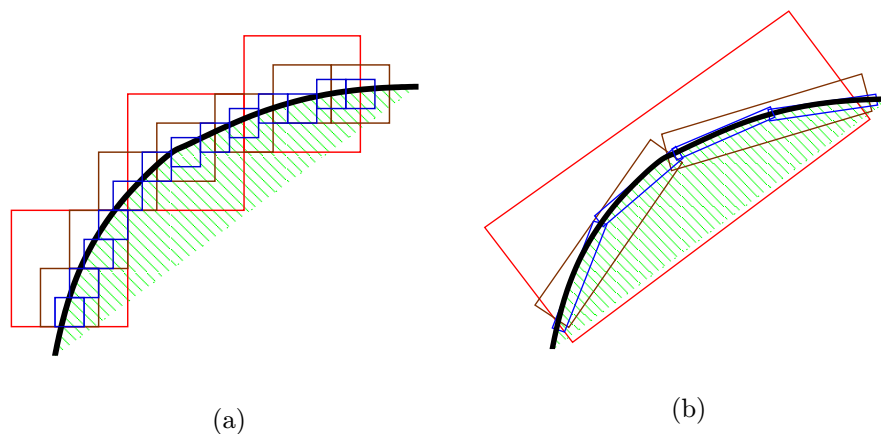


Obrázek 2.20: 2D znázornění ridgeletu. Obrázek převzat z [5]

Aplikováním ridgelet transformace na všechny okna v daném podpásmu potom získáme ridgelet (resp. curvelet) koeficienty. Tato transformace přináší redundanci $16J + 1$ (při J úrovních rozkladu), nicméně v případě vhodných vstupních obrazů obsahujících hrany lze většinu koeficientů vynulovat, a tím redundanci výrazně zredukovat. V roce 2006 Candes et al. [5] představili tzv. curvelet transformaci druhé generace, která už nevyužívá ridgelet transformace a také má nižší redundanci. Curvelet transformaci lze úspěšně využít např. pro redukci šumu, doostření nebo detekci změn v obraze.

Contourlety

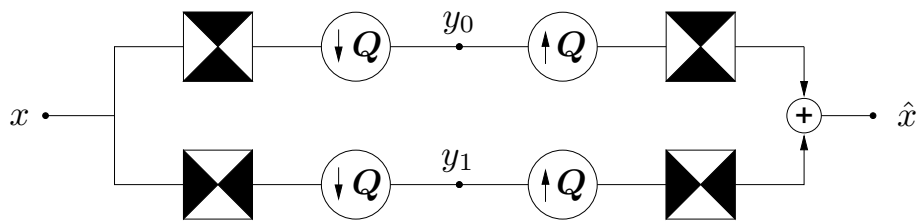
Cílem autorů countourlet transformace (Minh Do a Martin Vetterli [10]) bylo vytvořit pravou 2D transformaci, která dokáže zachytit geometrickou strukturu obrazu, která je jeho klíčovou vizuální informací. Autoři se přitom zaměřují na konstrukci transformace přímo v diskrétní oblasti, například narozdíl od curvelet transformace, která byla vyvinuta pro spojitou oblast a poté diskretizována pro navzorkovaná data. Cílem je, podobně jako u curvelet transformace, efektivně zachytit obrysy v obraze – tedy s co nejmenším počtem koeficientů. Obrázek 2.21 ilustruje rozdíl mezi efektivní a neefektivní reprezentací hrany v obraze.



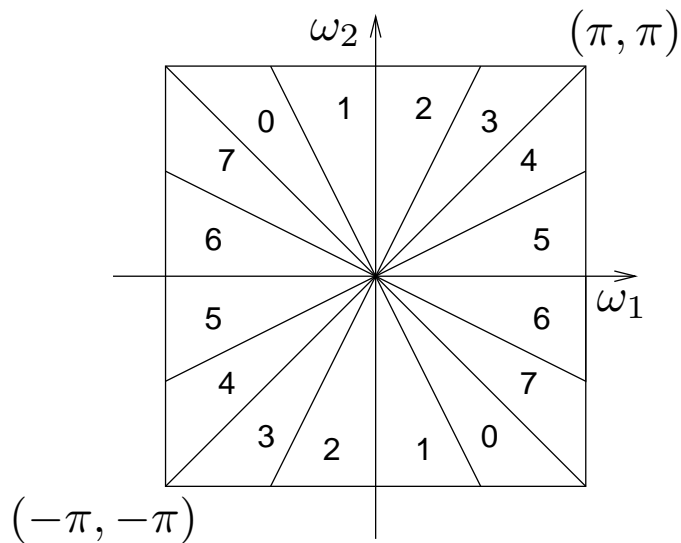
Obrázek 2.21: Reprezentace hrany (obrysu): (a) pomocí vlnkové transformace, (b) pomocí contourlet transformace. Obrázek převzat z [10].

Efektivní reprezentace obrysu je dosaženo seskupováním blízkých detailních koeficientů, které jsou vzájemně korelované díky hladkosti hran. Detailní koeficienty jsou získány dekompozicí pomocí Laplaceovy pyramidy (LP). [4] Touto dekompozicí získáme v každé úrovni podvzorkovanou verzi signálu tvořící nízkofrekvenční obraz a rozdíl mezi signálem a jeho predikcí, tvořící obraz s vysokými frekvencemi. Tento vysokofrekvenční obraz slouží jako vstup do druhé fáze algoritmu, tvořené tzv. bankou směrových filtrů (directional filter bank, DFB). DFB je tvořena ze dvou částí – z dvoukanalového quincunx³ filtru, který dělí 2D spektrum na horizontální a vertikální část a z operátoru zkosení, který pouze přeuspořádává vzorky signálu. DFB je znázorněna na obr. 2.22. Kombinací quincunx filtru a operátoru zkosení potom lze docílit rozdělení 2D spektra podle obr. 2.23. Množina posunutých impulzních odezví DFB potom tvoří bázi diskrétních signálů z $l^2(\mathbb{Z}^2)$. Tato báze má směrové i lokalizační vlastnosti.

³Nejjednodušší schéma neseeparabilního 2D filtru.

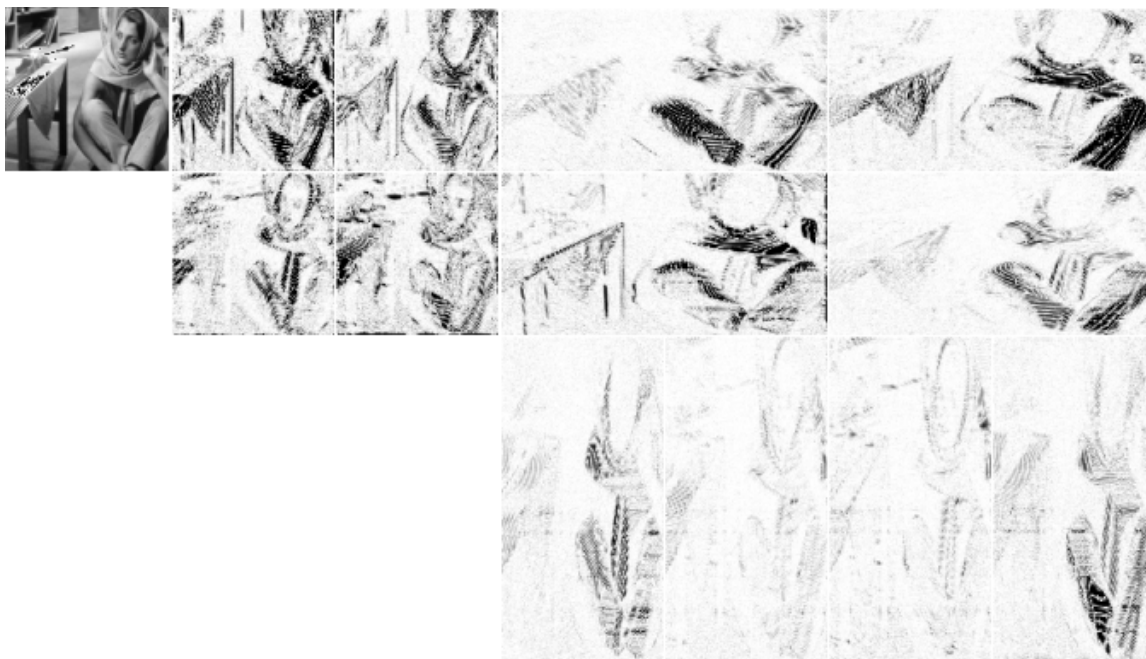


Obrázek 2.22: Banka směrových filtrů. Černé oblasti vymezují směry, se kterými filtr pracuje. Q je vzkovovací matice.



Obrázek 2.23: Rozdělení 2D spektra směrovými filtry.

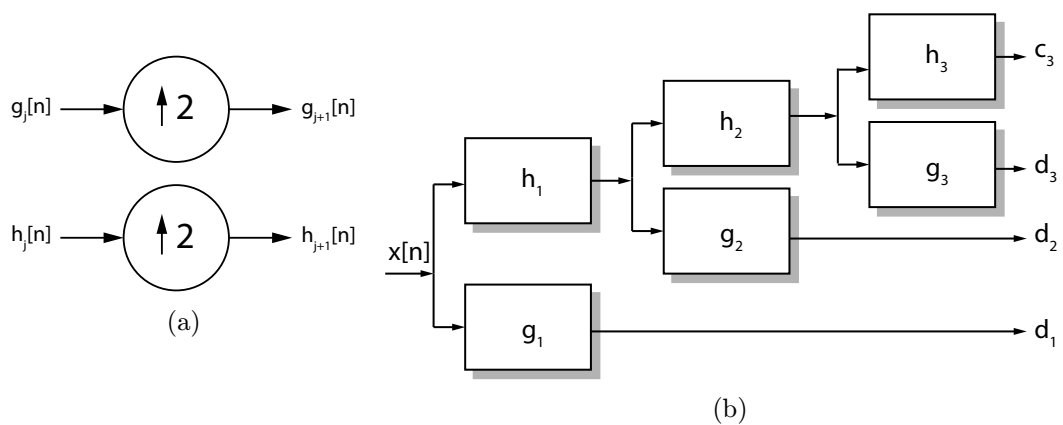
Spojením Laplaceovy pyramidy a DFB do dvojité banky filtrů tak, aby výstup z LP byl vstupem do DFB a celé schéma se dalo iterovat, získáme tzv. contourlet banku filtrů, která rozkládá obraz do směrových podpásem v různých měřících rozkladu. Tato transformace má lineární časovou složitost při redundanci až 33%. Příklad výstupu je znázorněn na obr. 2.24. Obrázky v této kapitole byly převzaty z [10].



Obrázek 2.24: Příklad výstupu contourlet transformace obrazu „Barbara“. Obraz je transformován ve dvou úrovních, které jsou poté rozděleny do čtyř a osmi směrových podpásem.

2.2 Stacionární vlnková transformace

Klasická diskrétní vlnková transformace je závislá na posunutí, tzn. že DWT posunutého signálu je různá od DWT signálu původního. V některých aplikacích, jako např. redukce šumu v obraze, je tato vlastnost nežádoucí. Jednou z metod, která se s tímto nedostatkem vypořádává, je stacionární vlnková transformace (SWT). Nezávislosti na posunutí je dosaženo odstraněním decimátorů a interpolátorů umístěných před a za filtry provádějící DWT a nadvzorkování koeficientů filtru faktorem $\frac{1}{2^{(j-1)}}$ pro j -tou úroveň rozkladu. SWT je tedy redundantní, počet výstupních koeficientů pro každou úroveň je stejný jako počet vzorků vstupního signálu. Obrázek 2.25 znázorňuje implementaci SWT pomocí banky filtrů. Filtry v každé vyšší úrovni jsou nadvzorkovanými verzemi filtrů z předchozí úrovně.



Obrázek 2.25: Stacionární vlnková transformace. (a) tvorba filtrů vyšších úrovní, (b) kaskáda filtrů implementující SWT.

Kapitola 3

Aplikace vlnkové transformace

V předchozích kapitolách byly popsány teoretické základy vlnkové transformace, jejích různých variant a možnosti její implementace. Kapitoly následující obsahují popis vybraných aplikací vlnkové transformace ve zpracování obrazu.

3.1 Detekce hran

Detekce hran je důležitý úkol ve zpracování obrazu. Kvalitně detekované hrany jsou potřeba například pro segmentaci obrazu nebo rozpoznání vzorů. Hranu lze definovat jako množinu bodů, obsahující ostré přechody v intenzitě obrazu. Tato definice ale není úplně přesná, častokrát se v obraze vyskytují ostré přechody, které nejsou způsobeny výskytem hrany ale texturou. Rozhodnutí mezi hranou a texturou také záleží na měřítku analýzy. Příklad výstupu algoritmu detekce hran je na obr. 3.1.



Obrázek 3.1: Příklad detekce hran (zde pomocí vlnkové transformace).

Tradiční metody

Většina z tradičních přístupů k detekci hran se dá rozdělit do dvou kategorií – metody založené na hledání nebo metody založené na průchodu nulou.

Metody založené na hledání obvykle určí sílu hrany vypočítáním gradientu pomocí aproximace první derivace. K tomu existují různé operátory, například Sobelův:

$$K_1 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \quad K_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

K_1 a K_2 jsou jádra pro konvoluci obrazu ve vertikálním a horizontálním směru. Následně jsou ve vypočítaném gradientu určeny hrany, většinou lokalizováním maxim absolutní hodnoty.

Příkladem metody využívající průchody nulou je Marr-Hildrethův operátor. Ten detekuje hrany pomocí druhé derivace obrazu vyhlazeného Gaussovou funkcí. Druhou derivaci lze aproximovat Laplaciánem (Laplacian of Gaussian, LoG). Výpočet se realizuje konvolucí obrazu s maskou nazývanou mexický klobouk:

$$K = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Z výsledného obrazu jsou získány hrany lokalizováním průchodů nulou. K nevýhodám této metody patří detekce velkého množství falešných hran kvůli citlivosti na šum a přílišné vyhlazení ostrých tvarů – tedy ztráta některých ostrých hran.

Za state-of-the-art metodu v detekci hran je stále považován Cannyho hranový detektor. John Canny v roce 1986 vyvinul metodu, která měla za cíl detekovat co nejvíce hran a to s co nejvyšší přesností a lokalizací a pokud možno bez falešných odezev. Postup se dá shrnout do následujících kroků:

- Vyhlazení obrazu Gaussovou funkcí
- Nalezení gradientu intenzity obrazu
- Potlačení nemaximálních hodnot gradientu
- Dvojitě prahování výsledku
- Hystereze

Gradient je vypočítán některým z operátorů aproximujících první derivaci obrazu (např. výše popsáním Sobelovým operátorem). Dále je vypočítán úhlový obraz:

$$\Theta = \arctan \left(\frac{G_y}{G_x} \right), \quad (3.1)$$

kde G_y a G_x je gradient obrazu ve vertikálním a horizontálním směru. Zaokrouhlením úhlového obrazu jsou detekovány vertikální, horizontální a diagonální hrany. Ty jsou následně zúženy potlačením nemaximálních hodnot ve směru gradientu. Dvojitým prahováním jsou

odfiltrovány slabé hrany, většinou způsobené šumem nebo variací barvy. Výsledek je nakonec vylepšen hysterezí, která zachová i některé slabší hrany, pokud jsou spojeny s hranou silnou.

Na obr. 3.2 je srovnání tradičních metod detekce hran včetně Cannyho detektoru.



Obrázek 3.2: Detekce hran pomocí (a) Sobelova operátoru, (b) laplaciánu, (c) Cannyho detektoru.

Detekce hran pomocí vlnkové transformace

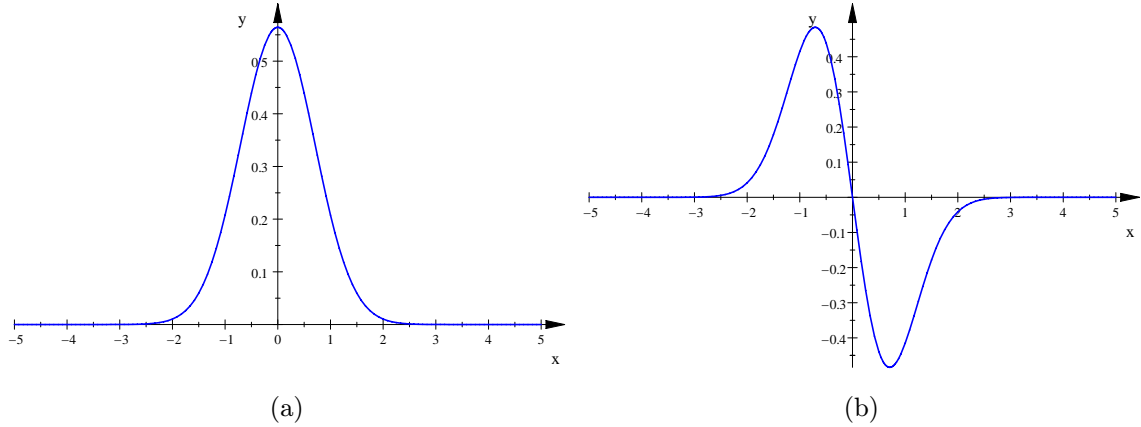
Vlnková transformace rozkládá signál na aproximační a detailní koeficienty. Detaily v podstatě představují hrany, z čehož vychází nejjednodušší metoda detekce hran pomocí vlnkové transformace – přepsání aproximačních koeficientů nulami. Tím po zpětné rekonstrukci vznikne obraz obsahující pouze hrany. Výsledek je ale hodně ovlivněn šumem a proto vznikly metody, které problematiku detekce hran řeší komplexněji. Jednou z takových je algoritmus Stéphane Mallata a Sifena Zhonga, kteří ve své práci [27] dokázali, že maxima absolutní hodnoty vlnkové transformace lze využít k detekci nepravidelných struktur. Tyto maxima také využívají k výpočtu tzv. Lipschitzova exponentu, který popisuje ostrost hrany.

V tomto algoritmu je využita dyadická vlnková transformace, hodnoty měřítka s jsou

tedy omezena sekvencí $(2^j)_{j \in \mathbb{Z}}$. Jako vlnková funkce je použita kvadratická spline křivka, konkrétněji první derivace Gaussovy funkce. Vlnková funkce dilatovaná faktorem 2^j je značena $\psi(x)_{2^j}$. Vlnková transformace obrazu proto odpovídá první derivaci původního signálu vyhlazeného Gaussovou funkcí. Její rovnice je definována pomocí konvoluce:

$$W_{2^j} f(x) = f * \psi_{2^j}(x) \quad (3.2)$$

Vlnková funkce a Gaussova funkce, ze které byla odvozena jsou znázorněny na obr. 3.3.



Obrázek 3.3: Gaussova funkce (a) a vlnková funkce (b) pro detekci hran.

V praxi má signál konečné rozlišení a proto není možné počítat transformaci v libovolném měřítku. Pro vypořádání se s tímto omezením se zavádí vyhlazovací funkce $\phi_{2^j}(x)$, kde faktor 2^j opět udává její dilataci. Opakovanou konvolucí signálu s různě dilatovanou verzí této funkce získáváme signál s různým stupněm vyhlazení, tedy v různém měřítku pro detekci hran:

$$S_{2^j} f(x) = f * \phi_{2^j}(x), \quad (3.3)$$

kde S_{2^j} se nazývá vyhlazovací operátor.

Sekvence signálů $S_{2^j} f$ a $W_{2^j} f$ pro měřítka $1 \leq j \leq J$ se potom nazývá diskrétní dyadická transformace signálu. Možnost volby měřítka je jednou z hlavních výhod a důvodů využití vlnkové transformace k detekci hran.

Pro transformaci 2D signálu stačí konvolovat s vlnkovou i vyhlazovací funkcí zvlášť v horizontálním a vertikálním směru, transformace je při použití vhodných funkcí separabilní.

Po transformaci následuje výpočet absolutní hodnoty vlnkové transformace:

$$M_s f(x, y) = \sqrt{|W_s^1 f(x, y)|^2 + |W_s^2 f(x, y)|^2}, \quad (3.4)$$

kde $W_s^1 f$ značí transformaci v horizontálním a $W_s^2 f$ ve vertikálním směru.

Dále je vypočítán úhlový obraz:

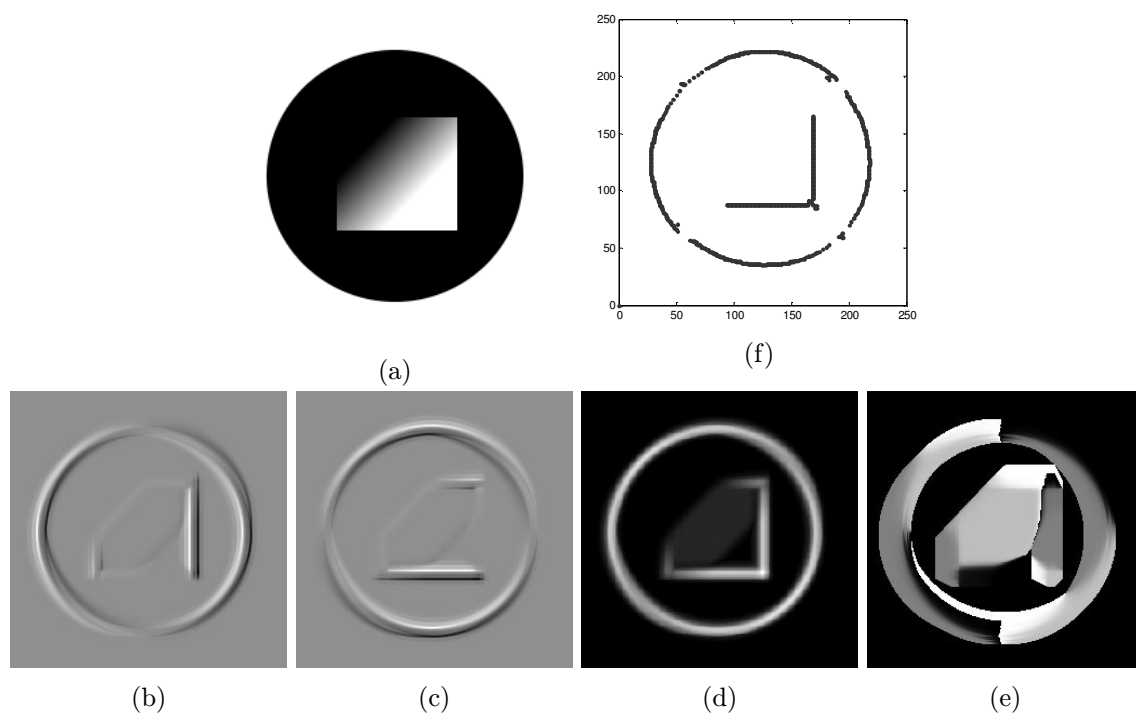
$$A_s f(x, y) = \arctan\left(\frac{W_s^2 f(x, y)}{W_s^1 f(x, y)}\right) \quad (3.5)$$

Z těchto dvou obrazů jsou nakonec spočítány maxima. O pixelu je rozhodnuto jako o maximu, pokud platí, že je jeho absolutní hodnota větší než jeho dvou sousedů podél úhlu

gradientu. Pro potlačení nevýrazných maxim jsou vykresleny jenom takové, jejichž hodnota je vyšší než určitý zvolený práh T :

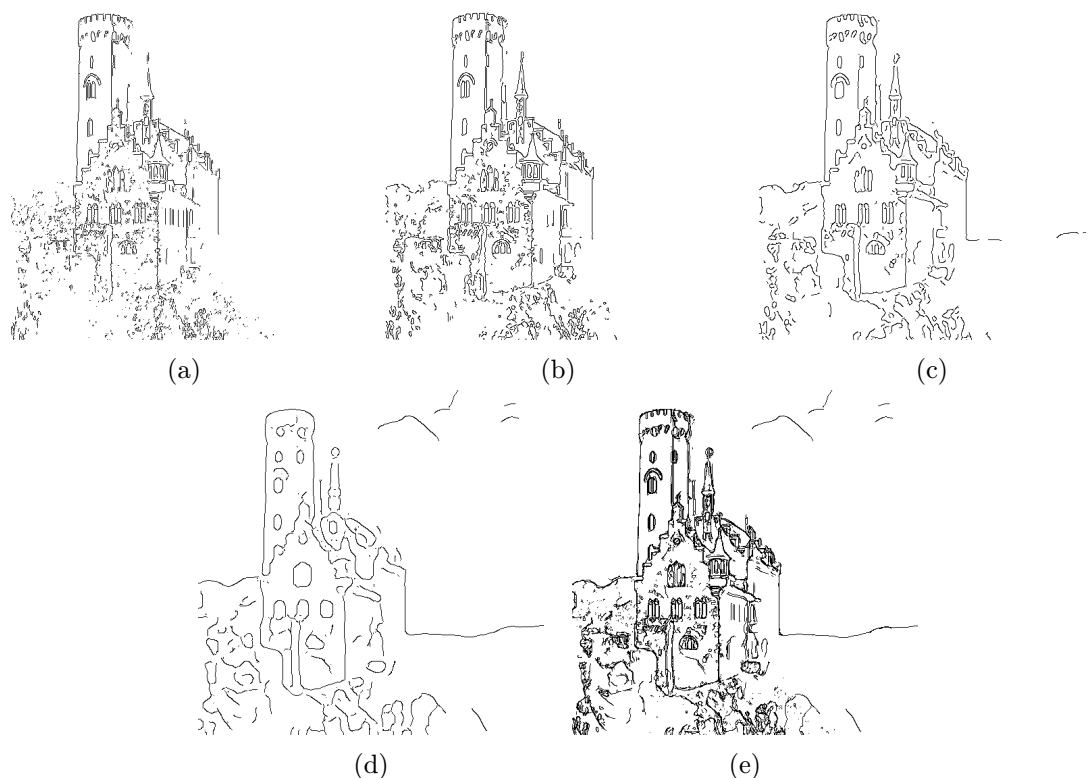
$$E_s(x, y) = \begin{cases} 0 & |W_s f(x, y)| < T \\ 1 & |W_s f(x, y)| \geq T \end{cases} \quad (3.6)$$

Hodnota prahu T může být zvolena například metodou Otsu [29]. Výstupy jednotlivých fází algoritmu jsou znázorněny na obr. 3.4.



Obrázek 3.4: (a) vstupní obraz. (b), (c) výstup vlnkové transformace pro horizontální a vertikální směr. (d) absolutní hodnota vlnkové transformace. (e) úhlový obraz. (f) maxima absolutní hodnoty vlnkové transformace. Obrázky převzaty z [20]

Výstupy detekce pro jednotlivé úrovně rozkladu lze potom sečtením spojit do jediného obrazu, jak je znázorněno na obr. 3.5.



Obrázek 3.5: Detekce hran. (a) - (d) Výstupy detekce pro první 4 úrovně rozkladu. (e) Obraz získaný sečtením hran ze všech úrovní.

Dyadickou vlnkovou transformaci lze tedy úspěšně použít pro detekci hran v 1D i 2D signálu. Problémem je zpracování signálů s vyšší úrovní šumu, kdy díky dyadické posloupnosti měřítek není vždy metoda schopna se přizpůsobit konkrétní úrovni šumu. Tento nedostatek se dá řešit například kombinací více měřítek s výpočtem Lipschitzovské spojitosti signálu. V místech, kde je spojitost nižší se použijí vlnky s větším měřítkem, čímž se potlačí vliv šumu, zatímco v místech s vyšší spojitostí se použijí vlnky s měřítkem malým, čímž zůstane lépe zachována přesná pozice hrany. [24]

3.2 Odstranění šumu

Při snímání, přenosu nebo jinými manipulacemi s obrazem dochází často k jeho narušení různými druhy šumu. Pro obnovení původní struktury obrazu je proto nutné tento šum lokalizovat a odstranit. Tradiční techniky pro odstranění šumu z obrazu využívají lineárního nebo nelineárního filtrování. Lineární filtrace může být dosaženo například použitím dolní propusti nebo rychlé Fourierovy transformace (FFT). Pro nelineární filtraci jsou typické statistické nebo morfologické filtry. Tyto techniky ale mohou mít negativní vliv na strukturu původního obrazu, často dochází ke ztrátě důležitých vlastností – např. k rozmazání hran. [36] Existují různé druhy šumu a pro každý z nich může být vhodná jiná metoda pro jeho odstranění. V této práci se předpokládá obraz poškozený aditivním Gaussovým šumem.

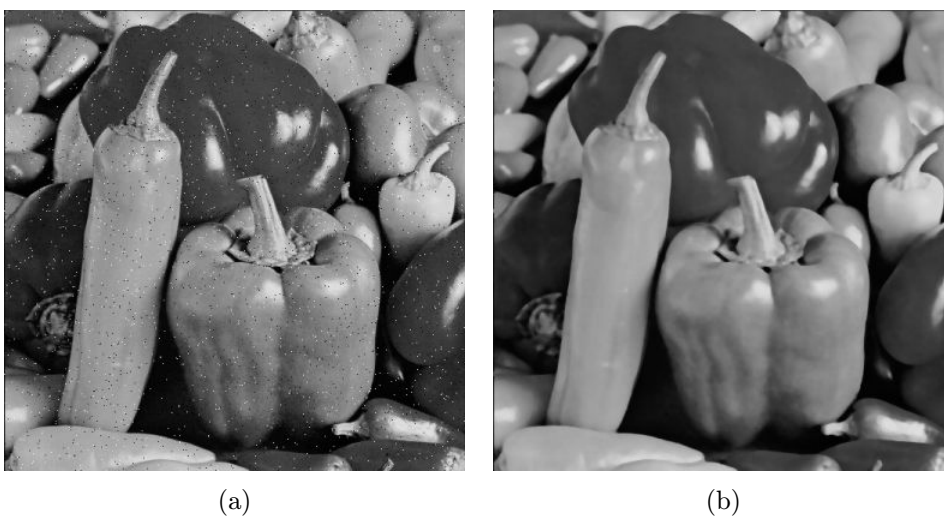
Tradiční metody

Jednou z nejjednodušších metod pro redukcí šumu v obraze je jeho filtrace Gaussovým filtrem. Jedná se v podstatě o pouhou frekvenční filtraci – dolní propust. Provádí se konvolucí obrazu s jádrem tvořeným hodnotami Gaussovy funkce:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.7)$$

Jedná se tedy o lineární filtraci. Tato metoda efektivně odstraňuje Gaussův šum, její nevýhodou je ale ztráta detailů vlivem rozmazání obrazu.

Základní nelineární metodou je mediánový filtr. Filtrace probíhá výběrem mediánu z okna kolem aktuálně sledovaného pixelu. Metoda je vhodná spíše pro odstranění náhodného šumu, např. typu sůl a pepř. Výsledek filtrace s použitím okna o velikosti 5x5 je na obr. 3.6.

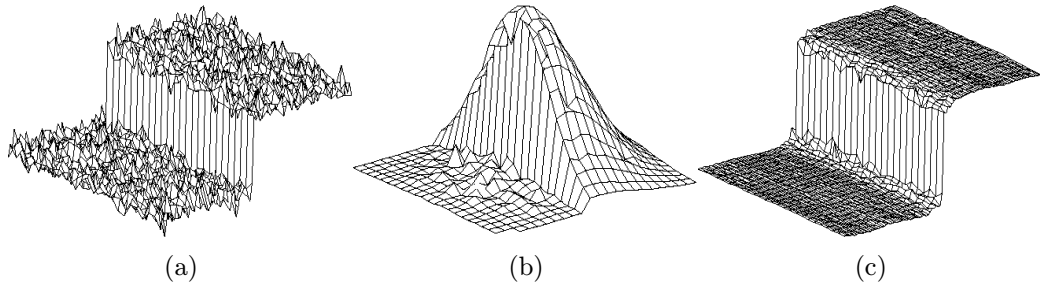


Obrázek 3.6: Mediánový filtr: (a) obraz poškozený šumem typu „sůl a pepř“, (b) výsledek filtrace.

Jednou z pokročilých nelineárních metod je bilaterální filtrace obrazu. Pixel je nahrazen váženým průměrem pixelů v jeho okolí, přičemž váhy nezávisí pouze na vzdálenosti pixelů, ale také na podobnosti jejich intenzit. Matematicky je bilaterální filtr definován jako:

$$I^B(x) = \frac{\sum_{y \in N(x)} I(y) g_{\sigma_S}(x, y) g_{\sigma_R}(I(x), I(y))}{\sum_{y \in N(x)} g_{\sigma_S}(x, y) g_{\sigma_R}(I(x), I(y))} \quad (3.8)$$

g_{σ_S} a g_{σ_R} jsou jádra Gaussova filtru, $N(x)$ množina pixelů okolo sledovaného pixelu a $I(x)$ značí intenzitu pixelu. g_{σ_S} ovlivňuje míru vyhlazení a g_{σ_R} míru zachování hran. Ve výsledku dostaneme vyhlazený obraz se zachovanými hranami. Na obr. 3.7 je znázorněno chování bilaterálního filtru.



Obrázek 3.7: Bilaterální filtr: (a) původní signál s hranou mezi světlou a tmavou oblastí, (b) jádro bilaterálního filtru, (c) výsledek filtrace. Obrázky převzaty z [35].

Další pokročilou nelineární technikou pro redukci šumu je metoda Non-local means. Na rozdíl od lokálních filtrů, které výslednou hodnotu počítají jako průměr hodnot z určitého malého okolí, metoda Non-local means počítá váženou hodnotu závislou na větším okolí sledovaného pixelu. Váha závisí na podobnosti pixelu s celým sledovaným oknem okolo něj. Matematická definice metody je následující:

$$u(p) = \frac{1}{C(p)} \int v(q) f(p, q) dq, \quad (3.9)$$

kde $u(p)$ je výsledná hodnota pixelu p po filtraci, $v(q)$ je původní hodnota pixelu q a $f(p, q)$ je vážící funkce. Více v [3].

Odstranění šumu pomocí vlnkové transformace

David Donoho a Iain Johnstone [11] navrhli řešení problému ztráty důležitých vlastností obrazu při redukci šumu. Jejich metoda kombinuje diskrétní vlnkovou transformaci s prahovací technikou. Oddělení signálu od šumu umožňuje schopnost vlnkové transformace soustředit energii signálu do malého počtu koeficientů s vysokou hodnotou. Naopak energie šumu je rozprostřena do velkého počtu malých koeficientů a proto je možné šum pomocí prahování oddělit od původního signálu. K získání odhadu původního obrazu je nakonec jednoduše použita inverzní transformace.

Předpokládejme signál konečné délky

$$y_i = x_i + \epsilon n_i, \quad i = 1, \dots, N \quad (3.10)$$

zašuměný Gaussovským šumem n_i . Potom

$$Y_i = X_i + N_i \quad (3.11)$$

představuje signál převedený DWT určenou maticí W :

$$Y_i = W_i y_i \quad (3.12)$$

do transformační domény. K matici W existuje inverzní matice W^{-1} . Celé schéma pro odstranění šumu je potom:

1. výpočet DWT $Y_i = W_i y_i$
2. provedení prahování v transformační doméně

(a) pomocí tvrdého prahování:

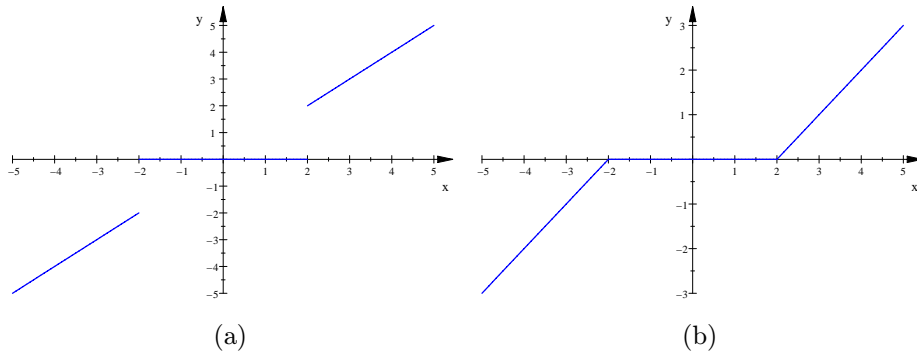
$$\hat{X} = T_h(Y, t) = \begin{cases} Y, & |Y| \geq T \\ 0, & |Y| < T \end{cases} \quad (3.13)$$

(b) nebo pomocí měkkého prahování:

$$\hat{X} = T_s(Y, t) = \begin{cases} \text{sgn}(Y)(|Y| - t), & |Y| \geq T \\ 0, & |Y| < T \end{cases} \quad (3.14)$$

3. výpočet inverzní DWT $\hat{x} = W^{-1}\hat{X}$

\hat{x} značí odhad původního signálu před znehodnocením šumem. Rozdíl mezi měkkým a tvrdým prahováním je znázorněn na obr. 3.8. Při použití měkkého prahování je odhad s vysokou pravděpodobností minimálně stejně hladký jako původní signál. Výpočetní náročnost metody je $\mathcal{O}(n)$.



Obrázek 3.8: (a) tvrdé prahování. (b) měkké prahování

Velký vliv na výsledek má výběr prahu t . Podobně jako při použití klasických metod redukce šumu je nutné volit kompromis mezi potlačením šumu a přílišným vyhlazením obrazu (se ztrátou detailů) – nicméně ne v tak velké míře. Donoho a Johnstone navrhli metodu VisuShrink, používající univerzální práh definovaný vztahem

$$t = \sigma\sqrt{2 \log M}, \quad (3.15)$$

kde σ značí rozptyl šumu a M je počet pixelů obrazu. Nicméně pro velká M bývá práh příliš velký, což má za následek odstranění mnoha koeficientů a přílišné vyhlazení obrazu. Stejní autoři proto ve své další práci [12] používají metodu SureShrink, ve které se práh hledá pro každé podpásmo zvlášť minimalizováním odhadu rizika metodou SURE (Stein's unbiased risk estimate). Značíme-li koeficienty transformace jako $\{X_i : i = 1, \dots, d\}$, odhad rizika je definován vztahem:

$$\text{SURE}(t; X) = d - 2\#\{i : |X_i| \leq t\} + \sum_{i=1}^d \min(|X_i|, t)^2 \quad (3.16)$$

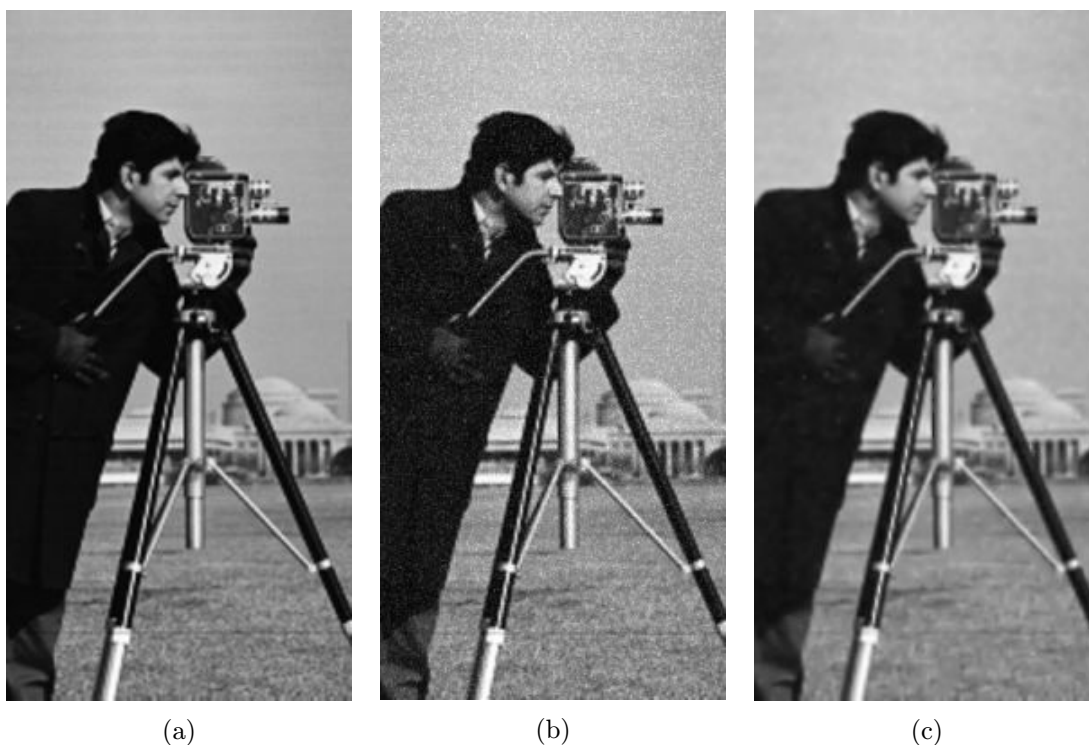
Jeho minimalizováním získáme hodnotu prahu:

$$t = \operatorname{argmin}_{0 \leq t \leq \sqrt{2 \log d}} \text{SURE}(t; X) \quad (3.17)$$

Prahováním koeficientů touto metodou získáme obraz se zachovaným větším počtem detailů a s nižší střední kvadratickou chybou.

Protože DWT není nezávislá na posunutí, mohou především kolem hran v obraze vznikat vizuální artefakty. Bylo zjištěno, že vícenásobným použitím výše zmíněné metody na různě posunutý signál a zprůměrováním výsledků se dosáhne podstatného zlepšení. Pokud bychom počítali a zprůměrovali všechny posunutí, výpočetní náročnost tohoto postupu by byla $\mathcal{O}(n \log n^2)$. Nicméně mezi koeficienty odpovídajícími různým posunutím je pouze $n \log n$ unikátních, ostatní se opakují. Tyto unikátní koeficienty lze spočítat pomocí stacionární vlnkové transformace. Namísto počítání DWT pro různá posunutí a průměrování výsledků tedy stačí spočítat SWT a po prahování koeficientů jednoduše použít inverzní transformaci. Výpočetní náročnost je potom $\mathcal{O}(n \log n)$.

Na obr. 3.9 je výsledek redukce šumu prahováním koeficientům SWT.



Obrázek 3.9: Redukce šumu. (a) původní obraz. (b) obraz s šumem ($\sigma = 25$). (c) Prahování koeficientů SWT.

3.3 Vylepšení kontrastu

Kontrast je jednou z klíčových vlastností obrazu, které ovlivňují jeho vizuální kvalitu. Existují dva základní typy kontrastu – globální, vyjadřující rozdíly v jasů v celém obraze a lokální, který představuje spíše výraznost či ostrost detailů. Bylo vyvinuto mnoho technik zpracování obrazu, které se zaměřují na vylepšení obou z nich.

Tradiční metody

Nejběžněji používané metody vylepšení kontrastu obvykle mění obraz globálně pomocí úprav histogramu, např. úpravou gamma křivky, logaritmickou kompresí nebo ekvalizací

histogramu. V případě vylepšení lokálního kontrastu je často používána tzv. unsharp maska. Ta provede odečtení rozmazaného obrazu od obrazu původního a tím dojde ke zvýraznění detailů. Podle nastavení velikosti jádra Gaussova filtru, kterým se rozmazání obvykle provádí, potom lze hovořit buďto o zostření obrazu nebo zlepšení lokálního kontrastu.

Tyto metody obecně mohou způsobit ztrátu některých vlastností obrazu, které mohou být důležité například pro detekci nebo rozpoznání objektů. Z tohoto důvodu bylo navrženo mnoho metod, které nejen mění dynamický rozsah obrazu, ale také se zaměřují na zlepšení lokálního kontrastu při zachování důležitých detailů. [37]

Vylepšení kontrastu pomocí vlnkové transformace

Jednou z metod zaměřujících se na zlepšení lokálního kontrastu je metoda Thomase Reevea a Eda Jernigana [30], která využívá koeficientů diskrétní vlnkové transformace zároveň pro vylepšení kontrastu a potlačení šumu v obrazu. Algoritmus je rozdělen na dvě části – první vylepšuje kontrast na hranách v obrazu, druhá vylepšuje kontrast mezi většími stejnorodými plochami. K transformaci jsou použity biortogonální spline vlnky, které díky své podobnosti LoG¹ filtru umožňují zesílením detailních koeficientů DWT zvýraznit kontrast na hranách obrazu. Detailní koeficienty $D_{2^j}^\theta$ jsou filtrovány upravenou verzí LLMSE (Local Linear Minimum Mean Square Error) filtru [23]:

$$\hat{D}_{2^j}^\theta = \kappa D_{2^j}^\theta + \frac{\sigma_n^2}{\sigma_D^2} \mu_D \quad (3.18)$$

$$\kappa = \begin{cases} \alpha_{2^j} - \frac{\sigma_n^2}{\sigma_D^2} & \text{pro } \sigma_n^2 \leq \sigma_D^2 \\ 0 & \text{pro } \sigma_n^2 > \sigma_D^2 \end{cases} \quad (3.19)$$

kde θ značí orientaci detailních koeficientů (horizontální, vertikální nebo diagonální). σ_D^2 je rozptyl signálu, spočítaný v lokálním okně o určité velikosti. Hodnoty σ_n^2 je založena na analýze rozptylu šumu v jednotlivých měřících DWT. μ_D je střední hodnota koeficientů v daném okně. Faktor α_{2^j} ovlivňuje zvýraznění kontrastu na hranách a stupeň potlačení šumu. Dle autorů této metody jsou pro dekompozici ve třech úrovních vhodné hodnoty $(\alpha_{2^1}, \alpha_{2^2}, \alpha_{2^3}) = (0,25; 0,5; 1)$ nebo $(0,5; 1; 1,5)$, tedy se snižující se úrovní detailů se zvyšuje důležitost koeficientů DWT. Pro zlepšení kontrastu mezi většími oblastmi, obsahujícími nízké frekvence autoři navrhují aplikovat ekvalizaci histogramu na aproximační koeficienty DWT. Protože vztah mezi dynamickým rozsahem původního a transformovaného obrazu není jednoduše odvoditelný, je obtížné zjistit rozsah koeficientů a šířku intervalů pro ekvalizaci histogramu. Jednoduchým řešením je omezit rozsah na aktuální rozsah detailních koeficientů a rovnoměrně rozdělit intervaly. Problémem je, pokud původní obraz plně nevyužívá dostupný dynamický rozsah, potom by bylo vhodné zvýšit rozsah koeficientů nad původní hodnoty.

Poněkud odlišný přístup zvolil ve své práci [18] Raanan Fattal. Ten nerozlišuje mezi kontrastem na hranách a mezi většími plochami v obrazu, ale pracuje s lokálním kontrastem v různých měřících. Správněji by se tedy dalo hovořit o zlepšení lokálního kontrastu (nebo zvýraznění detailů), nicméně při správné volbě počtu úrovní transformace a několika dalších parametrů se metoda chová podobně metodám zlepšujícím globální kontrast. Hlavní výhoda metody spočívá ve využití edge-avoiding transformace. Díky tomu jsou respektovány hrany v obraze a téměř nedochází k nežádoucím halo efektům a obrácením gradientů. Postup výpočtu je následující:

¹Laplacian of Gaussian

- Pracuje se s logaritmem jasové složky barevného modelu YUV.
- Po dopředné EAW transformaci obrazu jsou koeficienty WT upraveny podle následující rovnice:

$$\beta a^J, \{d^j\}_{j=1}^J = \beta a^J, \{\gamma^j d^j\}_{j=1}^J, \quad (3.20)$$

kde β a γ jsou koeficienty ovládající míru zvýraznění detailů a a^J , d^j jsou koeficienty WT.

- Obraz je zpět převeden z logaritmického měřítka a případně také do barevného modelu RGB, pracujeme-li s barevným obrazem.

Koeficienty $\{\gamma^j\}_{j=1}^J$ jsou určeny kubickým polynomem $p(j)$. Ten interpoluje uživatelem zadané hodnoty $p(1)$, $p(J/2)$ a $p(J)$, které určují míru zvýraznění jemných, středních, resp. hrubých detailů (neboli vysokých, středních, resp. nízkých frekvencí v obrazu). Při volbě hodnot vyšších než 1 dochází k zvýraznění detailů, naopak při volbě nízkých hodnot $p(1)$, případně i $p(J/2)$ a $p(J)$ dojde k rozmazání obrazu se zachováním hran. Koeficient β ovlivňuje aproximační koeficienty transformace, lze jej použít ke korekci celkové světlosti obrazu.

Výsledky algoritmu pro zvýraznění lokálního kontrastu nebo rozmazání obrazu se zachováním hran jsou na obr. 3.10.



Obrázek 3.10: Zvýraznění lokálního kontrastu. (a) původní obraz, (b) kombinované zvýraznění detailů, (c) rozmazání obrazu se zachováním hran, (d) zvýraznění jemných detailů, (e) zvýraznění středních detailů, (f) zvýraznění hrubých detailů.

3.4 Zobrazování HDR

Poslední dobou vzniká stále více zařízení vytvářející obraz s vysokým dynamickým rozsahem (HDR). Nicméně většina zobrazovacích zařízení takový rozsah stále neumí zobrazit. Proto se pro zobrazení HDR obrazu používá technika nazývaná komprese dynamického rozsahu (někdy také označovaná jako tone mapping). Jejím cílem je namapovat sadu barev s vysokým rozsahem na druhou, s rozsahem menším tak, aby výsledný obraz byl vizuálně podobný skutečné scéně. Další cíle mohou být například zachování co nejvíce detailů v obrazu nebo maximální zachování kontrastu.

Tradiční metody

Techniky tone mappingu se dají rozdělit do dvou hlavních kategorií. První jsou metody globální, které pomocí nelineární funkce namapují každý pixel v obraze stejným způsobem,

bez ohledu na jeho okolí. Jsou většinou jednoduché na implementaci ale můžou způsobit snížení kontrastu v obrazu. Příkladem je gamma komprese:

$$V_{out} = A \cdot V_{in}^\gamma, \quad (3.21)$$

kde V je jas obrazu $A > 0$ a $0 < \gamma < 1$. Jas je namapován z rozsahu $[0, 1/A^{1/\gamma}]$ do $[0, 1]$. Dochází ale ke ztrátě kontrastu v obrazu.

Druhou kategorií jsou metody lokální, u kterých se parametry nelineární funkce mění podle vlastností jeho okolí. Tyto metody jsou složitější a dosahují lepších výsledků. Nevýhodou těchto metod je častý výskyt artefaktů, např. halo efektu.

Jednou z pokročilých lokálních metod je adaptivní logaritmická komprese [13]. Její základní myšlenkou je přizpůsobení základu logaritmu v závislosti na záři (angl. radiance) pixelu. Výsledný jas pixelu je vypočítán následovně:

$$L_d = \frac{L_{dmax} \cdot 0,01}{\log_{10}(L_{wmax} + 1)} \cdot \frac{\log(L_w + 1)}{\log\left(2 + \left(\left(\frac{L_w}{L_{wmax}}\right)^{\frac{\log(b)}{\log(0,5)}}\right) \cdot 8\right)}, \quad (3.22)$$

kde L_w je původní jas pixelu, L_d jas výsledný, L_{wmax} je maximální jas ve scéně a L_{dmax} je konstanta určující maximální hodnotu jasu na použitém zobrazovacím zařízení. b je parametr ovlivňující jasovou křivku obrazu.

HDR komprese pomocí vlnkové transformace

Jednou z možností jak komprimovat dynamický rozsah obrazu je využít koeficientů vlnkové transformace. Raanan Fattal ve své práci [18] navrhuje použít tzv. edge avoiding wavelet transformaci (EAW, viz kapitola 2.1.6). Využitím EAW lze dosáhnout komprese dynamického rozsahu se zachováním detailů. Komprese se dosáhne změnou velikosti aproximačních a detailních koeficientů a^J a d^J . Pracuje se s logaritmem jasového kanálu barevného modelu YUV. Na základě EAW dekompozice, $\log Y \mapsto a^J, \{d^j\}_{j=1}^J$, se spočítá dynamicky komprimovaná jasová složka $Y'(x, y)$, a to změnou měřítka koeficientů:

$$f(a^J), \{f(d^j)\}_{j=1}^J \mapsto \log Y'(x, y), \quad (3.23)$$

kde

$$f(x) = \text{sgn}(x)|x|^\beta \quad (3.24)$$

$\beta < 1$ je parametr udávající míru komprese a \mapsto značí dopřednou a zpětnou EAW transformaci. Na obr. 3.11 je scéna „Stanford memorial church“ zobrazená pomocí EAW HDR komprese a stejná scéna zobrazená pomocí gamma komprese. Tato metoda produkuje obrázky s vysokou úrovní HDR komprese při zachování důležitých detailů v obraze a bez nežádoucích halo efektů.



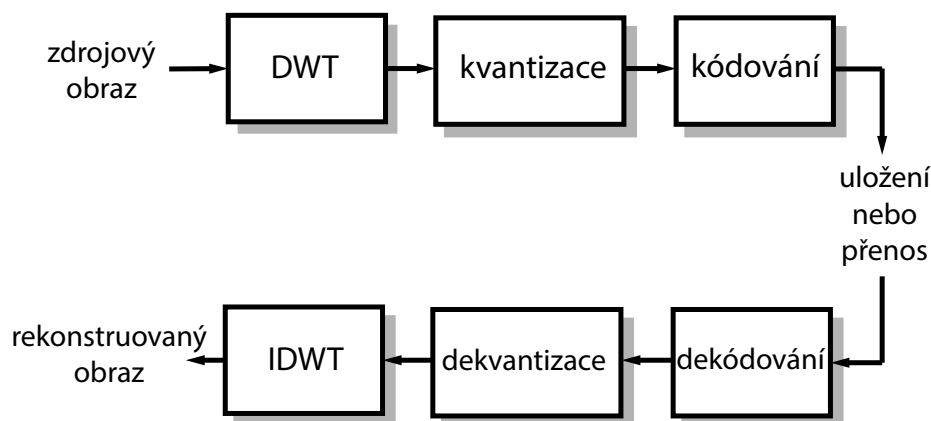
Obrázek 3.11: HDR komprese. (a) pomocí gamma komprese ($\gamma = 6$), (b) pomocí EAW transformace ($\beta = 0,7$).

3.5 Ostatní aplikace

Kromě výše zmíněných aplikací lze vlnkovou transformaci využít i k mnoha dalším. V této kapitole je představena jedna z neznámějších a nejpoužívanějších – komprese obrazu, a dále také kolorizace obrazu a tzv. inpainting.

Komprese obrazu

První obecně známé využití vlnkové transformace v kódování obrazu bylo její užití ve standardu FBI pro kompresi otisků prstů z roku 1993. Pro kompresi obecného obrazu se nejvíce ujal standard JPEG, poprvé představený v roce 1992. Komprese podle standardu JPEG je realizována kvantizací a kódováním koeficientů diskrétní kosinové transformace. Algoritmus je poměrně jednoduchý, lze dosáhnout dobrého kompresního poměru, nicméně je zde prostor k vylepšení, například v kvalitě obrazu nebo odolnosti vůči chybám. I z těchto důvodů byl v roce 2000 představen standard JPEG 2000. Schéma komprese a dekomprese (obr. 3.12) je velmi podobné schématu JPEG, vstupní obraz je tedy transformován a výsledné koeficienty jsou kvantovány a zakódovány. Hlavní rozdíl je ve využití diskrétní vlnkové transformace namísto kosinové, a v použití kódování založeném na algoritmu EBCOT [34].



Obrázek 3.12: Schéma komprese a dekomprese podle standardu JPEG 2000

Celý postup komprese dle standardu JPEG 2000 by se dal shrnout do následujících bodů [9]:

- Rozdělení obrazu na jednotlivé barevné komponenty. Při volbě ztrátové komprese se RGB obraz nejprve převede na $YCbCr$, pro bezztrátovou je nutné použít transformaci do barevného prostoru YUV, ke které existuje přesná zpětná transformace. Barevné prostory $YCbCr$ a YUV poskytují lepší dekorelaci barev, která je pro kompresi důležitá.
- Rozdělení komponent do stejně velkých pravoúhlých dlaždic. Malá velikost dlaždic snižuje výslednou kvalitu obrazu.
- Aplikace DWT (o několika úrovních) na každou dlaždici. Pro ireverzibilní (ztrátovou) kompresi se použije vlnka CDF 9/7, pro reverzibilní (bezeztrátovou) vlnka CDF 5/3. Transformaci je možné implementovat pomocí konvoluce nebo liftingu.
- Kvantizace koeficientů DWT a jejich rozdělení do čtvercových polí, tzv. code blocks. Kvantizační krok je vypočítán podle dynamického rozsahu daného podpásmu nebo je v případě bezztrátové komprese roven 1 (více v [9]). Tři skupiny několika polí code blocks z každého podpásmu na souhlasících místech tvoří tzv. precinct, důležitý pro následné kódování.
- Entropické kódování bitových rovin koeficientů v každém poli code block. Bitové roviny v rámci pole jsou čteny po 4 bitech po sloupcích zleva doprava. Kódování je provedeno tříprůchodovým kontextovým binárním aritmetickým kóděrem. Výstupem kóděru jsou 8 bitové pakety jdoucí v pořadí z podpásem LL, HL, LH a HH. Celkový výstupní datový tok je tvořen vrstvami, kde každá vrstva obsahuje proměnné množství informace (i nulové) od každého pole code block (více v [9]).
- Přidání značek (markerů) a hlavičky s informacemi pro dekompresi do výsledného bitového toku.

Na obr. 3.13 je srovnání komprese pomocí vlnkové transformace (JPEG 2000) s kompresí kosinovou transformací (JPEG). Vlnková transformace produkuje při stejném kompresním poměru vizuálně kvalitnější obraz.



Obrázek 3.13: Porovnání komprese JPEG 2000, využívající DWT s kompresí JPEG, využívající DCT. Kompresní poměr je 100:1. (a) Původní obraz, (b) JPEG 2000, (c) JPEG. Obrázky převzaty z [16]

Kolorizace obrazu

Vlnkovou transformaci, konkrétně její edge-avoiding variantu, využívá Raanan Fattal [18] mimo jiné i ke kolorizaci obrazu. V kolorizaci jde o přidání barevné informace k monochromatickému obrazu. Uživatel aplikace většinou vyznačí několik oblastí s konkrétním barevným odstínem a úkolem programu je tyto barvy roz distribuovat po celém obraze (nebo v případě selektivní kolorizace jen po jeho části). Fattal pro kolorizaci s využitím EAW navrhuje následující postup:

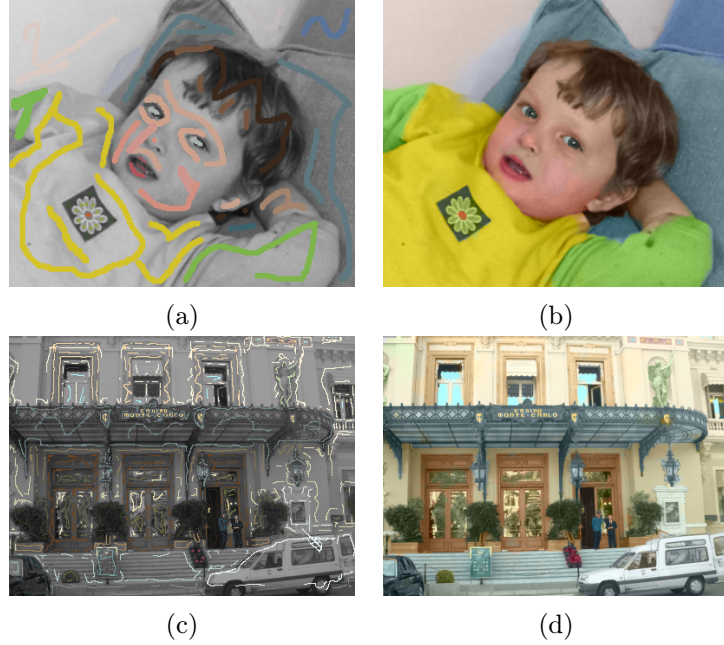
- Na vstupu je monochromatický obraz $Y(x, y)$ a obrazy s uživatelem zadanou barevnou informací $U(x, y)$ a $V(x, y)$. Tyto obrazy představují jednotlivé komponenty obrazu v barevném modelu YUV . Cílem navržené metody je interpolovat barevnou informaci se zohledněním hran v obraze. K tomu je definována normalizační funkce $N(x, y)$, která nabývá hodnoty 1 tam, kde uživatel zadal informaci o barvě, jinak 0.
- Prvním krokem je výpočet EAW transformace obrazu Y k získání vah pixelů. Stačí použít upravenou verzi této transformace, ve které jsou aproximační koeficienty zahozeny a detailní nejsou vůbec počítány.
- Na základě vypočítaných vah EAW transformace jsou z obrazů U , V a N vypočítány obrazy \tilde{U} , \tilde{V} a \tilde{N} . Matematická definice výpočtu je:

$$\bar{I}(x, y) = \sum_{j=1}^J 2^{-j} \sum_{\tilde{\phi} \in \tilde{V}^j} \tilde{\phi}(x, y) \langle I(x, y), \tilde{\phi}(x, y) \rangle, \quad (3.25)$$

kde $\langle \cdot, \cdot \rangle$ značí skalární součin a $\tilde{\phi}$ škálovací funkci. Jedná se tedy o vážený součet zpětných transformací obrazů, které průměrují pixely patřící ke stejným škálovacím funkcím. Váha 2^{-j} upřednostňuje jemnější měřítka, aby informace od pixelů, které jsou blíže k sobě, měly silnější efekt na výsledný obraz. Algoritmus je podrobněji popsán ve výše zmíněné práci.

- Nakonec jsou vypočítány normalizované barevné komponenty YUV obrazu $\tilde{U}(x, y) = \tilde{U}(x, y)/\tilde{N}(x, y)$ a $\tilde{V}(x, y) = \tilde{V}(x, y)/\tilde{N}(x, y)$, které v kombinaci se obrazem Y ze vstupu tvoří výsledný barevný obraz.

Příklady kolorizace monochromatických obrazů jsou na obr. 3.14.



Obrázek 3.14: Výstup algoritmu kolorizace využívajícího EAW transformaci. (a), (c) Vstupní obrazy s uživatelem vyznačenými barevnými oblastmi. (b), (d) Výsledek kolorizace. Obrázky převzaty z [18].

Inpainting

Inpainting je technika zpracování obrazu, která se snaží obnovit či doplnit porušené nebo ztracené části obrazu. Lze ji také použít k odstranění nechtěných objektů z obrazu. Existují dva základní přístupy k inpaintingu – strukturálně nebo texturově orientovaný. Metoda Tonyho Chana et al. [7], využívající vlnkovou transformaci patří k strukturálně orientovaným, snaží se zachovat konzistenci v geometrické struktuře obrazu. Jejich metoda, narozdíl od mnoha dalších, pracuje pouze ve vlnkové doméně – rekonstrukce je založena na obnovení koeficientů vlnkové transformace. K tomu autoři navrhli použít techniku minimalizace totální variace (total variation minimization, TVM). Představili dvě metody, jednu pro obraz bez nebo se zanedbatelným množstvím šumu a druhou pro zašuměný obraz. TVM schéma pro metodu pro obraz bez šumu je:

$$\min_{\beta_{j,k}:(j,k) \in I} F(U, z) = \int_{\mathbb{R}^2} |\nabla_x u(\beta, x)| dx = \text{TV}(u(\beta, x)), \quad (3.26)$$

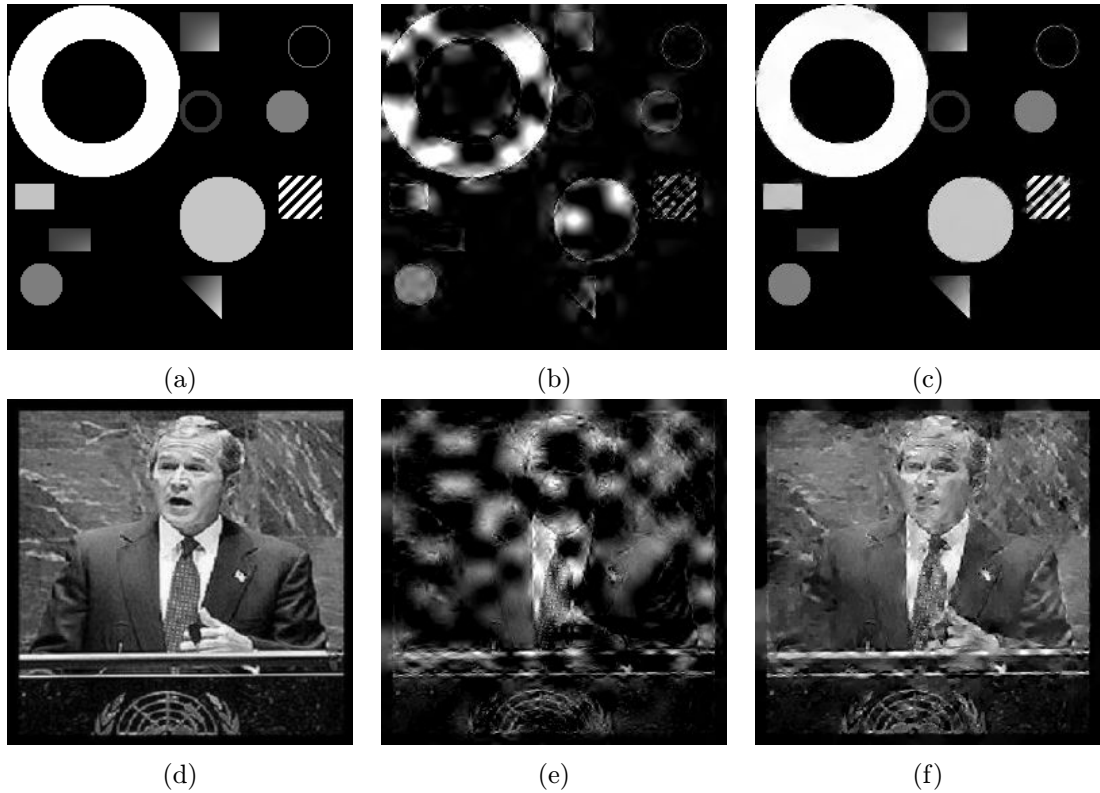
kde $u(\beta, x)$ je vlnková transformace:

$$u(\beta, x) = \sum_{j,k} \beta_{j,k} \psi_{j,k}(x), \quad j \in \mathbb{Z}, k \in \mathbb{Z}^{\neq}, \quad (3.27)$$

splňující podmínku $\beta_{j,k} = \alpha_{j,k}$, $(j, k) \notin I$ kde I je oblast pro inpainting (tedy s poškozenými koeficienty), α a β koeficienty WT (pro zjednodušení nejsou rozlišeny aproximační a detailní).

Minimalizace totální variace vede k řešení soustavy Euler-Lagrangeových rovnic². Numerický algoritmus pro jejich řešení, stejně jako TVM model pro obrazy se šumem jsou uvedeny ve výše citované práci.

Výsledky algoritmu jsou na obr. 3.15.



Obrázek 3.15: Výstup algoritmu inpaintingu využívajícího vlnkovou transformaci. (a) Syntetický vstupní obraz, (d) reálný vstupní obraz, (b), (e) poškozené obrazy – zasaženo 50% koeficientů WT, (c), (f) obnovené obrazy. Obrázky převzaty z [7].

²diferenciální rovnice umožňující nalezení extrému funkcionálu

Kapitola 4

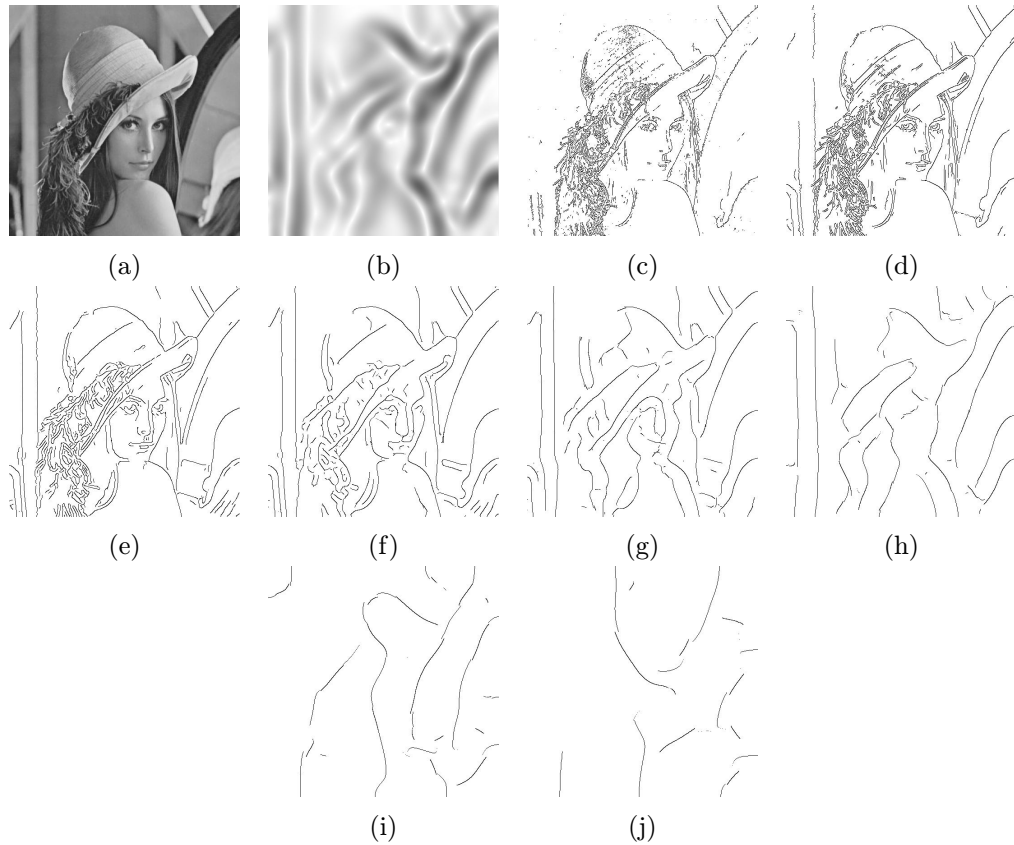
Testování a výsledky

Z aplikací vlnkové transformace, představených v této práci, jsem vybral k implementaci do C++ knihovny první čtyři z nich: detekci hran, odstranění šumu, vylepšení kontrastu a zobrazení HDR obrazu. V následujících kapitolách se budu věnovat testování vybraných metod implementujících jednotlivé aplikace WT.

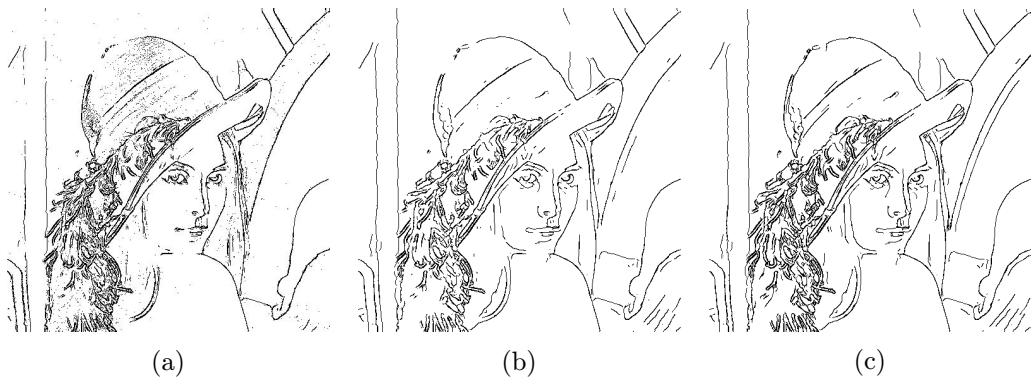
4.1 Detekce hran

Pro detekci hran jsem zvolil algoritmus Stéphane Mallata a Sifena Zhonga, popsany v kapitole 3.1. Tento algoritmus jsem implementoval v jeho původní podobě, jako vlnkovou funkci jsem tedy zvolil derivaci kvadratické spline křivky, která aproximuje derivaci Gaussovy funkce. Podstatnými parametry, ovlivňujícími detekci je počet úrovní transformace J (kde $j = 1$ značí první úroveň transformace) a volba prahu T_j pro každou z nich. Obrázek 4.1 znázorňuje výstup detekce testovacího obrázku Lena pro všech osm možných úrovní. Protože vyšší úrovně rozkladu již pracují s příliš rozmazaným obrazem (viz obr. 4.1b), detekované hrany již příliš neodpovídají skutečnosti. Osvědčilo se mi využívat tři, maximálně čtyři úrovně rozkladu.

Volba hodnoty prahu je buďto přenechána na uživateli aplikace – uživatel potom volí hodnoty prahů T_1 a T_J , ostatní jsou lineárně interpolovány, nebo lze zvolit metodu prahování OTSU. Dále je možné zvolit mezi prahováním se závěrečnou hysterezí, podobně jako u metody Canny, nebo bez ní. Metoda OTSU volí v nejnižší úrovni příliš nízký práh a produkuje proto i nežádoucí hrany odpovídající šumu. Hystereze některé hrany prodlužuje, výsledek detekce potom lépe popisuje původní obraz. Všechny varianty jsou na obr. 4.2.



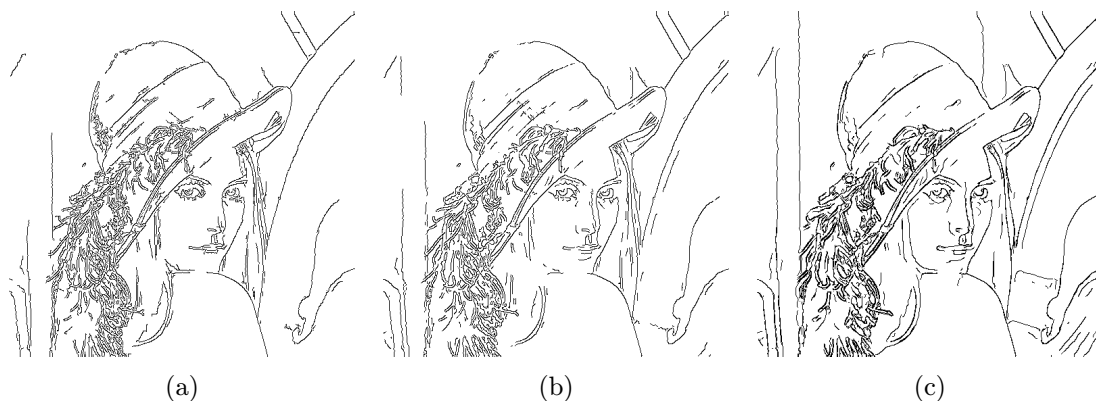
Obrázek 4.1: Výstup detekce hran obrázku Lena pro všechny úrovně rozkladu. (a) Původní obraz, (b) absolutní hodnota detailních koeficientů SWT 6. úrovně rozkladu, (c) - (j) detekované hrany.



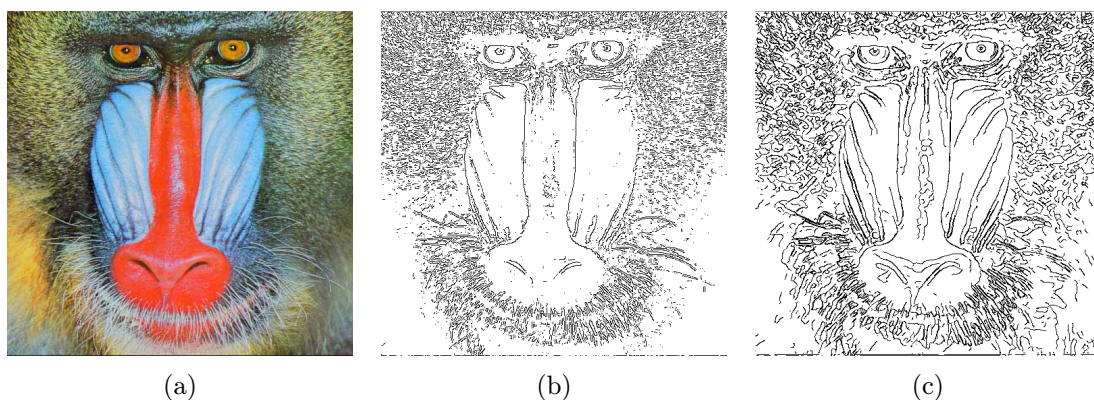
Obrázek 4.2: Možnosti prahování (sečteny výstupy pro $J = 3$). (a) Metoda OTSU, (b) prahování bez hystereze, (c) prahování s hysterezí.

Metody, které pouze vypočítají první či druhou derivaci obrazu (popsané v kapitole 3.1), a které dále hrany nijak nefiltrují, produkují obvykle příliš velké množství nežádoucích hran. I při použití prahování není výsledek příliš uspokojivý (viz obrázek 3.2). Jejich srovnání s metodou využívající vlnkovou transformaci proto příliš nemá smysl. Naopak Cannyho metoda používá velmi podobný postup výpočtu hran a měla by produkovat velmi podobné výsledky jako metody využívající WT. Její autoři dokonce ve své práci [27] uvádí, že jejich

metoda je s Cannyho metodou ekvivalentní. Na obr. 4.3 je srovnání s Cannyho metodou. Například v testovacím obrázku Lena při hodnotách $T_1^{Canny} = 80$ a $T_2^{Canny} = 200$ je výstup Cannyho metody velmi podobný výstupu metody využívající WT pro $j = 2$ a $T_2 = 35$. Z obrázků je ale patrné, že Cannyho metoda kvůli absenci víceúrovňového rozkladu nedokáže zachytit hrany z vyšších úrovní, odpovídající pozvolnějším přechodům mezi intenzitami obrazu. Tuto vlastnost metoda Stéphane Mallata má, pokud sečteme výsledky detekce z jednotlivých úrovní. Nevýhodou potom je, že se tím poruší vlastnost hranového detektoru, kdy se po něm požaduje unikátní odezva na hranu – hrany ve vyšších úrovních se můžou vlivem rozmazání posunout a tím vzniknout dvojité odezvy na hranu. Efekt vícenásobné odezvy lze vidět např. na obrázku 3.5. Na dalším testovacím obrázku – Baboon (4.4) – se výhoda víceúrovňové detekce projevuje ještě více. U Cannyho detektoru byly nastaveny prahy $T_1^{Canny} = T_2^{Canny} = 255$, aby byly co nejvíce potlačeny jemné hrany způsobené texturou, a byly zachovány pokud možno jen hrany odpovídající geometrické struktuře obrázku. Tento úkol lépe zvládá detekce využívající WT (prahy nastaveny na $T_1 = 140$, $T_3 = 30$).



Obrázek 4.3: Srovnání s Cannyho detektorem – Lena. (a) výstup Cannyho detektoru pro $T_1^{Canny} = 80$ a $T_2^{Canny} = 200$, (b) výstup Mallatovy metody pro $j = 2$ a $T_2 = 35$, (c) sečtené výstupy Mallatovy metody pro $J = 3$ a $T_1 = 80$, $T_3 = 30$.



Obrázek 4.4: Srovnání s Cannyho detektorem – Baboon. (a) původní obraz, (b) výstup Cannyho detektoru pro $T_1^{Canny} = T_2^{Canny} = 255$, (c) sečtené výstupy Mallatovy metody pro $J = 3$ a $T_1 = 140$, $T_3 = 30$.

4.2 Odstranění šumu

Odstranění šumu v obrazu jsem implementoval pomocí diskrétní i stacionární vlnkové transformace. V případě DWT jsem použil vlnky CDF 5/3 a Red-Black a jejich vážené varianty WCDF 5/3 a WRB. V případě SWT jsem použil vlnku odvozenou z B_3 spline křivky (filtr $h = [\frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16}]$). Předpokládám obraz poškozený aditivním Gaussovým šumem. Šum se obvykle projevuje v detailních koeficientech WT několika nejnižších úrovní, proto jsem v následujících testech omezil maximální úroveň rozkladu na $J = 3$. Implementoval jsem dvě automatické metody volby prahu – upravenou verzi New threshold od autorů B. Ismaila a Anjuma Khana [21] a Bayes Shrink od Hughha Chipmana [8]. Úprava první metody spočívá ve vynásobení prahu pro j -tou úroveň faktorem 2^j . Tyto metody lze využít například v automatizovaném zpracování obrazu. V tabulce 4.1 jsou uvedeny hodnoty PSNR¹ a SSIM² pro obě metody pro DWT vlnkou CDF 5/3 s měkkým prahováním a pro SWT s tvrdým prahováním. Test byl proveden na obrázku „Peppers“ s přidáním šumu s různými standardními odchylkami σ_N . Z výsledků je patrné, že s rostoucím množstvím šumu silně klesá úspěšnost odstranění šumu metodou Bayes shrink, přičemž při použití stacionární transformace se tento efekt projevuje ještě více. Lepších výsledků lze ale dosáhnout ručním nastavením prahu. V dalších testech se už proto budu zabývat pouze vlivem různých parametrů a vlnek při ručním nastavení prahu.

σ_N	práh	PSNR [dB]		SSIM	
		DWT	SWT	DWT	SWT
10	NT	30,4	31,16	83,86 %	84,1 %
	BS	30,65	31,87	85,64 %	85,24
25	NT	25,56	25,13	76,82 %	70,14 %
	BS	25,87	24,37	73,68 %	56,47 %
50	NT	20,91	20,07	66,57 %	53,37 %
	BS	20,92	18,71	56,99 %	31,83 %

Tabulka 4.1: Porovnání metod automatické detekce prahu New threshold (NT) a Bayes shrink (BS) pro $J = 3$.

V dalším testu (viz tabulka 4.2) jsou porovnány všechny kombinace vlnek a transformací pro obraz Peppers s $\sigma_N = 25$ a pro obraz Cameraman s $\sigma_N = 50$ s parametry nastavenými tak, aby byl maximalizován PSNR a poté i SSIM a byl tak zjištěn potenciál každé transformace. Pro zjednodušení jsou všechny vlnky pro DWT spolu s SWT transformací uvedeny v jednom sloupci.

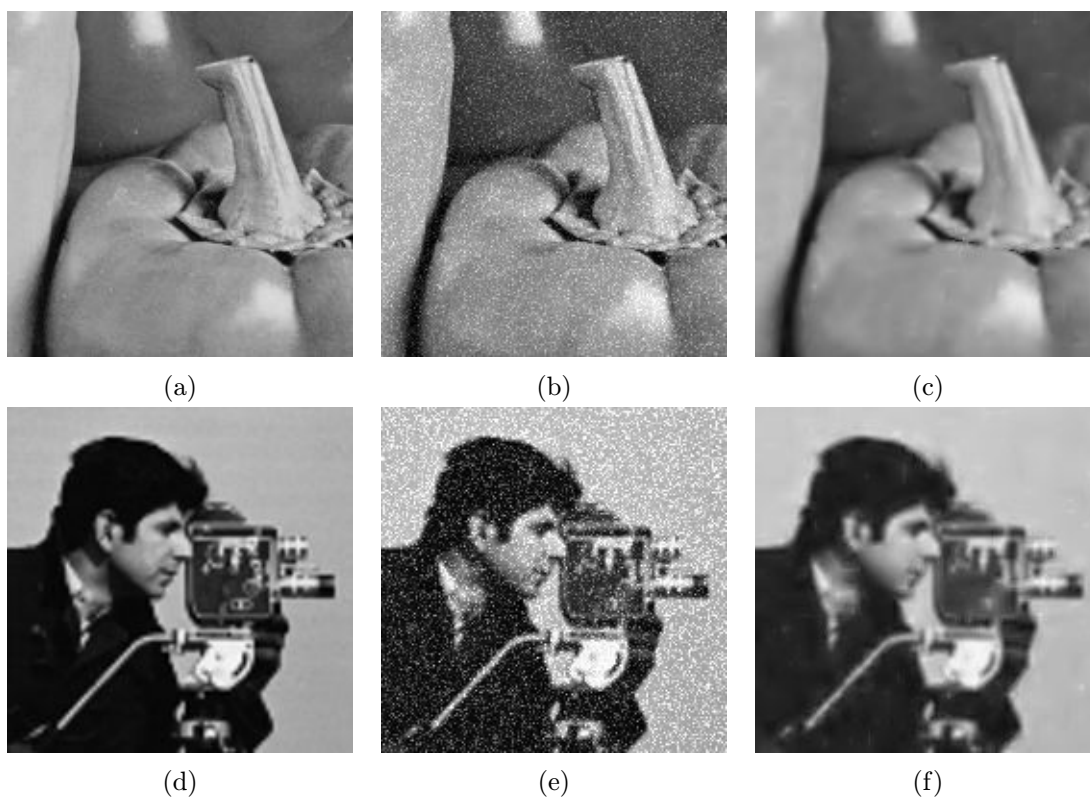
¹peak signal-to-noise ratio, špičkový odstup signálu k šumu

²structural similarity, metoda hodnocení podobnosti dvou obrazů repektující psychovizuální model lidského vidění

σ_N	transformace	PSNR _{max} [dB]	SSIM _{max}
25	CDF	26,08	77,73 %
	RB	25,89	76,58 %
	WCDF	24,88	74,67 %
	WRB	24,52	75,45 %
	SWT	26,18	88,37 %
50	CDF	21,21	70,53 %
	RB	21,06	68,19 %
	WCDF	19,25	66,19 %
	WRB	18,75	61,35 %
	SWT	21,51	78,21 %

Tabulka 4.2: Porovnání schopnosti odstranění šumu jednotlivých druhů transformace a vlnek.

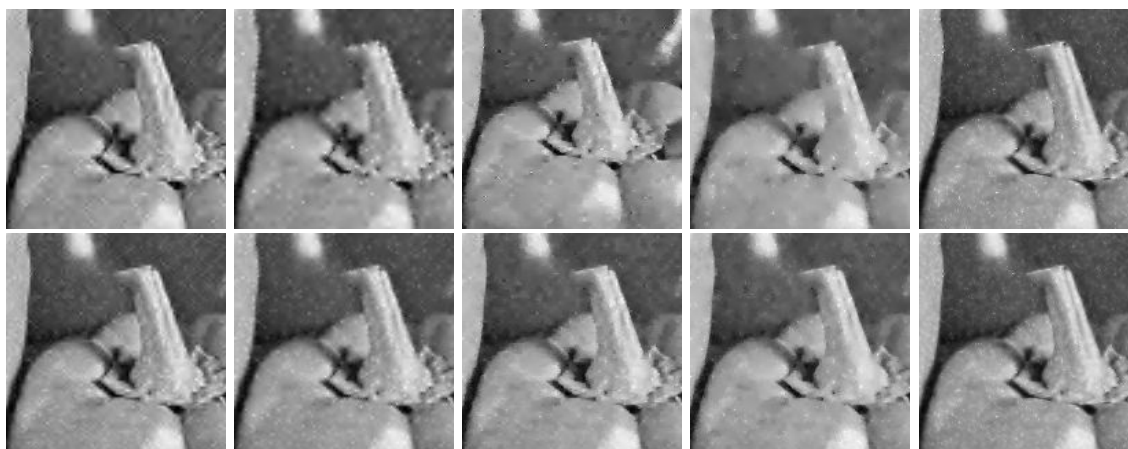
Z výsledků je patrné, že nejlepšího PSNR i SSIM dosahuje stacionární vlnková transformace. U DWT lze nejlépe odstranit šum prahováním koeficientů získaných vlnkami CDF. Vážené vlnky nedosahují dobrých výsledků ani v jednom z testů. Testovací obrazy a obrazy s přidáním a s nejlépe odstraněným šumem podle SSIM jsou na obr. 4.5.



Obrázek 4.5: Maximální PSNR a SSIM různých transformací. Obraz Peppers (detail): (a) původní obraz (b) přidáný šum se $\sigma = 25$, (c) šum odstraněný pomocí SWT, SSIM = 88,37 %. Obraz Cameraman (detail): (a) původní obraz, (b) přidáný šum se $\sigma = 50$, (c) šum odstraněný pomocí SWT, SSIM = 78,21 %.

Cílem dalšího testu je zjistit vliv použité vlnky na subjektivní kvalitu obrazu. K tomu

je použit testovací obraz Peppers s přidaným šumem o odchylce $\sigma = 30$. Parametry byly pokaždé nastaveny tak, aby výsledná podobnost podle SSIM byla 71 %. Výsledky jsou na obr. 4.6.



Obrázek 4.6: Vliv vlnky na kvalitu obrazu (SSIM = 71 %). V horní řadě použito tvrdé prahování, v dolní řadě měkké. Zleva doprava: CDF, RB, WCDF, WRB, SWT.

Při použití vlnek CDF a Red-Black vznikají artefakty ve tvaru vlnkové funkce. O něco méně jich vzniká u jejich vážených variant, které navíc díky schopnosti zachovat hrany produkují čistší obraz. Stacionární transformace netvoří žádné výraznější artefakty, jejím použitím docílíme odstranění šumu se vznikem minima rušivých elementů. Vliv artefaktů je méně znatelný při použití měkkého prahování, díky zmenšení koeficientů WT nad prahovou hodnotou.

Na obrázku 4.7 je nakonec srovnání prahování koeficientů SWT s ostatními používanými metodami pro odstranění šumu z obrazu – Gaussův filtr, mediánový filtr, bilaterální filtr a metoda non-local means. Původní obraz a obraz s přidaným šumem jsou stejné jako v 3.9.

PSNR, SSIM a doba výpočtu obrázků z 4.7 jsou shrnuty v následující tabulce:

metoda	PSNR [dB]	SSIM	doba výpočtu [ms]
SWT	26,54	87,11 %	132,18
Gaussův filtr	26,56	83,07 %	1,74
mediánový filtr	28,4	83,97 %	1,12
bilaterální filtr	27,22	82,75 %	22,47 ms
non-local means	27	86,47 %	14911

Tabulka 4.3: Srovnání PSNR, SSIM a doby výpočtu metod.

Prahování koeficientů SWT má spolu s frekvenční Gaussovou filtrací nejhorší odstup signálu od šumu, naopak má ale ze všech testovaných metod nejlepší výsledek SSIM, která bere v potaz psychovizuální model lidského vidění. Druhý nejlepší výsledek má metoda non-local means s SSIM 86,47 %. Nevýhodou může být v některých případech až příliš vyhlazený obraz, a také doba výpočtu, která je díky prohledávání velkého okolí pixelů mnohonásobně vyšší, než u ostatních metod. Prahování koeficientů dosahuje druhého nejhoršího času. Tento údaj je ale pouze orientační, algoritmus nebyl optimalizován na rychlost výpočtu.



(a)

(b)

(c)



(d)



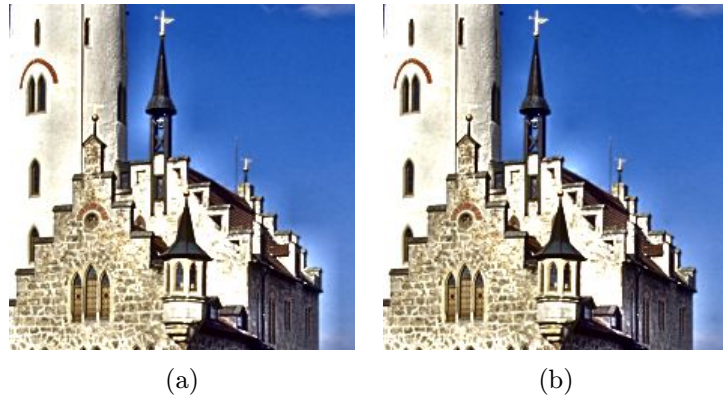
(e)

Obrázek 4.7: Porovnání s ostatními metodami redukce šumu. (a) Prahování koeficientů SWT, (b) Gaussův filtr, (c) mediánový filtr, (d) bilaterální filtr, (e) non-local means.

4.3 Vylepšení kontrastu

Pro vylepšení kontrastu v obrazu jsem zvolil metodu Raanana Fattala [18], která využívá EAW transformaci. Tu jsem implementoval pro vlnky WCDF a WRB, s volitelným parametrem váhové funkce α , pokud tedy nastavíme $\alpha = 0$, vlnky zjednodušíme na CDF, resp. klasické Red-Black. Dalšími volitelnými parametry jsou $p(1)$, $p(J/2)$ a $p(J)$, které ovlivňují míru zvýraznění kontrastu, β , který se dá použít pro korekci světlosti obrazu a maximální úroveň transformace J .

Rozdíl mezi separabilními vlnkami WCDF a neseperabilními WRB je nejvíce znatelný, pokud není zapojeno váhování ($\alpha = 0$). Protože tímto EAW transformaci degradujeme na klasickou vlnkovou transformaci, vznikají kolem hran halo efekty. Na jejich tvaru lze vidět rozdíl mezi oběma vlnkami – WCDF tvoří více hranaté artefakty. Také při rozmazání obrazu se zachováním hran je znatelný rozdíl mezi těmito vlnkami, WRB například lépe zvládá zachovat diagonální hrany. Rozdíly jsou znázorněny na obrázcích 4.8 a 4.9.



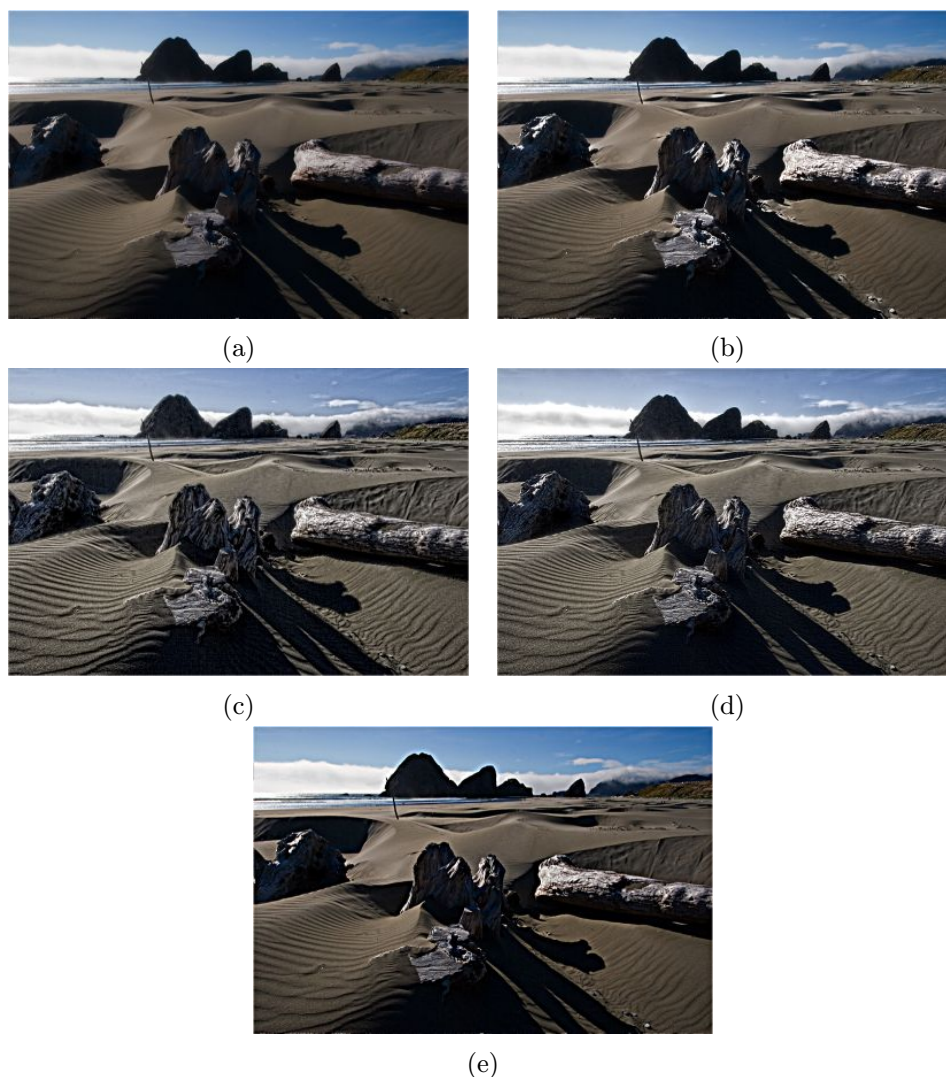
Obrázek 4.8: Srovnání vlnek WCDF a WRB při zvýraznění kontrastu ($J = 5$, $p(1) = 2,6$, $p(J/2) = 2,4$, $p(J) = 2,2$, $\beta = 1,2$). (a) CDF, (b) Red-Black.



Obrázek 4.9: Srovnání vlnek WCDF a WRB při rozmazání obrazu se zachováním hran ($J = 8$, $p(1) = 0$, $p(J/2) = 0,4$, $p(J) = 1$, $\beta = 0,5$). (a) původní obraz. (b) WCDF, (c) WRB.

Dále se zaměřím na srovnání s jinými metodami zvýšení kontrastu. Jednou ze základních metod zvyšujících lokální kontrast je tzv. unsharp masking. To efektivně zvýrazní detaily

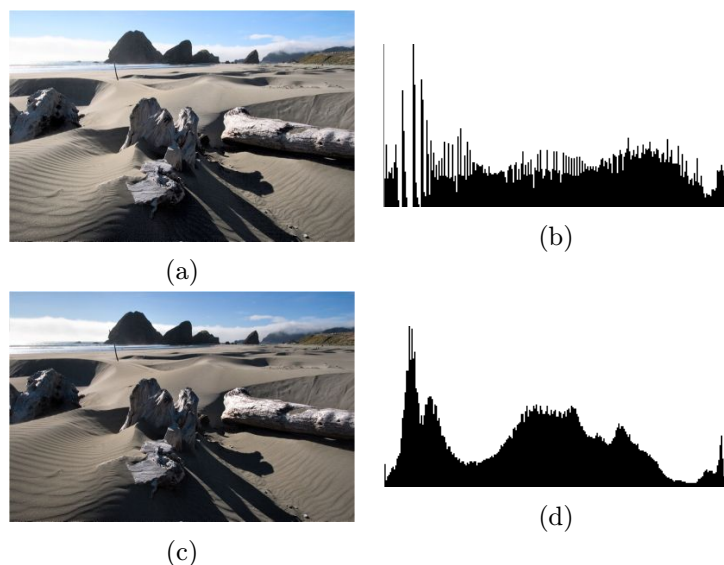
v měřítku, které odpovídá zvolené velikosti Gaussova filtru, který metoda využívá. Nevýhodou jsou výrazné halo efekty, jak je vidět kolem hory v pozadí na obr. 4.10. Další, již pokročilejší techniky, které jsem vybral pro srovnání je metoda Raanana Fattala [19] využívající bilaterální filtraci a metoda Zeeva Farbmana [17], využívající optimalizaci vážených nejmenších čtverců. Obě efektivně zvýrazňují detaily, přičemž druhá jmenovaná je téměř bez jakýchkoliv nežádoucích efektů. Výstupy obou z nich jsou spolu s metodou využívající EAW transformaci zobrazeny na obr. 4.10. Na tomto testovacím obrazu byla EAW metoda nastavena na zvýraznění jemných až středních detailů. Lokální kontrast byl zvýšen méně než u pokročilých metod Fattala a Farbmana. Na některých místech v obraze se také vyskytují „přepálené“ bílé plochy.



Obrázek 4.10: Srovnání metod vylepšení lokálního kontrastu. (a) původní obraz, (b) EAW metoda, $J = 7$, $p(1) = 2$, $p(J/2) = 1,5$, $p(J) = 1$, $\beta = 1$, (c) metoda Raanana Fattala využívající bilaterální filtr, (d) metoda Zeeva Farbmana využívající optimalizaci vážených nejmenších čtverců, (e) dvojitá aplikace unsharp masky. Obrázky (c) a (d) převzaty z [19], resp. z [17].

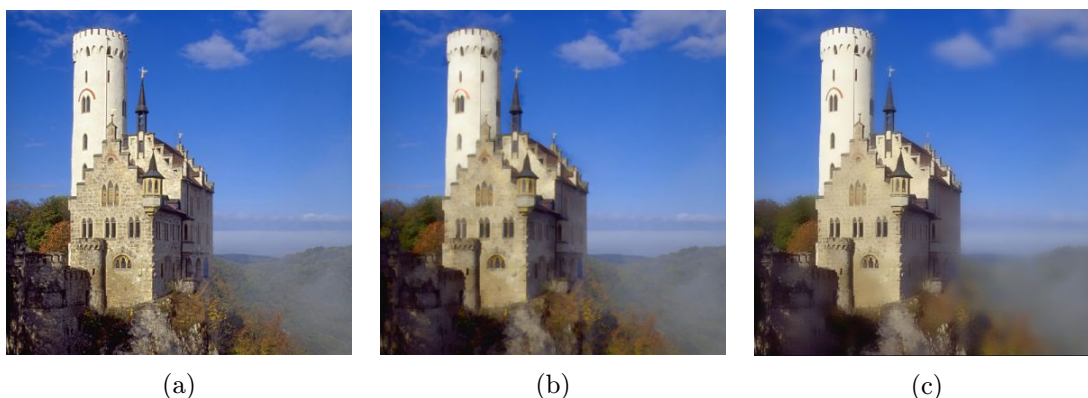
Také jsem zde zařadil srovnání s velmi často používanou technikou vylepšení globálního

kontrastu – ekvalizací histogramu. Na obr. 4.11 lze vidět testovaný obraz s ekvalizovaným histogramem a stejný obraz se zvýšeným lokálním kontrastem s parametry nastavenými tak, aby se výsledek podobal zvýšení globálního kontrastu.



Obrázek 4.11: Zvýšení globálního kontrastu. (a), (b) Obraz s ekvalizovaným histogramem, (c), (d) EAWT metoda, $J = 8$, $p(1) = 0,8$, $p(J/2) = 0,8$, $p(J) = 0,8$, $\beta = 0,7$.

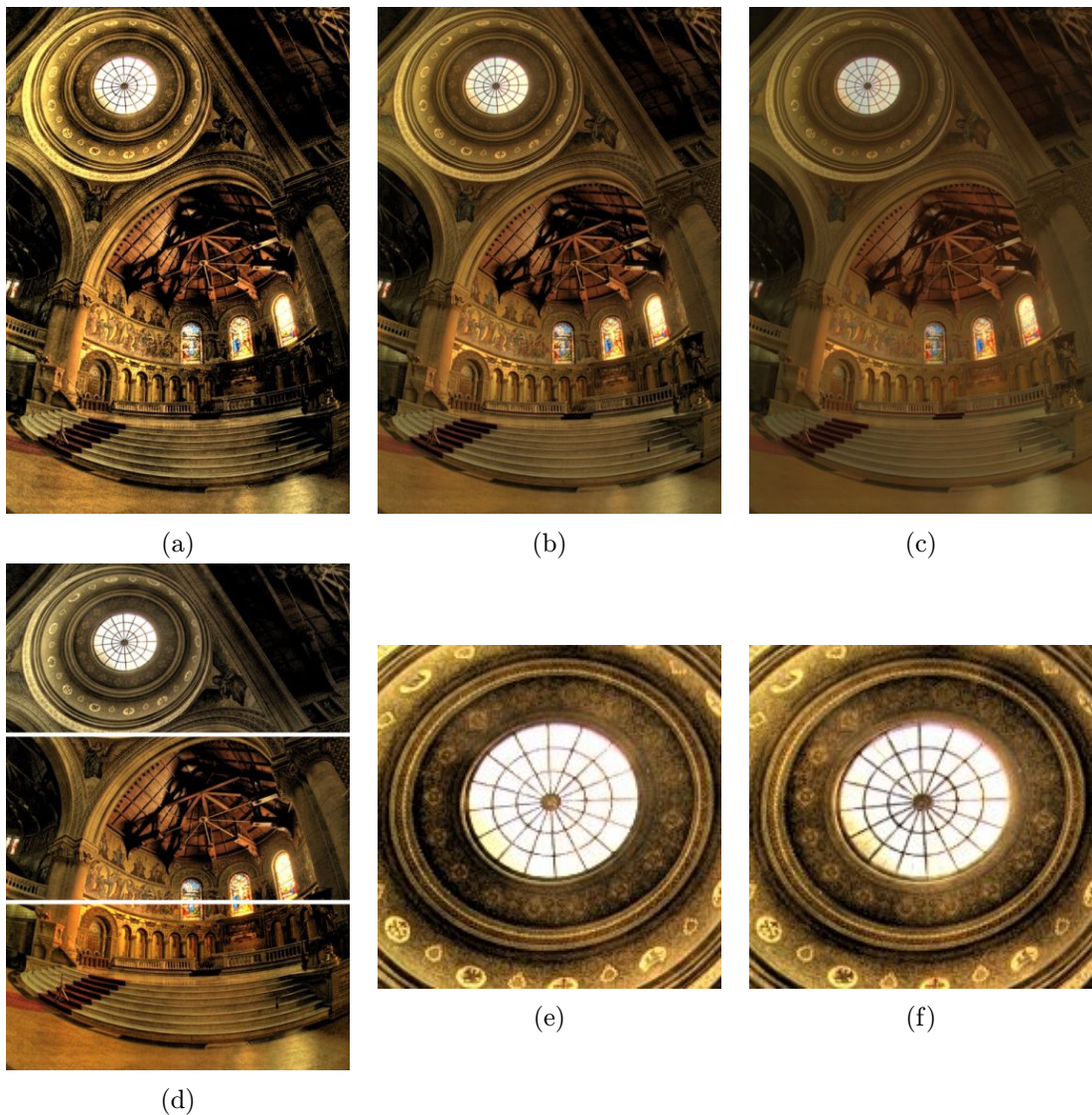
Rozmazání obrazu se zachováním hran nelze provést klasickou lineární filtrací. Proto jsem ke srovnání vybral bilaterální filtr, který je nelineární a má při správném nastavení velmi dobrou schopnost zachování hran. Srovnání metody využívající EAW transformaci (s parametry $J = 6$, $p(1) = 0,2$, $p(J/2) = 0,6$, $p(J) = 1$ a $\beta = 1$) s bilaterální filtrací je na obr. 4.12. Bilaterální filtrace zachovává čisté hrany, zatímco u první jmenované se vyskytují artefakty ve tvaru škálovací funkce EAW transformace.



Obrázek 4.12: Rozmazání obrazu se zachováním hran. (a) původní obraz, (b) EAW metoda, $J = 6$, $p(1) = 0,2$, $p(J/2) = 0,6$, $p(J) = 1$, $\beta = 1$, (c) bilaterální filtrace.

4.4 Zobrazování HDR

Pro zobrazení obrazů s vysokým dynamickým rozsahem jsem zvolil metodu Raanana Fattala [18] využívající koeficienty EAW transformace. Je možné využít vlnky CDF 5/3 a Red-Black, a to především v jejich vážených variantách respektujících hrany. Počet úrovní rozkladu J je vypočítán z rozměrů vstupního obrazu – aby algoritmus správně fungoval, musí být nastaven na maximální hodnotu. Volitelnými parametry jsou α , ovlivňující sílu edge-stopping funkce a β , která udává míru komprese. Oproti původní práci jsem přidal parametr γ , která ovládá míru barevnosti při konverzi mezi barevnými modely RGB a YUV, jak doporučuje Jiang Duan v [14]. Na obrázku 4.13 jsou výsledky tone mappingu Fattalovou metodou s různě nastavenými parametry α , β a γ .

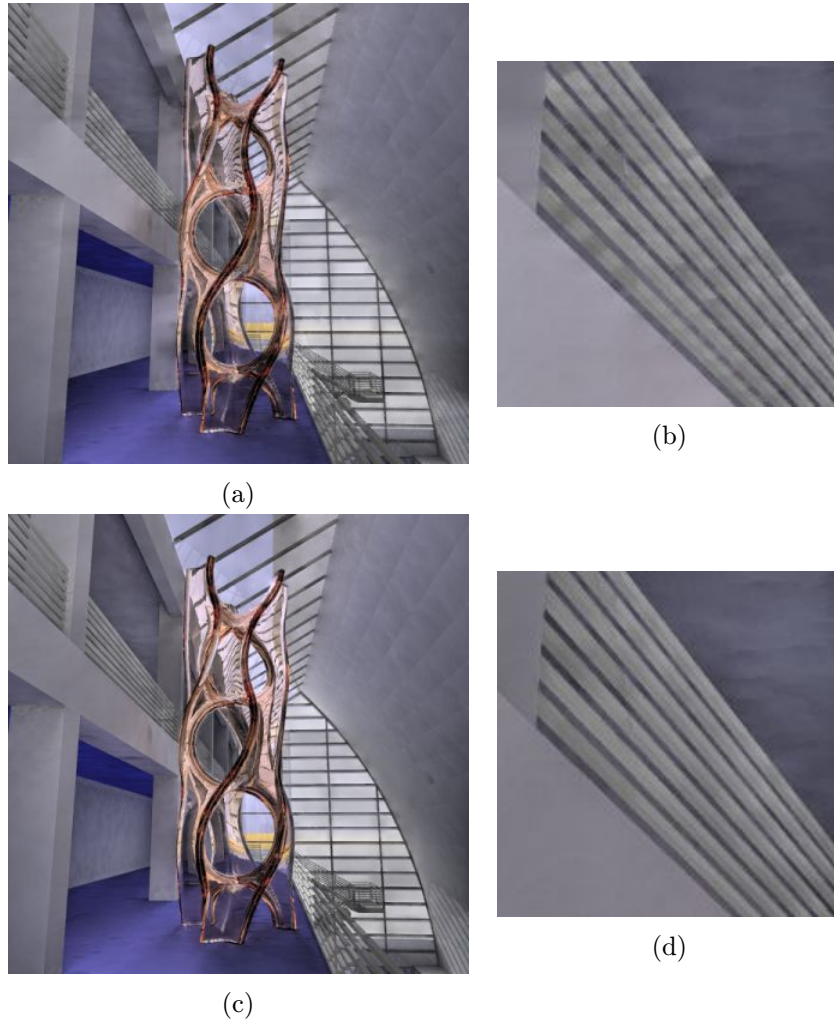


Obrázek 4.13: Tone mapping pomocí EAW transformace. (a) $\beta = 0,6$, (b) $\beta = 0,8$, (c) $\beta = 1,0$, (d) shora dolů: $\gamma = 0,3$, $\gamma = 0,5$, $\gamma = 0,7$, (e) $\alpha = 0$, (f) $\alpha = 0,8$.

Parametr β je vhodné nastavit v rozmezí 0,6 – 0,8, v závislosti na míře kontrastu původního obrazu. Vyhovující hodnoty parametrů α a γ pro většinu obrazů jsou $\alpha = 0,8$

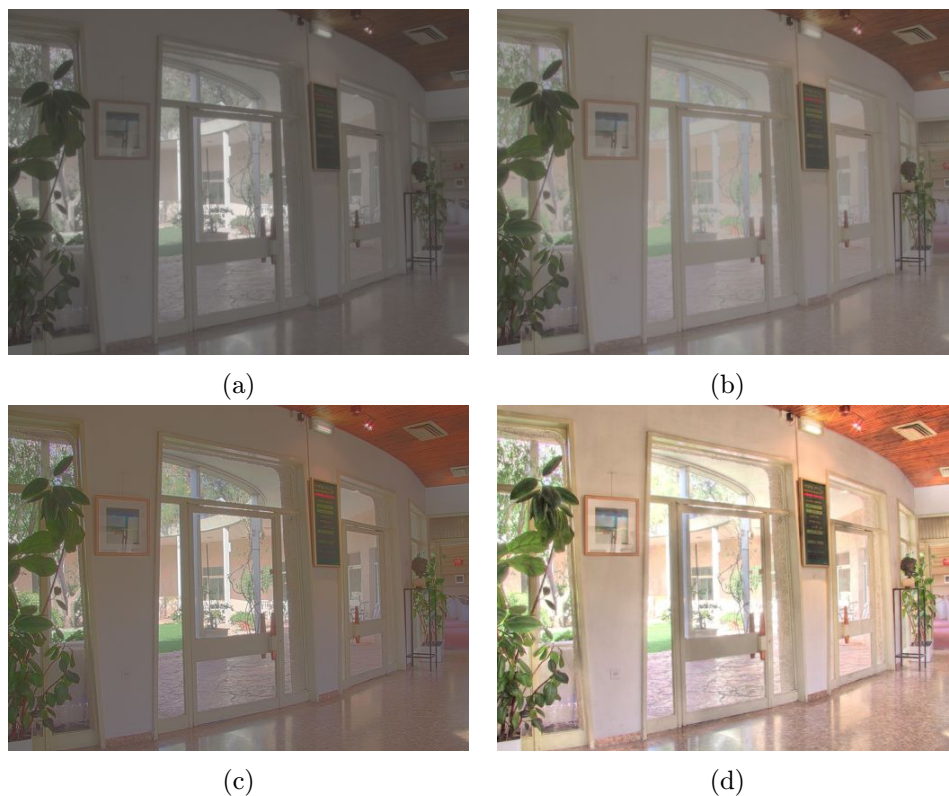
a $\gamma = 0,5$, pokud nechceme potlačit vliv případných artefaktů (v případě α) nebo upravit barevnost původního obrazu (parametr γ).

Na obrázku 4.14 je na uměle renderované scéně prezentován rozdíl mezi vlnkami WCDF a WRB. Vlnky WCDF obvykle nezvládnou zachovat diagonální hrany a na obrazu také mohou tvořit „mapy“ ve tvaru škálovací funkce ve vyšších úrovních rozkladu.



Obrázek 4.14: Porovnání vlnek WCDF a WRB při tone mappingu. $\alpha = 0,8$, $\beta = 0,6$, $\gamma = 0,5$. (a), (b): WCDF (c), (d): WRB.

Poslední test srovnává metodu Raanana Fattala s ostatními používanými technikami HDR komprese s vhodně nastavenými parametry. Zvolil jsem jednoduchou gamma kompresi, adaptivní logaritmickou kompresi Fréderica Draga [13] a metodu Frédo Duranda využívající bilaterální filtraci [15].



Obrázek 4.15: Porovnání metod tone mappingu. (a) gamma komprese, (b) adaptivní log. komprese, (c) metoda využívající bilaterální filtr, (d) EAW transformace, $\beta = 0,7$.

V tabulce 4.4 je uveden čas výpočtu každé metody a entropie výsledného obrazu, která vypovídá o jeho kontrastu.

metoda	doba výpočtu [ms]	entropie
gamma komprese	40 %	6,35
Drago	140	6,46
Durand	230	6,43
Fattal	1190	7,01

Tabulka 4.4: Srovnání doby výpočtu a entropie (kontrastu) metod tone mappingu.

Metoda využívající EAW transformaci dosahuje nejlepšího kontrastu ze všech běžně používaných technik tone mappingu a produkuje věrohodně vypadající scény. Doba výpočtu je opět pouze orientační, algoritmus nebyl optimalizován na rychlost výpočtu.

Kapitola 5

Závěr

V této diplomové práci jsem popsal teoretický základ vlnkové transformace a jejích variant, včetně jejich implementace. Dále jsem prozkoumal možnosti využití vlnkové transformace ve zpracování obrazu a popsal její použití pro detekci hran, odstranění šumu, vylepšení kontrastu a zobrazení HDR. Vlnková transformace se pro tyto účely ukázala jako velmi vhodná. Rozdělením obrazu na části s nízkými a vysokými frekvencemi a jejich zpracováním lze úspěšně provést úpravy v obraze, aniž by došlo ke ztrátě důležitých informací, které se v něm vyskytují, jak k tomu často dochází při použití některých jiných běžně používaných metod zpracování obrazu. Ukázala se například výhoda víceúrovňové analýzy při detekci hran, kdy metoda využívající WT dosahovala srovnatelných, a v některých případech i lepších výsledků, než nejčastěji používaný Cannyho hranový detektor. Při redukci šumu v obraze se zase projevila výhoda invariance vůči posunutí stacionární vlnkové transformace, kdy jejich prahováním lze dosáhnout velmi dobrého odstupů signálu od šumu a strukturální podobnosti s originálním nezašuměným obrazem. Vylepšení lokálního kontrastu úpravou koeficientů edge-avoiding transformace se ukázalo jako použitelné, nicméně v některých testovacích obrazech vznikaly při agresivnějším nastavení parametrů „přepálená“ místa a také míra zvýraznění kontrastu byla u ostatních testovaných metod vyšší. Zajímavou a dobře fungující alternativou k bilaterálnímu filtru je rozmazání obrazu se zachováním hran pomocí edge-avoiding transformace. Tato transformace se také ukázala jako velmi vhodná pro zobrazování HDR scén, výsledné obrazy měly ze všech testovaných většinou nejvyšší kontrast.

Příloha A

Obsah CD

- **aplikace** – testovací aplikace – projekt pro MS Visual Studio 2010, implementující všechny vybrané aplikace vlnkové transformace. V podadresáři x86 se nachází spustitelný binární soubor pro MS Windows spolu s dynamickými knihovnami potřebnými pro jeho běh.
- **knihovna** – hlavičkové a C++ soubory implementující vybrané metody zpracování obrazu pomocí vlnkové transformace. Do aplikace se vkládají: `contrast.h` – pro vylepšení kontrastu, `edges.h` – pro detekci hran, `hdr.h` – pro zobrazení HDR a `noise.h` – pro odstranění šumu.
- **plakát** – plakát prezentující práci ve formátech PDF a JPEG.
- **technická zpráva** – technická zpráva ve formátu PDF. V podadresáři tex se nachází zdrojové kódy pro sázeční systém \LaTeX .
- **testovací data** – obrázky použité v této práci včetně HDR obrazů.

Literatura

- [1] Abdul-Jauwad, S. H.: The Story of Wavelets Theory and Engineering Applications. http://faculty.kfupm.edu.sa/ee/samara/WT_lecture_7.ppt.
- [2] Bradley, J.; Brislawn, C.; Hopper, T.: WSQ Gray-Scale Fingerprint Image Compression Specification. *IAFIS-IC-0110(V3.1)*, Říjen 2010.
- [3] Buades, A.; Coll, B.; Morel, J.-M.: Non-Local Means Denoising. *Image Processing On Line*, ročník 1, 2011.
- [4] Burt, P.; Adelson, E.: The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, ročník 31, č. 4, Duben 1983: s. 532–540, ISSN 0090-6778.
- [5] Candes, E.; Demanet, L.; Donoho, D.; aj.: Fast Discrete Curvelet Transforms. 2005.
- [6] Candes, E. J.; Donoho, D. L.: Curvelets: a surprisingly effective nonadaptive representation of objects with edges. In *Curve and Surface Fitting: Saint-Malo*, University Press, 1999.
- [7] Chan, T. F.; Shen, J.; Zhou, H.-M.: Total Variation Wavelet inpainting. *Journal of Mathematical Imaging and Vision*, ročník 25, č. 1, 2006: s. 107–125.
- [8] Chipman, H. A.; Kolaczyk, E. D.; McCulloch, R. E.: Adaptive Bayesian Wavelet Shrinkage. *Journal of the American Statistical Association*, ročník 92, č. 440, 1997: s. 1413–1421, ISSN 01621459.
- [9] Christopoulos, C.; Skodras, A.; Ebrahimi, T.: The JPEG2000 still image coding system: an overview. *IEEE Transactions on Consumer Electronics*, ročník 46, č. 4, Listopad 2000: s. 1103–1127, ISSN 0098-3063.
- [10] Do, M.; Vetterli, M.: The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, ročník 14, č. 12, Prosinec 2005: s. 2091–2106, ISSN 1057-7149.
- [11] Donoho, D. L.; Johnstone, I. M.: Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, ročník 81, 1994: s. 425–455.
- [12] Donoho, D. L.; Johnstone, I. M.: Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 1995: s. 1200–1224.
- [13] Drago, F.; Myszkowski, K.; Annen, T.; aj.: Adaptive Logarithmic Mapping For Displaying High Contrast Scenes. *Computer Graphics Forum*, ročník 22, 2003: s. 419–426.

- [14] Duan, J.; Q., G.: Fast tone mapping for high dynamic range images. In *Proceedings of the 17th International Conference on Pattern Recognition.*, ročník 2, Srpen 2004, ISSN 1051-4651, s. 847–850.
- [15] Durand, F.; Dorsey, J.: Fast Bilateral Filtering for the Display of High-dynamic-range Images. *ACM Trans. Graph.*, ročník 21, č. 3, Červenec 2002: s. 257–266, ISSN 0730-0301.
- [16] Ebrahimi, F.; Chamik, M.; Winkler., S.: JPEG vs. JPEG2000: An objective comparison of image encoding quality. In *Proceedings of SPIE Applications of Digital Image Processing*, 2004.
- [17] Farbman, Z.; Fattal, R.; Lischinski, D.; aj.: Edge-preserving Decompositions for Multi-scale Tone and Detail Manipulation. *ACM Trans. Graph.*, ročník 27, č. 3, Srpen 2008: s. 67:1–67:10, ISSN 0730-0301.
- [18] Fattal, R.: Edge-avoiding Wavelets and Their Applications. *ACM Trans. Graph.*, ročník 28, č. 3, Červenec 2009: s. 22:1–22:10, ISSN 0730-0301.
- [19] Fattal, R.; Agrawala, M.; Rusinkiewicz, S.: Multiscale Shape and Detail Enhancement from Multi-light Image Collections. *ACM Transactions on Graphics*, ročník 26, č. 3, Červenec 2007, ISSN 0730-0301.
- [20] Hanov, S.: Wavelets and Edge Detection. *CS698 Final Project*, Duben 2006.
- [21] Ismail, B.; Khan, A.: Image De-noising with a New Threshold Value Using Wavelets. *Journal of Data Science*, ročník 10, č. 2, Duben 2012: s. 271–280, ISSN 1680-743X.
- [22] ISO: Information technology – JPEG 2000 image coding system – Part 1: Core coding system. ISO 15444-1:2000, International Organization for Standardization, Geneva, Switzerland, 2000.
- [23] Kuan, D. T.; Sawchuk, A. A.; Strand, T. C.; aj.: Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, ročník PAMI-7, č. 2, Březen 1985: s. 165–177, ISSN 0162-8828.
- [24] Li, J.: A wavelet approach to edge detection. *Sam Houston State University*, Srpen 2003.
- [25] Lischinski, D.; Farbman, Z.; Uyttendaele, M.; aj.: Interactive Local Adjustment of Tonal Values. *ACM Trans. Graph.*, ročník 25, č. 3, Červenec 2006: s. 646–653, ISSN 0730-0301.
- [26] Mallat, S.: *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, třetí vydání, 2008, ISBN 0123743702, 9780123743701.
- [27] Mallat, S.; Zhong, S.: Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, ročník 14, č. 7, Jul 1992: s. 710–732, ISSN 0162-8828.
- [28] Mallat, S. G.: A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, ročník 11, č. 7, Červenec 1989: s. 674–693, ISSN 0162-8828.

- [29] Niya, J.; Aghagolzadeh, A.; Tinati, M.; aj.: 2-step wavelet-based edge detection using gabor and cauchy directional wavelets. In *The 7th International Conference on Advanced Communication Technology, 2005, ICACT 2005.*, ročník 1, 2005, s. 115–120.
- [30] Reeves, T. H.; Jernigan, M. E.: Multiscale-based image enhancement. In *IEEE 1997 Canadian Conference on Electrical and Computer Engineering, 1997. Engineering Innovation: Voyage of Discovery.*, ročník 2, Květen 1997, ISSN 0840-7789, s. 500–503.
- [31] Starck, J.-L.; Candes, E. J.; Donoho, D. L.: The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, ročník 11, č. 6, Červen 2002: s. 670–684.
- [32] Sweldens, W.: The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions. In *Wavelet Applications in Signal and Image Processing III*, 1995, s. 68–79.
- [33] Sweldens, W.: Wavelets: what next? *Proceedings of the IEEE*, ročník 84, č. 4, Duben 1996: s. 680–685.
- [34] Taubman, D.: High performance scalable image compression with EBCOT. *IEEE Transactions on Image Processing*, ročník 9, č. 7, Červenec 2000: s. 1158–1170, ISSN 1057-7149.
- [35] Tomasi, C.; Manduchi, R.: Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision*, Leden 1998, s. 839–846.
- [36] Toufik, B.; Mokhtar, N.: The Wavelet Transform for Image Processing Applications, Physics and Technology. *Advances in Wavelet Theory and Their Applications in Engineering*, 2012.
- [37] Unaldi, N.; Temel, S.; Demirci, S.: Undecimated Wavelet Transform Based Contrast Enhancement. *International Journal of Computer, Information, Systems and Control Engineering*, ročník 7, č. 9, 2013: s. 571–574.
- [38] Uytterhoeven, G.; Roose, D.; Bultheel, A.: Wavelet Transforms Using the Lifting Scheme. 1997.
- [39] Uytterhoeven, G.; Roose, D.; Bultheel, A.: Integer wavelet transforms using the lifting scheme. In *Physics and Modern Topics in Mechanical and Electrical Engineering, 3rd World Multiconference on Circuits, Systems, Communications and Computers*, Červenec 1999, s. 198–200.