

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

**FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ**

**VYUŽITÍ WAVELETOVÉ TRANSFORMACE
A MATEMATICKÉ STATISTIKY PRO SEPARACI
SIGNÁLU A ŠUMU**

DIZERTAČNÍ PRÁCE

k získání akademického titulu Doktor (Ph.D.)

ve studijním oboru

TELEINFORMATIKA

Mgr. Pavel Rajmic

Školitel: Prof. Ing. Zdeněk Smékal, CSc.

Datum státní doktorské zkoušky: 30. 5. 2003

Datum odevzdání práce: 31. 8. 2004

Práce je k dispozici na Vědeckém oddělení děkanátu FEKT VUT v Brně,
Údolní 53, Brno, 602 00.

Upozornění

Tato verze dizertační práce byla vytvořena revizí a byla publikována dne 23. 10. 2007. Oproti původní dizertační práci jsou v tomto textu zejména zpracovány některé připomínky oponentů, zvláště Doc. RNDr. V. Veselého, CSc. Dále byla zvláštní pozornost věnována revizi kapitoly 8, nicméně aktuální informace právě ohledně segmentované waveletové transformace důrazně doporučuji čerpat z novějších pramenů.

Dizertační práci (původní) citujte v českých textech takto:

RAJMIC, P.: *Využití waveletové transformace a matematické statistiky pro separaci signálu a šumu*, Doktorská dizertační práce, Vysoké učení technické v Brně, Brno, 2004.

Děkuji Prof. Ing. Zdeňku Smékalovi, Csc. za vedení dizertační práce, Doc. RNDr. Vítězslavu Veselému, CSc., Prof. RNDr. Františku Melkesovi, CSc. a Prof. Brani Vidakovicovi, PhD. za odborné konzultace. Děkuji svým rodičům, své ženě Kateřině a dalším za vytrvalou podporu.

Prohlašuji, že tuto práci jsem vypracoval samostatně a že jsem uvedl všechnu použitou literaturu.

Pavel Rajmic

Abstrakt

Tato dizertační práce se zabývá metodou potlačování gaussovského (normálně rozloženého) šumu v signálu pomocí tzv. prahování waveletových koeficientů, s důrazem na řečové signály. Tento způsob separace užitečného a ruchového signálu vychází z metodologie *WaveShrink*. První část práce je věnována některým modifikacím této metodologie za účelem zvýšení její účinnosti a dále zejména statistické analýze a srovnání vlastností pěti prahovacích pravidel.

Druhá část práce je věnována nově vyvinuté metodě tzv. segmentované waveletové transformace, která, kromě dalších výhod, umožňuje implementovat waveletovou transformaci pro zpracování v reálném čase.

Abstract

The dissertation thesis considers a method for suppressing Gaussian (normally distributed) noise in signals utilizing the so-called thresholding of wavelet coefficients, with particular focus on speech signals. This type of separation of effective and noisy signals starts from the *WaveShrink* methodology. The first part of the thesis considers several modifications of this methodology in order to improve its performance, and, also, in particular the statistical analysis and comparison of five thresholding rules.

The second part is devoted to the newly-developed method of so-called segmented wavelet transform, which allows us to perform the wavelet transform, and thus also the “wavelet-type separation” in real time, in addition to other benefits.

Seznam použitých zkratek, symbolů a značení

\mathbb{R}	množina všech reálných čísel
\mathbb{R}^+	množina kladných reálných čísel
\mathbb{Z}	množina celých čísel
\mathbb{N}	množina přirozených čísel, tj. $\mathbb{N} = \{1, 2, 3, \dots\}$
$L^1(\mathbb{R})$	množina všech komplexních funkcí lebesgueovský integrovatelných s konečnou energií, $L^1(\mathbb{R}) = \left\{ f \mid \int_{-\infty}^{\infty} f(t) dt < \infty \right\}$
$L^2(\mathbb{R})$	množina všech komplexních funkcí lebesgueovský integrovatelných v kvadrátu s konečnou energií, $L^2(\mathbb{R}) = \left\{ f \mid \sqrt{\int_{-\infty}^{\infty} f(t) ^2 dt} < \infty \right\}$
$[a, b]$	uzavřený interval od a do b
(a, b)	otevřený interval od a do b
$A \cap B$	průnik množin A a B
$\mathbf{0}$	nulový vektor
\mathbf{I}_n	jednotková matice řádu n
\mathbf{A}^\top	matice transponovaná k matici \mathbf{A}
\mathbf{A}^{-1}	matice inverzní k (čtvercové regulární) matici \mathbf{A}
d	hloubka waveletové dekompozice
m	délka waveletových filtrů typu FIR
\mathbf{h}	waveletový filtr (tj. vektor) typu dolní propust
\mathbf{g}	waveletový filtr (tj. vektor) typu horní propust
$H(\cdot)$	přenosová funkce lineárního systému definovaného impulzní odezvou \mathbf{h}
$G(\cdot)$	přenosová funkce lineárního systému definovaného impulzní odezvou \mathbf{g}
$H(e^{i\omega})$	kmitočtová odezva filtru \mathbf{h} v závislosti na úhlové rychlosti ω
$G(e^{i\omega})$	kmitočtová odezva filtru \mathbf{g} v závislosti na úhlové rychlosti ω
$\downarrow 2$	označení operátoru dyadické decimace
$\uparrow 2$	označení operátoru dyadické interpolace
$\mathbf{a}^{(k)}$	vektor approximačních waveletových koeficientů v úrovni k
$\mathbf{d}^{(k)}$	vektor approximačních waveletových koeficientů v úrovni k
$A^{(k)}(\omega)$	amplitudová kmitočtová charakteristika příslušná uzlu s číslem k ve waveletovém paketu

ω	úhlový kmitočet (úhlová rychlosť)
s	délka segmentů v segmentované waveletové transformaci, SegWT
S	celkový počet segmentů signálu v SegWT
$P_{\min}(n)$	tzv. minimální délka pravého prodloužení n -tého segmentu v SegWT
$L_{\max}(n)$	tzv. maximální délka levého prodloužení n -tého segmentu v SegWT
χ_M	charakteristická funkce množiny $M \subseteq \mathbb{R}$, tj. funkce, pro kterou platí $\chi_M(x) = 1$ pro $x \in M$ a $\chi_M(x) = 0$ pro ostatní x
$\ \cdot\ $	norma prvku množiny; např. pro konečněrozměrné vektory definovaná $\ \mathbf{x}\ = \sqrt{ x_1 ^2 + \dots + x_n ^2}$, kde $\mathbf{x} = (x_1, \dots, x_n)^\top$
\bar{x}	komplexně sdružené číslo k číslu x
$\text{sgn}(\cdot)$	tzv. znaménková funkce: $\text{sgn}(x) = 1$ pro $x > 0$, $\text{sgn}(x) = -1$ pro $x < 0$ a $\text{sgn}(x) = 0$ pro $x = 0$
ϕ	měřítková funkce
ψ	waveletová funkce
\equiv	symbol pro identitu náhodných veličin
EX	střední hodnota náhodné veličiny X
$\text{var}X$	rozptyl náhodné veličiny X , $\text{var}X = E[(X - EX)^2]$
$X \sim N(\mu, \sigma^2)$	náhodná veličina X s normálním rozložením o střední hodnotě $EX = \mu$ a rozptylu $\text{var}X = \sigma^2$
$\mathbf{X} \sim N_p(\mathbf{m}, \boldsymbol{\Sigma})$	náhodný vektor \mathbf{X} s mnohorozměrným normálním rozložením o střední hodnotě \mathbf{m} a kovarianční matici $\boldsymbol{\Sigma}$
σ	standardizovaná odchylka dané náhodné veličiny, σ^2 její rozptyl
φ	hustota pravděpodobnosti standardizovaného normálního rozložení $N(0, 1)$, tj. funkce $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
Φ	distribuční funkce standardizovaného normálního rozložení, tj. $\Phi(x) = \int_{-\infty}^x \varphi(x) dx$
λ	prahová hodnota
$\delta(\cdot, \lambda), \delta(\cdot)$	označení prahovací funkce s prahovou hodnotou λ
$M_\lambda(\cdot)$	střední hodnota prahovacího pravidla při prahové hodnotě λ
$V_\lambda(\cdot)$	rozptyl prahovacího pravidla při prahové hodnotě λ
$R_\lambda(\cdot)$	riziko prahovacího pravidla při prahové hodnotě λ

\oplus	operátor ortogonálního součtu vektorových prostorů
\cap	operátor průniku množin
$\text{floor}(\cdot)$	zaokrouhlení na nejbližší celé číslo směrem k nule
$\text{ceil}(\cdot)$	zaokrouhlení na nejbližší celé číslo směrem k plus (pro kladná čísla) nebo minus (pro záporná čísla) nekonečnu
\square	symbol označující konec důkazu
FFT	rychlá Fourierova transformace (Fast Fourier Transform)
CWT	integrální waveletová transformace (Continuous Wavelet Transform)
DWT	diskrétní waveletová transformace (Discrete Wavelet Transform)
DTWT	konečná diskrétní waveletová transformace (Discrete Time Wavelet Transform)
FWT	„rychlá“ diskrétní waveletová transformace (Fast Wavelet Transform) – alternativní označení pro Mallatův pyramidový algoritmus
SWT	stacionární waveletová transformace (Stationary Wavelet Transform), [48]
SegWT	segmentovaná waveletová transformace (Segmented Wavelet Transform), zkratka pro novou metodu uvedenou v této práci
MRA	mnohoměřítková analýza (Multiresolution Analysis)
MR-analýza	viz MRA
SNR	Signal-to-Noise-Ratio, odstup signálu od šumu [dB]
FIR	Finite Impulse Response, konečná impulzní odezva (filtru)
IIR	Infinite Impulse Response, nekonečná impulzní odezva (filtru)

Matice značíme velkými tučnými písmeny, např. **A**. Vektory značíme malými tučnými písmeny, např. **a**. Prvky matic a vektorů značíme příslušnými malými písmeny italikou (kurzívou), $\mathbf{A} = (a_{ij})$, $\mathbf{a} = (a_1, \dots, a_n)^\top$. Pokud nebude uvedeno jinak, vektory uvažujeme jako sloupcové. Náhodné vektory značíme velkým tučným skloněným písmem, náhodné veličiny velkým skloněným netučným, např. $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$. Realizaci náhodného vektoru značíme příslušnými malými písmeny, tj. např. $\mathbf{y} = (y_1, \dots, y_n)^\top$.

V obrázcích vytvořených v systému Matlab je místo řádové čárky řádová tečka a také popisky v grafech jsou v anglickém jazyce.

Obsah

<i>Seznam použitých zkratek, symbolů a značení</i>	v
<i>Seznam obrázků</i>	x
<i>Seznam tabulek</i>	xii
Úvod	1
1 Přehled současné problematiky	3
1.1 Metody separace signálů od šumu a hluku	3
1.2 Waveletové metody zpracování signálu	7
1.2.1 Využití waveletových metod při zpracování řeči	8
1.2.2 Waveletové metody pro separaci signálu od šumu	8
1.2.3 Waveletová transformace realizovatelná v reálném čase	10
2 Cíle dizertační práce	11
3 Výchozí poznatky z matematiky a teorie zpracování signálů	12
3.1 Matematika	12
3.2 Teorie zpracování signálů	13
4 Waveletová transformace	15
4.1 Integrální waveletová transformace	15
4.2 MR – analýza	18
4.2.1 Wavelety s nulovými momenty v MR-analýze	21
4.3 Konečná diskrétní waveletová transformace	21
4.3.1 Souvislost DTWT s lineární filtrací signálu a Mallatův pyramidový algoritmus	22
4.3.2 Kmitočtová charakteristika waveletových filtrů	25
4.3.3 Waveletové filtry a aliasing	25
5 Waveletové pakety (wavelet packets) a výběr nejlepší báze	30
5.1 Waveletové pakety jako integrální transformace	30
5.2 Waveletové pakety jako konečná diskrétní transformace	32

5.3	Hledání nejlepší báze pro diskrétní signály	33
5.4	Kmitočtové vlastnosti waveletových paketů	35
6	Waveletové potlačování šumu	39
6.1	Princip waveletových metod potlačování šumu	39
6.2	Typy prahovacích pravidel	41
6.3	Statistická analýza prahovacích pravidel	44
6.4	Volba prahové hodnoty	53
6.4.1	Kvantilová volba prahu	53
6.4.2	Univerzální hodnota prahu	53
6.4.3	Kritérium zobecněné křížové validace	54
6.5	Závěr	55
7	Testování waveletových metod na signálech s aditivním šumem	56
7.1	Testovací a výsledné nahrávky	59
8	Segmentovaná konečná diskrétní waveletová transformace	63
8.1	Motivace a cíl modifikované metody	64
8.2	Rozbor pyramidového algoritmu DTWT	65
8.3	Nový algoritmus segmentované waveletové transformace	72
8.3.1	Algoritmus SegWT	80
8.3.2	Důsledky a omezení algoritmu SegWT	84
8.3.3	Algoritmus SegWT při prodloužení typu „zero padding“	87
8.4	Závěr	87
Závěr		90
Přílohy		93
A	Důkazy tvrzení	93
B	Odvození $M_\lambda^{\text{hy}}(b)$, $V_\lambda^{\text{hy}}(b)$, $R_\lambda^{\text{hy}}(b)$ pro hyperbolické prahování	95
C	Programový balík ThreshLab	99
D	Obsah přiloženého CD	101
Literatura		102

Seznam obrázků

4.1	Mateřský wavelet typu Haar	17
4.2	Příklad dilatace a translace mateřského waveletu	18
4.3	Ukázka MR-analýzy úseku hudebního signálu provedené s waveletem Daubechies řádu 2	20
4.4	Jeden krok waveletové dekompozice	23
4.5	Mallatův pyramidový algoritmus waveletové dekompozice o hloubce $d = 3$	24
4.6	Jeden krok waveletové rekonstrukce	24
4.7	Idealizované amplitudové kmitočtové charakteristiky filtrů \mathbf{h} a \mathbf{g} při jejich rekurzivní aplikaci v dekompozičním pyramidovém algoritmu	26
4.8	Skutečná amplitudová charakteristika dvou rozdílných waveletů typu Daubechies, řádu 2 (db2) a řádu 12 (db12)	27
4.9	MR-analýza stejného signálu jako v obrázku 4.3, tentokrát s waveletem Daubechies řádu 12	28
5.1	Průběh funkcí $\mathcal{W}_0, \dots, \mathcal{W}_5$ pro wavelet Daubechies 2	31
5.2	Rozklad signálu y do úplného binárního stromu (waveletového paketu) hloubky $d = 3$	34
5.3	Srovnání účinnosti vyjádření signálu pomocí waveletové transformace a nejlepší báze z waveletového paketu	36
5.4	Grafy amplitudových kmitočtových odezv vybraných uzlů waveletového paketu při dekompozici	38
6.1	Princip prahování	40
6.2	Grafy prahovacích pravidel	43
6.3	Srovnání kvadrátu výchylky, rozptylu a rizika pro tvrdé a měkké prahování	47
6.4	Prostorové znázornění rizika tvrdého a měkkého prahování pro různé hodnoty b a λ	48
6.5	Srovnání kvadrátu výchylky, rozptylu a rizika pro nezápornou garotu a hyperbolické prahování	49
6.6	Srovnání rizika prahovacích pravidel	50

6.7	Promítnutí neoptimální volby dvojice prahů λ_1, λ_2 do velikosti rizika poloměkkého prahování	52
6.8	Příklad typického průběhu kritéria GCV	54
7.1	Příklad prahování zarušeného signálu	58
8.1	„Waveletogram“	64
8.2	Znázornění kroků 2(a) až 2(d) v algoritmu 8.1 pro délku filtru $m = 4$. . .	67
8.3	Ilustrace algoritmu 8.1 ve třech úrovních dekompozice	68
8.4	Počet vzorků signálu potřebných pro výpočet jednoho waveletového koeficientu	70
8.5	Schéma segmentace signálu	74
8.6	Grafické znázornění prodlužování segmentů	75
8.7	Vývojové schéma algoritmu 8.16	82
8.8	Schéma algoritmu SegWT při prodloužení nulovými hodnotami („zero padding“)	88

Seznam tabulek

7.1	Výsledky testování na zarušeném signálu se SNR = 15 dB	60
7.2	Výsledky testování na zarušeném signálu se SNR = 10 dB	60
7.3	Výsledky testování na zarušeném signálu se SNR = 5 dB	61
7.4	Výsledky testování na zarušeném signálu se SNR = 0 dB	61
8.1	Počet vzorků vstupního signálu nutných pro výpočet jednoho waveletového koeficientu	69
8.2	Počty koeficientů vypočítaných algoritmem 8.1 ze vstupního signálu o délce q v různých hloubkách dekompozice	71
8.3	Velikosti prodloužení segmentů pro různé délky segmentů s při hloubce dekompozice $d = 3$ a délce filtru $m = 16$	79

Úvod

Waveletová neboli vlnková analýza je moderní disciplína v oblasti zpracování signálů, která má po krátkém, ale velmi prudkém a intenzivním vývoji velkou oblast uplatnění nejen v čistě technických oborech. Mezi nejvýznamnější aplikace waveletového zpracování signálů patří separace signálů od aditivního šumu z jejich směsí.

Tato dizertační práce se zabývá metodou potlačování gaussovského (normálně rozloženého) šumu v signálu tzv. prahováním waveletových koeficientů. Tento způsob separace užitečného a ruchového signálu vychází z metodologie *WaveShrink*, jejímiž autory jsou Donoho a Johnstone [17]. Druhou, neméně podstatnou částí této práce je kapitola pojednávající o nově vyvinuté metodě tzv. segmentované waveletové transformace, která umožňuje provádět waveletovou transformaci, a tedy v důsledku i tuto „waveletovou separaci“ kromě dalších výhod *v reálném čase*.

Text otevírá kapitola 1, ve které je uveden současný stav řešeného problému, tj. přehled používaných metod pro potlačování šumů a hluků v signálech, zejména řečových, a zhodnocení jejich vlastností. Je zde uvedena řada vzájemně odlišných přístupů, ať už jde o metody založené na spektrální reprezentaci, na adaptivní filtrace nebo jiných alternativních principech.

Na zhodnocení současného stavu problematiky v oblasti „waveletového“ zpracování dat přirozeně navazuje kapitola 2, obsahující vymezení konkrétních cílů této dizertační práce.

Po úvodních prerekvizitách týkajících se v textu používaných pojmu a tvrzení z oblasti matematiky a zpracování signálů (kapitola 3) je v kapitole 4 vyložena waveletová transformace, a to jak ve spojitém čase, tak zejména pro diskrétní signály. Je zde uveden i rychlý algoritmus pro výpočet waveletové transformace založený na konvoluci a decimaci a jsou rozebrány kmitočtové vlastnosti waveletových filtrů.

Kapitola 5 poté navazuje konceptem tzv. waveletových paketů (wavelet packets), které tvoří v jistém smyslu zobecnění waveletové transformace. Waveletové pakety lze s úspěchem využít pro účely potlačování šumu, jak bude ukázáno, s vyšší účinností nežli je tomu u standardní waveletové transformace.

Zásadní část textu je kapitola 6, která detailně rozebírá princip tzv. prahování waveletového spektra za účelem potlačení šumové složky signálu. Je uvedeno pět typů prahování (obvykle jsou v pracích podobného zaměření uváděny pouze dva základní typy) a zejména jsou odvozeny a porovnány jejich vlastnosti z pohledu matematické statistiky.

V kapitole 7 je provedeno testování uvedených metod na reálných nahrávkách a jejich objektivní i subjektivní zhodnocení.

Kapitola 8 tvoří druhou zásadní část práce. Je zde odvozen sofistikovaný algoritmus, který umožňuje waveletové zpracování dat v reálném čase (tedy pokud signál neznáme celý předem, ale přichází postupně jeho segmenty), případně umožňuje výpočetně náročnou úlohu rozložit na několik méně náročných úloh. Typ zpracování v reálném čase je často potřebný v komunikačních technologiích, zvláště u řečových signálů.

Na závěr je proveden souhrn a zhodnocení přínosů této dizertační práce. Rovněž jsou navrženy směry, kterými by se mohl ubírat další výzkum v této oblasti.

Důkazy nejdůležitějších tvrzení je možné nalézt v přílohách, s výjimkou kapitoly 8, ve které jsou naopak všechny důkazy uvedeny přímo v textu. Obsahem příloh je rovněž popis funkcí obsažených v programovém balíku ThreshLab pro MATLAB, který je rovněž součástí dizertační práce a který implementuje uvedené separační techniky a metody jejich statistické analýzy.

1 Přehled současné problematiky

Obsahem této kapitoly je uvedení a zhodnocení současného stavu problematiky, tj. dnes v praxi více či méně používaných metod pro separaci signálů (zejména řečových) od šumu, případně hluku v pozadí. V první části kapitoly budou uvedeny metody, které využívají principů Fourierovy transformace, znalostí charakteristických vlastností řečových signálů, sofistikovaných statistických přístupů apod. Pozornost druhé části kapitoly bude obrácena na waveletové metody zpracování signálů se zaměřením na používané metody separace deterministických signálů od šumu.

1.1 Metody separace signálů od šumu a hluku

V této části práce je pozornost zaměřena na výčet technik, které je v dnešní době možno použít pro separaci řečového signálu a aditivního šumu nebo hluku. Šumem zpravidla rozumíme náhodný proces, který má přesně definované statistické parametry (jako např. rozložení, střední hodnota, rozptyl, autokorelační funkce), přičemž tyto parametry nemusí být známé. Hlukem naopak rozumíme signál, který statistickým zákonitostem v tomto úzkém slova smyslu nepodléhá. Příkladem může být hluk motoru, vysavače, dopravního provozu atd. U hlukových signálů máme často možnost odhadovat jejich charakteristické parametry jako energie, spektrální výkonová hustota atd.

Jednoduší separační techniky jsou v současné době nedílnou součástí každého komunikačního zařízení pracujícího v zarušeném prostředí. Většinou se jedná o jednoduché algoritmy s malou účinností vlastní separace a s mnoha nedostatky. Ukazuje se, že nejobtížnější pro separaci jsou parazitní hluky, které kmitočtově spadají do pásma užitečné řeči. Takových zvuků je v praktickém provozu bohužel většina. Vhodný algoritmus by měl být proto schopen patřičně reagovat jak na širokopásmové šumy, impulsní hluky a trvalé rušící periodické signály, tak dokonce na nežádoucí lidskou řeč v pozadí.

Metody pro potlačení parazitního rušení v řečovém signálu lze rozdělit na *jednokanálové* a *vícekanálové*. V současné době jsou většinou v praxi používány jednoduché jednokanálové techniky založené na potlačení vyšších, případně nižších kmitočtů složek

spektra vstupního signálu v závislosti na velikosti odstupu signálu od rušení (SNR). Je zřejmé, že klasickou číslicovou filtrací rušivých signálů sice docílíme potlačení hluku, ale v užitečném pásmu kmitočtů dále přetrvává, pouze je vlivem omezené šířky přenášeného pásma pro lidský sluch méně nápadný.

V případě užití adaptivní lineární filtrace (například algoritmu typu LMS) vzniká nežádoucí přídavný efekt, který způsobí, že výstupní signál ztrácí svou dynamiku, hlas ztrácí barvu a zní strojově.

Podstatně dokonalejší jednokanálovou metodou, užívanou firmou Motorola v mobilních telefonech pro jednotky „handsfree“, je technika RASTA¹, která využívá pásmovou filtraci časově proměnného spektra signálu, tzv. časových trajektorií [25]. Na základě kmitočtové analýzy časového průběhu jednotlivých spektrálních složek modulových spekter (časových trajektorií), je vypočteno spektrum časových trajektorií, které je zvykem nazývat modulačním spektrem. Aplikací vhodného číslicového filtru dochází k omezení modulačního spektra parazitního signálu, přitom modulační spektrum užitečné řeči musí zůstat nezměněno. Potlačení trvalých nebo naopak velmi rychle se v čase měnících (rychleji než řeč) parazitních signálů je dostatečné, ovšem pro většinu běžných hluků není separace příliš účinná. Jinými slovy, separace je velmi neefektivní v případě, kdy spektrum hluku má charakter spektra řečového signálu, neboli rychlosť proměn spektrálních čar rušení je srovnatelná s rychlosťí změn ve spektru lidské promluvy. Navíc přídavnou modulací spekter modulu vstupního signálu se výstupní řeč stává spíše nepříjemnou.

K robustnějším technikám potlačování rušení z řečových signálů patří nová jednokanálová metoda nazvaná mapování spektrogramu, vyvinutá na FEKT VUT v Brně ve spolupráci s ApS Brno [35]. Algoritmus využívá ke své práci krátkodobých spekter zarušeného řečového signálu stejně jako metoda RASTA. Vlastní zpracování v časově-kmitočtovém prostoru je ale v této metodě zcela odlišné a také účinnější. Principem metody je adaptivní vytváření tzv. mapy (masky), kterou se následně násobí spektrogram řečového signálu, tím se signál „filtruje“, a poté se převede zpět do časové oblasti. Výsledky testování ukázaly, že metoda vykazuje excelentní separační vlastnosti v případě, kdy je možno optimálně stanovit velikost tzv. prahu pro dlouhé úseky signálu (až několik sekund). To samozřejmě v praxi není možné, neboť se požaduje nasazení metody v reálném čase. Tehdy je nutné

¹RelAtive SpecTrAl

práh nastavovat velmi rychle a to často vede k odchylkám od jeho optimální hodnoty, což se projevuje nežádoucími modulacemi řeči.

Pro rušení, které má charakter blízký stacionárnímu, je velmi vhodná metoda spektrálního odečítání [13, 15] a její modifikace. V řečových pauzách² je analyzován kmitočtový obsah rušení a v čase řečové aktivity je toto rušení ve spektrální oblasti odečítáno. Ukázalo se, že se snižujícím se poměrem SNR drasticky klesá účinnost metody [62]. To vedlo k modifikacím metody a ke vzniku metody nelineárního spektrálního odečítání [34] a metody, využívající maskovacího efektu lidského slyšení [52]. V těchto novějších verzích jsou již tzv. odečítací parametry určovány dynamicky. Spektrální odečítání má vynikající separační vlastnosti, pokud hluk v pozadí má v čase neměnný charakter (větrák, vysavač naprázdno, sprcha). Z jejího principu zároveň vyplývá, že pro impulzní rušení je naprosto nevhodná (údery kladiva apod.).

Odlišným přístupem k problematice separace je vývoj vícekanálových metod schopných pracovat v reálném čase. Nejstarší dvoukanálová metoda je již dlouhou dobu používána např. v televizních přenosech mezi moderátorem a přenosovým vozem. Hlas publicisty je v hlučném prostředí snímán směrovým mikrofonem. Hluk okolí se snímá ve vhodné vzdálenosti druhým mikrofonem s všeobecnou charakteristikou. Vzniklé signály se pak odečítají. Tato metoda však není příhodná pro univerzální použití, avšak ukazuje obecný princip vícekanálových metod.

Metody vícekanálového zpracování signálů mají daleko větší nároky na výpočetní výkon než metody jednokanálové. Současně se zvyšujícím se výkonem dnešních signálových procesorů se vícekanálovým metodám dostává stále více pozornosti. Můžeme je rozdělit do dvou skupin podle jejich principu, na

- metody pracující s exaktně determinovaným mikrofonním polem,
- „slepé“³ metody pracující přímo s několika směsmi, bez jakékoliv přídavné informace o poloze mikrofonů.

V prvním případě problém řešíme deterministicky, na základě souřadnic mikrofonů určíme dekompoziční parametry. Řešení druhého případu nabízí matematická statistika –

²pauzy řečové aktivity lze určovat několika typy detektorů řeč / pauza [38]

³z anglického Blind Source Separation, BSS

analýza nezávislých komponent [26]. Výhodou vícekanálového přístupu je značné zlepšení výsledného SNR (podle [62] v průměru o 6–7 dB) oproti jednokanálovým metodám, přičemž subjektivní hodnocení kvality je ještě výrazněji lepší. Další výhodou je, že kvalita separace není podmíněna stacionárním charakterem rušení. Naopak nevýhodou je, že nevhodná konfigurace mikrofonního pole vede téměř k nemožnosti separace signálů. Z toho plyne, že vícekanálové metody jsou vhodné k aplikaci na místech, kde nedochází ke změně uspořádání senzorů ani zdroje hluku (automobil, stroj v mechanické dílně).

Nejnovější přístupy k problematice separace vycházejí z originální myšlenky tzv. re-syntézy řeči. Algoritmy tohoto typu vycházejí z velmi dobré znalosti principu vzniku řeči v hlasovém traktu člověka. Nesnaží se v parazitním hluku nalézt řečový signál jako celek, ale detekují pouze základní parametry řečového signálu [63]. Těmito parametry jsou kmitočet základního tónu řeči F_0 , poloha a tvar formantů, energie řečových segmentů, informace o znělosti / neznělosti apod. Z uvedeného výčtu parametrů pak, pomocí standardních technik syntézy řeči, rekonstruujeme užitečnou promluvu. Teoreticky jsou tyto algoritmy schopné extrahat užitečný řečový signál z rušení, majícího v čase i kmitočtu charakter řečového signálu. Jde obecně o periodicky se opakující přechodné hluky s proměnným spektrem. Může jít například o zvuky zvířat, zvuky hudebních nástrojů nebo i nežádoucí řeč v pozadí. Taková specifická rušení, v lidském prostředí velice častá, jsou na potlačení nejnáročnější. Praktická realizace uvedeného principu je však zatím v počátcích. Nejprve bude nutné vyvinout robustní techniky schopné ze silně zarušené směsi řeči a hluku určit vyjmenované parametry, charakteristické pro řeč.

Celkově se dá shrnout, že žádná ze separačních metod není zcela univerzální, každá má své kladné i záporné rysy. Nicméně u všech metod lze vysledovat analogické vlastnosti – každá metoda při potlačování hluku zkresluje také samotnou řeč. Tento jev se samozřejmě stává čím dál výraznějším se snižujícím se odstupem signálu od šumu (hluku). Nastavení parametrů separace (a tím její výsledek) je proto *vždy nutné volit jako kompromis mezi mírou potlačení hluku a zkreslením samotné řeči*. Stejně tvrzení platí přirozeně i o waveletových metodách, které jsou uvedeny v části 1.2.2.

Na závěr je ještě je vhodné uvést poznámku, že pro posuzování kvality separace není možné využívat výhradně objektivní metody, zejména zlepšení poměru SNR. Rovněž subjektivní kvalita separace hraje důležitou roli. Dokonce nahrávka s objektivně horší

úrovní separace může člověku znít lépe a srozumitelněji, a naopak. Tento jev je způsoben faktem, že objektivní metody nezahrnují do svého hodnocení zkreslení samotné řeči.

1.2 Waveletové metody zpracování signálu

Zpracování signálů pomocí waveletové transformace je v posledních letech značně využívaná disciplína se stále se rozvíjející škálou aplikací při zpracování jak jednorozměrných, tak vícerozměrných signálů. Jde o relativně nový přístup k analýze signálů, jehož základní principy byly položeny již před několika desítkami let, avšak jehož bouřlivý vývoj odstartoval až počátkem osmdesátých let 20. století.

Své uplatnění při analýze a syntéze signálů našla waveletová transformace v oborech jako jsou telekomunikace, sonarová a radarová technika, seismologie, meteorologie, lékařství ad. V těchto oborech se používá buď pro časově-kmitočtovou analýzu signálů, pro rekonstrukci a restauraci neúplných či silně zarušených signálů, nebo pro kompresi dat. Na tomto posledně jmenovaném poli dosahuje waveletová transformace vynikajících výsledků, zejména při kompresi digitálních obrazů, což dokazuje např. nově zaváděný standard waveletové obrazové komprese JPEG 2000 [43].

Zmíněné obory tvoří těžiště aplikací waveletové transformace, nýbrž její okrajové (avšak neméně důležité) využití můžeme nalézt také v teoretické matematice (aproximační teorie, numerické řešení diferenciálních rovnic, [4, 29]) nebo dokonce v oblasti zpracování hudby [5, 11]. Pro přehled méně obvyklých aplikací viz [40].

Hlavní zásluhou waveletové analýzy je, že přinesla nový typ reprezentace signálů a rozšířila tak soubor nástrojů pro jejich zpracování. Ve srovnání s obvyklým fourierovským přístupem pozorujeme lepší, ale i horší vlastnosti. Použití waveletových technik tedy *není univerzální*, ale pro každý typ úlohy je třeba dobrě zvážit jeho vhodnost.

Matematická teorie waveletové transformace se vyvíjela dlouhá léta vedle prvních praktických inženýrských aplikačních experimentů, aniž by se tyto oblasti střetly. Počátkem osmdesátých let dvacátého století upozornil na vzájemnou úzkou provázanost těchto dvou oblastí francouzský vědec Stéphane Mallat a odstartoval tak „waveletovou revoluci“. Spoluprací mezi zmíněnými dvěma oblastmi se poté velmi intenzivním způsobem začala dotvářet jak sofistikovaná teorie, vycházející z funkcionálních prostorů, tak také stále nové

a nové aplikace. Tento trend stále ještě pokračuje, přestože v posledních letech ne tak bouřlivě.

Budoucí vývoj v této oblasti směřuje k systémům, které jsou po formální stránce zobecněním waveletů a jež jsou více přizpůsobeny konkrétním úlohám na úkor univerzálnosti jejich použití. To jsou např. dnes již velmi rozšířené biortogonální wavelety, frejmy (frames) a Rieszovy báze [9, 46], neortogonální a neúplné systémy [8, 47].

1.2.1 Využití waveletových metod při zpracování řeči

Waveletová transformace se uplatnila také v některých oblastech zpracování řečových signálů. Protože řeč má přibližně stacionární povahu pouze pro její velmi krátké úseky, fourierovský typ zpracování, který vychází z nekonečně dlouhých harmonických bázových funkcí, není pro určité aplikace vhodný. Waveletová transformace, umožňující ze své podstaty reprezentativnější pohled na signál z krátkodobého hlediska, se pokusila standardní metody zpracování řeči vylepšit.

Publikace [36, 61] ukazují, že při *detekci základního tónu řeči* F_0 má waveletová transformace několik výhod včetně odstranění problému dvojnásobení základního kmitočtu.

V oblasti *segmentace řeči na fonémy* [41] waveletová transformace nepředčila tradiční metody, zejména z toho důvodu, že řečový signál je velmi proměnlivý a není možné všechny jeho aspekty zachytit pomocí pevné waveletové báze.

Přínosem v oblasti *potlačování různých typů šumu* v pozadí řeči se stala metoda waveletového vyhlazování spektrogramu šumu [32, 42]. Použití takto vyhlazeného spektrogramu podle výzkumné zprávy [61] přineslo zlepšení SNR v průměru o 3–5 dB.

1.2.2 Waveletové metody pro separaci signálu od šumu

Základ pro všechny waveletové techniky separace byl položen v článku [17]. Zde byl uveden princip úpravy waveletového spektra po ortogonální transformaci za účelem odstranění šumu. Metodologie, kterou její tvůrci nazvali *Waveshrink*, vychází z vlastnosti waveletové transformace koncentrovat podstatnou část energie do několika málo koeficientů. V případě, že signál není pohlcen šumem, stačí vypustit „malé“ koeficienty, které reprezentují šumovou složku. Autoři také ukázali, že v určitém statistickém smyslu má tento přístup optimální

vlastnosti.

Na základě této metody byla postupně vystavěna různá zobecnění a vylepšení. Například Silverman a Johnstone [22] zobecnili tuto metodu i pro barevný (korelovaný) šum.

Dále byly zkoumány přínosy přístupu „wavelet shrinkage“ [48], což je ještě obecnější přístup k úpravě koeficientů, než bylo základní prahování použité v článku [17].

Původně pro kompresi otisků prstů pro FBI byla určena nová metoda, kterou přinesli Wickerhauser a Coifman [10] a která se nazývá „wavelet packets“. Tato metoda se ukázala být vhodnou pro daleko širší škálu problémů. Stejní autoři zároveň ukázali, jak vybírat optimální waveletovou bázi, která se co nejvíce přimyká charakteru zpracovávaného signálu. Při zvolení takovéto báze je možné dosáhnout lepších výsledků i v separačních úlohách.

U waveletové separace signálu od šumu pomocí prahování je nutné zvolit tzv. prahovací pravidlo, které říká, jakým způsobem bude zacházeno s koeficienty, které nebyly vynulovány. V základní metodě existují dvě pravidla – tvrdé a měkké prahování. Vzhledem k tomu, že obě tato základní pravidla kromě specifických výhod také trpí specifickými nedostatky, byla později uvedena ještě další pravidla, která se snaží příhodné vlastnosti obou typů kombinovat [48, 66, 70].

Na prahovací pravidla můžeme nahlížet z perspektivy matematické statistiky a v tomto smyslu posuzovat jejich vlastnosti. Statistickou technikou „Exact Risk Analysis“ lze odvodit několik veličin, které charakterizují základní znaky prahovacích pravidel [6, 23, 24]. V literatuře však chybí právě takový typ analýzy u tzv. hyperbolického prahování, což je jeden z nestandardních typů; zřejmě to je z důvodu, že je nutno analyticky vyjádřit velmi speciální integrály.

Rovněž chybí šířeji pojaté pojednání, které by se zabývalo srovnáním vlastností všech používaných pravidel v jednom celku.

Statistické vlastnosti bývají v literatuře uváděny pro speciální případ jednotkového rozptylu šumu, tj. $\sigma = 1$. Důkaz, že tento předpoklad není na újmu obecnosti, a zejména jakým způsobem vlastnosti přepočítat na případ šumu s jiným rozptylem, v dostupné literatuře také není uveden.

Nutno poznamenat, že waveletové metody separace signálů jsou omezeny na případ, kdy aditivní šum má známé (nebo alespoň odhadnutelné) statistické vlastnosti. Čím více se šum „blíží“ bílému, tím vyšší účinnost metody zaznamenávají. Pro separaci dvou signálů

deterministického charakteru (např. řeč a hluk motoru) metody selhávají.

1.2.3 Waveletová transformace realizovatelná v reálném čase

Za současného stavu vývoje problematiky není možné aplikovat waveletovou transformaci pro zpracování signálů v reálném čase. Autoři se zabývají teoretickými i praktickými aspekty waveletové transformace, avšak pouze pro zpracování signálu „off-line“⁴; v současném stavu poznání chybí sofistikovaný algoritmus, který by tento způsob zpracování umožnil.

Jestliže někteří autoři přistupují k sekvenčnímu waveletovému zpracování dat, tzn. po segmentech, např. [11, 68], volí jejich délku a překrytí ad hoc a kompenzují pak vzniklá zkreslení průměrováním či vážením výstupních segmentů. Jak bude ukázáno v této dizertační práci, lze však vyvinout algoritmus, pomocí něhož se určí překryvy segmentů a způsob jejich zpracování tak, že *k žádnému zkreslení nedochází*. Vývoj takového algoritmu je úkol značně obtížný a vyžaduje hlubokou znalost algoritmu waveletové transformace, neboť do úvahy se musí vzít proměnné veličiny jako délka segmentu, délka waveletového filtru a v neposlední řadě počet úrovní dekompozice.

⁴tzn. předpokládáme-li znalost celého průběhu signálu předem

2 Cíle dizertační práce

Na základě rozboru stavu současné problematiky uvedeného v kapitole 1 byly formulovány cíle této dizertační práce:

- Srovnat pět druhů prahovacích pravidel na základě statistické analýzy rizika. K tomu bude třeba nalézt vzorce střední hodnoty, rozptylu a rizikové funkce pro hyperbolické prahování, a v případě neúspěchu vyjádřit tyto statistiky alespoň numericky. Dále uvést důkaz, že předpoklad jednotkového rozptylu $\sigma = 1$ není na újmu obecnosti, a uvést z toho plynoucí výpočty statistických veličin pro případ $\sigma \neq 1$ pomocí předchozího případu.
- Prozkoumat, zda a jak velký přínos má změna waveletové báze (technika „wavelet packets“, která se v minulosti ukázala přínosnou v oblasti efektivní komprese obrázových dat.) pro řečové signály zatížené šumem. Tento přínos posoudit kvantitativně i kvalitativně.
- Vytvořit algoritmus, pomocí kterého bude možné waveletové zpracování diskrétního signálu po segmentech, tj. zejména v reálném čase. Sofistikovaný algoritmus tohoto typu, který by nevnášel do zpracovávaného signálu zkreslení, zatím nebyl vyvinut. Nový algoritmus bude přitom využívat postupy používané ve standardní waveletové transformaci.

3 Výchozí poznatky z matematiky a teorie zpracování signálů

Úkolem této kapitoly je uvést, příp. ozřejmit některé pojmy, jež budou v průběhu dizertační práce často používány. Abychom tuto kapitolu nečinili příliš podrobnou a zdlouhavou, v některých případech se odvoláváme na literaturu.

3.1 Matematika

Obvyklý základní aparát *vektorových prostorů* přebíráme z [37, 14]. V oblasti funkcionálních prostorů, ze které vychází také teorie waveletů, se často pracuje s *Hilbertovými prostory*. Jsou to úplné vektorové prostory s definovaným skalárním součinem. *Skalární součin* dvou prvků x, y z vektorového prostoru budeme značit $\langle x, y \rangle$. Pomocí skalárního součinu je možné vždy definovat tzv. indukovanou *normu* $\|x\| = \sqrt{\langle x, x \rangle}$.

Jako příklad vektorových prostorů můžeme uvést prostory funkcí

$$L^p(\mathbb{R}) = \left\{ f \left| \left(\int_{-\infty}^{\infty} |f(t)|^p dt \right)^{1/p} < \infty \right. \right\}, \quad 1 \leq p < \infty \quad (3.1)$$

(integrál uvažujeme v Lebesgueově smyslu [37]). V textu této práce budeme pracovat v prostoru $L^2(\mathbb{R})$, který je dokonce Hilbertovým prostorem; $L^1(\mathbb{R})$ však není! Dalšími typickými Hilbertovými prostory jsou \mathbb{R}^n (\mathbb{C}^n), prostory všech reálných (komplexních) vektorů délky n .

Jedním z nejdůležitějších pojmu nejen pro účely této práce je báze. *Báze vektorového prostoru* E je množina prvků E , které jsou lineárně nezávislé a uzávěr jejich lineárního obalu je roven prostoru E . Dva nenulové prvky $x, y \in E$ nazýváme *ortogonálními*, pokud pro ně platí $\langle x, y \rangle = 0$. Pokud navíc $\|x\| = \|y\| = 1$, řekneme, že x, y jsou *ortonormální*. Ortogonální, resp. ortonormální báze E je pak taková báze E , jejíž každé dva prvky jsou vzájemně ortogonální, resp. ortonormální.

Pokud A je regulární čtvercová matice (tj. má plnou hodnost), pak symbolem A^{-1} označujeme matici *inverzní* k A , pro niž platí $AA^{-1} = A^{-1}A = I$. Řekneme, že matice A

je *ortogonální*, pokud platí $\mathbf{A}^\top = \mathbf{A}^{-1}$. Taková matice má vlastnost, že vektory ve všech jejích sloupcích jsou navzájem ortonormální, neboť $\mathbf{A}^\top \mathbf{A} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$ a stejně tak ve všech řádcích, neboť $\mathbf{A} \mathbf{A}^\top = \mathbf{A} \mathbf{A}^{-1} = \mathbf{I}$.

Budeme-li hovořit o waveletových paketech (kap. 5), je vhodné k tomu využít terminologie teorie grafů. Základní pojmy jako *graf*, *uzel*, *hrana*, *cesta z uzlu do uzlu*, *následník*, *předchůdce* přebíráme z [31]. Matematické struktury teorie grafů obvykle vyjadřují intuitivní představu nějakého reálného objektu. Typ grafů, jimiž vyjadřujeme strukturu waveletového paketu, se díky svému charakteristickému tvaru nazývají *stromy*. Vzhledem k tomu, že každý uzel waveletového paketu má buď dva nebo žádného následníka, mluvíme o binárních stromech. Uzly, které nemají následníka, nazýváme přirozeně *koncové uzly*, příp. *listy*. *Hloubkou uzlu* ve stromu budeme rozumět délku cesty do tohoto uzlu z tzv. kořenového uzlu. *Hloubkou stromu* budeme rozumět délku nejdelší cesty v tomto stromu. *Úplný binární strom hloubky d* je binární strom, jehož všechny listy mají hloubku d .

Základní aparát pravděpodobnostního počtu přebíráme z učebnice [3]. Všechny náhodné veličiny, se kterými budeme pracovat, jsou definovány na stejném pravděpodobnostním prostoru (Ω, \mathcal{A}, P) . Skutečnost, že náhodný vektor \mathbf{Y} má rozložení o hustotě $f(\mathbf{y}, \Psi)$, budeme vyjadřovat zápisem $\mathbf{Y} \sim f(\Psi)$. Speciálně pro náhodný vektor, který má p -rozměrné normální rozložení o střední hodnotě \mathbf{m} a kovarianční matici Σ , příšeme $\mathbf{Y} \sim N_p(\mathbf{m}, \Sigma)$. Hustotu pravděpodobnosti standardizovaného normálního rozložení $N(0, 1)$ značíme $\varphi(x)$ a je definována vzorcem $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Distribuční funkci $N(0, 1)$ značíme $\Phi(x)$ a je definována $\Phi(x) = \int_{-\infty}^x \varphi(x) dx$. V dalším textu, zejména v důkazech a odvození, využíváme některé jejich vlastnosti:

$$\varphi(-x) = \varphi(x) \tag{3.2}$$

$$\Phi(-x) = 1 - \Phi(x) \tag{3.3}$$

$$\frac{d\varphi(x)}{dx} = -x \varphi(x). \tag{3.4}$$

3.2 Teorie zpracování signálů

Pokud budeme mluvit o signálu, budeme mít na mysli buď funkci v běžném slova smyslu (ve spojitém případě) nebo posloupnost (v diskrétním případě). Ze souvislosti bude vždy

zřejmé, o který případ se jedná. Posloupnosti a jejich složky budeme značit stejně jako konečné vektory, např. $\mathbf{x} = (\dots, x_{-1}, x_0, x_1, x_2, \dots)$.

Konvolucí dvou diskrétních signálů rozumíme signál $\mathbf{z} = \mathbf{x} * \mathbf{y}$, kde je definováno $z_n = \sum_k x_k y_{n-k}$. Pokud \mathbf{x} je vektor délky $n \in \mathbb{N}$ a \mathbf{y} vektor délky $m \in \mathbb{N}$, pak \mathbf{z} má délku $n + m - 1$.

(Diskrétním) lineárním filtrem rozumíme operátor F , který provádí zobrazení z prostoru všech posloupností P na tentýž prostor (tj. transformuje „vstupní“ signál na „výstupní“ signál) a platí pro něj

$$F(\alpha\mathbf{x} + \mathbf{y}) = \alpha F(\mathbf{x}) + F(\mathbf{y}) \quad \text{pro všechna } \mathbf{x}, \mathbf{y} \in P, \alpha \in \mathbb{R}. \quad (3.5)$$

Lze ukázat, že každý lineární časově invariantní filtr [51] je jednoznačně reprezentován posloupností \mathbf{f} ve smyslu $F(\mathbf{x}) = \mathbf{f} * \mathbf{x}$ pro každé $\mathbf{x} \in P$. Lineární časově invariantní filtry budeme dále nazývat jednoduše *filtry*. Posloupnost \mathbf{f} nazýváme *impulzní odezvou filtru* F , neboť položíme-li za \mathbf{x} jednotkový impulz, výstupem bude právě posloupnost \mathbf{f} . Vzhledem k předchozímu si dovolujeme nazývat *filtrem* i samotnou posloupnost \mathbf{f} . Pokud \mathbf{f} má konečný počet nenulových prvků, říkáme, že jde o *filtry s konečnou impluzní odezvou* (FIR), jinak říkáme, že jde o *filtry s nekonečnou impluzní odezvou* (IIR).

Funkci $F(z) = \sum_k f_k z^{-k}$, pokud součet existuje alespoň pro jedno $z \in \mathbb{C}$, nazýváme \mathcal{Z} -transformací signálu \mathbf{f} . Její hodnota vyjádřená pouze na jednotkové kružnici, tedy

$$F(e^{i\omega}) = \sum_k f_k e^{-i\omega k}, \quad (3.6)$$

nazýváme *kmitočtovou odezvou filtru* \mathbf{f} . $F(e^{i\omega})$ můžeme chápát jako komplexní funkci proměnné ω a lze ji tedy rozložit jako

$$F(e^{i\omega}) = A(\omega) \cdot e^{-i\theta(\omega)} \quad (3.7)$$

na *amplitudovou kmitočtovou odezvu* $A(\omega) = |F(e^{i\omega})|$ a *fázovou kmitočtovou odezvu* $\theta(\omega) = \arg F(e^{i\omega})$. Lze snadno ukázat, že $F(e^{i\omega})$ je 2π -periodická vzhledem k *úhlovému kmitočtu* ω .

Filtr \mathbf{f} nazveme *horní propustí*, pokud platí, že amplitudová kmitočtová odezva je na kmitočtu $\omega = \pi$ nenulová a na kmitočtu $\omega = 0$ nulová, tj. $|F(-1)| \neq 0, |F(1)| = 0$. Filtr \mathbf{f} nazveme *dolní propustí*, pokud platí $|F(1)| \neq 0, |F(-1)| = 0$ [39].

4 Waveletová transformace

V této kapitole jsou uvedeny nezbytné základy teorie waveletové transformace. Nejprve je princip rozkladu funkce (signálu) do množiny bázových funkcí (bázových signálů) vysvětlen v oblasti spojitého času (integrální waveletová transformace, CWT,¹ a diskrétní waveletová transformace, DWT), část 4.1. V části 4.2 bude zavedena tzv. měřítková funkce, důležitá pro odvození mnohoměřítkové analýzy. V části 4.3 je pak vysvětlena analogická podoba waveletové transformace v oblasti diskrétního času (tzv. konečná diskrétní waveletová transformace, DTWT²), které se budeme věnovat ve zbytku práce, a její úzká souvislost s lineární filtrací signálů.

4.1 Integrální waveletová transformace

Teorie waveletové analýzy je podobně jako Fourierova analýza signálu budována obvykle pro funkce z prostoru $L^2(\mathbb{R})$, což je množina všech (komplexních) funkcí definovaných na \mathbb{R}

$$L^2(\mathbb{R}) = \left\{ f \left| \sqrt{\int_{-\infty}^{\infty} |f(t)|^2 dt} < \infty \right. \right\},$$

kde integrál je uvažován v Lebesgueově smyslu [14, 37]. Tyto funkce představují signály s konečnou energií. Vnitřní součin, resp. normu v $L^2(\mathbb{R})$ definujeme

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t) \overline{g(t)} dt, \quad (4.1)$$

resp.

$$\|f\| = \sqrt{\int_{-\infty}^{\infty} |f(t)|^2 dt}. \quad (4.2)$$

Waveletem nazýváme funkci $\psi \in L^2(\mathbb{R})$ splňující tzv. podmínu přípustnosti

$$\int_{-\infty}^{\infty} \frac{|\widehat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad (4.3)$$

¹Continuous Wavelet Transform, proto v české literatuře najdeme i název spojitá waveletová transformace

²Discrete-Time Wavelet Transform, rovněž bývá používán název Finite Discrete Wavelet Transform

kde $\widehat{\psi}$ představuje integrální Fourierovu transformaci funkce ψ . Podmínka přípustnosti je postačující podmínkou pro existenci inverzní integrální waveletové transformace. Dá se ukázat, že v případě $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ díky (4.3) musí rovněž platit

$$\int_{-\infty}^{\infty} \psi(t) dt = 0, \quad (4.4)$$

podmínka nulové průměrné hodnoty. Dohromady s $\psi \in L^2(\mathbb{R})$ to lze fyzikálně interpretovat tak, že wavelety nutně musí mít oscilatorický charakter a jejich kmity musí být tlumeny směrem k $\pm\infty$.

(Integrální) waveletová transformace funkce (signálu se spojitým časem) f je definována

$$CWT_f(a, b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (4.5)$$

pro $a \in \mathbb{R} \setminus \{0\}$, $b \in \mathbb{R}$. Pro každou f je to funkce dvou spojitých proměnných; b vyjadřuje posunutí ψ v čase, resp. a vyjadřuje změnu měřítka³. Koeficient $|a|^{-\frac{1}{2}}$ je zde kvůli zachování normy v $L^2(\mathbb{R})$ (tj. energie), tzn. aby pro každé přípustné a platilo $\left\| |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \right\| = \|\psi(t)\|$. Podobně jako u transformací Fourierova typu však i zde stačí pro jednoznačnou existenci inverzní transformace brát v úvahu pouze vybrané z funkcí $\psi\left(\frac{t-b}{a}\right)$: definujeme-li spočetně mnoho funkcí

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k), \quad j, k \in \mathbb{Z}, \quad (4.6)$$

pak vzorcem

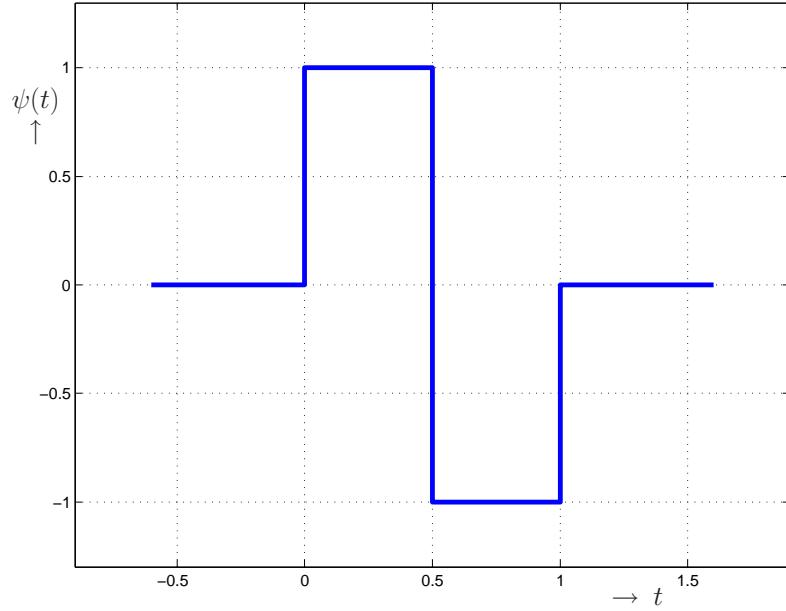
$$DWT_f(j, k) = \langle f, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} f(t) \overline{\psi_{j,k}(t)} dt, \quad j, k \in \mathbb{Z} \quad (4.7)$$

definujeme tzv. diskrétní waveletovou transformaci. Jinými slovy, DWT je vzorkovanou verzí CWT na síti parametrů $a = 2^j$, $b = 2^j k$, $j, k \in \mathbb{Z}$, neboť platí

$$\begin{aligned} CWT_f(2^j, 2^j k) &= |2^j|^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\psi(2^{-j}t - k)} dt \\ &= \int_{-\infty}^{\infty} f(t) \overline{2^{-j/2} \psi(2^{-j}t - k)} dt \\ &= \langle f, \psi_{j,k} \rangle = DWT_f(j, k). \end{aligned}$$

Název „diskrétní“ je odvozen od diskrétní (i když nekonečné) množiny parametrů transformace j, k .

³tzv. translace, resp. dilatace



Obr. 4.1: Mateřský wavelet typu Haar.

Byly nalezeny funkce ψ takové, že $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ je úplnou ortonormální bází $L^2(\mathbb{R})$. Pak libovolnou funkci $f \in L^2(\mathbb{R})$ lze vyjádřit součtem waveletové řady

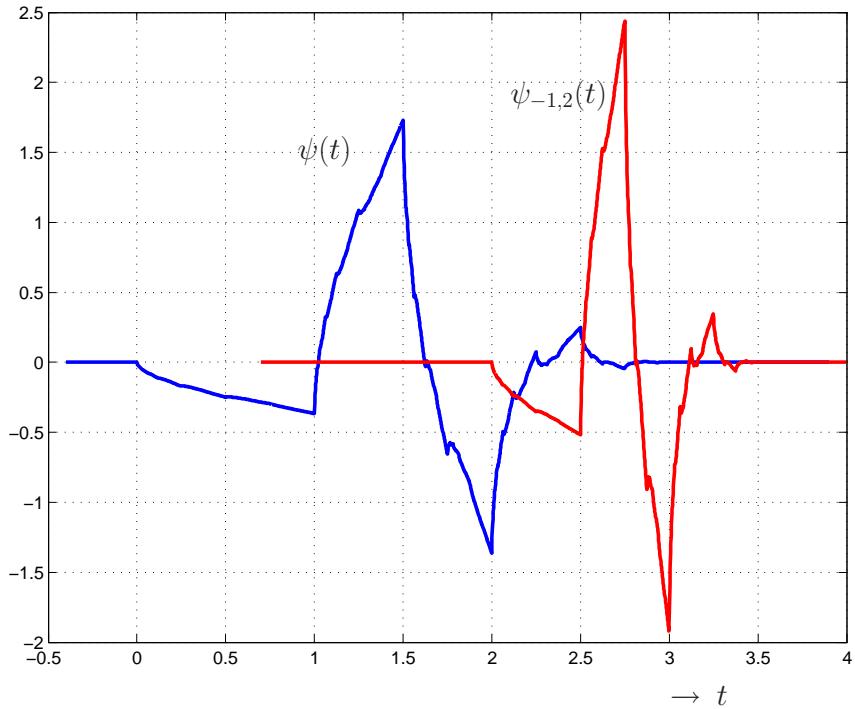
$$f(t) = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}(t), \quad (4.8)$$

přičemž rovnost platí pro skoro všechna t . Číslo $c_{j,k} = \langle f, \psi_{j,k} \rangle$ nazýváme j, k -tým waveletovým koeficientem funkce f . Pro množinu $\{c_{j,k}\}_{j,k \in \mathbb{Z}}$ používáme pojmenování waveletové spektrum funkce f , funkci ψ říkáme mateřský wavelet.

Byly dokonce nalezeny ortonormální báze $L^2(\mathbb{R})$ takové, že její prvky (funkce) mají kompaktní nosič, tj. lze nalézt ohraničený interval takový, že mimo něj má funkce nulovou hodnotu. Příkladem takových funkcí jsou wavelety typu Daubechies [12], obr. 4.2, nebo nejjednoduší tzv. Haarův wavelet (obr. 4.1).

Kompaktnost nosiče waveletu hraje často důležitou úlohu. Fakt, že wavelet ψ je vně nějakého ohraničeného intervalu nulový, znamená, že přínos každého jeho posunutí a smrštění $\psi_{j,k}$ k analyzované funkci je pouze lokální (délka intervalu záleží na zvoleném j , čím vyšší j , tím delší⁴ délka intervalu působnosti konkrétního waveletu).

⁴ 2^j -krát



Obr. 4.2: Příklad dilatace a translace mateřského waveletu. Funkce $\psi = \psi_{0,0}$ je originální wavelet typu Daubechies 2, funkce $\psi_{-1,2}$ je odvozený wavelet s dilatační konstantou $j = -1$ (tedy je $2^j = 2$ -krát „užší“ než originál) a konstantou translace $k = 2$ (wavelet je posunut doprava na číselné ose). Díky normalizačnímu koeficientu je zároveň $2^{-j/2} = \sqrt{2}$ -krát „vyšší“ než originál. Tyto dvě funkce jsou ortonormální.

4.2 MR – analýza

Pojem „MR-analýza“ neboli mnohoměřítková analýza pro rozklad signálu do bázových funkcí různých měřítek pochází z anglického názvosloví (multiresolution analysis, MRA).⁵ Koncept MR-analýzy vychází z toho, že velké hodnoty měřítka j korespondují s reprezentací trendu v signálu, zatímco pro $j \rightarrow -\infty$ bázové funkce $\psi_{j,k}$ vystihují detailnější chování signálu. To umožní rozložit jakýkoliv signál i s nespojitostmi nebo s ostrými hrotami.

Pro přípustný wavelet ψ lze formálně tento rozklad zapsat

$$L^2(\mathbb{R}) = \dots \oplus W_2 \oplus W_1 \oplus W_0 \oplus W_{-1} \oplus \dots, \quad (4.9)$$

kde \oplus značí ortogonální součet prostorů funkcí. Pokud navíc existuje tzv. měřítková funkce

⁵V české odborné literatuře se používají rovněž označení analýza o více úrovních rozlišení, analýza s několikanásobným rozlišením.

$\phi \in L^2(\mathbb{R})$, která generuje MR-analýzu $\{V_j\}_{j \in \mathbb{Z}}$ v $L^2(\mathbb{R})$ (přesnou matematickou definici viz [27, 14]), pak platí

$$V_j = V_{j+1} \oplus W_{j+1} \quad (4.10)$$

a rozklad (4.9) lze upravit na

$$L^2(\mathbb{R}) = V_j \oplus W_j \oplus W_{j-1} \oplus W_{j-2} \oplus \dots \quad (4.11)$$

pro libovolné pevně zvolené j . V uvedených vztazích

$$W_j = \left\{ w(t) \mid w(t) = \sum_{k=-\infty}^{\infty} s_{j,k} \psi_{j,k}(t), s_{j,k} \in \mathbb{C} \right\}, \quad j \in \mathbb{Z} \quad (4.12)$$

$$V_j = \left\{ x(t) \mid x(t) = \sum_{i=\infty}^{j+1} \sum_{k=-\infty}^{\infty} s_{i,k} \psi_{i,k}(t), s_{i,k} \in \mathbb{C} \right\} \quad (4.13)$$

$$= \left\{ x(t) \mid x(t) = \sum_{k=-\infty}^{\infty} s_{j,k} \phi_{j,k}(t), s_{j,k} \in \mathbb{C} \right\}, \quad j \in \mathbb{Z}, \quad (4.14)$$

kde definujeme $\phi_{j,k} = 2^{-j/2} \phi(2^{-j}t - k)$. Prostory V_j, W_j, W_{j-1}, \dots jsou (pro každé zvolené $j \in \mathbb{Z}$) navzájem ortogonální, tzn. pro libovolné dvě funkce z navzájem různých prostorů, f, g , platí $\langle f, g \rangle = 0$. Každou funkci $f \in L^2(\mathbb{R})$ tak lze rozložit pomocí MR-analýzy jako součet navzájem ortogonálních funkcí, např.

$$f(t) = \underbrace{\sum_{l \in \mathbb{Z}} a_{0l} \phi(t-l)}_{\in V_0} + \sum_{k,j \in \mathbb{Z}, j \leq 0} b_{j,k} \psi_{j,k}(t), \quad (4.15)$$

kde $a_{0l}, b_{j,k} \in \mathbb{C}$ jsou souřadnice f v bazích prostorů $V_0, W_0, W_{-1}, W_{-2}, \dots$. Příklad MR-analýzy, kdy f považujeme za prvek prostoru $V_0 = W_1 \oplus W_2 \oplus W_3 \oplus W_4 \oplus W_5 \oplus V_5$, ilustruje obrázek 4.3.

V základu MR-analýzy leží dva důležité vztahy vycházející z (4.10), tzv. měřítková rovnice

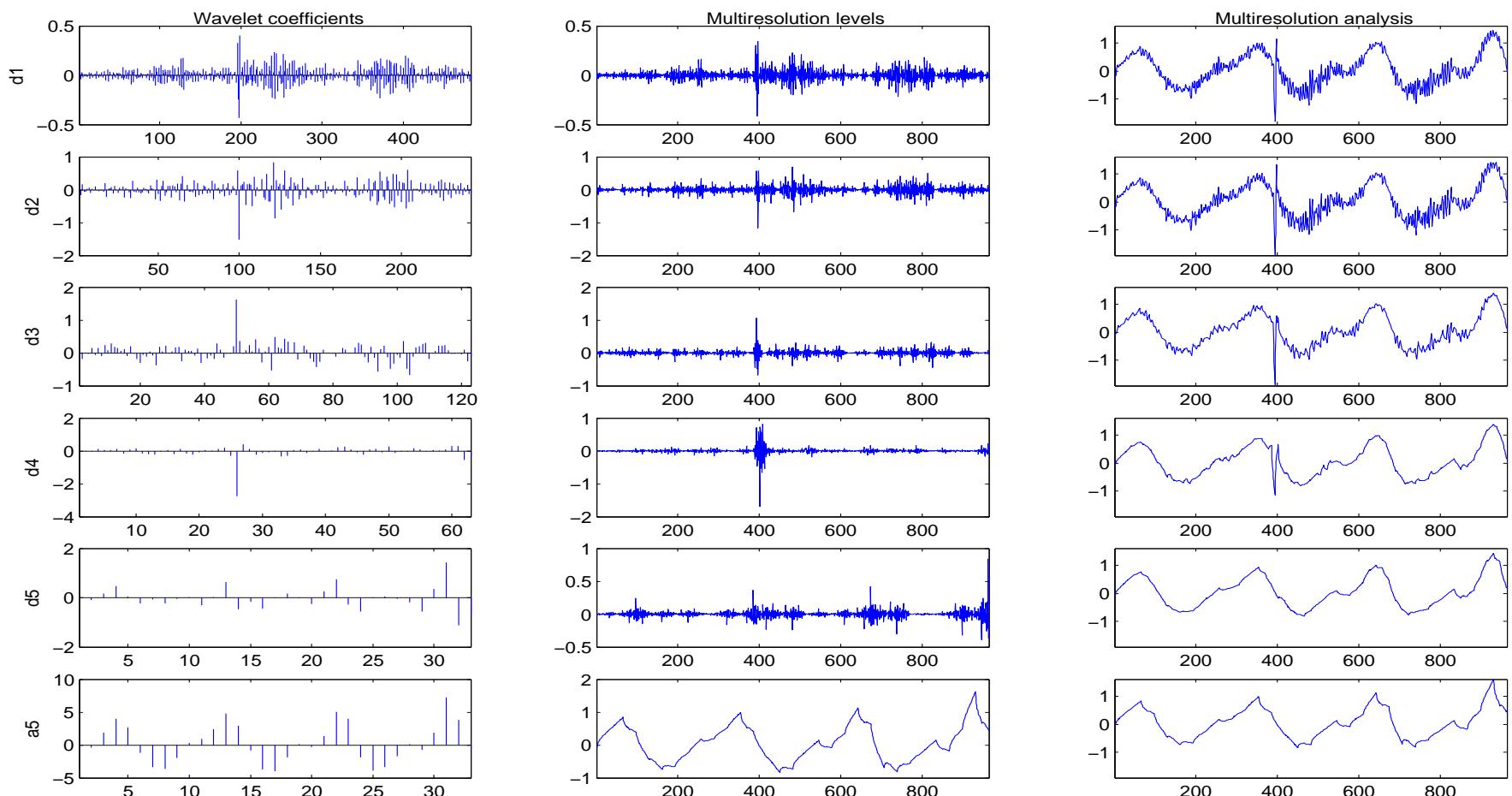
$$\phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} h_k \phi(2t-k) \quad (4.16)$$

a waveletová rovnice

$$\psi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} g_k \phi(2t-k) \quad (4.17)$$

pro dané posloupnosti (souřadnic) $\mathbf{h} = (\dots, h_{-1}, h_0, h_1, h_2, \dots)$ a $\mathbf{g} = (\dots, g_{-1}, g_0, g_1, g_2, \dots)$.

Tedy funkce ϕ a ψ se obě dají vyjádřit jako lineární kombinace posunutí dvakrát „smrštěné“ měřítkové funkce $\phi(2t)$.



Obr. 4.3: Ukázka MR-analýzy úseku hudebního signálu provedené s waveletem Daubechies řádu 2. V grafech v levém sloupci jsou detailní a approximační koeficienty jednotlivých úrovní, v prostředním sloupci jsou příspěvky jednotlivých podprostorů, od zdola nahoru V_5, W_5, \dots, W_1 , v pravém sloupci jim příslušné částečné součty. Vpravo nahoře je výchozí signál.

Pro každý wavelet jsou vektory \mathbf{h} a \mathbf{g} pochopitelně jiné.

4.2.1 Wavelety s nulovými momenty v MR-analýze

Důležitý koncept představují tzv. nulové momenty. V této sekci dále pracujeme pouze s wavelety, které mají kompaktní nosič. Říkáme, že wavelet ψ má u nulových momentů, jestliže

$$\int_{-\infty}^{\infty} \psi(t) t^i dt = 0 \quad \text{pro } i = 0, 1, \dots, u-1. \quad (4.18)$$

Každý přípustný wavelet má alespoň jeden nulový moment, jak plyne z podmínky (4.4).

Nulové momenty lze interpretovat takto: má-li wavelet ψ u nulových momentů, pak všechny polynomy až do řádu $u-1$ lze vyjádřit jako lineární kombinace celočíselných posunutí měřítkové funkce $\phi(t-l), l \in \mathbb{Z}$. To tedy znamená, že pokud podrobíme MR-analýze (viz oddíl 4.2) polynom $p(t)$ řádu nejvýše $u-1$, všechny detailní waveletové koeficienty

$$c_{j,k} = \int_{-\infty}^{\infty} \psi_{j,k}(t) p(t) dt \quad (4.19)$$

budou nulové pro $j \geq 0$.⁶ To je pro naše další úvahy velmi důležité, neboť pokud by se v signálu hladkého charakteru vyskytla „singularita“, odrazila by se tato skutečnost v MR-analýze tím, že pouze v místě příslušném této singularitě by koeficienty $c_{j,k}$ pro určitá j byly významně nenulové. Tento princip je velice dobře vidět v obrázku 4.3 v grafech příslušných třetí a čtvrté úrovni dekompozice.

4.3 Konečná diskrétní waveletová transformace

V praktických úlohách ovšem nejčastěji pracujeme s diskrétními signály konečné délky. Pro tento typ signálů je možno z integrální waveletové transformace odvodit její verzi pro diskrétní signály, obdobně, jako je tomu obvyklé např. u Fourierovy transformace. Tuto transformaci nazýváme konečnou diskrétní waveletovou transformací a značíme zkratkou DTWT. DTWT můžeme vyjádřit pomocí ortogonální⁷ matice \mathbf{W} řádu $n \times n$ [48, kap. 4].

⁶Bylo by formálně nesprávné zapsat $c_{j,k} = \langle p, \psi_{j,k} \rangle$, neboť $p \notin L^2(\mathbb{R})$.

⁷matice \mathbf{W} může být v případě volby určitých typů okrajových podmínek [39, kap. 8] při waveletové transformaci od ortogonality „nepatrně“ odchýlena

Tato matice obsahuje speciálně uspořádané prvky z vektorů \mathbf{h} a \mathbf{g} . Je-li $\mathbf{y} = (y_1, \dots, y_n)^\top$ vektor délky n , pak jeho konečnou diskrétní waveletovou transformací je vektor $\mathbf{d} = (d_1, \dots, d_n)^\top$, získaný jako součin

$$\mathbf{d} = \mathbf{W}\mathbf{y}. \quad (4.20)$$

Inverzní waveletovou transformaci je možno díky ortogonalitě \mathbf{W} vyjádřit jako

$$\mathbf{y} = \mathbf{W}^{-1}\mathbf{d} = \mathbf{W}^\top\mathbf{d}. \quad (4.21)$$

Z předchozího popisu vyplývá důležitá vlastnost waveletové transformace – její *linearity*.

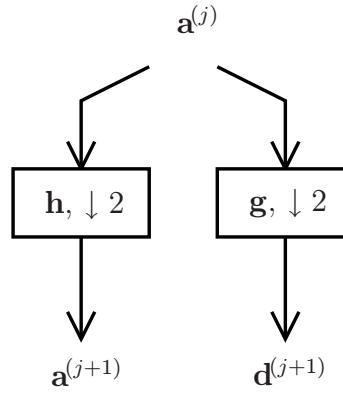
4.3.1 Souvislost DTWT s lineární filtrací signálu a Mallatův pyramidový algoritmus

Při provádění konečné diskrétní waveletové transformace lze násobení vektoru \mathbf{y} ortogonální maticí \mathbf{W} nahradit pyramidovým algoritmem, který pochází od S. Mallata [27, 28]. Mallat využil pevně daného řádu ve struktuře matice \mathbf{W} a díky tomu vyvinul následující rekurzivní algoritmus:⁸

Vstupní vektor \mathbf{y} filtrujieme filtrem typu dolní propust \mathbf{h} , resp. filtrem typu horní propust \mathbf{g} . Tyto filtry jsou speciální, tzv. kvadraturní zrcadlové filtry, viz [39], a jak již bylo naznačeno, jejich souvislost s měřítkovou funkcí ϕ , resp. s příslušným waveletem ψ je dána rovnicemi (4.16) a (4.17). Výsledné posloupnosti decimujeme, tj. vypustíme z nich každý druhý vzorek. Tím získáme nové posloupnosti (přibližně dvakrát kratší než vstupní signál). Koeficienty takto získané pomocí \mathbf{g} nazýváme detailními waveletovými koeficienty a koeficienty získané pomocí \mathbf{h} nazýváme aproximačními waveletovými koeficienty vstupního signálu. Tento jeden krok waveletové dekompozice je znázorněn na obrázku 4.4. Detailní koeficienty uchováme a právě popsaný způsob můžeme opakovat s tím, že místo originálního signálu \mathbf{y} podrobíme filtraci a decimaci vektoru aproximačních koeficientů.

Počet takto provedených filtrací je volitelný a nazýváme jej hloubkou dekompozice. Má-li vstupní signál délku s , maximální hloubka dekompozice d_{\max} je omezena vztahem

⁸někdy se Mallatovu algoritmu říká rychlá waveletová transformace (FWT, Fast Wavelet Transform), jako analogie k algoritmu FFT



Obr. 4.4: Jeden krok waveletové dekompozice. Aproximační koeficienty j -té úrovně (vektor $\mathbf{a}^{(j)}$) jsou pomocí filtrů \mathbf{h} , resp. \mathbf{g} a decimace (ozn. $\downarrow 2$) rozloženy na approximační, resp. detailní koeficienty úrovně $j + 1$.

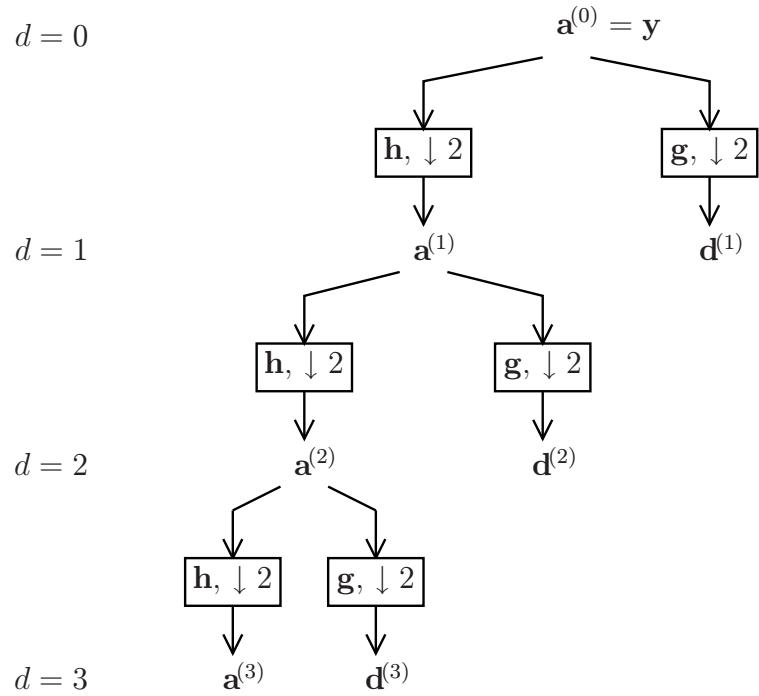
$$d_{\max} \leq \log_2 s [28].$$

Obr. 4.5 ilustruje Mallatův algoritmus DTWT pro hloubku dekompozice $d = 3$. Pro danou hloubku dekompozice d je waveletovou transformací signálu \mathbf{y} soubor koeficientů uložených ve vektorech $\mathbf{a}^{(d)}, \mathbf{d}^{(d)}, \mathbf{d}^{(d-1)}, \dots, \mathbf{d}^{(1)}$.

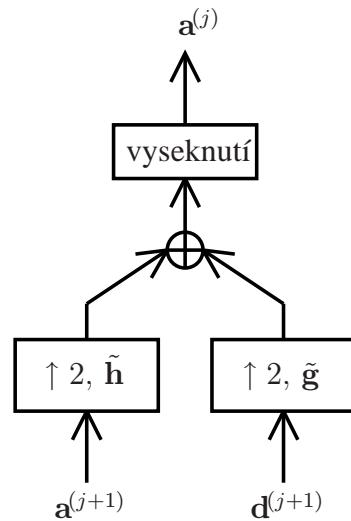
Podobně při inverzní transformaci není třeba násobit vektor waveletových koeficientů maticí \mathbf{W}^\top . Algoritmus rekonstrukce je podobný algoritmu dekompozice: Do posloupnosti approximačních, resp. detailních koeficientů vkládáme nuly, tj. provádíme interpolaci (nadvzorkování), a následně filtrujieme filtry $\tilde{\mathbf{h}}$, resp. $\tilde{\mathbf{g}}$ inverzními k \mathbf{h} , resp. \mathbf{g} . Výsledky sečteme, vysekneme jen potřebnou prostřední část a tím získáme approximační koeficienty příslušné k úrovni o jedna nižší. Tento postup provádíme rekurzivně tolikrát, kolik je hloubka dekompozice, až získáme výchozí signál v časové doméně. Jeden krok Mallatova pyramidového algoritmu waveletové rekonstrukce ilustruje obrázek 4.6.

Použití tohoto algoritmu znamená velké snížení výpočetní náročnosti waveletové transformace, neboť násobení signálu maticí podle vztahu (4.20), resp. (4.21) má složitost kvadratickou, a dále je naprostě nevhodné z hlediska paměťové náročnosti při výpočtu.

Detailní popis jednotlivých kroků Mallatova algoritmu a navazující rozbor je uveden v části 8.2, str. 65, jako potřebný úvod do metody segmentované waveletové transformace.



Obr. 4.5: Mallatův pyramidový algoritmus waveletové dekompozice o hloubce 3 signálu y . Vstupní signál je rekurzivní aplikací filtrů g, h rozložen na aproximační a detailní waveletové koeficienty. Po straně je označena hloubka dekompozice.



Obr. 4.6: Jeden krok waveletové rekonstrukce. Aproximační koeficienty j -té úrovně (vektor $\mathbf{a}^{(j)}$) jsou vypočítány z koeficientů úrovně o jednu vyšší ($\mathbf{a}^{(j+1)}, \mathbf{d}^{(j+1)}$) pomocí nadzorkování, filtrů \tilde{h} , resp. \tilde{g} , a vyseknutí užitečné části.

4.3.2 Kmitočtová charakteristika waveletových filtrů

Rozdílným waveletům přísluší rozdílné vektory \mathbf{h}, \mathbf{g} . Waveletům s neohraničeným nosičem přísluší vektory s nekonečným počtem koeficientů a tedy odpovídají IIR filtry. Waveletům s kompaktním nosičem, např. waveletům třídy Daubechies, odpovídají vektory s konečným počtem prvků, tedy FIR filtry. Na koeficientech filtrů \mathbf{h} , resp. \mathbf{g} záleží kmitočtová charakteristika dolní, resp. horní propusti. Idealizované kmitočtové charakteristiky filtrů při dekompozici výchozího signálu do jednotlivých pásem ukazuje obr. 4.7.

Filtry použité ve waveletové transformaci se někdy označují jako „oktávové“ filtry. Název vychází z principu pyramidového algoritmu, kdy kmitočtové pásmo signálu se rozpůlí, tj. horní oktava⁹ se oddělí od zbytku kmitočtů, a poté se stejná akce opakuje na dolní polovině kmitočtů. Při rekurzivním opakování algoritmu tímto způsobem dostaneme signál rozdelený do oktáv.

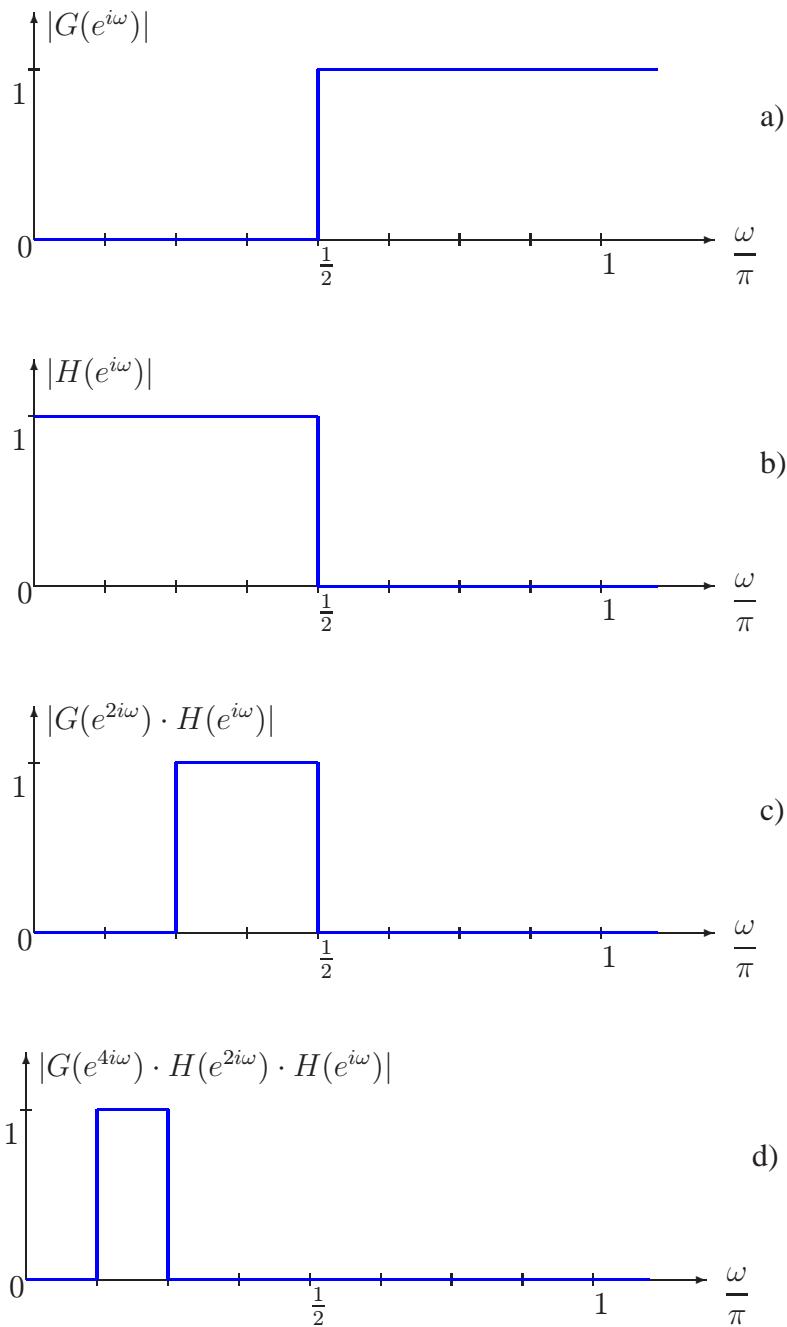
Žádné reálné filtry samozřejmě této ideální charakteristiky nedosahují. Na obrázku 4.8 je vyobrazeno srovnání amplitudové kmitočtové charakteristiky filtrů Daubechies řádu 2 ($2 \cdot 2 = 4$ koeficienty) a řádu 12 ($12 \cdot 2 = 24$ koeficientů). Vidíme, že čím vyšší řád filtru a tím více koeficientů, tím strmější je kmitočtová charakteristika a tím lépe jsou od sebe obě pásmá oddělena.

Praktický projev rozdílných vlastností různých waveletů lze pozorovat srovnáním obrázků 4.3 a 4.9. Na obou je uvedena MR-analýza stejného signálu, první provedená s waveletem Daubechies 2 a druhá s waveletem Daubechies 12. Je patrné, že v druhém případě dochází k lepší kmitočtové separaci (tzn. menší průsaky do ostatních pásem) v jednotlivých úrovních.

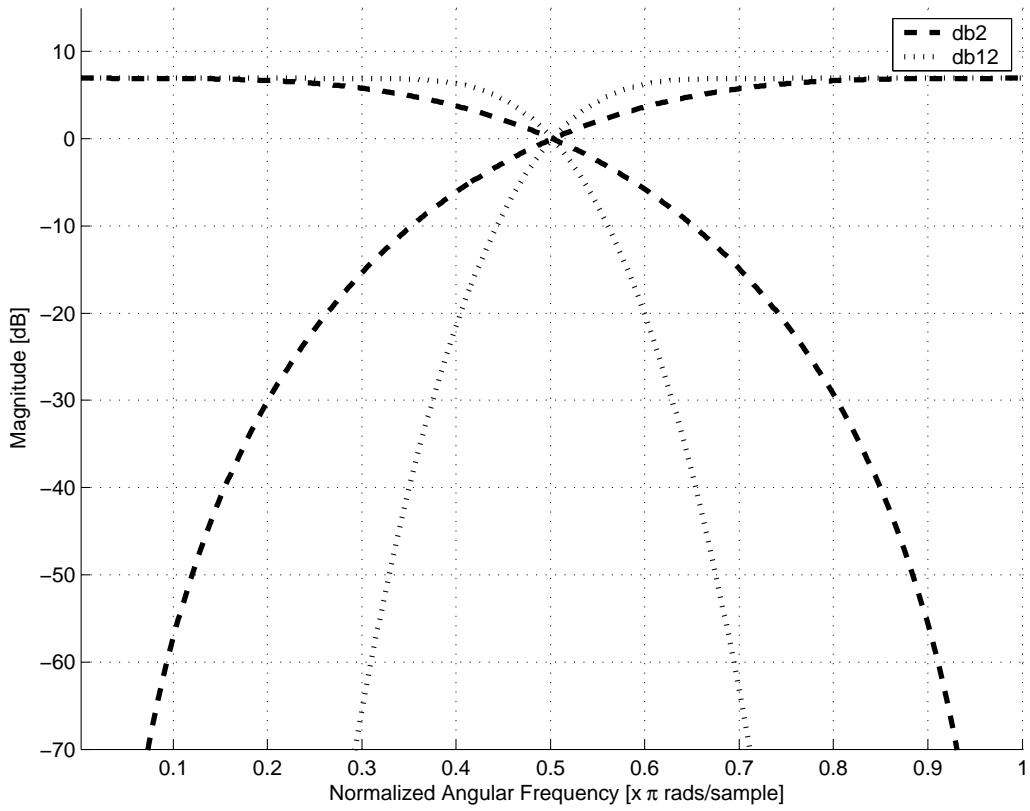
4.3.3 Waveletové filtry a aliasing

Žádný kmitočtový filtr nemá ideální charakteristiku tak jako na obr. 4.7. Pouze v případě ideálního filtru se decimace v každém kroku transformace neprojeví vznikem aliasingu. Je tedy jasné, že použitím každého reálného filtru k aliasingu dojde, a to tím menší měrou, čím strmější je amplitudová kmitočtová charakteristika filtru. Jako příklad uvedeme opět charakteristiku filtrů Daubechies 2 a Daubechies 12, tentokrát na obrázcích 5.4 (a),(b).

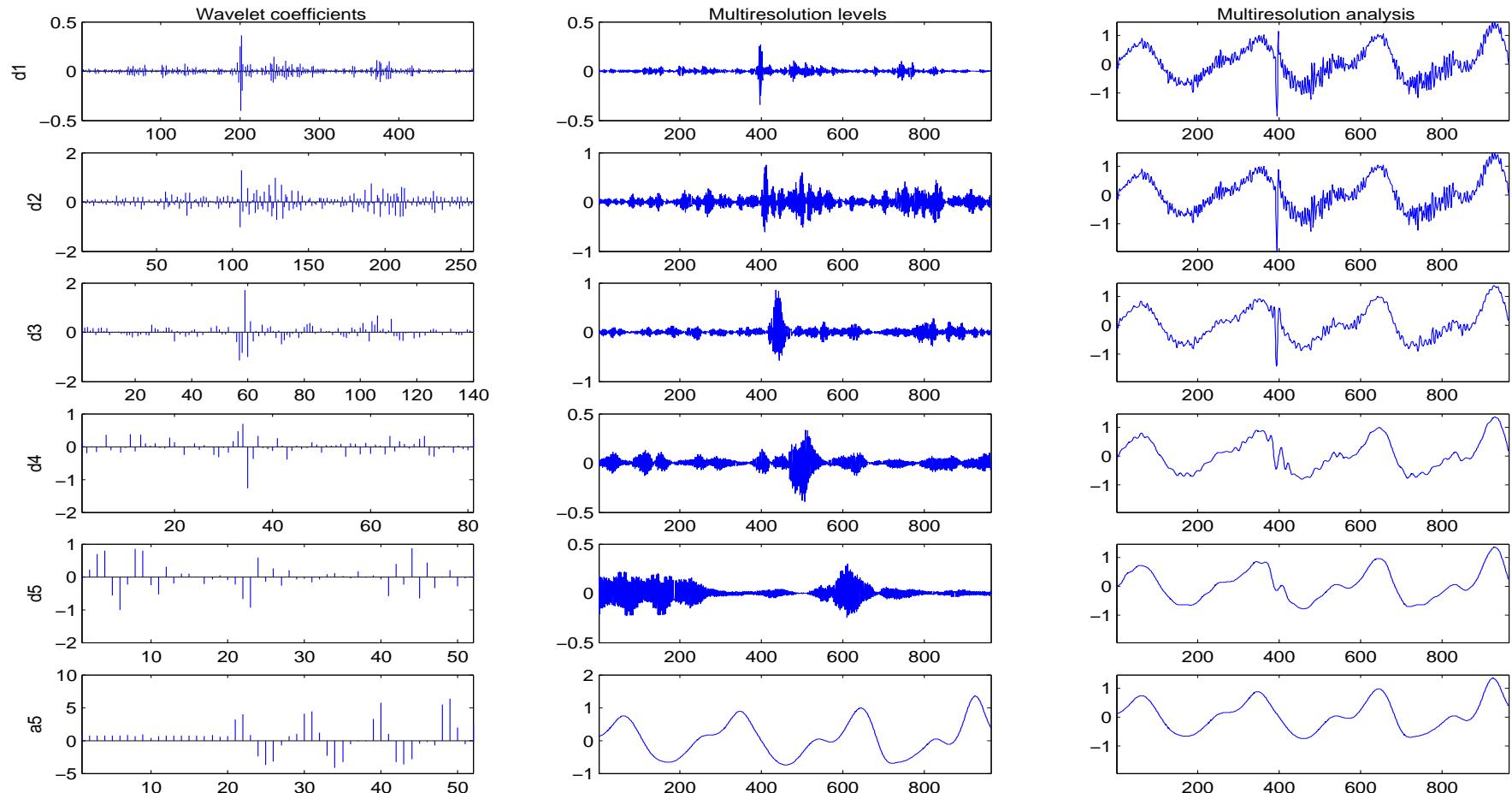
⁹koresponduje s názvoslovím hudební akustiky



Obr. 4.7: Idealizované amplitudové kmitočtové charakteristiky filtrů h a g při jejich rekurzivní aplikaci v dekompozičním pyramidovém algoritmu. Decimace v každé úrovni dekompozice způsobuje „smrštění“ kmitočtové osy na polovinu. Kmitočtová charakteristika a) odpovídá waveletovým koeficientům vektoru $d^{(1)}$, dále b) odpovídá $a^{(1)}$, c) odpovídá $d^{(2)}$, d) odpovídá $d^{(3)}$.



Obr. 4.8: Skutečná amplitudová charakteristika dvou rozdílných waveletů typu Daubechies, řádu 2 (db2) a řádu 12 (db12). Graf je převzat ze systému Matlab.



Obr. 4.9: MR-analýza stejného signálu jako v obrázku 4.3, tentokrát s waveletem Daubechies řádu 12. Je znatelná lepší separace jednotlivých kmitočtových pásem.

Waveletové filtry, které patří do skupiny kvadraturních zrcadlových filtrů, jsou však navrženy tak, že aliasing, který vznikne při dekompozici v obou větvích transformace, se při rekonstrukci vyruší [44]. Říkáme, že jde o tzv. banky filtrů s perfektní rekonstrukcí.

5 Waveletové pakety (wavelet packets) a výběr nejlepší báze

Waveletové pakety můžeme chápat jako zobecnění klasického pojetí waveletové transformace, která byla popsána výše. Pomocí obvyklých waveletů je vygenerována množina nových funkcí, ve které však již nejsou všechny její prvky (funkce) lineárně nezávislé, a tudíž z hlediska jednoznačné reprezentace prvků vektorového prostoru $L^2(\mathbb{R})$ je tato množina redundantní. Z tohoto systému lze však vhodným algoritmem vybrat tzv. nejlepší bázi, která v závislosti na zvoleném kritériu výběru nejlépe vystihuje chování signálu.

Mezi nejvýznamnější aplikace waveletových paketů můžeme zařadit [20] algoritmus komprese otisků prstů, který vyvinul M. Wickerhauser pro FBI, možnost restaurace např. poškozených hudebních nahrávek, nebo zlepšení kvality separace šumu a signálů, kterým se v této práci zabýváme.

5.1 Waveletové pakety jako integrální transformace

Popišme nyní stručně konstrukci waveletového paketu. Předpokládejme, že existuje MR-analýza v $L^2(\mathbb{R})$, tedy máme funkce ϕ a ψ , pro které platí rovnice (4.16) a (4.17). Definujme nyní pomocí vektorů \mathbf{h} a \mathbf{g} posloupnost funkcí rekurzivně takto:

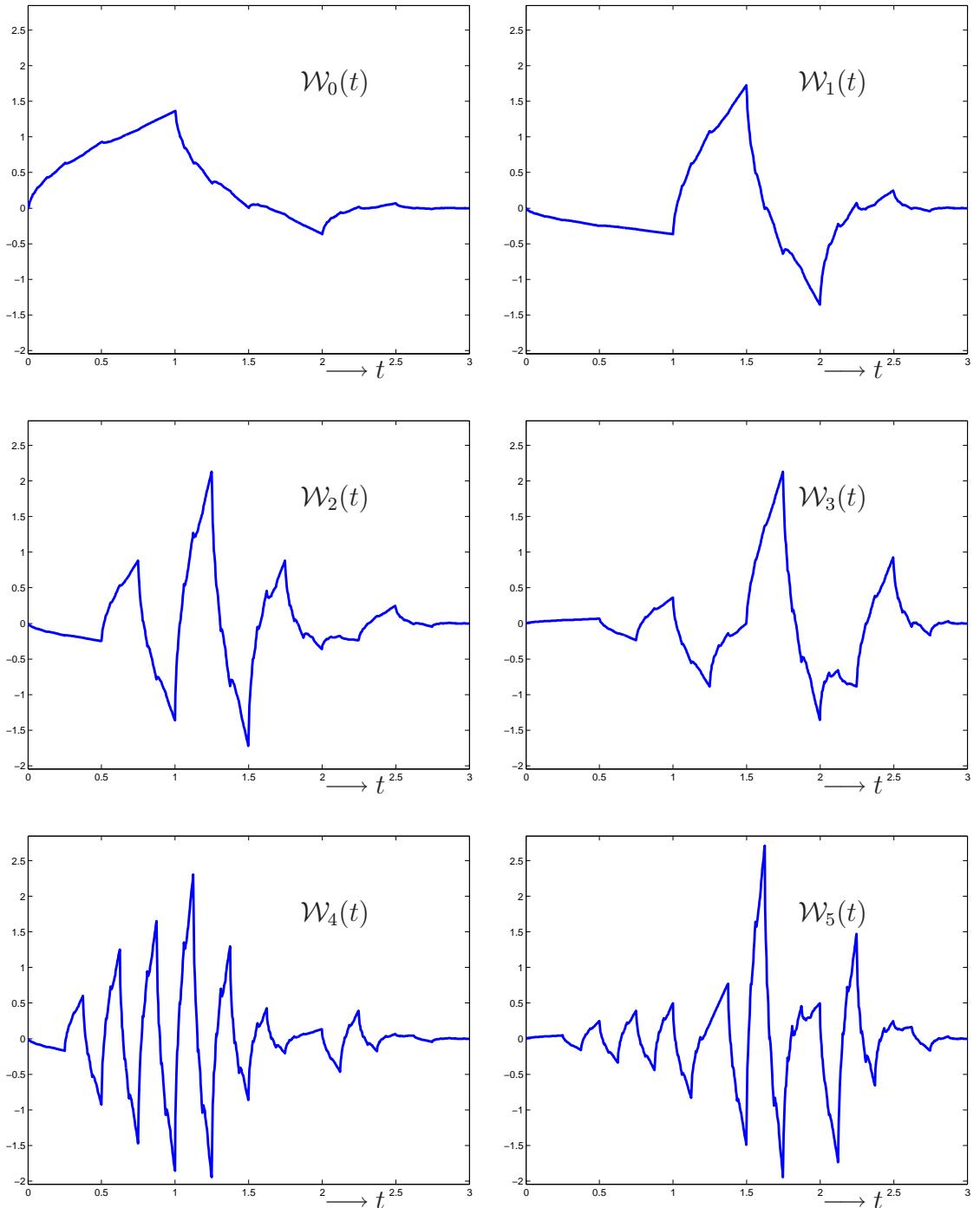
$$\mathcal{W}_{2n}(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} h_k \mathcal{W}_n(2t - k), \quad (5.1)$$

$$\mathcal{W}_{2n+1}(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} g_k \mathcal{W}_n(2t - k), \quad n = 0, 1, 2 \dots, \quad (5.2)$$

přičemž $\mathcal{W}_0 = \phi$ a $\mathcal{W}_1 = \psi$. Všimněme si v prvé řadě, že rovnice (4.16), (4.17) tak zůstaly v platnosti.

Na obrázku 5.1 jsou zobrazeny funkce $\mathcal{W}_0, \dots, \mathcal{W}_5$ vygenerované z waveletu Daubechies 2.

Poznámka: Nosič všech funkcí \mathcal{W}_n je obsažen v nosiči měřítkové funkce $\phi = \mathcal{W}_0$. Například u waveletů Daubechies třídy N je to interval $[0, 2N - 1]$ (viz obr. 5.1).



Obr. 5.1: Průběh funkcí $\mathcal{W}_0, \dots, \mathcal{W}_5$ pro wavelet Daubechies 2.

Waveletový paket je množina všech funkcí

$$\mathcal{W}_{j,n,k}(t) = 2^{-j/2} \mathcal{W}_n(2^{-j}t - k), \quad (j, n, k) \in \mathbb{Z} \times \mathbb{N} \times \mathbb{Z}. \quad (5.3)$$

Z každého paketu lze vybrat ortonormální bázi prostoru $L^2(\mathbb{R})$, a tím se zřejmě zbavit redundantní informace. Označíme-li bázi $\{\mathcal{W}_{j,n,k}\}_{j \in J, n \in N, k \in K}$ pro vhodné množiny indexů J, N, K , pak (integrální) diskrétní transformaci funkce f ve smyslu waveletových paketů můžeme psát

$$WP_f(j, n, k) = \int_{-\infty}^{\infty} f(t) \overline{\mathcal{W}_{j,n,k}(t)} dt \quad \text{pro } j \in J, n \in N, k \in K. \quad (5.4)$$

Poznámka: Lze ukázat, že např. systém $\{\mathcal{W}_{0,n,k}\}_{n \in \mathbb{N}, k \in \mathbb{Z}}$ je ortonormální bází $L^2(\mathbb{R})$. Důkaz viz [48, str. 137].

5.2 Waveletové pakety

jako konečná diskrétní transformace

Bylo již ukázáno výše, že při práci s diskrétním signálem odpovídá jeden krok waveletové transformace filtraci horní a dolní propustí a decimaci (viz obrázky 4.4, 4.5). V dalším kroku transformace se stejně filtraci podrobí tzv. approximační koeficienty. To odpovídá tomu, že v každém následujícím kroku pracujeme pouze s dolní polovinou kmitočtového pásma oproti předchozímu kroku.

Transformace typu wavelet packets se od klasické waveletové transformace liší tím, že rekurzivní filtraci nepodrobujeme pouze approximační koeficienty, nýbrž filtrujieme i detailní koeficienty předchozí úrovně. Tím prudce narůstá počet nových možných reprezentací původního signálu ve waveletové oblasti.

Zvolíme-li maximální hloubku dekompozice d , pak každý z možných rozkladů signálu je jednoznačně reprezentován jedním binárním podstromem tzv. úplného binárního stromu hloubky d . Tento strom nazýváme waveletovým paketem. Každému uzlu tohoto úplného stromu přísluší jeden vektor waveletových koeficientů.¹ Úplný binární strom hloubky d má

¹které můžeme i nadále považovat za detailní, resp. approximační, v závislosti na tom, zda na cestě do příslušného uzlu byla poslední v řadě filtrací horní propust g, resp. dolní propust h

$2^{d+1} - 1$ uzlů, které budeme značit $\mathbf{c}^{(0)}, \mathbf{c}^{(1)}, \dots, \mathbf{c}^{(2^{d+1}-2)}$. Signál je například možné vždy rekonstruovat z koeficientů obsažených v listech úplného binárního stromu.

Waveletový paket pro hloubku dekompozice $d = 3$ je vidět na obr. 5.2.

Je zřejmé, že konečná diskrétní waveletová transformace je speciálním případem transformace paketového typu, neboť waveletový strom je podstromem waveletového paketu. Také paketovou transformaci lze vyjádřit pomocí ortogonální nebo téměř ortogonální matic, jak plyne z její konstrukce [48, kap. 4].

5.3 Hledání nejlepší báze pro diskrétní signály

Podstromů úplného binárního stromu je však mnoho, proto vyvstává otázka, který z nich je „nejlepším“ reprezentantem signálu ve waveletové oblasti. Přirozeně se nabízí možnost každý podstrom ohodnotit (ocenit) podle nějakého kritéria a vybrat ten s nejlepším ohodnocením.

V praxi se používají zejména následující kritéria [48, kap. 5], která minimalizujeme. Vektor $\mathbf{y} = (y_1, \dots, y_n)^\top$ obsahuje waveletové koeficienty vstupního signálu v pevné bázi, zvolené z waveletového paketu.

- Shannonova entropie

$$\mathcal{C}(\mathbf{y}) = - \sum_{i=1}^n \frac{|y_i|^2}{\|\mathbf{y}\|^2} \log \frac{|y_i|^2}{\|\mathbf{y}\|^2} \quad (5.5)$$

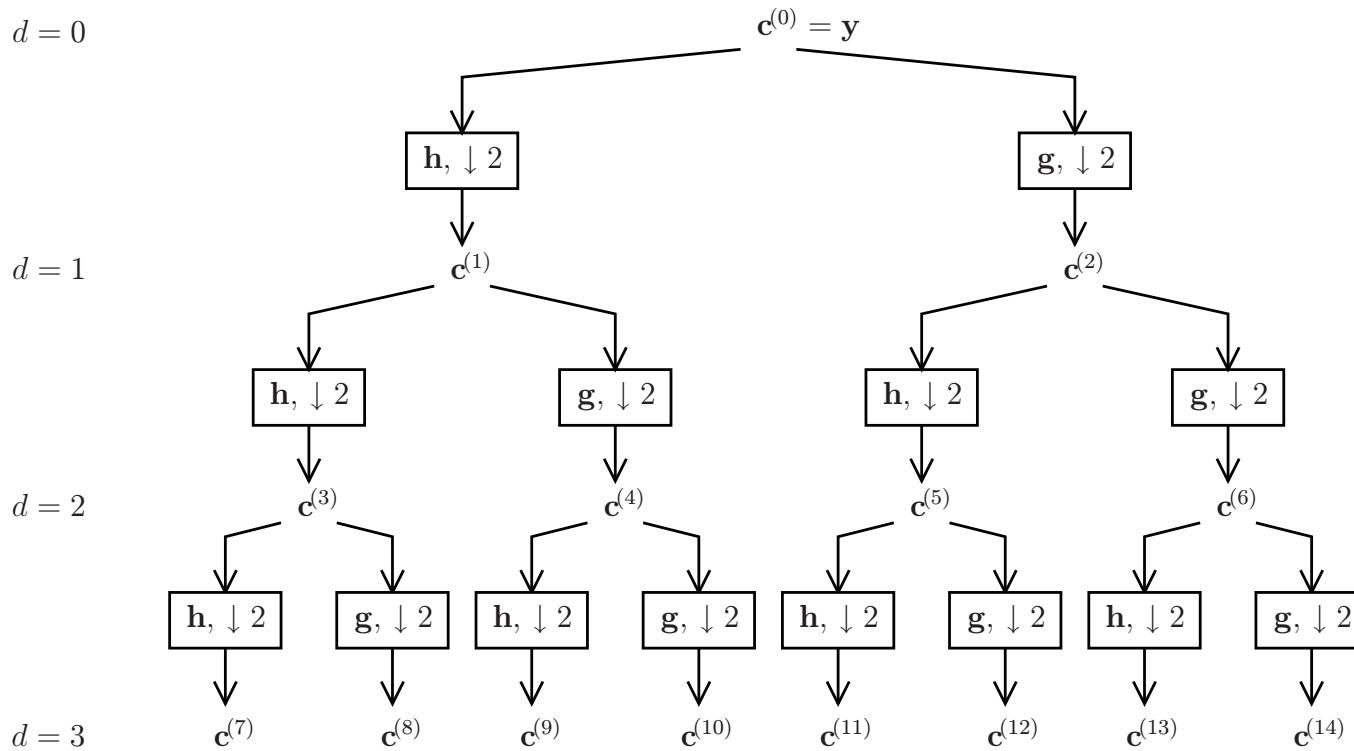
nebo její pro výpočet vhodnější ekvivalent z hlediska minimalizace

$$\tilde{\mathcal{C}}(\mathbf{y}) = - \sum_{i=1}^n |y_i|^2 \log |y_i|^2. \quad (5.6)$$

- Ohodnocení pomocí prahu – pro zvolenou hodnotu prahu $\lambda > 0$ je ohodnocení $\mathcal{C}(\mathbf{y})$ rovno počtu prvků \mathbf{y} , pro které platí $|y_i| \geq \lambda$.
- SURE² ohodnocení

$$\mathcal{C}(\mathbf{y}) = \sigma^2 \left[n - 2 \cdot \sum_{i \in \{j \mid |y_j| > \sigma\lambda\}} 1 + \sum_{i=1}^n \min \left\{ \frac{y_i^2}{\sigma^2}, \lambda^2 \right\} \right], \quad (5.7)$$

²Stein Unbiased Risk Estimator, viz [16]



Obr. 5.2: Rozklad signálu y do úplného binárního stromu (waveletového paketu) hloubky $d = 3$. Vektory $c^{(1)}, \dots, c^{(14)}$ obsahují waveletové koeficienty. Po straně je označena hloubka dekompozice.

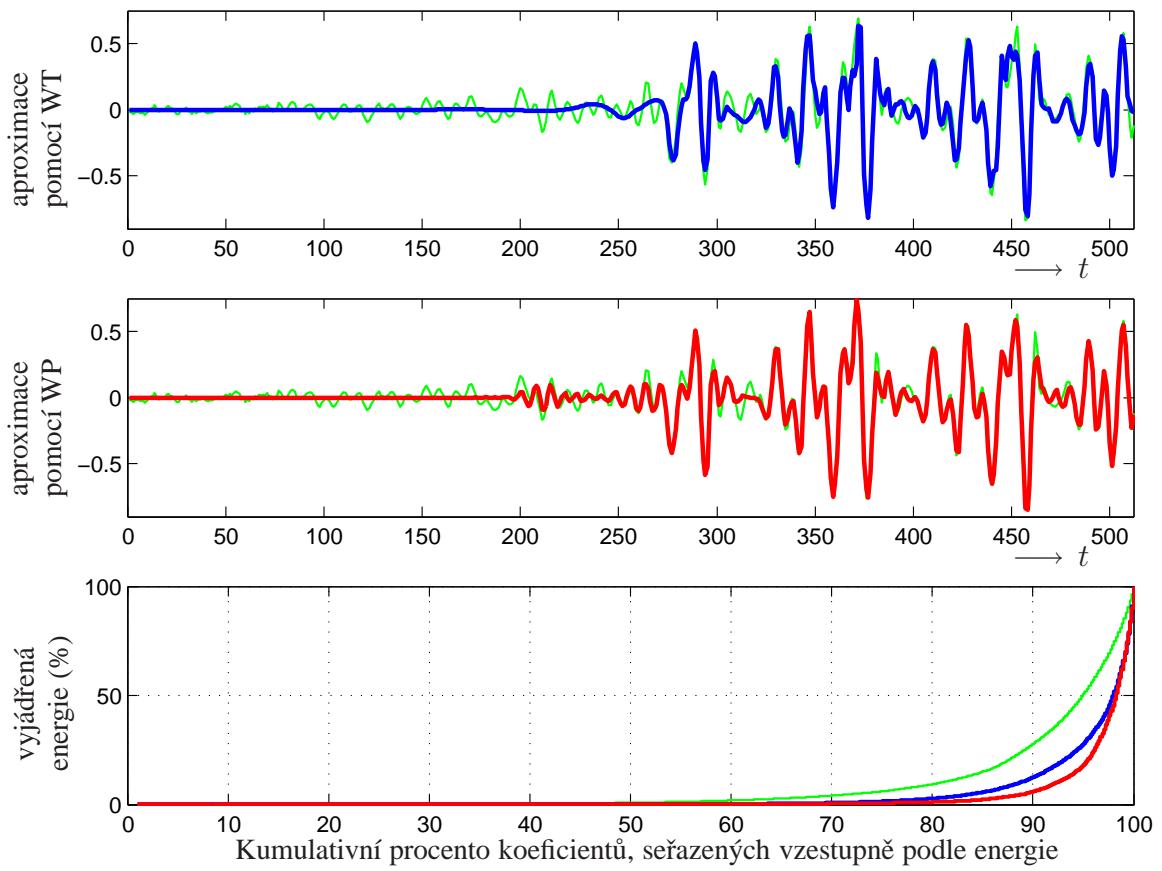
kde σ^2 je rozptyl aditivního šumu s normálním rozložením. Toto ohodnocení je vhodné pro účely redukce šumu. Donoho a Johnstone [18] nalezli v jistém smyslu optimální volbu prahu pro toto ohodnocení, $\lambda = \sqrt{2 \ln(n \log_2 n)}$.

Nalézt nevhodnější binární podstrom pro daný signál a dané ohodnocení obecně znamená ohodnotit všechny podstromy, které tvoří bázi, tzn. pro každý z nich vypočítat veličinu $\mathcal{C}(\mathbf{y})$, a pak vybrat podstrom s minimálním ohodnocením. V případě, že kritérium má navíc vlastnost aditivity, lze použít rychlý algoritmus hledání nejlepší báze (Best Basis Algorithm, [10, 18]). Výše uvedená ohodnocení vlastnost aditivity mají.

Na obrázku 5.3 je uvedeno srovnání klasické waveletové transformace a transformace nejlepší bází, ze kterého je zřejmá vyšší účinnost vystižení signálu druhým způsobem. Akustický signál o délce 512 vzorků byl rozložen těmito dvěma způsoby do hloubky 6 pomocí waveletu Daubechies 5. V obou případech bylo pro approximaci použito pouze 10 % koeficientů s nejvyšší energií. V případě waveletové transformace (WT, nahoře) bylo těmito koeficienty vyjádřeno 87 % energie originálního signálu, v případě nejlepší paketové báze (WP, uprostřed) téměř 94 % energie. Dolní graf ukazuje, jak v obou případech závisí procento zachované energie na procentu zachovaných koeficientů. Je zřejmé, že transformace paketového typu koncentruje většinu energie signálu do ještě menšího počtu koeficientů než obvyklá transformace. V tomto příkladu bylo použito ohodnocení binárního stromu pomocí prahu, hloubka waveletového stromu $d = 6$, nejlepší báze vybraná z paketu má hloubku pouze 4. Pro doplnění, 10 % vzorků originálního signálu s nejvyšší energií obsahuje necelých 72 % energie celého signálu.

5.4 Kmitočtové vlastnosti waveletových paketů

V transformaci „wavelet packet“ používáme stejné waveletové filtry jako u „obyčejné“ waveletové transformace, popsané v kapitole 4. Rozdíl je v tom, že v transformaci paketového typu máme na výběr z množství binárních podstromů, z nichž pouze jediný reprezentuje waveletovou transformaci v obvyklém smyslu. Každý bázový binární strom zřejmě reprezentuje jedno dělení osy kmitočtů, přičemž nejmenší délka jednotlivých dělicích intervalů je $\pi/2^d$ rad. Již nemusíme spektrum dělit po oktavéch, ale je možné dělit libovolnou část



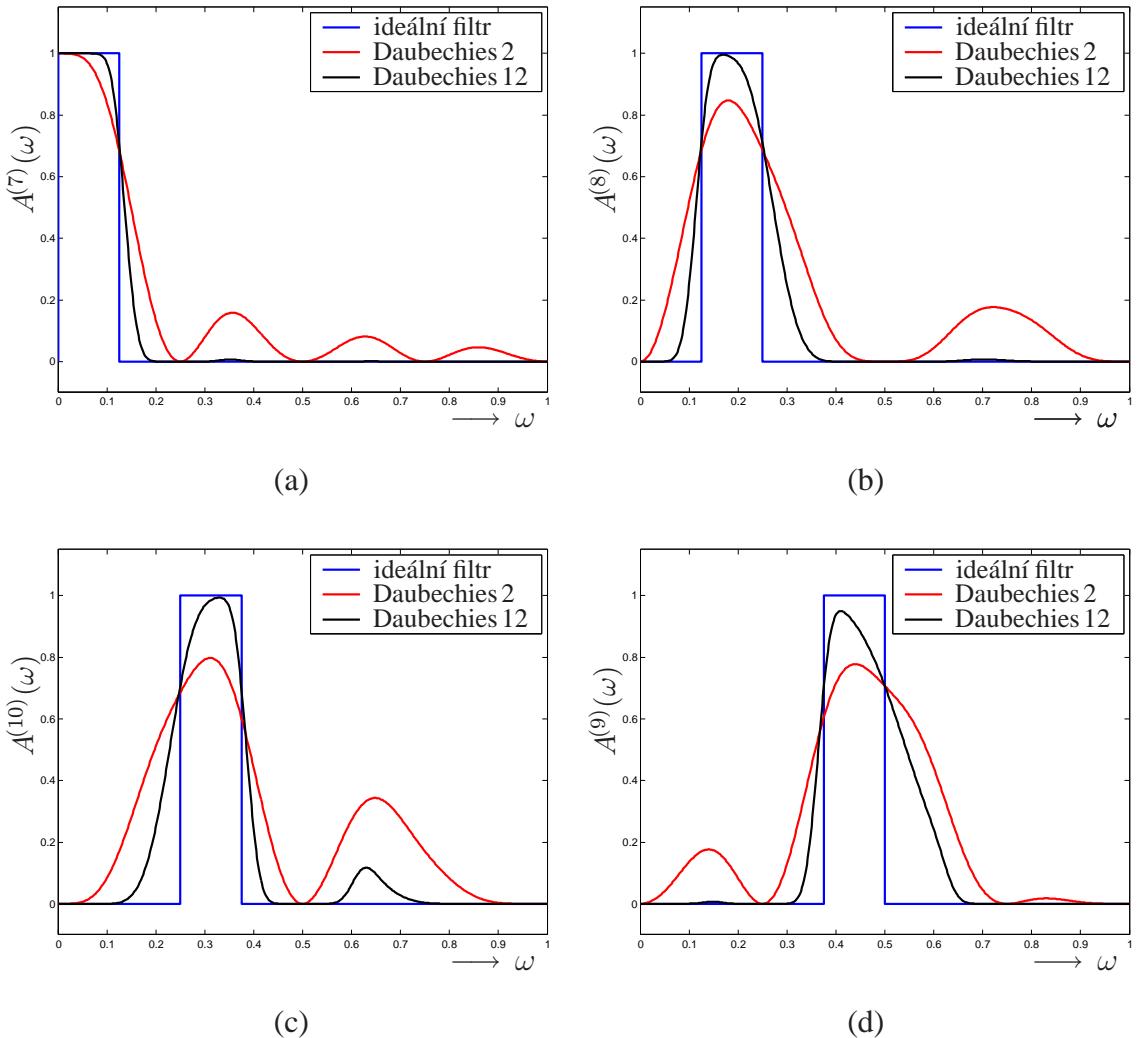
Obr. 5.3: Srovnání účinnosti vyjádření signálu pomocí waveletové transformace (WT) a nejlepší báze z waveletového paketu (WP). V obou případech bylo využito pouze 10 % koeficientů s nejvyšší energií.

kmitočtového spektra.

Na obrázku 5.4 jsou vidět amplitudové kmitočtové charakteristiky čtyř uzelů z waveletového paketu o hloubce $d = 3$ (tentotéž paket je na obrázku 5.2).

U waveletových paketů se objevuje artefakt, se kterým se u klasické waveletové transformace nesetkáme, a který je vhodné krátce zmínit. Vlivem decimace v každém kroku transformace dojde zejména k převrácení orientace kmitočtového obsahu signálu po průchodu větví horní propusti. To má za následek, že hned v následujícím kroku větev dolní propusti ponechává horní polovinu spektra a větev horní propusti naopak ponechává dolní polovinu spektra. To vede v důsledku k vlastnosti, že uzel s číslem o jedna vyšším nemusí nutně reprezentovat přilehlý úsek kmitočtové osy. Tomuto uspořádání uzelů, které je použito na obrázku 5.2, říkáme *přirozené*. Jiný typ číslování, kdy následující uzel ve stejně úrovni dekompozice vždy odpovídá následujícímu úseku na ose kmitočtů, se nazývá *kmitočtové*. Oba typy číslování je možné mezi sebou jednoznačně převádět. Pro účely této práce však rozdílné značení waveletových paketů není důležité. Pro podrobný výklad problematiky odkazujeme na [33, kap. 6].

Pro paketový typ transformace zůstává v platnosti vše, co bylo řečeno o aliasingu v části 4.3.3.



Obr. 5.4: Grafy amplitudových kmitočtových odezv vybraných uzlů waveletového paketu při dekompozici. Obrázek (a) ukazuje amplitudovou odezvu uzlu $c^{(7)}$, (b) uzlu $c^{(8)}$, (c) uzlu $c^{(10)}$, (d) uzlu $c^{(9)}$. Příslušné amplitudové charakteristiky (srv. obr. 4.7) jsou $A^{(7)}(\omega) = |H(e^{4i\omega}) \cdot H(e^{2i\omega}) \cdot H(e^{i\omega})|$, $A^{(8)}(\omega) = |G(e^{4i\omega}) \cdot H(e^{2i\omega}) \cdot H(e^{i\omega})|$, $A^{(10)}(\omega) = |H(e^{4i\omega}) \cdot G(e^{2i\omega}) \cdot H(e^{i\omega})|$, $A^{(9)}(\omega) = |G(e^{4i\omega}) \cdot G(e^{2i\omega}) \cdot H(e^{i\omega})|$. Všimněme si, že uzel $c^{(7)}$ odpovídá uzlu $a^{(3)}$ a podobně uzel $c^{(8)}$ odpovídá uzlu $d^{(3)}$ ve stromu waveletové transformace. Zbylé dva uzly, $c^{(10)}$ a $c^{(9)}$ již v klasické waveletové transformaci nejsou obsaženy. V části 5.4 je vysvětleno, proč uzel s číslem 10 odpovídá oproti očekávání nižším kmitočtům než uzel s číslem 9.

6 Waveletové potlačování šumu

Tato kapitola tvoří těžiště dizertační práce. Nejprve je v části 6.1 detailně rozebrán princip, na němž staví waveletové metody potlačování šumu. Nejpoužívanějším typem modifikace waveletového spektra za tímto účelem je tzv. prahování. Typy prahovacích pravidel jsou uvedeny dále v části 6.2. Část 6.3 pak navazuje statistickou analýzou těchto pravidel a jejich kvalitativním srovnáním.

6.1 Princip waveletových metod potlačování šumu

Nežádoucí šum v signálu má obvykle aditivní charakter a proto je vhodné uvažovat model

$$\mathbf{X} = \mathbf{f} + \boldsymbol{\varepsilon}, \quad (6.1)$$

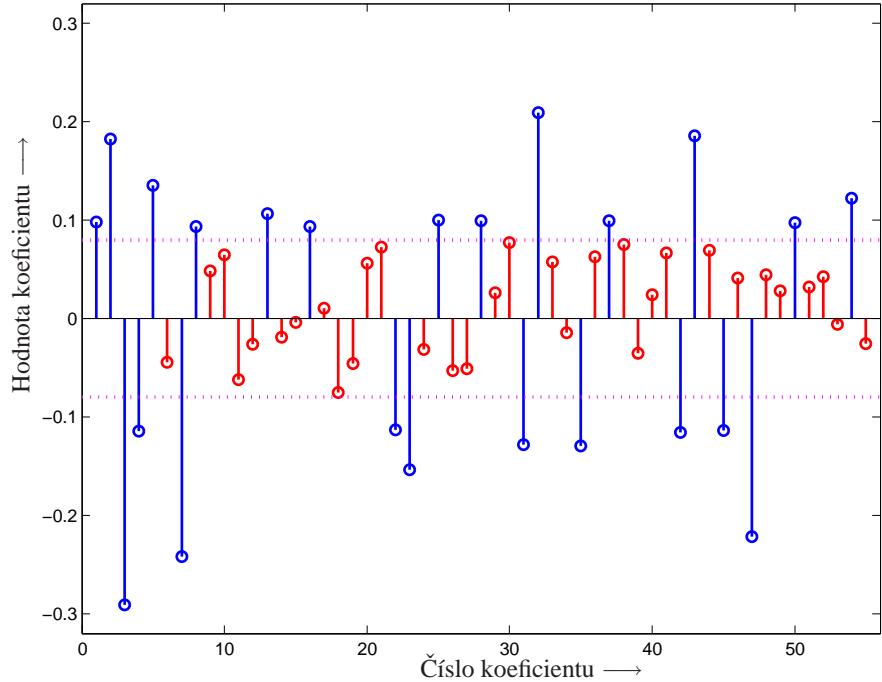
kde $\mathbf{X} = (X_1, \dots, X_n)^\top$ je náhodný vektor (konečný diskrétní náhodný signál), $\mathbf{f} = (f_1, \dots, f_n)^\top = (f(t_1), \dots, f(t_n))^\top$ pro nějakou funkci f a ekvidistantně vzdálené body t_1, \dots, t_n (deterministický, analytický signál), a $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^\top$ je náhodný vektor s n -rozměrným normálním rozložením, $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$.

Waveletovou transformací vyjádřenou pomocí matice \mathbf{W} přejde vztah (6.1) s využitím linearity v

$$\mathbf{Y} = \mathbf{b} + \boldsymbol{\nu}, \quad (6.2)$$

kde $\mathbf{Y} = \mathbf{W}\mathbf{X}$, $\mathbf{b} = \mathbf{W}\mathbf{f}$, $\boldsymbol{\nu} = \mathbf{W}\boldsymbol{\varepsilon}$. Díky ortogonalitě transformace platí $\boldsymbol{\nu} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, tedy rozložení pravděpodobnosti šumu zůstalo stejné [3]. Označme pozorování náhodných vektorů $\mathbf{X}, \boldsymbol{\varepsilon}, \mathbf{Y}, \boldsymbol{\nu}$, pořadě symboly $\mathbf{x}, \mathbf{e}, \mathbf{y}, \mathbf{v}$. Vektory $\mathbf{y} = (y_1, \dots, y_n)^\top$, $\mathbf{b} = (b_1, \dots, b_n)^\top$, $\mathbf{v} = (v_1, \dots, v_n)^\top$ tedy představují waveletové koeficienty příslušné vektorům $\mathbf{x}, \mathbf{f}, \mathbf{e}$.

Cílem waveletového potlačování šumu je nalézt vhodný modifikační předpis $\delta(\cdot)$ tak, aby odhad $\hat{b}_i = \delta(y_i) \approx b_i$ byl v nějakém smyslu dobrým odhadem b_i . Koeficienty b_i totiž jsou waveletové koeficienty příslušící signálu *nezatíženého šumem*. Po úpravě waveletových koeficientů navrátíme pomocí matice \mathbf{W}^\top signál do výchozí domény. Tím získáme



Obr. 6.1: Princip prahování. Hodnoty pod prahovou úrovni (červená barva) jsou vynulovány a zbylé (modrá barva) se dále zpracovávají.

signál s potlačenou šumovou složkou. Tuto v principu jednoduchou metodologii nazvali její autoři Donoho a Johnstone [17] *WaveShrink*.

Popsaný přístup je analogií klasické fourierovské filtrace, kde podobně upravujeme Fourierovy koeficienty signálu. Avšak u waveletové analýzy je příspěvek každého koeficientu pouze lokální (neboť wavelet je tlumený směrem k $\pm\infty$, případně má dokonce kompaktní nosič), takže waveletová reprezentace dovoluje tímto způsobem konstruovat *lokálně adaptivní filtry*, což je vynikající rys ve srovnání se zmíněnými fourierovskými filtry, kde efekt každého koeficientu je obecně globální [45].

Byla vyvinuta řada technik, jak konstruovat odhady $\hat{b}_i = \delta(y_i)$. Mezi nejpoužívanější metody patří tzv. wavelet shrinkage [48, kap. 6]. Speciálním typem těchto technik jsou tzv. *prahovací techniky*. Jejich princip je následující: hodnoty y_i , které jsou v absolutní hodnotě menší než zvolený práh $\lambda > 0$, jsou vynulovány a ostatní jsou buď ponechány nebo ještě dále upraveny. Z toho je zřejmé, že prahování je nelineární operací na datech. Prahování lze formálně pojmenovat jako testování hypotézy, že jednotlivý waveletový koeficient reprezentuje deterministický signál oproti alternativě, že reprezentuje aditivní šum [48]. Princip

prahování je znázorněn na obr. 6.1.

Waveletová transformace má tu vynikající vlastnost, že koncentruje podstatnou část energie obsažené ve vektoru \mathbf{f} do pouze malého počtu waveletových koeficientů b_i (viz obr. 5.3 dole). Tento fakt spolu s vlastností $\nu \sim N_n(\mathbf{0}, \sigma^2 I_n)$ implikuje a vysvětluje princip waveletového potlačování šumu. Při vhodně zvolené hodnotě prahu λ se totiž prahování (vynulování) nedotkne koeficientů y_i , které reprezentují především deterministický signál, ale naopak zasáhne „malé“ koeficienty y_i reprezentující aditivní šum.

Prahování podrobujeme pouze detailní koeficienty, neboť approximační koeficienty vyjadřují trend v signálu, a ten není vhodné prahováním modifikovat. Jednotlivé úrovně detailních koeficientů lze navíc prahovat s navzájem různými hodnotami prahu.

Podle způsobu úprav koeficientů y_i rozlišujeme různé druhy prahování (viz níže).

Poznámka: Pro transformaci paketového typu platí vše, co bylo uvedeno výše v tomto oddílu, neboť tato transformace je rovněž reprezentována ortogonální maticí \mathbf{W} . Důvod pro možné použití waveletového paketu, přesněji jeho nejlepší báze, je ten, že může dojít k ještě větší koncentraci energie celého signálu do několika málo waveletových koeficientů než u klasické waveletové transformace (viz příklad na obr. 5.3). Tak může dojít ještě ke zvýšení účinnosti separace signálu a šumu. Praktické výsledky viz kapitolu 7.

6.2 Typy prahovacích pravidel

V této části je vysvětleno, jakým způsobem jednotlivá prahovací pravidla zacházejí s waveletovými koeficienty:

1. Tvrdé prahování

Všechny waveletové koeficienty y_i pod prahovou hodnotou $\lambda > 0$ jsou vynulovány a ostatní jsou ponechány beze změny: $\hat{b}_i = \delta^h(y_i, \lambda)$, kde

$$\delta^h(x, \lambda) = \begin{cases} 0 & \text{pro } |x| \leq \lambda \\ x & \text{pro } |x| > \lambda. \end{cases}$$

Viz obrázek 6.2(a).

2. Měkké prahování

Waveletové koeficienty y_i pod prahovou hodnotou λ jsou vynulovány a velikost

ostatních se o tuto prahovou hodnotu sníží: $\hat{b}_i = \delta^s(y_i, \lambda)$, kde

$$\delta^s(x, \lambda) = \operatorname{sgn}(x) \max\{0, |x| - \lambda\}.$$

Viz obrázek 6.2(b).

3. Poloměkké prahování

Toto pravidlo závisí na dvou parametrech $0 < \lambda_1 < \lambda_2$; $\hat{b}_i = \delta^{ss}(y_i, \lambda_1, \lambda_2)$, kde

$$\delta^{ss}(x, \lambda_1, \lambda_2) = \begin{cases} 0 & \text{pro } |x| \leq \lambda_1 \\ \operatorname{sgn}(x) \frac{\lambda_2(|x| - \lambda_1)}{\lambda_2 - \lambda_1} & \text{pro } \lambda_1 < |x| \leq \lambda_2. \\ x & \text{pro } |x| > \lambda_2. \end{cases}$$

Viz obrázek 6.2(c).

Toto pravidlo je zobecněním bodů 1. a 2.; skutečně, neboť pro $\lambda_2 \rightarrow \infty$ přechází $\delta^{ss}(x, \lambda_1, \lambda_2)$ v $\delta^s(x, \lambda_1)$ a pro $\lambda_2 \rightarrow \lambda_1$ přechází $\delta^{ss}(x, \lambda_1, \lambda_2)$ v $\delta^h(x, \lambda_1)$ (v bodovém smyslu pro x).

4. Nezáporná garota¹ (Non-negative garotte)

$\hat{b}_i = \delta^{\text{nng}}(y_i, \lambda)$, kde

$$\delta^{\text{nng}}(x, \lambda) = \begin{cases} 0 & \text{pro } |x| \leq \lambda \\ x - \frac{\lambda^2}{x} & \text{pro } |x| > \lambda. \end{cases}$$

Viz obrázek 6.2(d).

5. Hyperbolické prahování

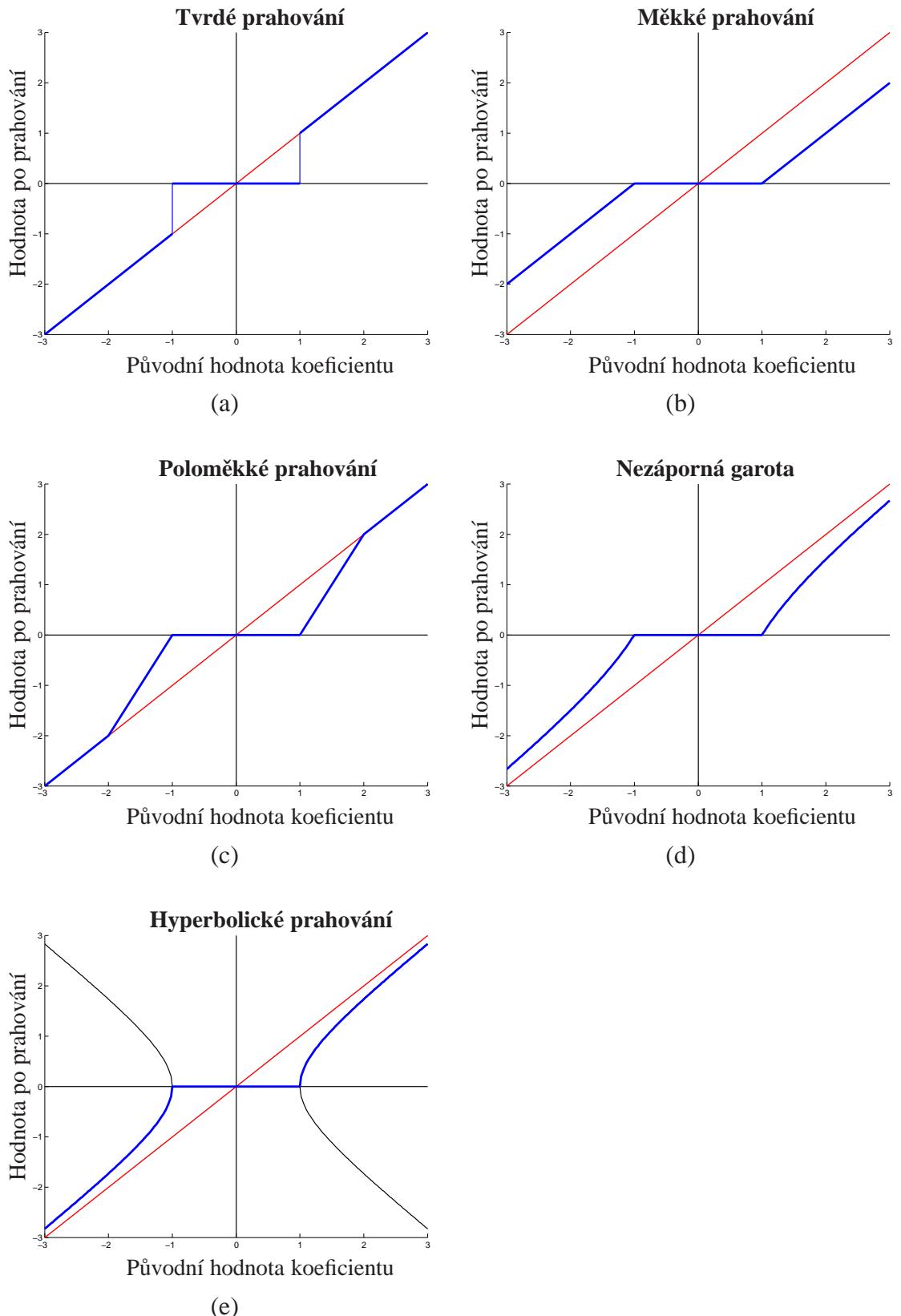
Hyperbolické prahovací pravidlo lze odvodit měkkým prahováním energie waveletových koeficientů y_i^2 . Pokud $y_i \geq \lambda$ (tj. $y_i^2 \geq \lambda^2$), odečteme od y_i^2 prahovou energii λ^2 . Formálně tedy $\hat{b}_i = \delta^{\text{hy}}(y_i, \lambda)$, kde

$$\delta^{\text{hy}}(x, \lambda) = \begin{cases} 0 & \text{pro } |x| < \lambda \\ \operatorname{sgn}(x) \sqrt{x^2 - \lambda^2} & \text{pro } |x| \geq \lambda. \end{cases}$$

Viz obrázek 6.2(e).

Kromě uvedených metod prahování je možné použít i další sofistikované metody založené na bayesovském statistickém přístupu. Této problematice je věnována kapitola 8 v [48], dále např. [49].

¹garota – železné škrtidlo



Obr. 6.2: Grafy prahovacích pravidel. Pro všechna prahovací pravidla zvoleno $\lambda = 1$, kromě poloměkkého prahování, kde $\lambda_1 = 1$ a $\lambda_2 = 2$.

6.3 Statistická analýza prahovacích pravidel

Matematická statistika nabízí pro naše účely dva způsoby, jak posuzovat vlastnosti odhadů. Prvním z nich je tzv. statistická analýza rizikové funkce (Exact Risk Analysis) a druhým je tzv. „minimax“ přístup. Tento text se věnuje prvnímu přístupu, který je pro naše účely názornější.

Poznámka: Připomeňme na tomto místě, že hodnota prahovacího pravidla δ je pozorování (transformované) náhodné veličiny.

Vyjdeme z modelu (6.2) a uvažujme *jedinou* náhodnou veličinu Y_i , kterou pro jednoduchost budeme dále značit pouze Y . Bez újmy na obecnosti (viz tvrzení 6.7) nadále předpokládejme jednotkový rozptyl $\sigma^2 = 1$, tedy $Y \sim N(b_i, 1)$. Pokud kontext nebude vyžadovat jinak, nebudeme rovněž dále psát $\delta(Y, \lambda)$, nýbrž pouze $\delta(Y)$.

Posoudíme prahovací pravidla $\delta^h, \delta^s, \delta^{ss}, \delta^{nng}, \delta^{hy}$ z hlediska:

1. **střední hodnoty** $M_\lambda(b) = E[\delta(Y)]$,
2. **rozptylu** $V_\lambda(b) = \text{var} [\delta(Y)] = E[\delta(Y) - M_\lambda(b)]^2$,
3. **rizika**² $R_\lambda(b) = E[\delta(Y) - b]^2$.

pro pevně zvolenou hodnotu prahu $\lambda > 0$. Pro nás nejdůležitější informaci přitom obsahuje hodnota rizika, neboť udává odchylku od „správné“ hodnoty b , nikoliv od její střední hodnoty $M_\lambda(b)$.

Tvrzení 6.1: *Riziko lze rozložit pomocí výchylky³ a rozptylu takto:*

$$R_\lambda(b) = V_\lambda(b) + [M_\lambda(b) - b]^2. \quad (6.3)$$

Důkaz tvrzení je uveden v příloze A.

Označme nyní příslušnost výše uvedených statistik k prahovacím pravidlům horním indexem, tj. např. rozptyl měkkého prahování bude označen $V_\lambda^s(b)$.

Tvrzení 6.2: [6, 23, 24] *Pro náhodnou veličinu $Y \sim N(b, 1)$ platí:*

²někdy též L^2 -risk, L^2 -riziková funkce

³výchylku (bias) definujeme jako $[M_\lambda(b) - b]$

Tvrdé prahování:

$$\begin{aligned}
 M_\lambda^h(b) &= b + b[1 - \Phi(\lambda - b) - \Phi(\lambda + b)] + \varphi(\lambda - b) - \varphi(\lambda + b) \\
 V_\lambda^h(b) &= (b^2 + 1)[2 - \Phi(\lambda - b) - \Phi(\lambda + b)] + \\
 &\quad + (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) - M_\lambda^h(b)^2 \\
 R_\lambda^h(b) &= 1 + (b^2 - 1)[\Phi(\lambda - b) - \Phi(-\lambda - b)] + (\lambda + b)\varphi(\lambda + b) + (\lambda - b)\varphi(\lambda - b)
 \end{aligned}$$

Měkké prahování:

$$\begin{aligned}
 M_\lambda^s(b) &= M_\lambda^h(b) - \lambda[\Phi(\lambda + b) - \Phi(\lambda - b)] \\
 V_\lambda^s(b) &= V_\lambda^h(b) - \lambda[v_1(\lambda, b) + v_1(\lambda, -b)] \\
 R_\lambda^s(b) &= 1 + \lambda^2 + (b^2 - \lambda^2 - 1)[\Phi(\lambda - b) - \Phi(-\lambda - b)] - \\
 &\quad - (\lambda - b)\varphi(\lambda + b) - (\lambda + b)\varphi(\lambda - b)
 \end{aligned}$$

Poloměkké prahování:

$$\begin{aligned}
 M_{\lambda_1, \lambda_2}^{ss}(b) &= M_{\lambda_2}^h(b) + \frac{\lambda_2}{\lambda_2 - \lambda_1} \{m_1(\lambda_1, \lambda_2, b) - m_1(\lambda_1, \lambda_2, -b)\} \\
 V_{\lambda_1, \lambda_2}^{ss}(b) &= V_{\lambda_2}^h(b) - \frac{\lambda_2}{\lambda_2 - \lambda_1} \{v_2(\lambda_1, \lambda_2, b) + v_2(\lambda_1, \lambda_2, -b)\} \\
 R_{\lambda_1, \lambda_2}^{ss}(b) &= b^2[\Phi(\lambda_1 - b) - \Phi(-\lambda_1 - b)] - \Phi(\lambda_2 - b) + \Phi(-\lambda_2 - b) + \\
 &\quad + \Psi(\lambda_1 - b, \lambda_2 - b, r_2, r_1(b - \lambda_2)) + \Psi(\lambda_1 + b, \lambda_2 + b, r_2, -r_1(b + \lambda_2)) + \\
 &\quad + (\lambda_2 - b)\varphi(\lambda_2 - b) + (\lambda_2 + b)\varphi(\lambda_2 + b) + 1
 \end{aligned}$$

Nezáporná garota:

$$\begin{aligned}
 M_\lambda^{nng}(b) &= M_\lambda^h(b) - \lambda^2 A_\lambda(b) \\
 V_\lambda^{nng}(b) &= V_\lambda^h(b) + \lambda^4 [B_\lambda(b) - A_\lambda(b)^2] + 2\lambda^2 A_\lambda(b) M_\lambda^h(b) - \\
 &\quad - 2\lambda^2 [1 - \Phi(\lambda - b) + \Phi(-\lambda - b)] \\
 R_\lambda^{nng}(b) &= R_\lambda^h(b) + 2\lambda^2 b A_\lambda(b) + \lambda^4 B_\lambda(b) - 2\lambda^2 [1 - \Phi(\lambda - b) + \Phi(-\lambda - b)]
 \end{aligned}$$

Hyperbolické prahování:

$$\begin{aligned}
 M_\lambda^{hy}(b) &= \int_\lambda^\infty \sqrt{x^2 - \lambda^2} [\varphi(x - b) - \varphi(x + b)] dx \\
 V_\lambda^{hy}(b) &= (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) + \\
 &\quad + (1 + b^2 - \lambda^2)[2 - \Phi(\lambda - b) - \Phi(\lambda + b)] - [M_\lambda^{hy}(b)]^2
 \end{aligned}$$

$$\begin{aligned} R_\lambda^{\text{hy}}(b) &= (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) + \\ &\quad + (1 + b^2 - \lambda^2)[2 - \Phi(\lambda - b) - \Phi(\lambda + b)] - 2bM_\lambda^{\text{hy}}(b) + b^2, \end{aligned}$$

kde

$$m_1(\lambda_1, \lambda_2, b) = (b - \lambda_1)[\Phi(\lambda_2 - b) - \Phi(\lambda_1 - b)] - \varphi(\lambda_2 - b) + \varphi(\lambda_1 - b),$$

$$\begin{aligned} m_2(\lambda_1, \lambda_2, b) &= [1 + (b - \lambda_1)^2][\Phi(\lambda_2 - b) - \Phi(\lambda_1 - b)] - \\ &\quad - (\lambda_2 - 2\lambda_1 + b)\varphi(\lambda_2 - b) - (\lambda_1 - b)\varphi(\lambda_1 - b), \end{aligned}$$

$$v_1(\lambda, b) = [1 + \Phi(\lambda - b) - \Phi(\lambda + b)]\{(2b - \lambda)[1 - \Phi(\lambda - b)] + 2\varphi(\lambda - b)\},$$

$$\begin{aligned} v_2(\lambda_1, \lambda_2, b) &= m_1(\lambda_1, \lambda_2, b) \{2M_{\lambda_2}^h(b) + r_2[m_1(\lambda_1, \lambda_2, b) - m_1(\lambda_1, \lambda_2, -b)]\} \\ &\quad - r_2 m_2(\lambda_1, \lambda_2, b), \end{aligned}$$

$$\begin{aligned} A_\lambda(b) &= \int_\lambda^\infty \frac{\varphi(x - b) - \varphi(x + b)}{x} dx, \\ B_\lambda(b) &= \int_\lambda^\infty \frac{\varphi(x - b) + \varphi(x + b)}{x^2} dx, \end{aligned}$$

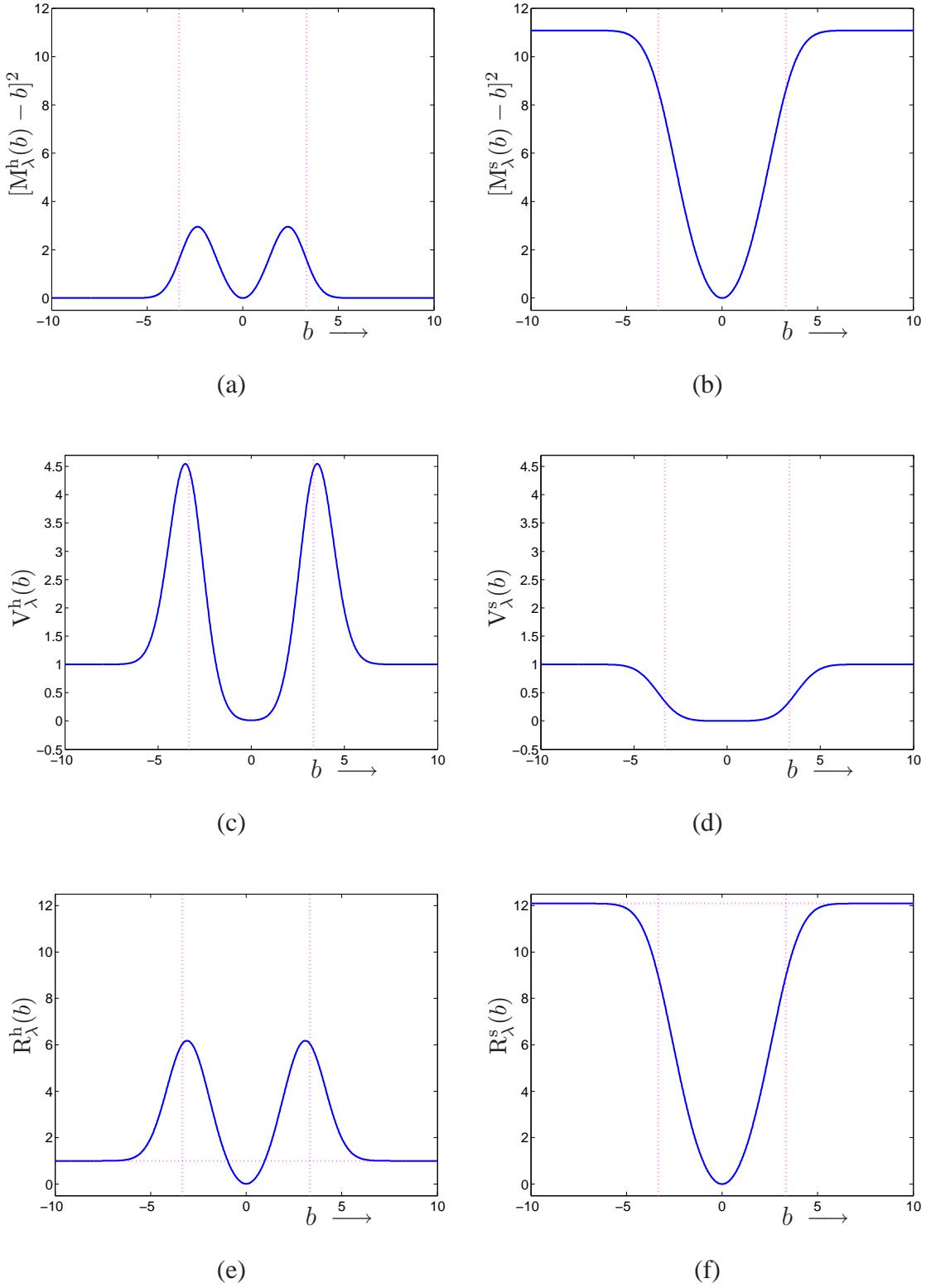
$$\Psi(a, b, c, d) = (c^2 + d^2)[\Phi(b) - \Phi(a)] + c\varphi(a)(ac + 2d) + c\varphi(b)(bc + 2d),$$

$$r_1 = \frac{\lambda_1}{\lambda_2 - \lambda_1}, \quad r_2 = \frac{\lambda_2}{\lambda_2 - \lambda_1}.$$

Důkaz je uveden v příloze A, str. 93 a odvození vzorců pro hyperbolické prahování je uvedeno zvlášť v příloze B, str. 95.

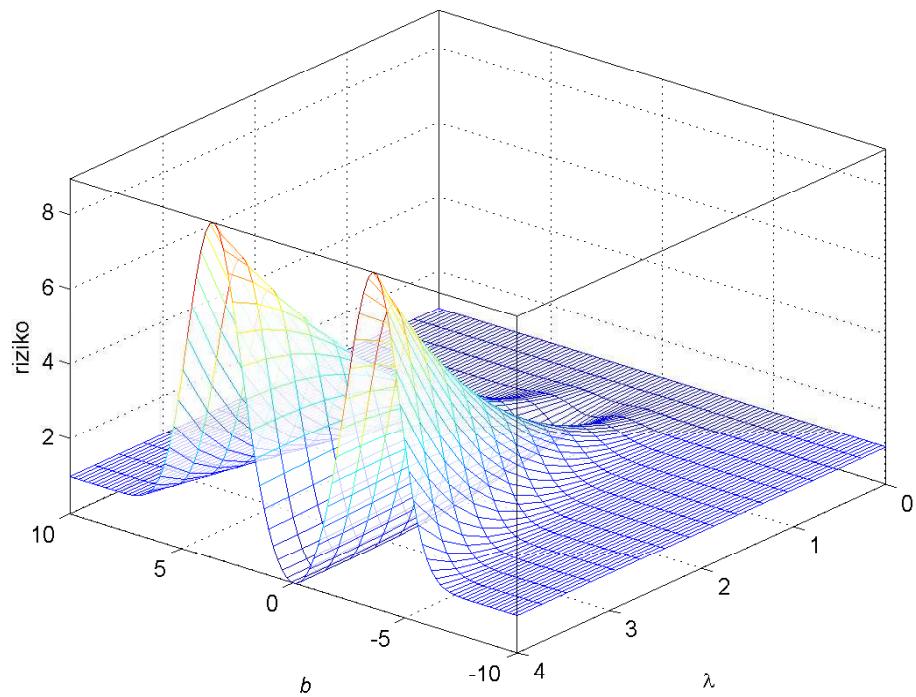
Pomocí uvedených vzorců můžeme vyčíslit hodnoty statistik v libovolném bodě $b \in \mathbb{R}$. Ve vzorcích pro nezápornou garotu vystupují integrály A_λ, B_λ . Abychom se vyhnuli vyčíslení přímo z tohoto tvaru standardními numerickými metodami, které jsou závislé na zvolené integrační síti bodů a také na zvolené konečné horní hranici namísto nekonečné, převedeme tyto integrály na konvergentní nekonečnou řadu s rekurzivními členy [23]. Výpočet hodnoty z tohoto tvaru již závisí pouze na zvolené přesnosti konvergence.

Obrázek 6.3 znázorňuje chování kvadrátu výchylky, rozptylu a rizika tvrdého a měkkého prahování pro různé hodnoty b za předpokladu pevně zvoleného prahu λ . Obrázek 6.4 ukazuje prostorově hodnoty rizika pro tvrdé a měkké prahování pro různé hodnoty λ .

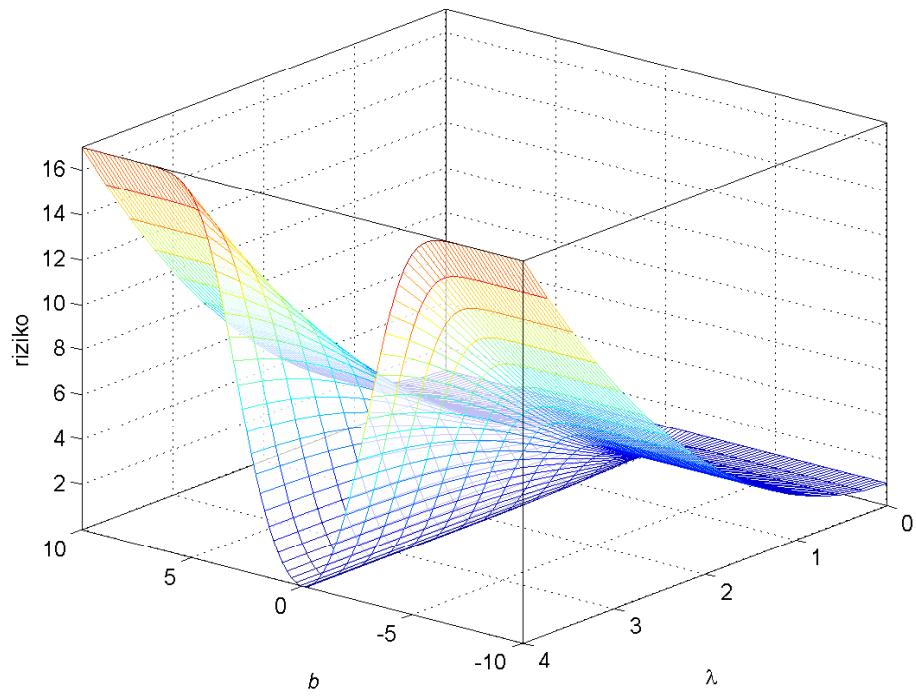


Obr. 6.3: Srovnání kvadrátu výchylky, rozptylu a rizika (od shora dolů) pro tvrdé a měkké prahování. Levý sloupec odpovídá tvrdému, zatímco pravý sloupec měkkému prahování. Vertikální linky naznačují polohu prahu, pro tyto grafy bylo zvoleno $\lambda = 3,33$. Horizontální linky u grafů rizika jsou asymptoty. Riziko tvrdého prahování se asymptoticky blíží k 1, zatímco riziko měkkého se blíží k $1 + \lambda^2$.

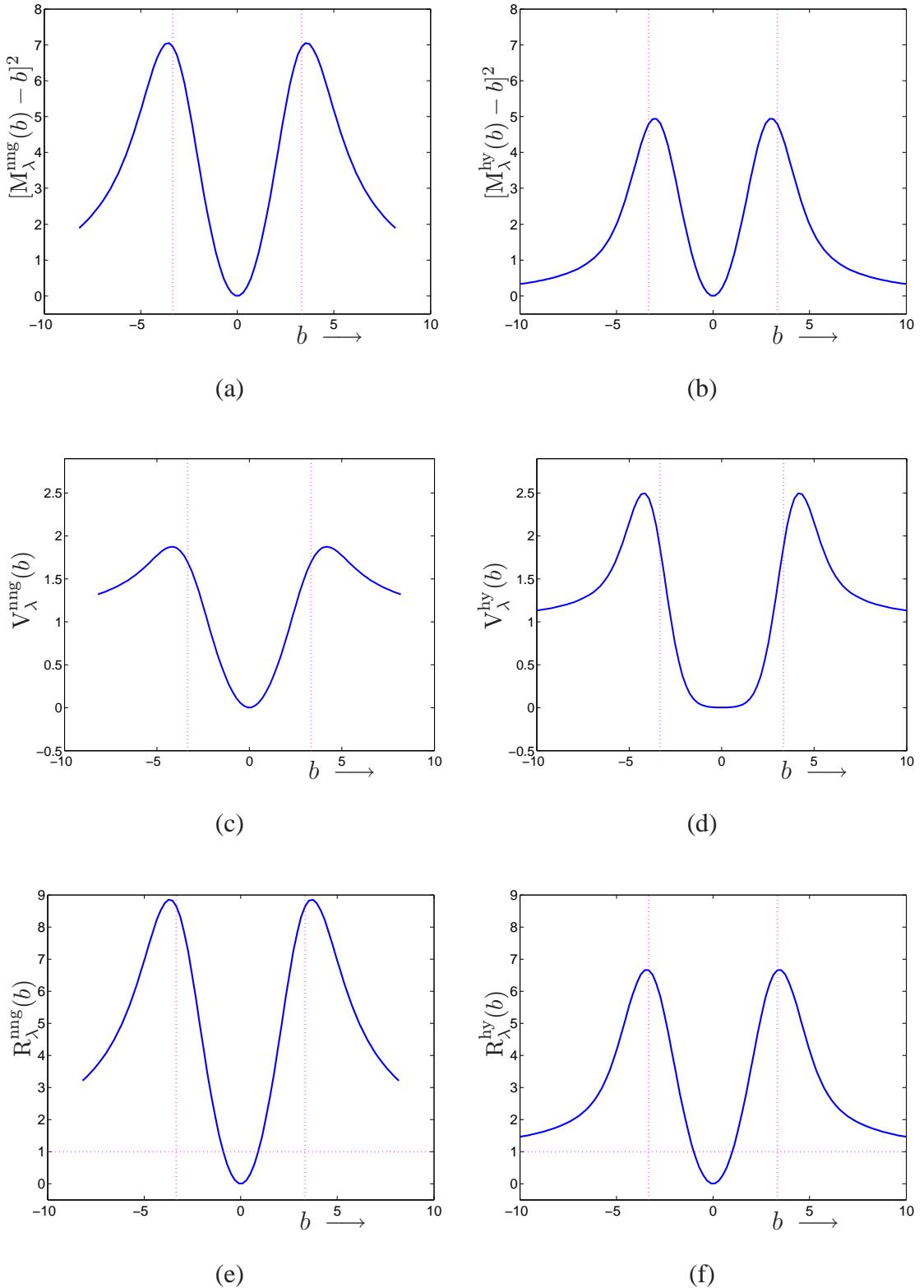
Riziko tvrdého prahovani pro ruzne hodnoty prahu



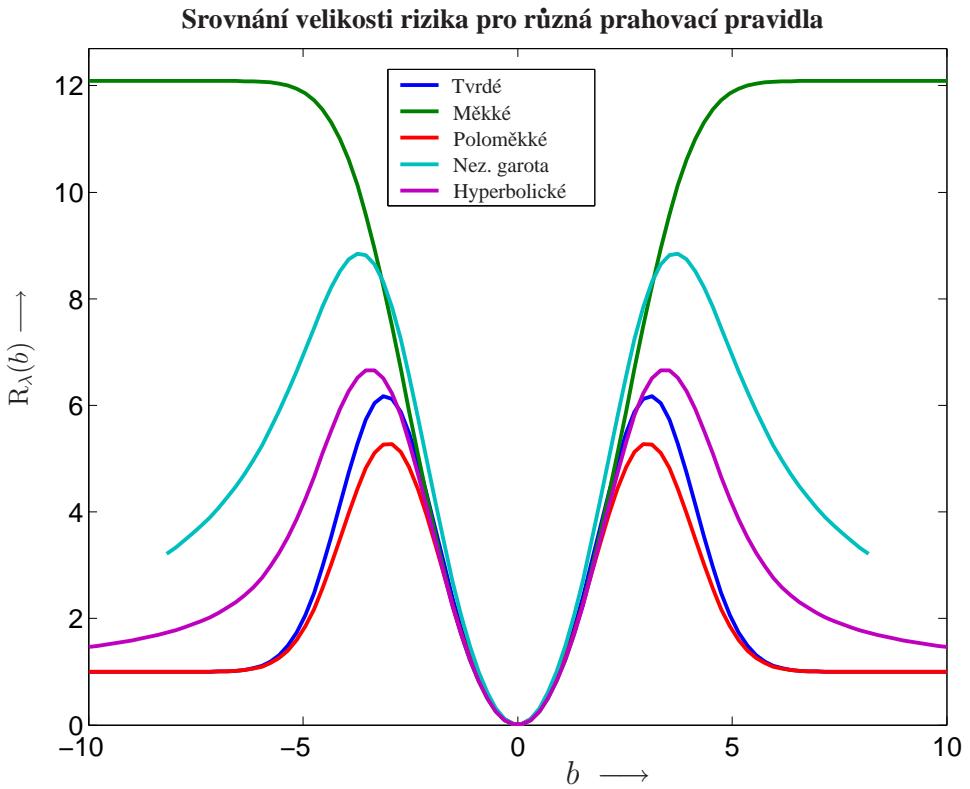
Riziko měkkého prahovani pro ruzne hodnoty prahu



Obr. 6.4: Prostorové znázornění rizika tvrdého (nahoře) a měkkého prahování (dole) pro různé hodnoty b a λ .



Obr. 6.5: Srovnání kvadrátu výchylky, rozptylu (od shora dolů) a rizika pro nezápornou garotu a hyperbolické prahování. Levý sloupec odpovídá nezáporné garotě, pravý hyperbolickému prahování. Vertikální linky naznačují polohu prahu, pro tyto grafy bylo zvoleno $\lambda = 3,33$.



Obr. 6.6: Srovnání rizika pěti prahovacích pravidel. Zde $\lambda = 3,33$, $\lambda_1 = 2,86$, $\lambda_2 = 3,72$. Je zřetelné, že stejnoměrně nejmenšího rizika dosahuje poloměkké prahování, avšak jen v případě správné volby λ_1, λ_2 , jak je uvedeno v tomto oddílu textu.

Obrázek 6.5 srovnává nezápornou garotu a hyperbolické prahování podobným způsobem jako obr. 6.3. Konečně obrázek 6.6 porovnává riziko všech pěti zmíněných pravidel.

Tvrzení 6.2 spolu s uvedenými obrázky nám umožňuje vyvodit následující důsledky [6]:

Důsledek 6.3: Pro pevné $\lambda \in \mathbb{R}^+$ má měkké prahování menší rozptyl než tvrdé prahování pro všechna b , tj. $V_\lambda^s(b) \leq V_\lambda^h(b)$. Plyne to z vlastnosti $v_1(\lambda, b) \geq 0$ pro všechna b při pevně zvoleném $\lambda \geq 0$.

Důsledek 6.4: Pro pevné $\lambda \in \mathbb{R}^+$ má měkké prahování mnohem větší výchylku než tvrdé prahování pro $|b|$ dostatečně velké. Také platí, že pro $b \rightarrow \pm\infty$: $R_\lambda^s(b) \rightarrow 1 + \lambda^2$, zatímco $R_\lambda^h(b) \rightarrow 1$. To znamená, že průměrná kvadratická odchylka od správné hodnoty je pro rostoucí b vyšší u měkkého prahování, a to tím větší, čím větší je zvoleno λ . Riziko tvrdého prahování v limitě na hodnotě prahu nezávisí. Celkově to znamená, že k detailním koeficientům, které představují významné rychlé změny v signálu, a mají tedy relativně

velké b , se tvrdé prahování chová obezřetně – ponechá jím svou původní velikost, kdežto měkké prahování jejich velikost sníží. Tato vlastnost se odráží v tom, že měkké prahování má tendenci „přehladit“ rychlé změny, skoky a špičky v signálu.

Důsledek 6.5: Pro pevné $\lambda \in \mathbb{R}^+$ má tvrdé prahování maximální výchylku, rozptyl a riziko pro hodnoty $b \approx \pm\lambda$. Měkké prahování má maximální výchylku, rozptyl a riziko pro velké hodnoty $|b|$.

Důsledek 6.6: Pro pevné $\lambda \in \mathbb{R}^+$ má měkké prahování menší riziko v počátku $b = 0$. Obě statistiky $R_\lambda^s(0)$ a $R_\lambda^h(0)$ jsou monotónně klesající pro rostoucí λ .

Ostatní typy prahování tvoří jakýsi kompromis mezi tvrdým a měkkým prahováním. Gao a Bruce [24] ukázali, že pro dané λ existují $\lambda_1 < \lambda < \lambda_2$ tak, že pro všechna b_i platí

$$E[\delta^{ss}(y_i, \lambda_1, \lambda_2) - b_i]^2 < E[\delta^h(y_i, \lambda) - b_i]^2, \quad (6.4)$$

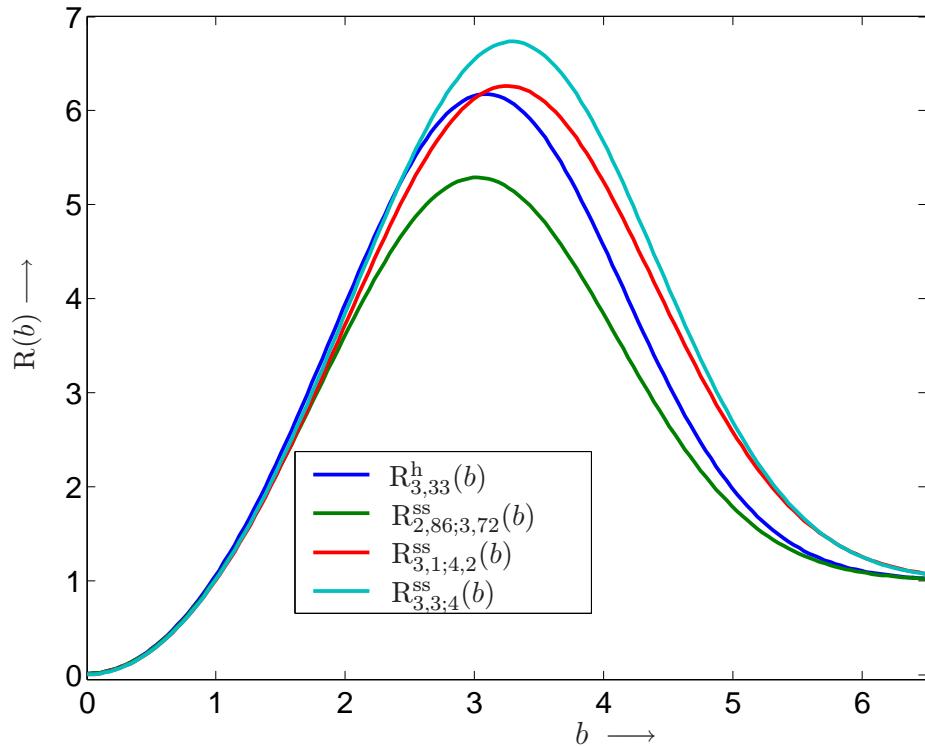
tedy že poloměkké prahování δ^{ss} má (stejnoměrně pro b) menší riziko než δ^h . Tato skutečnost je vidět na obrázku 6.6, kde byla zvolena vhodná λ_1, λ_2 . Volbou parametrů λ_1, λ_2 můžeme δ^{ss} podle potřeby více „přiblížit“ vlastnostem tvrdého nebo měkkého prahování. Poloměkké prahování má výhodu v tom, že překonává citlivost tvrdého prahování na malé změny ve vstupních datech, a zároveň zamezuje vzniku velké výchylky pro velká $|b|$, jak se tomu děje u měkkého prahování. Pro poloměkké prahování rovněž platí, že $\lim_{b \rightarrow \pm\infty} R_{\lambda_1, \lambda_2}^{ss}(b) = 1$.

Nutnost volby dvou prahových hodnot je však velkou nevýhodou tohoto prahování. Postup určení optimálních prahů je algoritmicky i výpočetně velmi náročný [24], je potřeba použít dvojrozměrnou Newtonovu iterační metodu.

Zbylá dvě prahovací pravidla se snaží s tímto hendikepem vyrovnat, a to tak, že napodobují tvar poloměkkého prahování funkcí s jedním volitelným parametrem – prahem. Statistické vlastnosti poloměkkého prahování jsou přitom v asymptotickém smyslu zachovány.

Gao [23] ukázal, že nezáporná garota je lepší než tvrdé i měkké prahování ve smyslu velikosti rizika a také citlivosti na malé změny ve výchozích datech. Také hyperbolické prahování kombinuje vlastnosti tvrdého a měkkého prahování, a z obrázku 6.6 je zřejmé, že z hlediska rizika je úspěšnější než nezáporná garota.

Neexistuje takové prahovací pravidlo, jehož riziková funkce by asymptoticky nabyla hodnoty menší než σ^2 pro $|b_i| \rightarrow \infty$.



Obr. 6.7: Promítnutí neoptimální volby dvojice prahů λ_1, λ_2 do velikosti rizika poloměkkého prahování. Z obrázku je patrné, že nesprávná volba může vést ke zvýšení rizika. Modrou barvou je pro srovnání znázorněno riziko tvrdého prahování, zelená barva přísluší „ideální“ volbě prahů.

Obrázek 6.7 ukazuje, jak se neoptimální volba dvojice prahů u poloměkkého prahování promítne do jejího rizika.

Následující tvrzení nám umožňuje všechny předchozí úvahy rozšířit na náhodnou veličinu s obecným normálním rozložením a zároveň dává návod, jak v takovém případě spočítat střední hodnotu, rozptyl a riziko.

Tvrzení 6.7: Pokud $Y \sim N(b, \sigma^2)$, pak $\frac{Y}{\sigma} \sim N(b/\sigma, 1)$ a platí

$$1. \quad \delta(Y, \lambda) \equiv \sigma \delta\left(\frac{Y}{\sigma}, \frac{\lambda}{\sigma}\right),$$

$$2. \quad M_\lambda(b) = \sigma M_{\lambda/\sigma}\left(\frac{b}{\sigma}\right),$$

$$3. \quad V_\lambda(b) = \sigma^2 V_{\lambda/\sigma}\left(\frac{b}{\sigma}\right),$$

$$4. \quad R_\lambda(b) = \sigma^2 R_{\lambda/\sigma}\left(\frac{b}{\sigma}\right).$$

Symbol \equiv zde představuje identitu náhodných veličin ve smyslu identity jejich distribučních funkcí. Tvrzení platí i pro poloměkké prahování, přičemž v tom případě chápeme λ jako vektor $\lambda = [\lambda_1, \lambda_2]$.

Důkaz je uveden v příloze A.

6.4 Volba prahové hodnoty

Volba optimální prahové hodnoty je pro dobrý výsledek separace rovněž důležitá. Tato oblast však již přesahuje záměr této práce, a tak se v této části textu omezíme pouze na stručné uvedení základních technik volby vhodné prahové hodnoty.

6.4.1 Kvantilová volba prahu

Kvantilová volba prahu se vyznačuje tím, že seřadíme waveletové koeficienty podle jejich absolutní hodnoty a zvolíme takovou hodnotu prahu, aby bylo po prahování ponecháno jen zvolené procento koeficientů. Jinak řečeno, jako λ použijeme zvolený kvantil množiny všech waveletových koeficientů, tj. např. tak, že se vynuluje 30 % nejmenších koeficientů.

Je zřejmé, že tento způsob volby prahu je závislý na zpracovávaných datech.

6.4.2 Univerzální hodnota prahu

Univerzální metodu stanovení prahu navrhli Donoho a Johnstone [17]. Na základě délky dat n a směrodatné odchylky aditivního bílého šumu σ vypočteme hodnotu λ_{univ} , která minimalizuje riziko odchylky od optimální (avšak neznámé) hodnoty prahu.

$$\lambda_{\text{univ}} = \sigma \sqrt{2 \log n} \quad (6.5)$$

Prahování se pak provede pro všechny waveletové koeficienty $c_{j,k}$, pro něž $j > j_0$. Hodnotu j_0 volíme podle charakteru zpracovávaných dat. Metodu lze použít pro měkké a tvrdé prahování.

V článku [22] je rozvinut tento přístup obecněji i pro korelovaný šum.

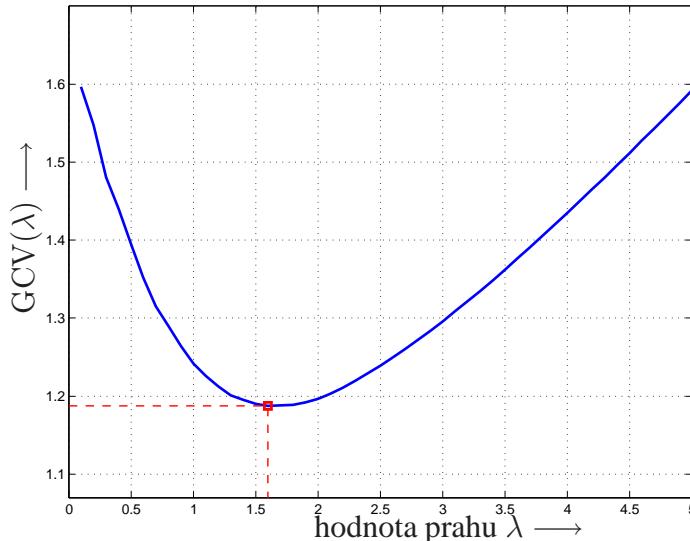
6.4.3 Kritérium zobecněné křížové validace

Technika křížové validace spadá pod klasické statistické přístupy. Její zobecněnou podobu (generalized cross-validation, GCV) lze úspěšně využít i pro nalezení vhodné prahové hodnoty [48, 54].

Hodnotu kritéria GCV vypočteme pro zvolené λ jako

$$GCV(\lambda) = \frac{\frac{1}{n} \|\mathbf{y} - \mathbf{y}_\lambda\|^2}{\left(\frac{n_\lambda}{n}\right)^2}. \quad (6.6)$$

V tomto vzorci \mathbf{y} představuje pozorovaný signál o délce n , \mathbf{y}_λ představuje signál po prahování s hodnotou prahu λ , n_λ je počet waveletových koeficientů, které byly během prahování vynulovány. Symbol $\|\cdot\|$ zde značí eukleidovskou normu.



Obr. 6.8: Příklad typického průběhu kritéria GCV (zde optimální prahová hodnota 1,6).

Vektor \mathbf{y} je dle (6.1) pozorováním náhodného vektoru. Proto nahradíme-li formálně ve vzorci (6.6) \mathbf{y} příslušným náhodným vektorem \mathbf{Y} a podobně pro \mathbf{y}_λ a n_λ , můžeme uvažovat o celé $GCV(\lambda)$ jako o náhodné veličině (pro dané λ). Optimální hodnotu pro λ pak definujeme jako

$$\lambda^* = \arg \min_{\lambda \in \mathbb{R}^+} E[GCV(\lambda)]. \quad (6.7)$$

Pokud máme v praxi možnost vícečetného pozorování signálu \mathbf{y} , vypočteme funkci GCV pro každé pozorování (pro vhodně volenou diskrétní síť parametrů λ), následně tyto

funkce zprůměrujeme a nalezneme optimální hodnotu λ^* . Pokud máme k dispozici pouze jedno pozorování y , nalezneme λ^* přímo z příslušné GCV. Typický průběh funkce GCV s vyznačením optimální hodnoty prahu je na obr. 6.8.

6.5 Závěr

V kapitole 6 byl vysvětlen princip waveletového potlačení šumu signálů a byly uvedeny obvyklé i méně používané používané prahovací techniky. Tyto techniky byly srovnány pomocí statistické analýzy rizika. Z tohoto srovnání plyne několik závěrů: tvrdé prahování hůře potlačuje šum, ale lépe zachovává rychlé změny v signálu (skoky a špičky), zatímco měkké prahování je velmi úspěšné při potlačování šumu, ale má tendenci „přehladit“ skoky a špičky. Tyto dva typy prahování tvoří dva protilehlé póly z hlediska charakteru separace. Víc než jen kompromis mezi nimi tvoří poloměkké prahování, které, jak bylo ukázáno, má při správné volbě dvou parametrů nejnižší riziko chyby. Ovšem dva parametry metody zároveň tvoří velkou nevýhodu oproti jiným metodám, které pracují pouze s parametrem jedním. Další uvedené metody, nezáporná garota a hyperbolické prahování, se snaží tuto nevýhodu zmírnit. Pracují pouze s jedním parametrem, avšak ve statistickém smyslu nedosahují takové kvality jako poloměkké prahování. Jak bude ale ukázáno v kapitole pojednávající o testování na reálných signálech, srovnávat pravidla pouze podle kritéria teoreticky nejnižšího rizika by bylo chybou, subjektivnímu vnímání nevyhovuje vždy signál odšuměný takovýmto prahovacím pravidlem.

7 Testování waveletových metod na signálech s aditivním šumem

Úspěšnost waveletového potlačování šumu byla testována a srovnávána na řečovém signálu. Šlo o promluvu mužského mluvčího „*vysyp všechny pytle s pšenicí*“. Tento signál, vzorkovaný s kmitočtem 16 kHz, byl uměle zarušen gaussovským šumem s různými rozptyly tak, že k dispozici byly testovací signály se SNR 15 dB, 10 dB, 5 dB, 0 dB a –5 dB. Odstup signálu od šumu SNR přitom počítáme jako

$$\text{SNR} [\text{dB}] = 10 \log_{10} \frac{\|\mathbf{y}\|}{\|\mathbf{e}\|}, \quad (7.1)$$

kde $\|\mathbf{y}\|$ je norma (odmocnina energie) nezarušeného signálu a $\|\mathbf{e}\|$ je norma šumu. Pro reálný signál $\mathbf{y} = (y_1, \dots, y_n)^\top$ normu definujeme $\|\mathbf{y}\| = \sqrt{y_1^2 + \dots + y_n^2}$. Po separaci řeči a šumu (viz níže) byl odhadnut výstupní SNR následujícím vztahem

$$\widehat{\text{SNR}} [\text{dB}] = 10 \log_{10} \frac{\|\mathbf{y}\|}{\|\mathbf{y} - k\hat{\mathbf{y}}\|}, \quad (7.2)$$

kde $\hat{\mathbf{y}}$ je „odšuměný“ signál a $k = \frac{\|\mathbf{y}\|}{\|\hat{\mathbf{y}}\|}$ je kompenzační konstanta.

Při všech analýzách byl použit wavelet Daubechies rádu 8. Další zvyšování rádu waveletu již nepřinášelo markantní zvýšení kvality separace. Hloubka dekompozice (a také hloubka prahování) byla zvolena $d = 5$. Testování probíhalo v systému Matlab pomocí vytvořeného balíku ThreshLab (viz příloha C). Příklad grafického znázornění prahování je na obr. 7.1.

Pro každý z pěti testovacích SNR byla nejprve stanovena prahová hodnota λ , přičemž výchozím odhadem byla univerzální hodnota prahu λ_{univ} . Použitá prahová hodnota byla nakonec zvolena subjektivně (i vzhledem k tomu, že tato dizertační práce není zaměřena na optimální volbu prahové hodnoty).

Pro tuto hodnotu prahu bylo provedeno tvrdé, měkké, hyperbolické prahování a nezáporná garota, a to:

[WT] detailních koeficientů klasické waveletové transformace do hloubky $d = 5$,

[WP] koeficientů odpovídajících nejlepší bázi ve waveletovém paketu hloubky $d = 5$.

Nejlepší báze (část 5.3) byla vyhledána pomocí prahové entropie (tamtéž) s prahem λ .

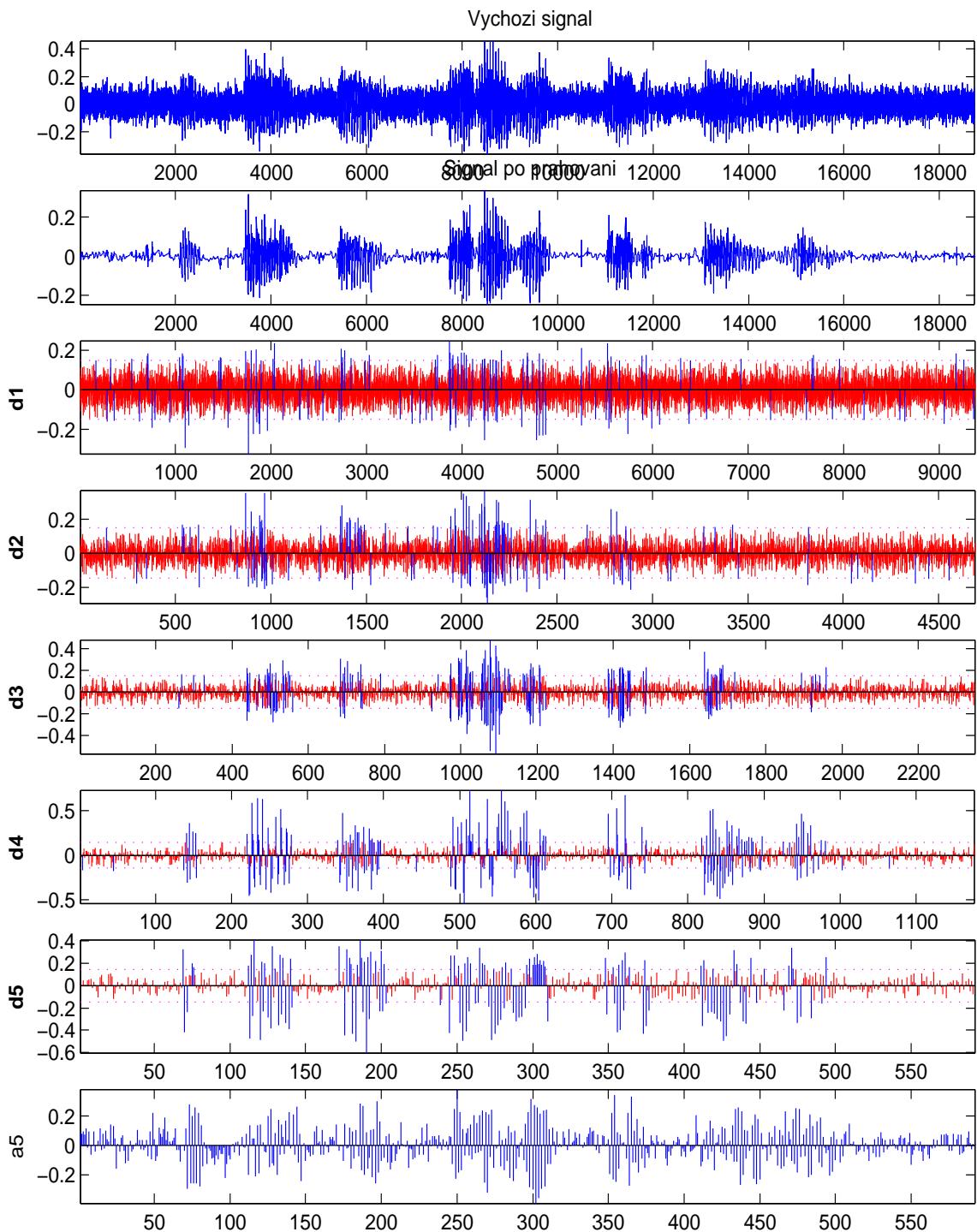
Operace [WT] byla provedena procedurou `thresh.m` a operace [WP] procedurou `wp_denoise.m`. Výsledky testování jsou obsaženy v tabulkách 7.1 až 7.4. V těchto tabulkách lze nalézt kromě objektivního parametru \widehat{SNR} subjektivní hodnocení kvality separace (stanovení pořadí úspěšnosti). Kritérii subjektivního posuzování byly stejnou měrou *příjemnost poslechu* (množství rušivých artefaktů v signálu) a *srozumitelnost* odrušené nahrávky.

Celkově lze k výsledkům říci:

Zlepšení objektivního \widehat{SNR} při přechodu od WT k nejlepší bázi WP tvoří průměrně pouze 0,2 dB. Zlepšení \widehat{SNR} lze vysvětlit tím, že WP lépe koncentruje energii signálu do jeho waveletových koeficientů (viz obr. 5.3). Rozdíl v subjektivním vnímání je však podstatný: objevují-li se v odrušeném signálu nežádoucí artefakty, pak u WT se většinou jedná o osamocené „lupance“ nebo jejich rychlý sled, na rozdíl od WP, kde dochází k rozostření šumu a k efektu „bublání“ na pozadí, což je lidskému sluchu mnohem příjemnější na poslech. Rovněž u WP pozorujeme lepší srozumitelnost.

Nejméně úspěšné pro zpracování *řečového* signálu je měkké prahování, a to jak z objektivního, tak subjektivního hlediska. Nejhorší výsledky \widehat{SNR} jsou paradoxem, neboť bylo řečeno v závěru kapitoly 6, že právě měkké prahování nejlépe potlačuje šum. Toto tvrzení ale zároveň musíme potvrdit ze subjektivního pohledu – skutečně je tomu tak, avšak u řečového signálu tím zanikají neznělé souhlásky jako zejména *s, š, c, č, z, ž*, které jsou právě pro srozumitelnost řeči nejdůležitější.

V subjektivním hodnocení dosáhlo obecně nejlepších výsledků hyperbolické prahovací pravidlo, i když je třeba upozornit, že se snižujícím se vstupním SNR mu významně konkuруje tvrdé prahování. To lze vysvětlit tím, že směrem k malým hodnotám SNR přirozeně klesá srozumitelnost odrušených signálů a právě tvrdé prahování i přes velké množství nepříjemných artefaktů srozumitelnost zachovává nejdéle. Ve chvíli, kdy kvalita poslechu je již pro všechny typy prahování velmi nízká, proto přirozeně dáváme přednost



Obr. 7.1: Příklad prahování zarušeného signálu. Shora dolů jsou vykresleny: původní zarušený signál, separovaný signál a prahované waveletové koeficienty.

tvrdému, které dává nejsrozumitelnější výsledek.

V případě $\text{SNR} = -5 \text{ dB}$ ($\lambda_{\text{univ}} = 0,4592$) se výsledný $\widehat{\text{SNR}}$ u všech typů prahování pohyboval kolem $-1,5 \text{ dB}$ u metody WT, přičemž ale mluvenému slovu již nebylo rozumět. Pro zvolený práh $\lambda = 0,85 \cdot \lambda_{\text{univ}}$ metoda nejlepší báze WP selhala a proto není možné uvést srovnávací tabulku. Volbou $\lambda = 0,57 \cdot \lambda_{\text{univ}}$ však již metoda dává ucházející výsledky, zatímco WT je pro tuto hodnotu již nepoužitelná.

Na závěr ještě poznamenejme, že poloměkké prahování nebylo do tohoto testu zahrnuto, neboť v balíku Threshlab není funkce, která by pro zvolené λ našla optimální λ_1, λ_2 ve smyslu (6.4). Jak bylo zmíněno, to vyžaduje implementaci dvojrozměrné Newtonovy iterační metody. Dá se však očekávat, že při správné volbě obou prahů bude dosahovat tak dobrých výsledků jako tvrdé a hyperbolické prahování.

Ukázalo se, že nastavení hodnoty prahu je vždy záležitostí kompromisu mezi mírou potlačení šumu a zachováním co nejpřirozenějšího charakteru řeči, tak jak je tomu u všech metod jednokanálové separace. Výsledný poměr $\widehat{\text{SNR}}$ lze umělým snižováním prahu zpravidla ještě zlepšovat, avšak subjektivní ohodnocení se rychle zhorší.

V budoucím vývoji se zřejmě uplatní adaptivní prahování [1, 50]. Konkrétně u řečových signálů by se postup separace měl doplnit o indikátor znělosti / neznělosti aktuálního fonému. V případě znělého by se pak provedlo normální prahování s nejlepší bází WP a v případě neznělého by se práh uměle nadhodnotil, tím by se lokálně signál „méně prahoval“ a srozumitelnost by se díky vyšší míře uchování problematických souhlásek *s, š, c, č, z, ž* zlepšila.

7.1 Testovací a výsledné nahrávky

Vstupní i výstupní nahrávky je možno nalézt na přiloženém CD v adresáři `/testovani`. Vstupní nahrávky jsou obsaženy v souboru `psenice_input.mat` pro Matlab. Vektor s originální nahrávkou má jméno `psenice` a jména zarušených signálů jsou vždy opatřena příznamem odpovídajícího SNR, např. `psenice05` odpovídá $\text{SNR} = 5 \text{ dB}$. Výstupní nahrávky jsou pak podle příznamu SNR rozděleny do souborů `psenice15_output.mat` až `psenice-5_output.mat`. Každý z těchto souborů obsahuje 8 nahrávek, jejichž jména jsou logicky opatřena příznamy WT nebo WP a zkratkou typu použitého prahování,

Prahování	WT		WP	
	\widehat{SNR} [dB]	Pořadí subjekt. hodnocení	\widehat{SNR} [dB]	Pořadí subjekt. hodnocení
tvrdé	26,4	3	26,3	3
měkké	22,3	3	22,6	4
nezáp. garota	24,6	2	24,7	2
hyperbolické	25,6	1	25,6	1

Tab. 7.1: Výsledky testování na zarušeném signálu se $SNR = 15$ dB. Univerzální práh $\lambda_{univ} = 0,0046$, použitá hodnota prahu $\lambda = 1,2 \cdot \lambda_{univ}$.

Prahování	WT		WP	
	\widehat{SNR} [dB]	Pořadí subjekt. hodnocení	\widehat{SNR} [dB]	Pořadí subjekt. hodnocení
tvrdé	17,5	2	17,6	3
měkké	15,2	4	15,4	4
nezáp. garota	16,5	2	16,7	1
hyperbolické	17,0	1	17,2	1

Tab. 7.2: Výsledky testování na zarušeném signálu se $SNR = 10$ dB. Univerzální práh $\lambda_{univ} = 0,0145$, použitá hodnota prahu $\lambda = 1,1 \cdot \lambda_{univ}$.

Prahování	WT		WP	
	$\widehat{\text{SNR}}$ [dB]	Pořadí subjekt. hodnocení	$\widehat{\text{SNR}}$ [dB]	Pořadí subjekt. hodnocení
tvrdé	9,2	3	9,2	1
měkké	8,1	4	8,2	4
nezáp. garota	8,7	1	8,8	1
hyperbolické	9,0	1	9,0	1

Tab. 7.3: Výsledky testování na zarušeném signálu se $\text{SNR} = 5 \text{ dB}$. Univerzální práh $\lambda_{\text{univ}} = 0,0459$, použitá hodnota prahu $\lambda = \lambda_{\text{univ}}$.

Prahování	WT		WP	
	$\widehat{\text{SNR}}$ [dB]	Pořadí subjekt. hodnocení	$\widehat{\text{SNR}}$ [dB]	Pořadí subjekt. hodnocení
tvrdé	2,1	1	2,1	1
měkké	1,6	4	1,7	1
nezáp. garota	1,9	2	1,9	1
hyperbolické	2,0	2	2,0	1

Tab. 7.4: Výsledky testování na zarušeném signálu se $\text{SNR} = 0 \text{ dB}$. Univerzální práh $\lambda_{\text{univ}} = 0,1452$, použitá hodnota prahu $\lambda = 0,9 \cdot \lambda_{\text{univ}}$. Subjektivně je mezi jednotlivými typy prahování u WP zanedbatelný rozdíl.

např. data pod názvem `psenice10_thresh_WT_nng` odpovídají prahovanému signálu `psenice10` pomocí WT s nezápornou garotou. Pro případ $\text{SNR} = -5 \text{ dB}$ byl výstup pro WP a WT uložen do zvláštních souborů, z důvodu odlišné volby λ (viz výše).

Do stejného adresáře byl nakopírován také soubor `testovani.m`, který byl pro všechny testy používán.

Zvukový výstup v Matlabu umožňuje příkaz `sound`, příp. výstupu do souboru „wav“ lze dosáhnout příkazem `wavwrite`.

8 Segmentovaná konečná diskrétní waveletová transformace

V praktických aplikacích je často třeba signál zpracovávat v „reálném čase“, tzn. s minimálním zpožděním. Předem neznámý signál přichází na vstup nějakého systému po úsecích – na sobě nezávislých segmentech – které musejí být efektivně zpracovány a modifikované poslány na výstup systému. Typickým příkladem je např. zpracování akustických signálů, zejména řečových signálů v telekomunikacích.

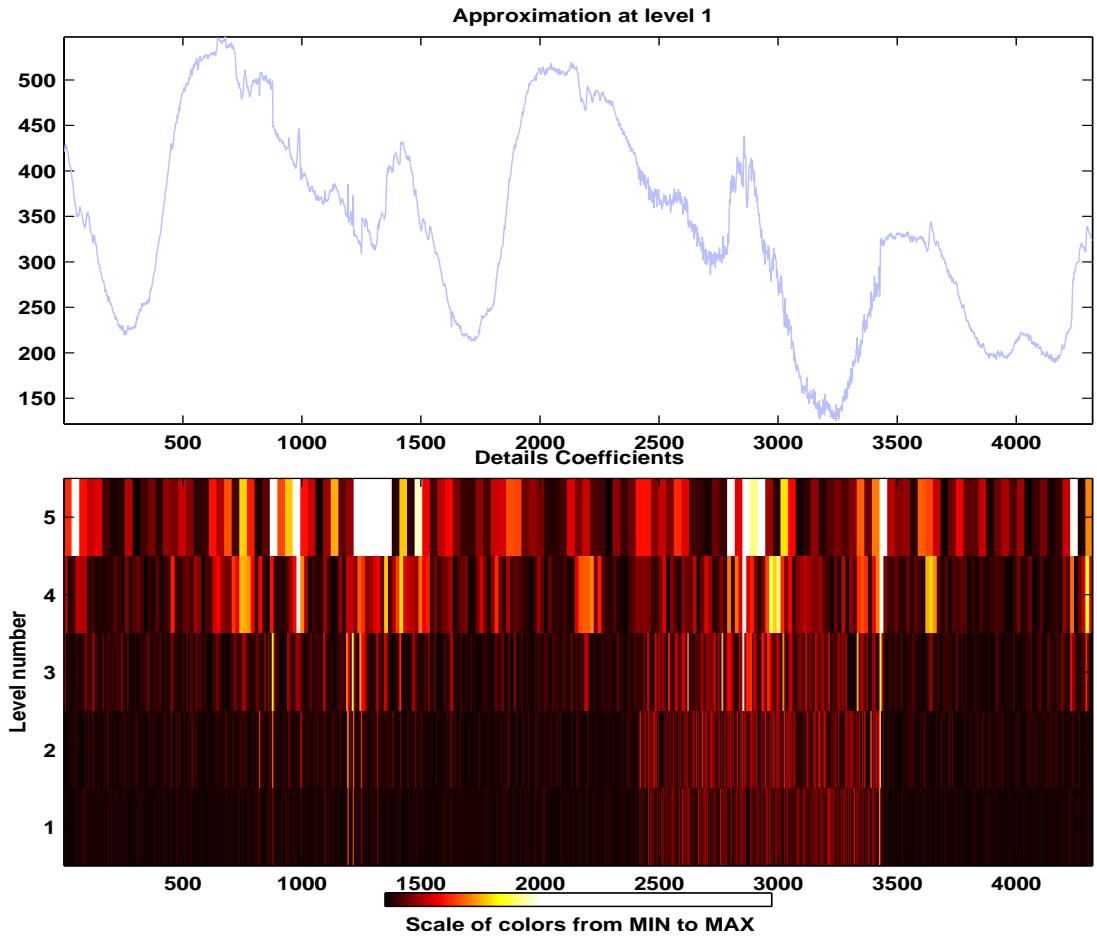
Segmentovaná konečná waveletová transformace (SegWT¹), která tento typ zpracování umožňuje, má velké potencionální využití také v případě, že je třeba zpracovat dlouhý signál (ne nutně v reálném čase), ale přitom není k dispozici dostatečná výpočetní síla nebo paměťový prostor. Tehdy je možné touto novou metodou signál ekvivalentně zpracovat po částech a tím ušetřit výpočetní výkon i paměťový prostor. V tomto smyslu algoritmus SegWT koresponduje s algoritmy přičtení přesahu (overlap-add) a úschovy přesahu (overlap-save) [21] u fourierovské lineární filtrace.

Dalším možným využitím algoritmu SegWT je okamžitá vizualizace signálu pomocí zobrazení, které nazýváme „waveletogram“. Waveletogram zobrazuje graficky hodnoty waveletových koeficientů, viz obr. 8.1. Je to technika příbuzná vykreslování spektrogramu v reálném čase. U waveletové transformace máme však výhodu, že signál nemusíme váhovat okny, čímž dochází ke zkreslení kmitočtové informace, jak je tomu právě u spektrogramu. Waveletogram vytvořený pomocí SegWT je navíc zcela nezávislý na zvolené délce segmentu.

V dostupné literatuře se tomuto způsobu waveletové transformace prakticky nevěnuje žádná pozornost a to bylo důvodem, proč bylo vyvinuto úsilí odvodit tento modifikovaný algoritmus. Tato kapitola podrobně uvádí tuto modifikaci pro segmentovaný signál, včetně odvození algoritmu a důsledků pro implementaci na libovolném zařízení.

Po definování účelu modifikované metody (část 8.1) je nejprve zapotřebí podrobně

¹Zavádíme zkratku SegWT (Segmented Wavelet Transform), zkratka SWT je již vyhrazena pro stacionární waveletovou transformaci (Stationery Wavelet Transform)



Obr. 8.1: „Waveletogram“. Horní část představuje analyzovaný signál a spodní část jeho detailní waveletové koeficienty do hloubky dekompozice $d = 5$. Čím větší absolutní hodnota koeficientu, tím světlejší barva jemu příslušného políčka v grafu (převzato z programu Matlab).

rozebrat klasický algoritmus waveletové transformace (část 8.2) a na základě tohoto rozboru sestavit algoritmus waveletové transformace modifikovaný pro zpracování po nezávislých segmentech (část 8.3).

8.1 Motivace a cíl modifikované metody

Běžně používaná konečná diskrétní transformace (DTWT, část 4.3) je vhodná pro zpracování signálů „off-line“, tzn. předem známých, přestože libovolně dlouhých. Úkolem segmentované konečné diskrétní waveletové transformace je tedy přirozeně zpracovat sig-

nál po segmentech, a to tak, abyhom po postupném zpracování všech segmentů obdrželi stejný výsledek (tj. stejné waveletové koeficienty), jako kdybychom signál znali celý předem a zpracovali ho pomocí běžného algoritmu DTWT.

Jak bude ukázáno, v této úloze hrají zásadní roli zvolená hloubka waveletové dekompozice, délka waveletového filtru a délka segmentu signálu. Označme:

- m délka waveletového filtru, $m > 0$,
- d hloubka waveletové dekompozice, $d > 0$,
- s délka segmentu diskrétního signálu, $s > 0$.

Symboly d , d sice používáme též pro označení detailních waveletových koeficientů, ze souvislosti však bude vždy zřejmé, zda se jedná o tento případ nebo o hloubku dekompozice.

8.2 Rozbor pyramidového algoritmu DTWT

Nejprve uvedeme detailní popis algoritmu DTWT a poté na něj navážeme s vyvozenými tvrzeními a důsledky.

Poznámka: V části 8.2 budeme značit délku celého signálu také s . Při tomto označení vycházíme z toho, že při DTWT zpracováváme signál naráz, tj. signál je zároveň jakoby jediným segmentem délky s .

Algoritmus 8.1: (dekompoziční pyramidový algoritmus DTWT)

Je dán vstupní signál \mathbf{x} délky s , dva waveletové dekompoziční filtry délky m , horní propust \mathbf{g} a dolní propust \mathbf{h} , je zvolena hloubka dekompozice d a rovněž je dán tzv. typ prodloužení signálu.²

1. Označíme vstupní signál \mathbf{x} jako $\mathbf{a}^{(0)}$. Nastavíme $j = 0$.
2. Jeden krok dekompozice:
 - (a) *Rozšíření vstupního vektoru.* Prodloužíme vektor $\mathbf{a}^{(j)}$ nalevo i napravo o $(m - 1)$ vzorků, podle zvoleného typu prodloužení.
 - (b) *Filtrace.* Filtrujeme prodloužený signál filtrem \mathbf{g} , což můžeme vyjádřit jejich konvolucí.

²ohledně typů prodloužení odkazujeme např. na [39, kap. 8].

- (c) *Vyseknutí.* Z výsledku filtrace vybereme jen její prostřední část, tak, že „zbytky“ vlevo a vpravo jsou shodně dlouhé $m - 1$ vzorků.
- (d) *Podvzorkování (decimace).* Vyseknutý vektor decimujeme, tj. ponecháváme pouze jeho vzorky se sudými indexy (předpokládáme indexování vektorů počínaje jedničkou).

Uložíme tento vektor jako $\mathbf{d}^{(j+1)}$.

Opakujeme body (b)–(d), nyní s filtrem \mathbf{h} , a výsledek uložíme jako $\mathbf{a}^{(j+1)}$.

3. Zvýšíme j o jedničku. Pokud nyní $j < d$, jdeme na bod 2., jinak algoritmus končí.

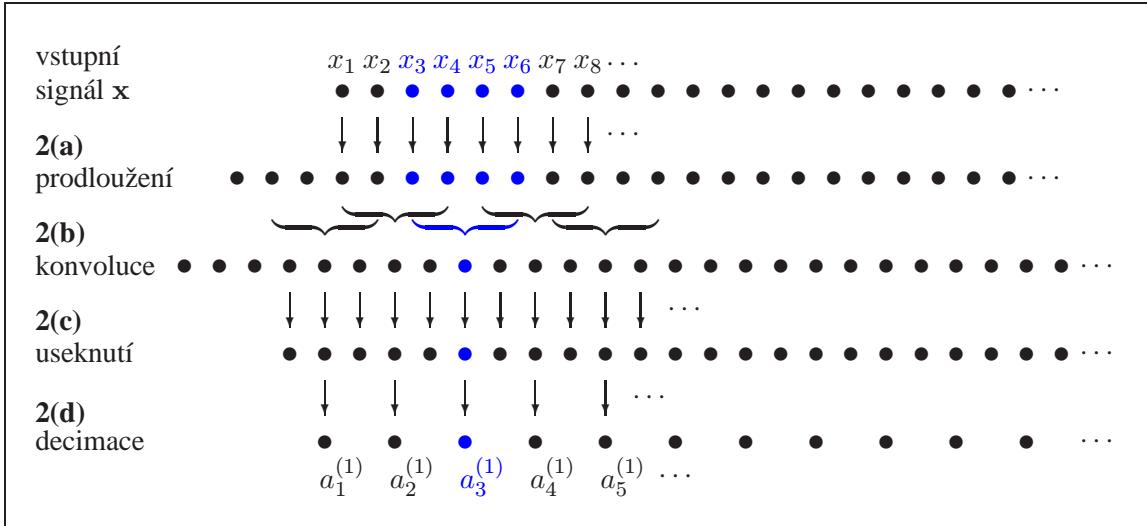
Poznámky:

1. Po skončení algoritmu máme uloženy waveletové koeficienty v $d + 1$ vektorech $\mathbf{a}^{(d)}, \mathbf{d}^{(d)}, \mathbf{d}^{(d-1)}, \dots, \mathbf{d}^{(1)}$.
2. Ad 2(a): Prodloužení je nutné vykonat kvůli jednoznačnosti transformace, a to v každém kroku dekompozice.
3. Ad 2(a): Ve speciálním případě tzv. periodického prodloužení je délka potřebného prodloužení signálu menší než $m - 1$. Periodické prodloužení však pro naše účely, v návaznosti na část 8.3, není relevantní, a proto jej dále neuvažujeme.
4. Ad 2(a): V prvním kroku algoritmu tedy prodloužíme $\mathbf{a}^{(0)}$ na délku $s + 2(m - 1)$.
5. Ad 2(b): Délka výsledné konvoluce v prvním kroku je $[s + 2(m - 1) + m - 1] = s + 3(m - 1)$.
6. Ad 2(c): Prostřední část má tedy délku, jakou by měla konvoluce neprodlouženého vektoru $\mathbf{a}^{(j)}$ s waveletovým filtrem. V prvním kroku je vyseknutá část délky $s + m - 1$.
7. Ad 2(d): Touto operací dostáváme v prvním kroku vektor o délce $\text{floor}\left(\frac{s+m-1}{2}\right)$.

Připomeňme konvenci, že vstupní signál ztotožňujeme s (aproximačními) waveletovými koeficienty nejnižší úrovně, obvykle 0, tzn. v dalším textu platí $\mathbf{x} = \mathbf{a}^{(0)}$. Dále také předpokládáme $k \in \mathbb{N}$ a pevně zvolenou délku waveletových filtrů, m .

Lemma 8.2: *Pro výpočet jednoho waveletového koeficientu v úrovni k je třeba m koeficientů úrovně $k - 1$.*

Důkaz: To je zřejmé z popisu algoritmu, neboť jeden waveletový koeficient z úrovně k je jedním prvkem konvoluce vektoru $\mathbf{a}^{(k-1)}$ s waveletovým filtrem. Viz obr. 8.2. \square



Obr. 8.2: Znázornění kroků 2(a) až 2(d) v algoritmu 8.1 pro délku filtru $m = 4$. Každý kroužek představuje jeden element vektoru (signálu).

Lemma 8.3: Pro výpočet $q \geq 1$ po sobě jdoucích waveletových koeficientů v k -té úrovni je třeba $m + 2(q - 1)$ po sobě jdoucích koeficientů z $(k - 1)$ -ní úrovni.

Důkaz: Pro výpočet jednoho waveletového koeficientu v úrovni k je třeba m koeficientů z úrovni $k - 1$. Pro každý další koeficient v úrovni k je (díky decimaci) třeba uvažovat další dva koeficienty z úrovni $k - 1$, viz obr. 8.2, 8.3. \square

Tvrzení 8.4: Pro výpočet $q \geq 1$ po sobě jdoucích waveletových koeficientů v k -té úrovni je třeba

$$p_q(k) = (2^k - 1)m + 2^k q - 2(2^k - 1) \quad (8.1)$$

vzorků vstupního signálu.

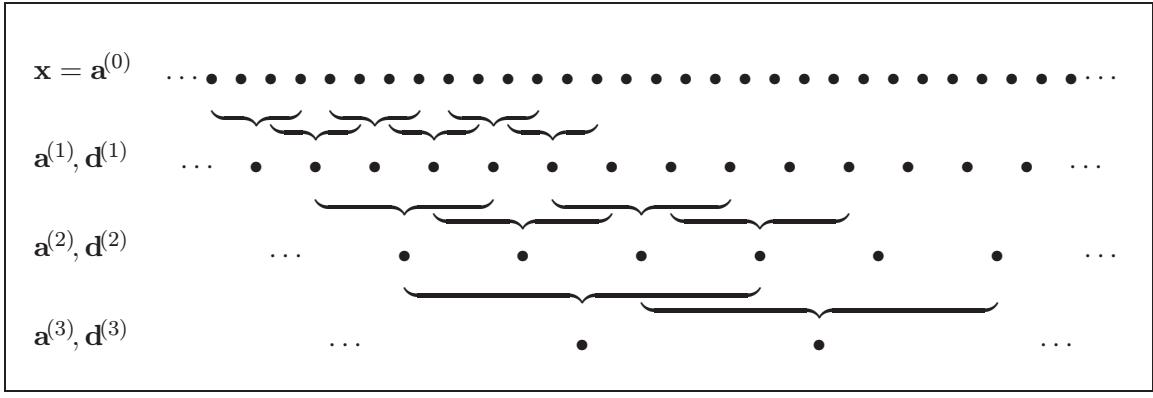
Důkaz: Rekurzivní aplikací lemmatu 8.3 lze odvodit vzorec

$$p_q(k) = m + 2[p_q(k - 1) - 1], \quad \text{kde definujeme } p_q(0) = q. \quad (8.2)$$

Ukážeme indukcí vzhledem ke k , že vzorec (8.1) vyjadřuje (8.2) v nerekurzivním tvaru.

Pro $k = 1$ podle (8.2) platí $p_q(1) = m + 2[p_q(0) - 1] = m + 2(q - 1)$ a podle (8.1) platí $p_q(1) = m + 2q - 2 = m + 2(q - 1)$. Předpokládejme dále, že (8.1) platí pro zvolené $k \geq 1$. S využitím (8.2) pak

$$p_q(k + 1) = m + 2[p_q(k) - 1]$$



Obr. 8.3: Ilustrace algoritmu 8.1 ve třech úrovních dekompozice, tj. $d = 3$. Každý koeficient v k -té úrovni je vypočítán z m koeficientů v úrovni $k - 1$. Zde opět délka filtru $m = 4$.

$$\begin{aligned}
 &= m + 2[(2^k - 1)m + 2^k q - 2(2^k - 1) - 1] \\
 &= m + 2(2^k - 1)m + 2^{k+1}q - 2(2^{k+1} - 1) \\
 &= (2^{k+1} - 1)m + 2^{k+1}q - 2(2^{k+1} - 1). \quad \square
 \end{aligned}$$

Důsledek 8.5: Pro výpočet jednoho waveletového koeficientu v k -té úrovni je třeba $p_1(k)$ vzorků vstupního signálu. Platí

$$p_1(k) = (2^k - 1)(m - 1) + 1. \quad (8.3)$$

Důkaz: Dosazením $q = 1$ do (8.1) a úpravou.

Tabulka 8.1 ukazuje počty vzorků vstupního signálu potřebné pro výpočet jednoho waveletového koeficientu v různé hloubce dekompozice. Obrázek 8.4 znázorňuje data z tabulky v grafech. Je zřejmé, že pro rostoucí délku filtru se $p_1(k)$ zvyšuje lineárně a pro rostoucí hloubku dekompozice téměř přesně exponenciálně o základu 2.

Před uvedením dalších úvah je nutné zavést pojmy, které budeme dále využívat:

Sousední koeficienty jsou dva koeficienty, které v dané úrovni sousedí; např. $a_{23}^{(3)}, a_{24}^{(3)}$ sousedí v úrovni 3.

Koeficienty vzdálené ob q koeficientů jsou koeficienty v dané úrovni, mezi nimiž leží právě q koeficientů; např. $a_{19}^{(3)}, a_{26}^{(3)}$ jsou vzdálené ob 6 koeficientů.

Působnost koeficientu z k -té úrovně v u -té úrovni, $u < k$ je množina koeficientů v úrovni

$k \setminus m$	2	3	4	5	6	8	10	12	14	16	18
1	2	3	4	5	6	8	10	12	14	16	18
2	4	7	10	13	16	22	28	34	40	46	52
3	8	15	22	29	36	50	64	78	92	106	120
4	16	31	46	61	76	106	136	166	196	226	256
5	32	63	94	125	156	218	280	342	404	466	528
6	64	127	190	253	316	442	568	694	820	946	1072
7	128	255	382	509	636	890	1144	1398	1652	1906	2160
8	256	511	766	1021	1276	1786	2296	2806	3316	3826	4336

Tab. 8.1: Počet vzorků vstupního signálu nutných pro výpočet jednoho waveletového koeficientu, $p_1(k)$. Horizontálně délka waveletového filtru m , vertikálně hloubka koeficientu k .

u , jejichž hodnota spoluúspobí na hodnotě daného koeficientu (ovlivňuje tuto hodnotu) z úrovně k .

Posun ve vzorcích vstupního signálu pro nějaké dva koeficienty z k -té úrovně je počet vzorků vstupního signálu, o který se liší působnosti těchto dvou koeficientů ve vstupním signálu. Např. pokud působnost prvního z nich je x_8, x_9, \dots, x_{23} a druhého $x_{12}, x_{13}, \dots, x_{27}$, pak posun ve vzorcích vstupního signálu je 4.

Společné koeficienty úrovně u pro dané dva koeficienty z úrovně k , $u < k$ jsou koeficienty v úrovni u , které zároveň patří do působnosti obou koeficientů z úrovně k . V příkladu výše uvedeném jsou společné koeficienty ve vstupním signálu (tj. pro $u = 0$) $x_{12}, x_{13}, \dots, x_{23}$.

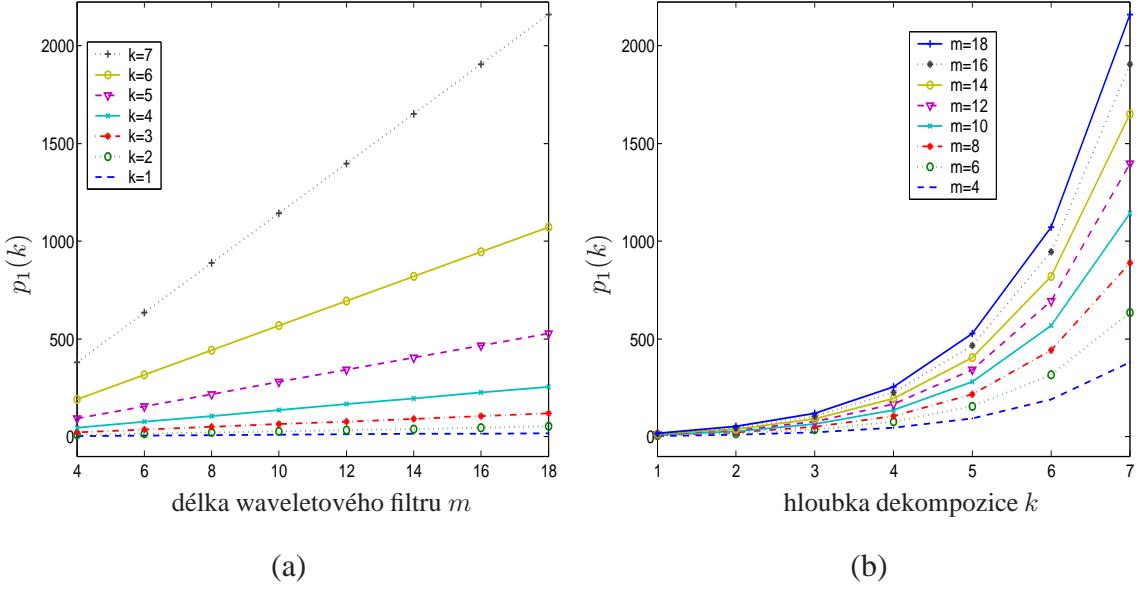
Lemma 8.6: Posun ve vzorcích vstupního signálu u dvou sousedních koeficientů v k -té úrovni je $a(k) = 2^k$. Sousední koeficienty v k -té úrovni mají společných

$$b(k) = (2^k - 1)(m - 2) \quad (8.4)$$

vzorků vstupního signálu.

Důkaz: První část je zřejmá (viz obr. 8.3). Dva sousední koeficienty v k -té úrovni mají společných právě $m - 2$ koeficientů v $(k - 1)$ -ní úrovni. Nyní stačí využít tvrzení 8.4 a získáváme vztah

$$\begin{aligned} b(k) &= p_{m-2}(k-1) \\ &= (2^{k-1} - 1)m + 2^{k-1}(m-2) - 2(2^{k-1} - 1) \end{aligned}$$



Obr. 8.4: Počet vzorků signálu potřebných pro výpočet jednoho waveletového koeficientu: (a) hloubka dekompozice k je fixní, (b) délka waveletového filtru m je fixní.

$$\begin{aligned}
 &= (2^k - 1)m - 2(2^k - 1) \\
 &= (2^k - 1)(m - 2). \quad \square
 \end{aligned}$$

Lemma 8.7: Posun ve vzorcích vstupního signálu u dvou koeficientů v k -té úrovni, které jsou od sebe vzdáleny ob q koeficientů, je $c(k) = a(k) \cdot (q+1) = 2^k(q+1)$. Dva koeficienty v k -té úrovni, vzdálené ob q koeficientů, mají ve vstupním signálu společných $d(k)$ vzorků, kde

$$d(k) = \begin{cases} 0 & \text{pro } b(k) \leq q \cdot a(k) \\ b(k) - q \cdot a(k) & \text{jinak.} \end{cases} \quad (8.5)$$

Důkaz: První část zřejmá (pohybujeme se díky decimaci po mocninách dvou, viz obr. 8.3). Dále z druhé části lemmatu 8.6 a z první části tohoto lemmatu plyne, že koeficienty z úrovni k vzdálené ob q koeficientů musí mít společných $d(k) = b(k) - q \cdot a(k)$ vzorků vstupního signálu (pokud toto číslo není záporné, jinak samozřejmě klademe $d(k) = 0$). \square

Poznámky:

1. Lemma 8.6 je speciálním případem lemmatu 8.7 při volbě $q = 0$.

$m = 8$

$k \setminus q$	50	80	110	140	170	200	230	260	290	320	350
1	28	43	58	73	88	103	118	133	148	163	178
2	17	25	32	40	47	55	62	70	77	85	92
3	12	16	19	23	27	31	34	38	42	46	49
4	9	11	13	15	17	19	20	22	24	26	28
5	8	9	10	11	12	13	13	14	15	16	17
6	7	8	8	9	9	10	10	10	11	11	12
7	7	7	7	8	8	8	8	8	9	9	9

$m = 16$

$k \setminus q$	50	80	110	140	170	200	230	260	290	320	350
1	32	47	62	77	92	107	122	137	152	167	182
2	23	31	38	46	53	61	68	76	83	91	98
3	19	23	26	30	34	38	41	45	49	53	56
4	17	19	20	22	24	26	28	30	32	34	35
5	16	17	17	18	19	20	21	22	23	24	25
6	15	16	16	16	17	17	18	18	19	19	20
7	15	15	15	15	16	16	16	16	17	17	17

Tab. 8.2: Počty koeficientů vypočítaných algoritmem 8.1 ze vstupního signálu o délce q v různých hloubkách dekompozice k .

2. Všimněme si, že $a(k)$ ani $c(k)$ nezávisí na délce filtru, m .

Lemma 8.8: Nechť je dána délka filtrů m . Z vektoru o délce q vypočítáme pomocí algoritmu 8.1

$$\text{floor} \left(2^{-k} q - (2^{-k} - 1)(m - 1) \right) \quad (8.6)$$

waveletových koeficientů úrovně k .

Důkaz: Důkaz viz str. 85 využitím lemmatu 8.21.

Tabulka 8.2 uvádí pro ilustraci počty koeficientů, které vypočítáme algoritmem 8.1 při různé volbě q a m .

8.3 Nový algoritmus

segmentované waveletové transformace

Jak již bylo řečeno v úvodu této kapitoly, chceme použít algoritmus DTWT (nebo nějakou jeho modifikaci) na každý segment, a to tak, abychom vhodným sloučením (spojením) waveletových transformací jednotlivých segmentů dostali stejný výsledek jako obyčejnou DTWT celého signálu. Pro tyto účely je nutno vstupní segmenty před provedením transformace upravit.

Konkrétně jde o to, že jednotlivé segmenty musíme vhodným způsobem prodloužit, aby se optimálně překrývaly. Rovněž musíme samostatně řešit zacházení s okrajovými segmenty (tj. s prvním a posledním). O těchto modifikacích detailně pojednává následující část textu.

Předpokládejme dále, že vstupní signál \mathbf{x} je rozdelen na $S \geq 1$ segmentů stejně délky s , které budeme označovat ${}^1\mathbf{x}, {}^2\mathbf{x}, \dots, {}^S\mathbf{x}$. Poslední ze segmentů může mít délku nižší než s , viz obr. 8.5.

Návazností koeficientů v k-té úrovni dekompozice budeme mít na mysli stav, kdy dva na sebe navazující segmenty (tj. bloky vzorků) jsou vhodně prodlouženy, viz obr. 8.6, a to tak, že když aplikujeme DTWT³ hloubky k zvlášť na první prodloužený segment a zvlášť na druhý prodloužený segment a takto vypočtené dvě množiny koeficientů správně ořízneme a spojíme dohromady, dostáváme tím v místě přechodu segmentů stejnou množinu koeficientů, jako kdybychom aplikovali DTWT na oba segmenty již sloučené.

Tvrzení 8.9: *Pro návaznost koeficientů v k-té úrovni dekompozice musejí sousedící segmenty mít*

$$r(k) = (2^k - 1)(m - 1) \quad (8.7)$$

společných vzorků vstupního signálu.

Důkaz: Pro návaznost koeficientů z první úrovni je třeba, aby sousední segmenty po prodloužení měly $m - 1$ společných vzorků. To plyne ze vztahu (8.4), když přičteme jedničku kvůli tomu, že při sudém decimování se neuplatní („vypadne“) první vzorek v každém segmentu.

³s vynecháním kroku 2(a), důvod bude vysvětlen později

Analogicky pro $k > 1$ je pro návaznost potřeba $m - 1$ koeficientů z $(k - 1)$ -ní úrovni. To vede na rekurzivní vzorec

$$r(k) = m + 2r(k-1) - 1, \quad k \geq 1, \quad \text{kde definujeme } r(0) = 0. \quad (8.8)$$

Ten je přirozeně podobný s vzorcem (8.2), neboť z něj vychází; oba vzorce se liší ve své rekurzivní části tím, že u $r(k)$ přičteme v každém kroku rekurze jedničku, což je zdůvodněno zmíněným „vypadnutím“ prvního koeficientu v každé úrovni.

Podobně jako v důkazu tvrzení 8.4 lze matematickou indukcí ukázat, že $r(k) = (2^k - 1)(m - 1)$. \square

Všimněme si, že veličina $r(k)$ se liší od $p_1(k)$ pouze o jedna, a proto chování $r(k)$ pro různou volbu k, m můžeme vysledovat z tabulky 8.1 a obrázku 8.4.

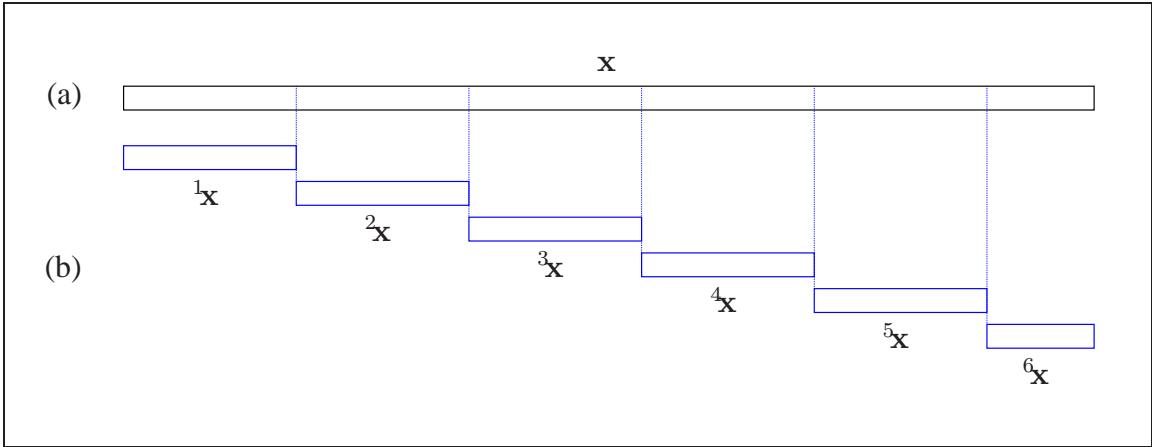
Lemma 8.10: *Pro m sudé je $r(k)$ liché, pro m liché je $r(k)$ sudé, $k \in \mathbb{N}$.*

Důkaz: Snadno plyne z rovnice (8.7) díky tomu, že $(2^k - 1)$ je vždy liché a součin lichého čísla se sudým je sudé číslo, součin lichého čísla s lichým je liché číslo. \square

Ve tvrzení 8.9 byl odvozen počet společných vzorků dvou prodloužených sousedních segmentů, nutný pro návaznost waveletových koeficientů vypočtených z těchto segmentů v hloubce dekompozice d , a to $r(d) = (2^d - 1)(m - 1)$. Je zřejmé, že na nižších úrovniích dekompozice k , $0 < k < d$, proto při výpočtech ze dvou po sobě následujících segmentů dochází k redundancii vypočtených koeficientů, neboť některé z nich jsou vypočteny dvakrát – poprvé z jednoho a podruhé z následujícího segmentu. Přitom platí, že čím menší k , tím větší je počet takových koeficientů. Přesně, v hloubce k je dvakrát vypočteno vždy $r(d-k)$ koeficientů. Z toho plyne, že pro návaznost koeficientů v těchto úrovniích musíme tento počet koeficientů „odmazat“, ať už z jednoho nebo druhého segmentu. Tato operace je implementována v bodě 6 algoritmu 8.16.

Cílem úvah v následující části textu je najít vhodný způsob překrytí dvou po sobě jdoucích segmentů signálu. Ukážeme, že překrytí musí splňovat přísná pravidla vyplývající z uvedených tvrzení. Omezení vyplývá z toho, že DTWT není časově invariantní transformace (v podstatě je 2^d -invariantní).

Překrytí dvojice sousedních prodloužených segmentů o celkové délce $r(d)$ můžeme rozdělit na pravé prodloužení prvního z nich o délce P a na levé prodloužení následujícího



Obr. 8.5: Schéma segmentace signálu. Vstupní signál x (a) je rozdělen na segmenty stejné délky, s výjimkou posledního segmentu, který může být kratší (b); n -tý segment x označujeme symbolem $^n x$.

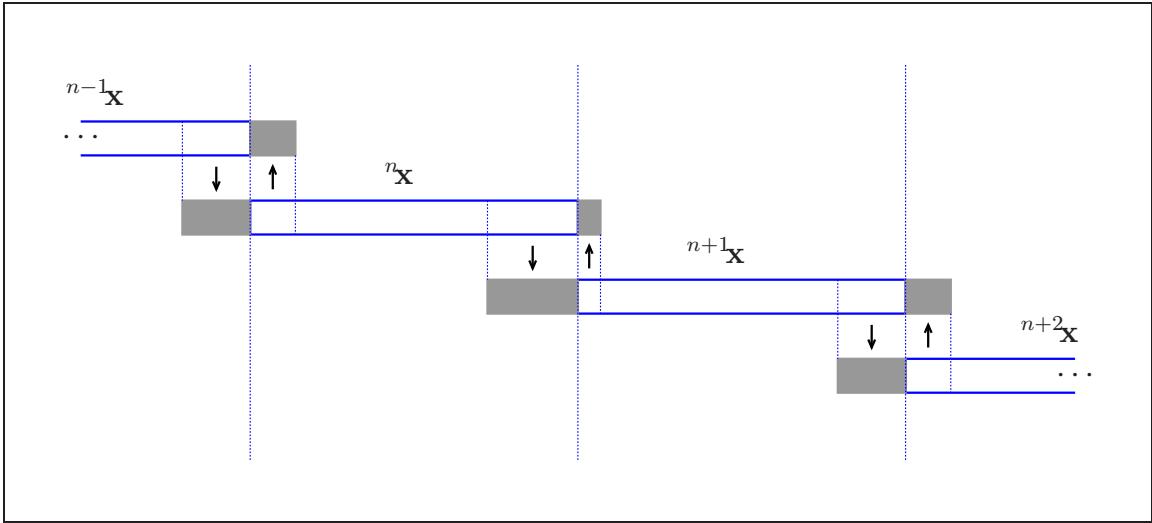
o délce L , přičemž samozřejmě platí $r(d) = P + L$. Čísla $L \geq 0, P \geq 0$ však nemůžeme volit libovolně. Tvrzení 8.11 říká, podle jakého pravidla mohou být L, P vybrána. Protože tato volba je obecně nejednoznačná, příkláníme se k volbě co největšího přesahu na levé straně každého segmentu. Snaha o co největší levé prodloužení přirozeně vyplývá z principu (potenciálního) využití algoritmu při zpracování signálu v reálném čase – usilujeme o to, aby algoritmus pracoval s co nejvíce vzorky, které již jsou známé, a tím zároveň s co nejméně vzorky, které ještě známy nejsou – teprve „přijdou“. Vzorec pro výpočet tohoto maximálního přesahu L_{\max} pak uvádí tvrzení 8.12.

Tvrzení 8.11: Nechť je dán segment, který má včetně svého levého prodloužení délku l . Pak délka levého překrytí v následujícím segmentu musí splňovat

$$L = l - 2^d(k-1) \quad \text{pro nějaké } k \in \left[\frac{l}{2^d} + 1 - \frac{r(d)}{2^d}, \frac{l}{2^d} + 1 \right], \quad k \in \mathbb{N}. \quad (8.9)$$

Délka pravého překrytí v daném segmentu je tedy $P = r(d) - L$.

Důkaz: Působnost prvního waveletového koeficientu z d -té úrovni 0 (tj. v úrovni vstupního signálu), vypočítaného z následujícího prodlouženého segmentu (zatím neznáme délku tohoto prodloužení, tu chceme určit), v tomto následujícím segmentu začíná na pozici 2^d (viz lemma 8.20). Působnost k -tého koeficientu z d -té úrovni vypočítaný z předcházejícího segmentu dle stejného lemmatu začíná v tomto segmentu na pozici $2^d k$. Aby na sebe koeficienty z obou segmentů v d -té úrovni navazovaly, je tedy třeba, aby délka levého



Obr. 8.6: Grafické znázornění prodloužování segmentů. Každý segment je prodloužen zleva o určitý počet posledních vzorků předchozího segmentu a zprava o určitý počet prvních vzorků segmentu následujícího. Délka levého a pravého prodloužení nemusí být pro každý segment stejná.

překrytí následujícího segmentu, L , právě splňovala

$$l - L + 2^d = 2^d k \quad (8.10)$$

neboli

$$L = l - 2^d(k - 1) \quad \text{pro nějaké } k \in \mathbb{N}. \quad (8.11)$$

Zároveň zřejmě musí platit

$$\begin{aligned} 0 &\leq L \leq r(d) \\ 0 &\leq l - 2^d(k - 1) \leq r(d) \\ \frac{l}{2^d} + 1 &\geq k \geq \frac{l - r(d)}{2^d} + 1. \end{aligned}$$

Tedy vyhovující k jsou všechna $k \in \mathbb{N}$, ležící v intervalu $\left[\frac{l}{2^d} + 1 - \frac{r(d)}{2^d}, \frac{l}{2^d} + 1 \right]$. \square

Tvrzení 8.12: Necht' je dán segment, který má včetně svého levého prodloužení délku l . Největší možné levé prodloužení následujícího segmentu, L_{\max} , se vypočítá vzorcem

$$L_{\max} = l - 2^d \operatorname{ceil}\left(\frac{l - r(d)}{2^d}\right). \quad (8.12)$$

Nejmenší pravé prodloužení daného segmentu je tedy zřejmě

$$P_{\min} = r(d) - L_{\max}. \quad (8.13)$$

Důkaz: Délka levého překrytí L definovaná v tvrzení 8.11 dosáhne své maximální přípustné hodnoty pro nejmenší celé $k \geq \frac{l}{2^d} + 1 - \frac{r(d)}{2^d}$, což znamená $k = \text{ceil}\left(\frac{l-r(d)}{2^d}\right) + 1$. Dosazením do (8.9) dostaváme tvrzení. \square

Pro účely dalšího textu je vhodné, opatříme-li proměnné L_{\max}, P_{\min}, l příznakem sobě příslušného segmentu, tj. n -tý segment bude mít levé prodloužení délky $L_{\max}(n)$, pravé prodloužení délky $P_{\min}(n)$ a délku výchozího n -tého segmentu včetně svého levého prodloužení budeme značit $l(n)$. Celkovou délku n -tého segmentu po prodloužení budeme značit $\sum(n)$. Zejména s tímto značením můžeme přepsat rovnici (8.13) jako

$$P_{\min}(n) = r(d) - L_{\max}(n+1). \quad (8.14)$$

Současně se zavedením tohoto označení je třeba poznamenat, že navzájem různé segmenty mají obecně různé délky svých pravých a levých prodloužení. Jak ale bude ukázáno dále, jistá pravidla je zde možno nalézt.

Ještě předtím ale rozebereme situaci prvního a posledního segmentu: Tyto segmenty přirozeně reprezentují „okraje“ signálu. Konečná diskrétní waveletová transformace zná několik způsobů, jak zacházet s okraji signálu – a tyto módy musíme zachovat i v našem modifikovaném algoritmu. S prvním a posledním segmentem musíme tedy zacházet zvlášť, poněkud odlišně od zbylých segmentů. Je třeba, abychom v naší nové metodě simulovali situaci klasické waveletové transformace, kdy v každé úrovni dekompozice prodlužujeme koeficienty před provedením konvoluce o $m-1$ vzorků, viz bod 2(a) algoritmu 8.1. To nám zaručí prodloužení prvního segmentu zleva o $r(d)$ nulových vzorků⁴, tj. $L_{\max}(1) = r(d)$, a postup jeho zpracování podle algoritmu 8.18. Podobně poslední segment je nutno zprava prodloužit o $r(d)$ nulových vzorků a zpracovat jej dle algoritmu 8.19. Algoritmy 8.18 a 8.19 jsou malé modifikace algoritmu 8.1 (DTWT).

Lemma 8.13: (o prodloužení prvního segmentu)

První segment (původní délky s) musí být zprava prodloužen o

$$P_{\min}(1) = 2^d \text{ceil}\left(\frac{s}{2^d}\right) - s \quad (8.15)$$

vzorků a má celkovou délku $\sum(1) = r(d) + 2^d \text{ceil}\left(\frac{s}{2^d}\right)$.

Důkaz: Zdůvodnění volby levého prodloužení $L_{\max}(1) = r(d)$ bylo uvedeno v předchozích

⁴to můžeme zdůvodnit jako nutnou délku překryvu s neexistujícím předchozím segmentem

odstavcích. Celkově má tedy původní segment a jeho levé prodloužení délku $l(1) = r(d) + s$.

Pak díky tvrzení 8.12 je možno vypočítat

$$L_{\max}(2) = l(1) - 2^d \operatorname{ceil} \left(\frac{l(1) - r(d)}{2^d} \right) = r(d) + s - 2^d \operatorname{ceil} \left(\frac{s}{2^d} \right) \quad (8.16)$$

a

$$P_{\min}(1) = r(d) - L_{\max}(2) = 2^d \operatorname{ceil} \left(\frac{s}{2^d} \right) - s. \quad (8.17)$$

Celková délka je tedy $L_{\max}(1) + s + P_{\min}(1) = r(d) + 2^d \operatorname{ceil} \left(\frac{s}{2^d} \right)$. \square

Tvrzení 8.14: Pro délku pravého prodloužení n -tého segmentu, $n = 1, 2, \dots, S-2$, platí

$$P_{\min}(n) = 2^d \operatorname{ceil} \left(\frac{ns}{2^d} \right) - ns \quad (8.18)$$

a pro levé prodloužení $(n+1)$ -ního segmentu platí

$$L_{\max}(n+1) = r(d) - P_{\min}(n) \quad (8.19)$$

Důkaz: Matematickou indukcí vzhledem k n . Pro $n = 1$ vztah platí, jak je evidentní z (8.15). Předpokládejme dále, že (8.18) platí pro zvolené $n \geq 1$. Pak díky (8.14) a (8.12) vypočítáme

$$L_{\max}(n+1) = r(d) - P_{\min}(n) \quad (8.20)$$

$$l(n+1) = s + L_{\max}(n+1) = s + r(d) - P_{\min}(n) \quad (8.21)$$

$$\begin{aligned} L_{\max}(n+2) &= l(n+1) - 2^d \operatorname{ceil} \left(\frac{l(n+1) - r(d)}{2^d} \right) \\ &= s + r(d) - P_{\min}(n) - 2^d \operatorname{ceil} \left(\frac{s - P_{\min}(n)}{2^d} \right) \\ &= s + r(d) - \left[2^d \operatorname{ceil} \left(\frac{ns}{2^d} \right) - ns \right] - 2^d \operatorname{ceil} \left(\frac{s - P_{\min}(n)}{2^d} \right) \\ &= (n+1)s + r(d) - 2^d \operatorname{ceil} \left(\frac{ns}{2^d} \right) - 2^d \operatorname{ceil} \left(\frac{s - [2^d \operatorname{ceil} \left(\frac{ns}{2^d} \right) - ns]}{2^d} \right) \\ &= (n+1)s + r(d) - 2^d \operatorname{ceil} \left(\frac{ns}{2^d} \right) - 2^d \operatorname{ceil} \left(\frac{(n+1)s - \operatorname{ceil} \left(\frac{ns}{2^d} \right)}{2^d} \right) \\ &= (n+1)s + r(d) - 2^d \operatorname{ceil} \left(\frac{(n+1)s}{2^d} \right) \\ P_{\min}(n+1) &= r(d) - L_{\max}(n+2) = 2^d \operatorname{ceil} \left(\frac{(n+1)s}{2^d} \right) - (n+1)s. \end{aligned} \quad \square$$

Poznámka: Ze vzorce (8.18) plyne, že P_{\min} je periodická vzhledem k s s periodou 2^d , tzn. $P_{\min}(n + 2^d) = P_{\min}(n)$. Zároveň z něj plyne, že pokud s je dělitelné 2^k , $k = 1, 2, \dots, d$, pak perioda P_{\min} klesá na 2^{d-k} . Speciálně tedy pokud s je dělitelné 2^d , perioda P_{\min} je 1 a délky pravých, resp. levých prodloužení jsou u všech segmentů shodné. Ze vzorce je rovněž jasné, že délka filtru m na délku periody P_{\min} nemá žádný vliv, ani na velikost P_{\min} , na velikost L_{\max} ale ano, neboť to závisí na $r(d)$, viz (8.19). Dále je také jasné, že také $\sum(n)$ má stejnou periodu jako P_{\min} a také, že $\sum(n)$ je na délce filtru také závislá, viz např. (8.23). Všechny tyto zákonitosti je možné vidět v tabulce 8.3.

Lemma 8.15: (celková délka segmentu)

V pořadí n -tého segmentu (původní délky s) má po prodloužení délku

$$\sum(n) = r(d) + 2^d \left[\text{ceil} \left(\frac{ns}{2^d} \right) - \text{ceil} \left(\frac{(n-1)s}{2^d} \right) \right]. \quad (8.22)$$

Tento výraz může nabývat pouze dvou hodnot, a to bud'

$$r(d) + 2^d \text{ceil} \left(\frac{s}{2^d} \right) \quad \text{nebo} \quad r(d) + 2^d \text{ceil} \left(\frac{s}{2^d} \right) - 2^d. \quad (8.23)$$

Důkaz: Délka n -tého segmentu je

$$\begin{aligned} \sum(n) &= L_{\max}(n) + s + P_{\min}(n) \\ &= r(d) - P_{\min}(n-1) + s + P_{\min}(n) \end{aligned}$$

pro $n = 1, 2, \dots$, přičemž ustanovíme $P_{\min}(0) = 0$. Uvedený výraz můžeme dále rozepsat podle (8.18) jako

$$\begin{aligned} \sum(n) &= r(d) - 2^d \text{ceil} \left(\frac{(n-1)s}{2^d} \right) + (n-1)s + s + 2^d \text{ceil} \left(\frac{ns}{2^d} \right) - ns \\ &= r(d) + 2^d \left[\text{ceil} \left(\frac{ns}{2^d} \right) - \text{ceil} \left(\frac{(n-1)s}{2^d} \right) \right]. \end{aligned} \quad (8.24)$$

Označme výraz v hranaté závorce jako $z(n, s, d)$. Tento výraz může nabývat různých hodnot v závislosti na dělitelnosti čísel $s, ns, (n-1)s$ číslem 2^d . Omezíme se pouze na důkaz, že tyto hodnoty jsou pouze dvě: Protože platí

$$\begin{aligned} \frac{ns}{2^d} &\leq \text{ceil} \left(\frac{ns}{2^d} \right) < \frac{ns}{2^d} + 1 \\ \frac{(n-1)s}{2^d} &\leq \text{ceil} \left(\frac{(n-1)s}{2^d} \right) < \frac{(n-1)s}{2^d} + 1, \end{aligned}$$

s	n	1	2	3	4	5	6	7	8	9	10	11	12	13	...
512	$L_{\max}(n)$	105	105	105	105	105	105	105	105	105	105	105	105	105	...
	$P_{\min}(n)$	0	0	0	0	0	0	0	0	0	0	0	0	0	...
	$\sum(n)$	617	617	617	617	617	617	617	617	617	617	617	617	617	...
513	$L_{\max}(n)$	105	98	99	100	101	102	103	104	105	98	99	100	101	...
	$P_{\min}(n)$	7	6	5	4	3	2	1	0	7	6	5	4	3	...
	$\sum(n)$	625	617	617	617	617	617	617	617	625	617	617	617	617	...
514	$L_{\max}(n)$	105	99	101	103	105	99	101	103	105	99	101	103	105	...
	$P_{\min}(n)$	6	4	2	0	6	4	2	0	6	4	2	0	6	...
	$\sum(n)$	625	617	617	617	625	617	617	617	625	617	617	617	625	...
515	$L_{\max}(n)$	105	100	103	98	101	104	99	102	105	100	103	98	101	...
	$P_{\min}(n)$	5	2	7	4	1	6	3	0	5	2	7	4	1	...
	$\sum(n)$	625	617	625	617	617	625	617	617	625	617	625	617	617	...
516	$L_{\max}(n)$	105	101	105	101	105	101	105	101	105	101	105	101	105	...
	$P_{\min}(n)$	4	0	4	0	4	0	4	0	4	0	4	0	4	...
	$\sum(n)$	625	617	625	617	625	617	625	617	625	617	625	617	625	...
517	$L_{\max}(n)$	105	102	99	104	101	98	103	100	105	102	99	104	101	...
	$P_{\min}(n)$	3	6	1	4	7	2	5	0	3	6	1	4	7	...
	$\sum(n)$	625	625	617	625	625	617	625	617	625	625	617	625	625	...
518	$L_{\max}(n)$	105	103	101	99	105	103	101	99	105	103	101	99	105	...
	$P_{\min}(n)$	2	4	6	0	2	4	6	0	2	4	6	0	2	...
	$\sum(n)$	625	625	625	617	625	625	625	617	625	625	625	617	625	...
519	$L_{\max}(n)$	105	104	103	102	101	100	99	98	105	104	103	102	101	...
	$P_{\min}(n)$	1	2	3	4	5	6	7	0	1	2	3	4	5	...
	$\sum(n)$	625	625	625	625	625	625	625	617	625	625	625	625	625	...
520	$L_{\max}(n)$	105	105	105	105	105	105	105	105	105	105	105	105	105	...
	$P_{\min}(n)$	0	0	0	0	0	0	0	0	0	0	0	0	0	...
	$\sum(n)$	625	625	625	625	625	625	625	625	625	625	625	625	625	...
521	$L_{\max}(n)$	105	98	99	100	101	102	103	104	105	98	99	100	101	...
	$P_{\min}(n)$	7	6	5	4	3	2	1	0	7	6	5	4	3	...
	$\sum(n)$	633	625	625	625	625	625	625	625	633	625	625	625	625	...
522	$L_{\max}(n)$	105	99	101	103	105	99	101	103	105	99	101	103	105	...
	$P_{\min}(n)$	6	4	2	0	6	4	2	0	6	4	2	0	6	...
	$\sum(n)$	633	625	625	625	633	625	625	625	633	625	625	625	633	...
523	$L_{\max}(n)$	105	100	103	98	101	104	99	102	105	100	103	98	101	...
	$P_{\min}(n)$	5	2	7	4	1	6	3	0	5	2	7	4	1	...
	$\sum(n)$	633	625	633	625	625	633	625	625	633	625	633	625	625	...
524	$L_{\max}(n)$	105	101	105	101	105	101	105	101	105	101	105	101	105	...
	$P_{\min}(n)$	4	0	4	0	4	0	4	0	4	0	4	0	4	...
	$\sum(n)$	633	625	633	625	633	625	633	625	633	625	633	625	633	...

Tab. 8.3: Velikosti prodloužení segmentů pro různé délky segmentů s při hloubce dekompozice $d = 3$ a délce filtru $m = 16$. Z tabulky je patrná periodicitu levých a pravých prodloužení a tím také táž periodicitu celkové délky prodlouženého segmentu.

platí pro $z(n, s, d)$:

$$\frac{s}{2^d} - 1 < z(n, s, d) < \frac{s}{2^d} + 1, \quad (8.25)$$

a protože $z(n, s, d) \in \mathbb{Z}$, může nabývat pouze hodnoty $\text{ceil}\left(\frac{s}{2^d}\right)$ nebo $\text{ceil}\left(\frac{s}{2^d}\right) - 1$. Dosazením $z(n, s, d)$ zpět do rovnice (8.24) získáváme druhou část tvrzení. \square

Zákonitosti uvedené v tomto lemmatu jsou rovněž patrné v tabulce 8.3.

8.3.1 Algoritmus SegWT

V tomto oddílu navazujeme na předchozí úvahy uvedením samotného algoritmu SegWT. Algoritmus pracuje tak, že postupně načítá (přijímá) segmenty vstupního signálu, vhodným způsobem je prodlužuje – dalo by se říct navzájem překrývá, z těchto prodloužených segmentů pak počítá modifikovaným způsobem waveletové koeficienty, které nakonec jednoduchým způsobem spojuje dohromady. S prvním a posledním segmentem, jak již bylo vysvětleno výše, musí algoritmus zacházet odlišně od ostatních segmentů.

Algoritmus 8.16: Jsou dány waveletové filtry g, h délky m , hloubka dekompozice d a typ prodloužení na okrajích signálu. Dále jsou dány segmenty délky $s > 0$ vstupního signálu x , které označíme ${}^1x, {}^2x, {}^3x, \dots$. Poslední segment může mít délku menší než s .

1. Nastavíme $N = 1$.
2. Načteme první segment vstupního signálu, 1x , a označíme jej za „aktuální“. Prodloužíme jej zleva o $r(d)$ nulových vzorků.
3. **Pokud** první segment je zároveň posledním
 - (a) Jde o klasickou waveletovou transformaci. Vypočítáme DTWT tohoto jediného segmentu algoritmem 8.1.
 - (b) Algoritmus končí.
4. Načteme $(N + 1)$ -ní segment a označíme ho za „následující“.
5. **Pokud** tento segment je posledním segmentem
 - (a) Spojíme aktuální a následující segment dohromady a označíme tento vektor jako aktuální. (Aktuální se tedy stává posledním.)
 - (b) Prodloužíme aktuální vektor zprava o $r(d)$ nulových vzorků.

- (c) Vypočítáme DTWT hloubky d z prodlouženého aktuálního segmentu podle algoritmu 8.19.

Jinak

- (d) Spočítáme L_{\max} pro následující segment a P_{\min} pro aktuální segment (viz tvrzení 8.12).
- (e) Prodloužíme aktuální segment zprava o P_{\min} vzorků z následujícího segmentu. Prodloužíme následující segment zleva o L_{\max} vzorků z aktuálního segmentu.
- (f) **Pokud** aktuální segment je první segment, vypočítáme DTWT hloubky d z prodlouženého aktuálního segmentu podle algoritmu 8.18. **Jinak** vypočítáme DTWT hloubky d z prodlouženého aktuálního segmentu podle algoritmu 8.17.
6. Upravíme jednotlivé vektory s vypočtenými waveletovými koeficienty tak, že odstraníme zleva určitý počet redundantních (přebytečných) koeficientů, konkrétně takto: v k -té úrovni, $k = 1, 2, \dots, d - 1$ odstraníme zleva $r(d - k)$ koeficientů.
 7. **Pokud** je aktuální segment posledním, pak stejným způsobem jako v předchozím bodu odstraníme redundantní koeficienty, tentokrát však zprava.
 8. Výsledek uložíme jako ${}^N\mathbf{a}^{(d)}, {}^N\mathbf{d}^{(d)}, {}^N\mathbf{d}^{(d-1)}, \dots, {}^N\mathbf{d}^{(1)}$.
 9. **Pokud** aktuální segment není posledním
 - (a) Označíme následující segment jako aktuální.
 - (b) Zvýšíme N o 1 a vrátíme se na bod 4.

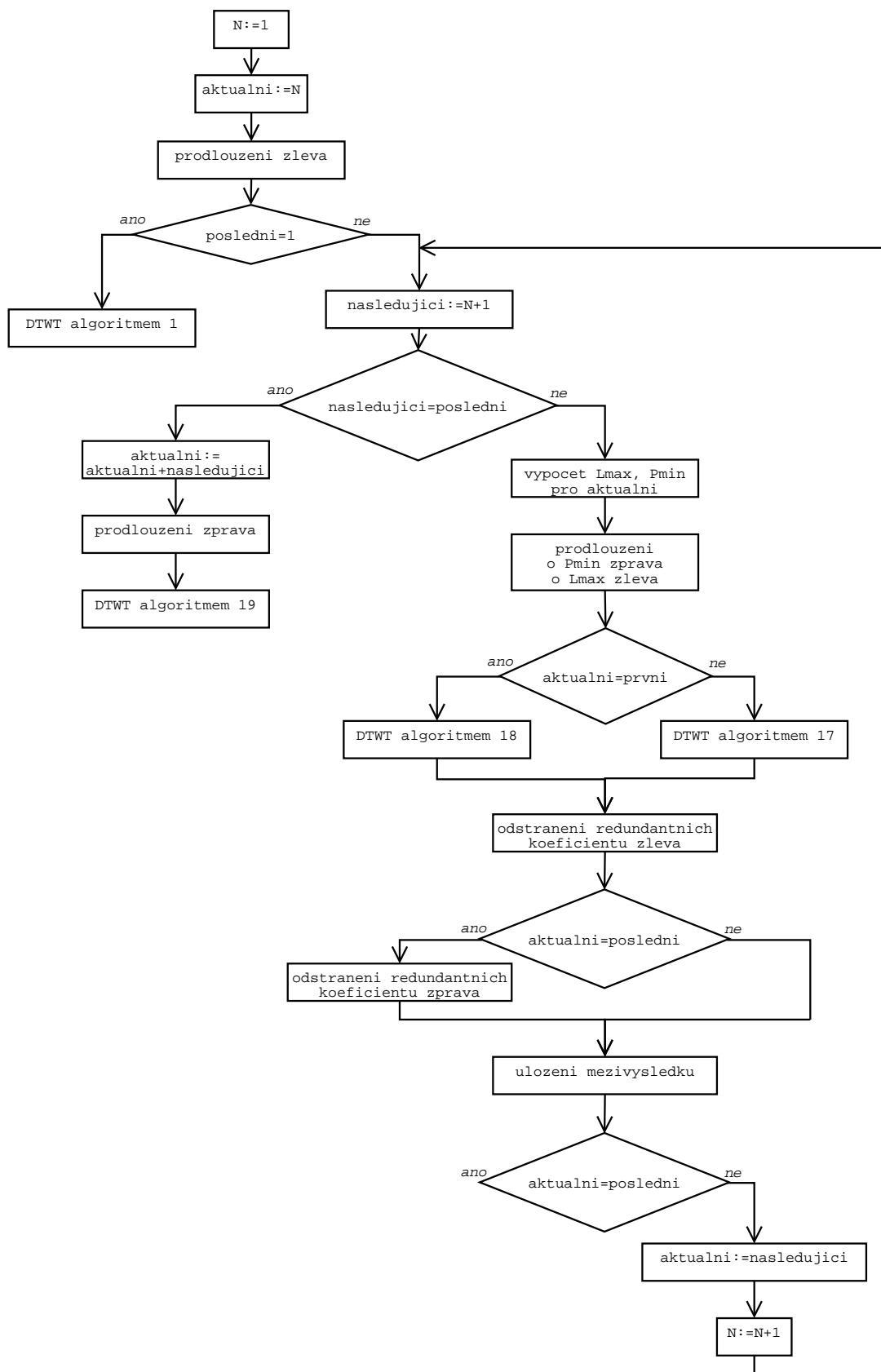
Schématický diagram algoritmu je na obrázku 8.7.

Poznámky:

1. Pokud vstupní signál byl rozdělen na $S > 1$ segmentů, pak výstupem algoritmu je celkem $(S - 1)(d + 1)$ vektorů s waveletovými koeficienty:

$$\{ {}^i\mathbf{a}^{(d)}, {}^i\mathbf{d}^{(d)}, {}^i\mathbf{d}^{(d-1)}, \dots, {}^i\mathbf{d}^{(1)} \}_{i=1}^{S-1}.$$

Číslo $(S - 1)$ je zde proto, že poslední dva segmenty jsou v algoritmu vždy počítány jako jeden segment.



Obr. 8.7: Vývojové schéma algoritmu 8.16.

Spojíme-li tyto vektory dohromady, tzn. za sebe seřadíme vektory approximačních koeficientů ${}^1\mathbf{a}^{(d)}, {}^2\mathbf{a}^{(d)}, \dots, {}^{S-1}\mathbf{a}^{(d)}$ a obdobně seřadíme vektory všech detailních koeficientů na stejně úrovni dekompozice, dostáváme tím soubor $d+1$ vektorů, které jsou totožné s waveletovými koeficienty signálu \mathbf{x} .

2. Ad 5(a): sloučení posledních dvou segmentů do jednoho je nutné udělat, neboť nemáme jistotu, že poslední segment je dostatečně dlouhý, aby z něj bylo možné vypočítat alespoň jeden waveletový koeficient hloubky d .

Dále uvádíme „vnitřní“ algoritmy metody SegWT, na které se odvolává „hlavní“ algoritmus 8.16.

Algoritmus 8.17: Tento algoritmus je totožný s algoritmem 8.1, s tou výjimkou, že zde neprovádíme krok 2(a), tj. prodlužování vektoru.

Vynechání kroku 2(a) je záměrem, neboť dostatečné prodloužení (celkové délky $\sum(n)$ namísto $(m-1)$) se v algoritmu 8.16 provede již před aplikací výpočtu transformace na prodloužený segment. Vyseknutí užitečné části koeficientů je v algoritmu SegWT implementováno až později, v obecnější podobě, v kroku 6.

Následující dva algoritmy využíváme pro zpracování prvního, resp. posledního segmentu. Zajišťují, aby tyto segmenty byly zpracovány včetně zvoleného typu prodloužení na okrajích signálu. Protože toto prodloužení je nutné provést v každé úrovni dekompozice, využíváme algoritmu 8.1, kde modifikujeme krok 2(a), který se týká právě prodloužení:

Algoritmus 8.18: Algoritmus je totožný s algoritmem 8.1, s tou výjimkou, že v tomto algoritmu nahradíme krok 2(a) krokem:

„Změníme koeficienty vektoru $\mathbf{a}^{(j)}$ na pozicích $r(d-j)-m+2, \dots, r(d-j)$, tak, jak to odpovídá typu prodloužení.“

Algoritmus 8.19: Algoritmus je totožný s algoritmem 8.1, s tou výjimkou, že v tomto algoritmu nahradíme krok 2(a) krokem:

„Změníme koeficienty vektoru $\mathbf{a}^{(j)}$ na pozicích $r(d-j)-m+2, \dots, r(d-j)$, nyní ale počítaných z pravé strany $\mathbf{a}^{(j)}$, tak, jak to odpovídá typu prodloužení.“

Poznámka: Při určování, které koeficienty změnit v případě algoritmu 8.19, nevíme předem, jak dlouhý bude poslední sloučený segment včetně prodloužení. Kvůli tomu počítáme

v algoritmu 8.19 indexy koeficientů, které máme měnit, z pravé strany.

8.3.2 Důsledky a omezení algoritmu SegWT

V tomto oddílu bude odvozeno, kolik koeficientů jsme schopni z každého segmentu vypočítat metodou SegWT a dále bude odvozena minimální délka segmentu.

V následujících dvou tvrzeních předpokládáme, že vstupní signál zpracováváme algoritmem 8.17. Ve všech tvrzeních pak uvažujeme libovolnou délku prodloužení, jednoduše řečeno, dovolujeme si prodloužit vstupní signál $\mathbf{x} = \mathbf{a}^{(0)}$ z pravé a z levé strany o libovolný počet vzorků.

Lemma 8.20: *Necht' $\tilde{\mathbf{x}}$ je nějaké prodloužení vstupního vektoru \mathbf{x} . Pak působnost n -tého waveletového koeficientu v řadě z k -té úrovni, pokud takový lze z $\tilde{\mathbf{x}}$ vypočítat, začíná ve vektoru $\tilde{\mathbf{x}}$ na pozici (vzorku) $2^k n$ a končí na pozici $[2^k n + (2^k - 1)(m - 1)]$.*

Důkaz: Pro $n = 1$, tzn. pro první koeficient v k -té úrovni začíná působnost v prodlouženém signálu $\tilde{\mathbf{x}}$ díky sudé decimaci na pozici 2^k . Obecně působnost n -tého začíná v $\tilde{\mathbf{x}}$ na pozici $2^k + 2^k(n - 1) = 2^k n$. Konec působnosti pak určíme jednoduše s využitím důsledku 8.5 jako $[2^k n + p_1(k) - 1] = 2^k n + (2^k - 1)(m - 1)$. \square

Lemma 8.21: *Necht' $\tilde{\mathbf{x}}$ je nějaké prodloužení vstupního vektoru \mathbf{x} . Nechť' l je délka $\tilde{\mathbf{x}}$. Pak z takto prodlouženého vektoru \mathbf{x} je možno vypočítat nejvýše q_{\max} waveletových koeficientů úrovni k , kde q_{\max} je největší $q \in \{0, 1, 2, \dots\}$, splňující nerovnost*

$$q \leq 2^{-k}l + (2^{-k} - 1)(m - 1). \quad (8.26)$$

Důkaz: Z lemmatu 8.20 víme, že v k -té úrovni působnost q -tého waveletového koeficientu v řadě končí v prodlouženém signálu $\tilde{\mathbf{x}}$ na pozici $[2^k q + (2^k - 1)(m - 1)]$. Stačí tedy spočítat, pro jaké nejvyšší q je splněna nerovnost

$$\begin{aligned} 2^k q + (2^k - 1)(m - 1) &\leq l \\ 2^k q &\leq l - (2^k - 1)(m - 1) \\ q &\leq 2^{-k}[l - (2^k - 1)(m - 1)] \\ q &\leq 2^{-k}l + (2^{-k} - 1)(m - 1). \end{aligned} \quad \square$$

Důkaz lemmatu 8.8: Provedení algoritmu DTWT (algoritmus 8.1) o hloubce k na signálu \mathbf{x} o délce v odpovídá provedení algoritmu 8.17 na signálu $\tilde{\mathbf{x}}$, který je z \mathbf{x} vytvořen prodloužením zleva i zprava o $r(k)$ vzorků. V lemmatu 8.8 jde pouze o počet koeficientů, nikoliv o jejich hodnoty, proto můžeme nyní zanedbat typ prodloužení okrajů signálu. Proto počet koeficientů získaných algoritmem 8.1 získáme užitím lemmatu 8.21 s volbou $l = v + 2r(k)$. Počet koeficientů je pak nejvyšší přirozené q splňující

$$\begin{aligned} q &\leq 2^{-k}[v + 2r(k)] + (2^{-k} - 1)(m - 1) \\ q &\leq 2^{-k}[v + 2(2^k - 1)(m - 1)] + (2^{-k} - 1)(m - 1) \\ q &\leq 2^{-k}v + 2^{-k+1}(2^k - 1)(m - 1) + (2^{-k} - 1)(m - 1) \\ q &\leq 2^{-k}v + (m - 1)[2^{-k+1}(2^k - 1) + (2^{-k} - 1)] \\ q &\leq 2^{-k}v - (m - 1)(2^{-k} - 1), \end{aligned}$$

což je právě číslo floor $(2^{-k}v - (2^{-k} - 1)(m - 1))$. \square

Tvrzení 8.22: Nechť je dána hloubka dekompozice d a nechť ${}^n\tilde{\mathbf{x}}$ je prodloužení n -tého segmentu ${}^n\mathbf{x}$ původní délky s . Pak ze segmentu \mathbf{x} vypočítáme algoritmem 8.17

$$q_{\max}(n) = \begin{cases} \text{ceil}\left(\frac{s}{2^d}\right) & \text{pro } \sum(n) = r(d) + 2^d \text{ceil}\left(\frac{s}{2^d}\right) \\ \text{ceil}\left(\frac{s}{2^d}\right) - 1 & \text{pro } \sum(n) = r(d) + 2^d \text{ceil}\left(\frac{s}{2^d}\right) - 2^d \end{cases} \quad (8.27)$$

waveletových koeficientů úrovně d .

Důkaz: Užitím lemmatu 8.21 při volbě $k = d$ a dosazením $l = \sum(n) = r(d) + 2^d \text{ceil}\left(\frac{s}{2^d}\right)$:

$$\begin{aligned} q &\leq 2^{-d}l + (2^{-d} - 1)(m - 1) \\ q &\leq 2^{-d}r(d) + \text{ceil}\left(\frac{s}{2^d}\right) + (2^{-d} - 1)(m - 1) \\ q &\leq 2^{-d}(2^d - 1)(m - 1) + \text{ceil}\left(\frac{s}{2^d}\right) + (2^{-d} - 1)(m - 1) \\ q &\leq [2^{-d}(2^d - 1) + (2^{-d} - 1)](m - 1) + \text{ceil}\left(\frac{s}{2^d}\right) \\ q &\leq \text{ceil}\left(\frac{s}{2^d}\right) \end{aligned}$$

Největší přirozené q splňující tuto nerovnost je zřejmě $q_{\max} = \text{ceil}\left(\frac{s}{2^d}\right)$.

V případě kratšího segmentu, $l = \sum(n) = r(d) + 2^d \text{ceil}\left(\frac{s}{2^d}\right) - 2^d$, vyjde analogickým postupem $q_{\max} = \text{ceil}\left(\frac{s}{2^d}\right) - 1$. \square

Jako přirozený se jeví požadavek, aby z každého segmentu bylo možné vypočítat alespoň jeden waveletový koeficient maximální hloubky. To vede na následující tvrzení.

Důsledek 8.23: (minimální délka segmentu)

Nechť je dána hloubka dekompozice d . Za předpokladu $S > 2^d + 1$ musí délka neprodloženého segmentu, s , splňovat podmínu $s > 2^d$.

Důkaz: Za předpokladu $S > 2^d + 1$ jistě nastane v průběhu algoritmu 8.16 situace, že budeme zpracovávat segment o délce včetně prodloužení $\sum = r(d) + 2^d \text{ceil}(\frac{s}{2^d}) - 2^d$, viz (8.23). Tvrzení 8.22 říká, kolik koeficientů jsme z tohoto segmentu schopni vypočítat.

Pro toto číslo q_{\max} přirozeně požadujeme

$$\begin{aligned} q_{\max} &\geq 1 \\ \text{ceil}\left(\frac{s}{2^d}\right) - 1 &\geq 1 \\ \frac{s}{2^d} &> 1 \\ s &> 2^d. \quad \square \end{aligned}$$

Poznámky:

1. V případě $S \leq 2^d + 1$ nemáme jistotu, že budeme v průběhu algoritmu pracovat se segmentem o celkové délce $\sum = r(d) + 2^d \text{ceil}(\frac{s}{2^d}) - 2^d$.
2. Všimněme si, že minimální délka segmentu nezávisí na délce waveletového filtru m .

Co se týká zpoždění, z popisu algoritmu 8.16 je zřejmé, že ve chvíli, kdy je znám segment číslo n a P_{\min} koeficientů segmentu číslo $n + 1$, je možno vypočítat waveletové koeficienty příslušející segmentu n . Vzhledem k tomu, že algoritmus SegWT vychází z předpokladu segmentového zpracování signálu, to znamená, že zpoždění algoritmu SegWT je jeden segment (tzn. s příchozích vzorků) plus čas potřebný na výpočet waveletových koeficientů z aktuálního segmentu. Ovšem ve speciálním případě, kdy s je dělitelné číslem 2^d , však platí dokonce $P_{\min}(n) = 0$ pro všechna $n \in \mathbb{N}$ (viz tvrzení 8.14), a to znamená, že zpoždění algoritmu je determinováno pouze délkou výpočtu!

Algoritmus 8.16 byl vyvinut i k obecnějšímu použití pro nestejnou délku segmentů. V tomto případě samozřejmě ztrácíme některé informace, které jsme v případě shodné délky segmentů mohli vypočítat předem a využít pak pro zrychlení algoritmu; např. ztrácíme periodicitu levých a pravých prodloužení segmentů.

Vzhledem k tomu, že základní krok transformace typu wavelet packet je totožný se základním krokem klasické waveletové transformace, je zřejmé, že algoritmus je možno téměř bez úprav využít i pro tento typ transformace.

8.3.3 Algoritmus SegWT při prodloužení typu „zero padding“

Jak bylo uvedeno, při waveletové transformaci hraje důležitou úlohu typ prodloužení okrajů signálu. Ve speciálním případě tzv. zero padding neboli prodloužení nulami je možné algoritmus 8.16 výrazně zjednodušit. Schéma upraveného algoritmu je uvedeno na obrázku 8.8.

Zjednodušení je umožněno tím, že při transformaci okrajových segmentů signálu provádíme pouze prodloužení pouze nulami. Díky tomu v algoritmu z obr. 8.8 již není třeba využívat výpočty pomocí algoritmů 8.1, 8.18, 8.19, ale pouze pomocí jediného algoritmu 8.17.

Uvedený algoritmus je implementován ve funkci `segwt.m`, která je k nalezení na přiloženém médiu. Tato funkce na základě parametrů s a d provede waveletovou transformaci signálu po segmentech a na výstupu dává stejné vektory waveletových koeficientů, jaké bychom vypočítali klasickým dekompozičním pyramidovým algoritmem při prodloužení „zero padding“⁵.

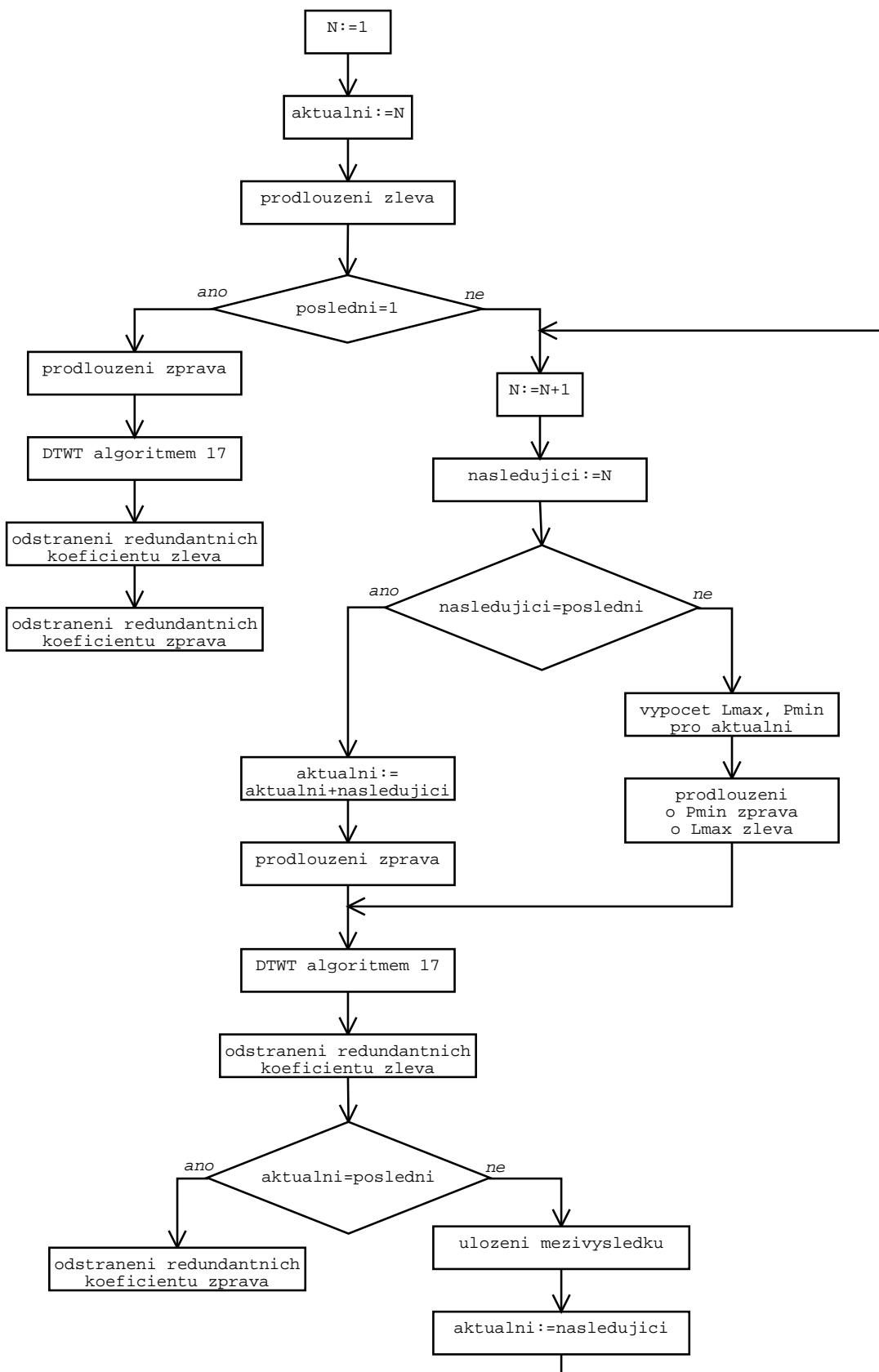
8.4 Závěr

Obsahem této kapitoly je algoritmus, který umožnuje provádět waveletovou transformaci diskrétních signálů v reálném čase. Algoritmus pracuje na principu určení optimální způsobu překrývání signálových segmentů a následném provedení modifikovaného transformačního postupu.

Uvedeme na závěr ještě několik poznámek ohledně možného budoucího vylepšení představeného algoritmu.

Vzhledem k výše uvedené nevýhodě redundance vypočtených koeficientů se nabízí možnost pokusit se tuto redundanci odstranit. V principu by šlo o to, neprodlužovat v úrovni $k = 0$ segment o $r(d)$ vzorků, nýbrž namísto toho v každé úrovni ($k = 0, \dots, d - 1$) pouze

⁵ve Wavelet Toolboxu Matlabu do tohoto módu přejdeme příkazem `dwtmode('zpd')`



Obr. 8.8: Schéma algoritmu SegWT při prodloužení nulovými hodnotami („zero padding“).

o nutných ($m - 1$) vzorků, resp. koeficientů. Toto sice sníží celkovou výpočetní náročnost algoritmu, nicméně tento přístup má také své nevýhody, a to jednak, že algoritmus nutně zesložití, a dále, že vzroste paměťová náročnost algoritmu, neboť vypočtené koeficienty si musí algoritmus „zapamatovat“ do doby zpracování dalšího segmentu.

Pokud bychom netrvali na provádění algoritmu v reálném čase, ale kritériem optimality by zůstala naprostá eliminace redundantních výpočtů, lze metodu použít iterativně pokaždé pouze s hloubkou $d = 1$; poté, co po segmentech vypočítáme waveletové koeficienty první úrovně, opakujeme segmentovanou metodu na těchto koeficientech atd., až se dostaneme do původně požadované hloubky dekompozice. Tím vlastně napodobujeme chování Mallatova algoritmu, přičemž hospodárně využíváme paměťové prostředky.

V budoucnu je třeba pro úplnost doplnit algoritmus, který je schopen po segmentech provádět také inverzní waveletovou transformaci. Po základních úvahách se zdá, že tento úkol zdaleka nebude tak komplikovaný, neboť již bude možné využít aparátu uvedeného v této kapitole. Po odvození tohoto algoritmu bude pak možno skutečně zpracovávat signál modifikací jeho waveletových koeficientů v reálném čase.

Pro budoucnost se rovněž nabízí šance zobecnit algoritmus SegWT také pro biortogonální wavelety⁶ a pro obecnější způsoby decimace [19, 30].

⁶algoritmus umožňuje použití libovolné, tedy i liché délky filtru m , takže v tomto smyslu je SegWT pro biortogonální wavelety připraven

Závěr

V dizertační práci byla prezentována metoda potlačování šumu obsaženého v signálu pomocí waveletové transformace a některé její modifikace provedené za účelem zlepšení separačních vlastností. Rovněž byla uvedena nová metoda tzv. segmentované waveletové transformace.

Po uvedení do problematiky a stanovení cílů dizertační práce byly v kapitolách 4 a 5 uvedeny základy waveletové analýzy a teorie tzv. waveletových paketů, pomocí nichž lze nalézt „nejlepší“ bázi, která nejvíce odpovídá konkrétnímu analyzovanému signálu.

Těžiště vědeckého přínosu práce tvoří kapitoly 6 a 8.

V kapitole 6 byl vysvětlen známý princip waveletového odšumování signálů a byly uvedeny používané prahovací techniky včetně jejich statistických vlastností (pomocí analýzy tzv. vychýlení, rozptylu a rizikové funkce) a tyto vlastnosti byly porovnány. Srovnáním dvou hlavních typů prahování, tvrdého a měkkého, byly vyvozeny závěry, že tvrdé prahování hůře potlačuje šum, ale lépe zachovává rychlé změny v signálu (skoky a špičky), zatímco měkké prahování je velmi úspěšné při potlačování šumu, ale má tendenci „přehladit“ skoky a špičky (např. v obraze se jedná o hrany objektů). Víc než jen kompromis tvoří poloměkké prahování, které má při správné volbě dvou parametrů nejnižší riziko chyby. Ovšem dva parametry metody zároveň tvoří velkou nevýhodu oproti jiným metodám, které pracují pouze s parametrem jedním. Další uvedené metody, a to nezáporná garota a hyperbolické prahování, se snaží tuto diskvalifikaci změnit. Pracují pouze s jedním parametrem, avšak ve statistickém smyslu nedosahují takové kvality jako poloměkké prahování.

Toto srovnání prahovacích pravidel pomocí matematické statistiky se nám podařilo publikovat na významných zahraničních konferencích [58, 65, 66], na tuzemských konferencích [59, 60], odborných seminářích [70] a také v řadě výzkumných zpráv.

Na základě testování waveletových metod na řečových signálech (kap. 7) se ukázalo, že waveletové potlačování šumu vykazuje při správně nastavených parametrech velmi dobré výsledky, zejména při hodnotách SNR přesahujících 0 dB. V objektivním i subjektivním srovnání přitom metoda nejlepší báze waveletového paketu (WP) předčila metodu prahování v klasické waveletové transformaci (WT), ovšem za cenu zvýšení výpočetní

náročnosti. Nahrávky separované pomocí WP obsahují mnohem méně rušivých vjemů a jsou srozumitelnější. V rozsahu SNR od ∞ přibližně do 3 dB se jako nejvhodnější pro řečový signál jeví hyperbolické prahovací pravidlo. Pro nižší odstup signálu od šumu je pak vhodnější použít tvrdé prahování – přes všechny rušivé artefakty poskytuje nejlepší srozumitelnost. V navazujícím výzkumu bude vhodné implementovat také algoritmus nalezení optimálních prahů λ_1, λ_2 pro poloměkké prahování.

Praktickým testováním se potvrdily závěry teoretické analýzy kapitoly 6, že měkké prahování nejlépe potlačuje šum na úkor zachování detailů, a tvrdé prahování naopak rychlé změny v signálu zachovává při současné nižší míře účinnosti separace. Měkké prahování se však ukázalo být nevhodným pro řečový signál, neboť potlačuje složky řeči, které jsou podstatné pro její srozumitelnost.

V kapitole 8 byla prezentována zcela nová metoda segmentované waveletové transformace, SegWT, která umožňuje zpracovávat signál po segmentech zvolené délky, a tudíž umožňuje jakýkoliv typ waveletového zpracování signálů v reálném čase. To má velký význam pro budoucí vývoj v širokém spektru aplikací, jako např. pro separaci řeči a šumu při komunikaci v reálném čase, zpracování hudebních signálů pomocí systému plug-in modulů nebo pro modulace waveletového typu, které se experimentálně začínají používat v systémech ADSL a SDR⁷.

Vývoj metody SegWT započal v říjnu roku 2003 a současná verze algoritmu vznikla v květnu 2004, a to je důvodem, proč tato část práce zatím nebyla publikována. Stručný popis metody bude prezentován ve sborníku [69] a rozšířenou verzi na některé významné konferenci nebo v mezinárodním časopise.

V rámci dizertační práce byl rovněž vytvořen programový balík ThreshLab pro prostředí MATLAB. Tento balík obsahuje funkce pro zpracování signálů pomocí prahování, pro numerickou i grafickou analýzu prahovacích pravidel, pro provedení a zobrazení MRA atd.

V budoucím výzkumu bude vhodné se zaměřit, jako pokračování této dizertační práce, na metody obecného návrhu waveletových bazí [56], na tzv. frejmy (frames) [9], biortogonální systémy [40], nebo dokonce na reprezentaci signálu v neúplných systémech [47]. Předmětem zájmu by mohlo být také přenesení waveletových metod separace na směsi

⁷Software-defined Radio, systém, kdy veškeré zpracování signálu pro jeho vysílání a přijímání probíhá softwarovými prostředky, ať už v počítači nebo např. na signálovém procesoru

signálu se šumem, který již nemá gaussovské rozdělení (barevné šumy apod.). Rovněž hodláme optimalizovat algoritmus SegWT, jelikož, jak bylo uvedeno v závěru kapitoly 8, jeho současná verze obsahuje duplicitu výpočtů na přechodech jednotlivých segmentů.

Přílohy

A Důkazy tvrzení

Důkaz tvrzení 6.1: Ukážeme, že součet rozptylu a kvadrátu výchylky je rizikovou funkcí:

$$\begin{aligned}
 V_\lambda(b) + [M_\lambda(b) - b]^2 &= E[\delta(D) - M_\lambda(b)]^2 + [M_\lambda(b) - b]^2 = \\
 &= E[\delta(D)]^2 - 2M_\lambda(b)E\delta(D) + [M_\lambda(b)]^2 + [M_\lambda(b)]^2 - 2bM_\lambda(b) + b^2 = \\
 &= E[\delta(D)]^2 - 2[M_\lambda(b)]^2 + 2[M_\lambda(b)]^2 - 2bM_\lambda(b) + b^2 = \\
 &= E[\delta(D)]^2 - 2bM_\lambda(b) + b^2 = E[\delta(D)]^2 - 2bE\delta(D) + b^2 = \\
 &= E[\delta(D) - b]^2 = R_\lambda(b). \quad \square
 \end{aligned}$$

Důkaz tvrzení 6.2: Nechť $X \sim N(b, 1)$ a $Y = \delta(X, \lambda)$ je transformovaná náhodná veličina. Všechny vzorce lze odvodit přímým výpočtem z definice střední hodnoty EY , rozptylu $\text{var}Y = E(Y - EY)^2$ a rizika $E(Y - b)^2$.

Dále ukážeme odvození např. vztahu pro střední hodnotu tvrdého prahování $M_\lambda^h(b)$. Vzorce pro hyperbolické prahování jsou uvedeny zvlášť v příloze B. $EY = M_\lambda^h(b)$ vypočítáme jako

$$\begin{aligned}
 EY &= E(Y\chi_{\{x| |x|\leq\lambda\}} + Y\chi_{\{x| |x|>\lambda\}}) = E(Y\chi_{\{x| |x|\leq\lambda\}}) + E(Y\chi_{\{x| |x|>\lambda\}}) \\
 &= E(Y\chi_{\{x| |x|>\lambda\}}) = E(X\chi_{\{x| |x|>\lambda\}}) \\
 &= \int_{-\infty}^{\infty} x\varphi(x-b) dx - \int_{-\lambda}^{\lambda} x\varphi(x-b) dx \\
 &= \int_{-\infty}^{\infty} (x+b)\varphi(x) dx - \int_{-\lambda-b}^{\lambda-b} (x+b)\varphi(x) dx.
 \end{aligned}$$

První integrál

$$\int_{-\infty}^{\infty} (x+b)\varphi(x) dx = \int_{-\infty}^{\infty} x\varphi(x) dx + b \int_{-\infty}^{\infty} \varphi(x) dx = 0 + b \cdot 1 = b$$

a druhý integrál

$$\int_{-\lambda-b}^{\lambda-b} (x+b)\varphi(x) dx = \int_{-\lambda-b}^{\lambda-b} x\varphi(x) dx + b \int_{-\lambda-b}^{\lambda-b} \varphi(x) dx$$

$$\begin{aligned}
&= - \int_{-\lambda-b}^{\lambda-b} \frac{d\varphi(x)}{dx} dx + b [\Phi(x)]_{-\lambda-b}^{\lambda-b} = - [\varphi(x)]_{-\lambda-b}^{\lambda-b} + b [\Phi(x)]_{-\lambda-b}^{\lambda-b} \\
&= - [\varphi(\lambda-b) - \varphi(-\lambda-b)] + b [\Phi(\lambda-b) - \Phi(-\lambda-b)] \\
&= [\varphi(\lambda+b) - \varphi(\lambda-b)] + b [\Phi(\lambda-b) - 1 + \Phi(\lambda+b)]
\end{aligned}$$

s využitím $\frac{d\varphi(x)}{dx} = -x \varphi(x)$, $\Phi(-x) = 1 - \Phi(x)$ a $\varphi(-x) = \varphi(x)$. Pak EY je dohromady rovno

$$EY = b + b[1 - \Phi(\lambda-b) - \Phi(\lambda+b)] + \varphi(\lambda-b) - \varphi(\lambda+b).$$

Obdobně ostatní výpočty vedou na (již složitější) integrály obsahující standardizované normální rozložení, z nichž některé (nezáporná garota, hyperbolické prahování) ani nemají analytický tvar a musejí být vypočítány iterativně numericky. \square

Důkaz tvrzení 6.7: Platnost bodu 1. je zřejmá a následující body z něj plynou takto:

$$2. M_\lambda(b) = E[\delta(D, \lambda)] = \sigma E[\delta(\frac{D}{\sigma}, \frac{\lambda}{\sigma})] = \sigma M_{\lambda/\sigma}(\frac{b}{\sigma}),$$

$$3. V_\lambda(b) = \text{var}[\delta(D, \lambda)] = \sigma^2 \text{var}[\delta(\frac{D}{\sigma}, \frac{\lambda}{\sigma})] = \sigma^2 V_{\lambda/\sigma}(\frac{b}{\sigma}),$$

$$\begin{aligned}
4. R_\lambda(b) &= V_\lambda(b) + [M_\lambda(b) - b]^2 = \sigma^2 V_{\lambda/\sigma}(\frac{b}{\sigma}) + [\sigma M_{\lambda/\sigma}(\frac{b}{\sigma}) - b]^2 = \\
&= \sigma^2 \left\{ V_{\lambda/\sigma}(\frac{b}{\sigma}) + \left[M_{\lambda/\sigma}(\frac{b}{\sigma}) - \frac{b}{\sigma} \right]^2 \right\} = \sigma^2 R_{\lambda/\sigma}(\frac{b}{\sigma}) \quad \text{s využitím 2. a 3.} \quad \square
\end{aligned}$$

B Odvození $M_\lambda^{\text{hy}}(b)$, $V_\lambda^{\text{hy}}(b)$, $R_\lambda^{\text{hy}}(b)$ pro hyperbolické prahování

Při výpočtu statistických charakteristik hyperbolického prahování vycházíme jako u ostatních pravidel (viz důkaz tvrzení 6.2) přímo z definice střední hodnoty a rozptylu transformované náhodné veličiny. Střední hodnota

$$\begin{aligned}
M_\lambda^{\text{hy}}(b) &= E[\delta^{\text{hy}}(X, \lambda)] = EY = \\
&= E(Y\chi_{\{x| |x|\leq\lambda\}}) + E(Y\chi_{\{x| |x|>\lambda\}}) \\
&= E(Y\chi_{\{x| |x|>\lambda\}}) = E\left[\operatorname{sgn}(X)\sqrt{X^2 - \lambda^2}\chi_{\{x| |x|>\lambda\}}\right] \\
&= \int_{\lambda}^{\infty} \sqrt{x^2 - \lambda^2} \varphi(x-b) dx - \int_{-\infty}^{-\lambda} \sqrt{x^2 - \lambda^2} \varphi(x-b) dx \\
&= \int_{\lambda}^{\infty} \sqrt{x^2 - \lambda^2} \varphi(x-b) dx + \int_{-\infty}^{\lambda} \sqrt{x^2 - \lambda^2} \varphi(x+b) dx \\
&= \int_{\lambda}^{\infty} \sqrt{x^2 - \lambda^2} [\varphi(x-b) - \varphi(x+b)] dx.
\end{aligned}$$

Rozptyl

$$\begin{aligned}
V_\lambda^{\text{hy}}(b) &= E\left[\delta^{\text{hy}}(X, \lambda) - M_\lambda^{\text{hy}}(b)\right]^2 \\
&= E\left[\delta^{\text{hy}}(X, \lambda)\right]^2 - 2M_\lambda^{\text{hy}}(b)E[\delta^{\text{hy}}(X, \lambda)] + [M_\lambda^{\text{hy}}(b)]^2 \\
&= E\left[\delta^{\text{hy}}(X, \lambda)\right]^2 - [M_\lambda^{\text{hy}}(b)]^2. \tag{B.1}
\end{aligned}$$

Hodnota výrazu $M_\lambda^{\text{hy}}(b)$ už je známá a proto nyní vypočítějme, čemu je roven první člen v (B.1). Označíme-li nyní $Y^2 = Z$ a

$$C_\lambda(b) = \int_{\lambda}^{\infty} x^2 [\varphi(x-b) - \varphi(x+b)] dx, \tag{B.2}$$

pak

$$\begin{aligned}
E\left[\delta^{\text{hy}}(X, \lambda)\right]^2 &= EY^2 = EZ \\
&= E(Z\chi_{\{x| |x|\leq\lambda\}} + Z\chi_{\{x| |x|>\lambda\}}) \\
&= E(Z\chi_{\{x| |x|>\lambda\}}) \\
&= E\left\{\left[\operatorname{sgn}(X)\sqrt{X^2 - \lambda^2}\right]^2 \chi_{\{x| |x|>\lambda\}}\right\}
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[(X^2 - \lambda^2) \chi_{\{x \mid |x| > \lambda\}} \right] \\
&= \int_{-\infty}^{-\lambda} (x^2 - \lambda^2) \varphi(x - b) dx + \int_{\lambda}^{\infty} (x^2 - \lambda^2) \varphi(x - b) dx \\
&= \int_{\lambda}^{\infty} (x^2 - \lambda^2) [\varphi(x - b) + \varphi(x + b)] dx \\
&= C_{\lambda}(b) - \lambda^2 \int_{\lambda}^{\infty} \varphi(x - b) dx - \lambda^2 \int_{\lambda}^{\infty} \varphi(x + b) dx \\
&= C_{\lambda}(b) - \lambda^2 [1 - \Phi(\lambda - b)] - \lambda^2 [1 - \Phi(\lambda + b)] \\
&= C_{\lambda}(b) - \lambda^2 [2 - \Phi(\lambda - b) - \Phi(\lambda + b)].
\end{aligned}$$

Dohromady tedy

$$\begin{aligned}
V_{\lambda}^{\text{hy}}(b) &= C_{\lambda}(b) - \lambda^2 [2 - \Phi(\lambda - b) - \Phi(\lambda + b)] - [M_{\lambda}^{\text{hy}}(b)]^2 \\
&= (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) + \\
&\quad + (1 + b^2 - \lambda^2)[2 - \Phi(\lambda - b) - \Phi(\lambda + b)] - [M_{\lambda}^{\text{hy}}(b)]^2
\end{aligned} \tag{B.3}$$

s využitím vyjádření $C_{\lambda}(b)$ pomocí lemmatu 9.1.

Riziko hyperbolického prahování nyní můžeme napsat s využitím tvrzení 6.1 jako

$$\begin{aligned}
R_{\lambda}^{\text{hy}}(b) &= V_{\lambda}^{\text{hy}}(b) + [M_{\lambda}^{\text{hy}}(b)]^2 - 2b M_{\lambda}^{\text{hy}}(b) + b^2 \\
&= (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) + \\
&\quad + (1 + b^2 - \lambda^2)[2 - \Phi(\lambda - b) - \Phi(\lambda + b)] - 2b M_{\lambda}^{\text{hy}}(b) + b^2.
\end{aligned} \tag{B.4}$$

Lemma 9.1: Platí

$$C_{\lambda}(b) = (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) + (1 + b^2)[2 - \Phi(\lambda + b) - \Phi(\lambda - b)].$$

Důkaz:

$$\begin{aligned}
C_{\lambda}(b) &= \int_{\lambda}^{\infty} x^2 [\varphi(x - b) + \varphi(x + b)] dx \\
&= \int_{\lambda}^{\infty} x^2 \varphi(x - b) dx + \int_{\lambda}^{\infty} x^2 \varphi(x + b) dx \\
&= \int_{\lambda-b}^{\infty} (x + b)^2 \varphi(x) dx + \int_{\lambda+b}^{\infty} (x - b)^2 \varphi(x) dx.
\end{aligned} \tag{B.5}$$

Přitom pro integrály v (B.5) platí

$$\int_{\lambda \mp b}^{\infty} (x \pm b)^2 \varphi(x) dx = \int_{\lambda \mp b}^{\infty} x^2 \varphi(x) dx \pm 2b \int_{\lambda \mp b}^{\infty} x \varphi(x) dx + b^2 \int_{\lambda \mp b}^{\infty} \varphi(x) dx. \tag{B.6}$$

Jednotlivé integrály v (B.6) upravíme:

$$\begin{aligned}
\int_{\lambda \mp b}^{\infty} x^2 \varphi(x) dx &= \Gamma_1(\lambda \mp b) \\
&= (\lambda \mp b)\varphi(\lambda \mp b) + \Gamma_0(\lambda \mp b) \\
&= (\lambda \mp b)\varphi(\lambda \mp b) + 1 - \Phi(\lambda \mp b)
\end{aligned} \tag{B.7}$$

s využitím lemmatu 9.2, dále

$$\begin{aligned}
\int_{\lambda \mp b}^{\infty} x \varphi(x) dx &= - \lim_{a \rightarrow \infty} \int_{\lambda \mp b}^a \frac{d\varphi(x)}{dx} dx \\
&= - \left[\lim_{a \rightarrow \infty} \varphi(a) - \varphi(\lambda \mp b) \right] \\
&= \varphi(\lambda \mp b)
\end{aligned} \tag{B.8}$$

s využitím $\frac{d\varphi(x)}{dx} = -x \varphi(x)$ a dále

$$\begin{aligned}
\int_{\lambda \mp b}^{\infty} \varphi(x) dx &= \lim_{a \rightarrow \infty} \int_{\lambda \mp b}^a \varphi(x) dx \\
&= \lim_{a \rightarrow \infty} [\Phi(x)]_{\lambda \mp b}^a \\
&= 1 - \Phi(\lambda \mp b).
\end{aligned} \tag{B.9}$$

Díky těmto vztahům tedy můžeme dohromady napsat (B.6) jako

$$\begin{aligned}
\int_{\lambda \mp b}^{\infty} (x \pm b)^2 \varphi(x) dx &= (\lambda \mp b)\varphi(\lambda \mp b) + 1 - \Phi(\lambda \mp b) \\
&\quad \pm 2b\varphi(\lambda \mp b) + b^2[1 - \Phi(\lambda \mp b)]
\end{aligned} \tag{B.10}$$

a po úpravách pak vyjde konečně

$$\begin{aligned}
C_{\lambda}(b) &= (\lambda + b)\varphi(\lambda - b) + (\lambda - b)\varphi(\lambda + b) \\
&\quad + (1 + b^2)[2 - \Phi(\lambda - b) - \Phi(\lambda + b)]. \quad \square
\end{aligned}$$

Lemma 9.2: Označíme-li $\Gamma_n(\lambda) = \int_{\lambda}^{\infty} x^{2n} \varphi(x) dx$, pak tento integrál lze vypočítat rekurencí

$$\Gamma_n(\lambda) = \lambda^{2n-1}\varphi(\lambda) + (2n-1)\Gamma_{n-1}(\lambda), \quad n \geq 1, \tag{B.11}$$

kde $\Gamma_0(\lambda) = 1 - \Phi(\lambda)$.

Důkaz: Integrováním $\Gamma_n(\lambda)$ per partes pro $n \geq 0$ dostáváme

$$\begin{aligned}
\Gamma_n(\lambda) &= \left[\frac{x^{2n+1}}{2n+1} \varphi(x) \right]_{\lambda}^{\infty} + \int_{\lambda}^{\infty} \frac{x^{2n+1}}{2n+1} x \varphi(x) dx \\
&= \frac{1}{2n+1} \left[\lim_{a \rightarrow \infty} a^{2n+1} \varphi(a) - \lambda^{2n+1} \varphi(\lambda) \right] + \int_{\lambda}^{\infty} \frac{x^{2n+1}}{2n+1} x \varphi(x) dx \\
&= -\frac{\lambda^{2n+1}}{2n+1} \varphi(\lambda) + \frac{1}{2n+1} \int_{\lambda}^{\infty} x^{2(n+1)} x \varphi(x) dx \\
&= \frac{1}{2n+1} [-\lambda^{2n+1} \varphi(\lambda) + \Gamma_{n+1}(\lambda)]
\end{aligned}$$

a z této rovnosti plyne

$$\Gamma_{n+1}(\lambda) = \lambda^{2n+1} \varphi(\lambda) + (2n+1)\Gamma_n(\lambda), \quad n \geq 0,$$

což po změně indexů je přesně tvrzení, které chceme dokázat.

Dále $\Gamma_0(\lambda) = \int_{\lambda}^{\infty} \varphi(x) dx = \int_{-\infty}^{\infty} \varphi(x) dx - \int_{-\infty}^{\lambda} \varphi(x) dx = 1 - \Phi(\lambda)$. □

C Programový balík ThreshLab

ThreshLab je autorský soubor programů (tzv. m-funkcí) pro prostředí MATLAB¹, který navazuje na Wavelet Toolbox² firmy MathWorks. Tento balík zahrnuje zejména:

- implementaci všech pěti druhů prahování waveletových koeficientů,
- numerickou i grafickou statistickou analýzu každého z prahovacích pravidel, a to pro obecný rozptyl šumu $\sigma \neq 1$,
- vytvoření potřebných typů grafů, což zahrnuje např. přehledné zobrazení prahových koeficientů, zhodnocení výsledků prahování, mnohoměřítkovou analýzu (MRA), těsnost přiléhavosti waveletové báze k analyzovanému signálu
- to vše dostupné pomocí několika hlavních funkcí s volitelnými parametry.

M-funkce obsažené v tomto balíku jsou k nalezení na přiloženém médiu. Návod k funkčím a popisky v jimi produkovaných grafech jsme přepracovali do anglického jazyka, kvůli zamýšlené prezentaci a zveřejnění na mezinárodní úrovni. Hlavní m-funkce ThreshLabu jsou:

`mydwt.m` provede waveletovou transformaci (dekompozici) signálu na základě řady volitelných parametrů a vykreslí přehledný graf waveletových koeficientů;

`myidwt.m` provede inverzní waveletovou transformaci (rekonstrukci) signálu z waveletových koeficientů a vykreslí přehledný graf;

`thresh.m` ústřední procedura balíčku. Na základě řady parametrů provede prahování waveletového spektra vstupního signálu a přehledně zobrazí výsledný signál a statistiku prahování jednotlivých úrovní waveletových koeficientů. Praktická ukázka viz obr. 7.1. Implementováno je pět druhů prahovacích pravidel: tvrdé, měkké, poloměkké, hyperbolické a nezáporná garota.

¹verze 6 a vyšší

²verze 2.2

`plot_wav.m`..... funkce umožňující vykreslení různých typů grafů obsahujících waveletové koeficienty, příspěvky jednotlivých úrovní v MRA apod.
Praktická ukázka viz obr. 4.3, 4.9.

`era_*_+.m`..... patnáct funkcí pro statistickou analýzu prahovacích pravidel. Symbol * zastupuje označení typu prahování, symbol + zastupuje název statistiky. Např. funkce `era_hy_variance.m` vypočítá rozptyl hyperbolického prahování.

`denoise_tree.m`... odšumování signálu reprezentovaného binárním stromem.

`wp_denoise.m`.... provede odšumování signálu reprezentovaného binárním stromem pomocí nejlepší báze waveletového paketu a tzv. „prahovou“ entropií (str. 33). Využívá funkci `denoise_tree.m`.

Funkce `era_*_+.m` využívají soubory `distchck.m`, `normpdf.m` a `normcdf.m` ze Statistics Toolbox for Matlab. Tyto soubory jsou nakopírovány do adresáře balíku Threshlab, aby uživatel tento balík využívat i bez zmíněného toolboxu.

D Obsah přiloženého CD

Přiložené CD obsahuje tři adresáře:

/SegWT

obsahuje m-funkce Matlabu, sloužící k provedení algoritmu SegWT pro prodloužení okrajů signálu nulami (schéma na obr. 8.8). Hlavní funkce je `segwt.m`. Adresář také obsahuje funkci `compare.m`, která kontroluje, zda waveletové koeficienty vypočtené `segwt.m` jsou shodné s koeficienty vypočtenými klasickým způsobem, funkcí `wavedec.m` Wavelet Toolboxu pro Matlab.

/threshlab

obsahuje všechny funkce balíku ThreshLab pro Matlab (popis viz přílohu C).

/testovani

obsahuje vstupní a výstupní soubory získané během testování účinnosti waveletových metod (viz kapitolu 7).

Literatura

- [1] ABRAMOVICH, F., BENJAMINI, Y.: Adaptive thresholding of wavelet coefficients. *Computational Statistics & Data Analysis*, 22, 351–361, 1996.
- [2] ABRAMOWITZ, M., STEGUN, I. A.: *Handbook of Mathematical Functions – Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 1972.
- [3] ANDĚL, J.: *Matematická statistika*. SNTL, Praha, 1985.
- [4] BEYLKIN, G., COIFMAN, R., ROKHLIN, V.: Fast Wavelet Transform and Numerical Algorithms. *Comm. Pure Appl. Math.*, 44, 141–183, 1991.
- [5] BERGER, J., COIFMAN, R., POPOVIC, I.: Aspects of Pitch-Tracking and Timbre Separation: Feature Detection in Digital Audio Using Adapted Local Trigonometric Bases and Wavelet Packets. In *Proceedings of the 1995 International Computer Music Conference*, Banff, Canada, pp. 280–283, 1995.
- [6] BRUCE, A. G., GAO, H.-Y.: Understanding WaveShrink: Variance and Bias Estimation. *Research Report 36*, StatSci Division of MathSoft, Inc., 1996.
- [7] BUČEK, J.: Segmentace řeči na fonemy metodou sledování rozdílnosti příznaků. In *Sborník prací studentů a doktorandů*. VUT Brno, 1997.
- [8] CHEN, S. S., DONOHO, D. L., SAUNDERS, M. A.: Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.* 20 (1998), no. 1, 33–61, reprinted in *SIAM Review*, 43 (2001), no. 1, pp. 129–159.
- [9] CHRISTENSEN, O.: *An Introduction to Frames nad Riesz Bases*. Birkhäuser, Boston-Basel-Berlin, 2003.
- [10] COIFMAN, R.R., WICKERHAUSER, M.V.: Entropy-based Algorithms for best basis selection. *IEEE Trans. on Inf. Theory*, vol. 38, 2, pp. 713–718. 1992.
- [11] DARLINGTON, D., DAUDET, L., SANDLER, M.: Digital Audio Effects in the Wavelet Domain. In *Proc. of the 5th Int. Conf. on Digital Audio Effects (DAFX-02)*, Hamburg, 2002.
- [12] DAUBECHIES, I.: *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Pennsylvania, 1992.
- [13] DAVIS, G.M.: *Noise reduction in speech applications*. Davis, G.M., editor, CRC Press LLC, 2002. ISBN 0-8493-0949-2
- [14] DEBNATH, L., MIKUSIŃSKI, P.: *Introduction to Hilbert Spaces with Applications*. 2nd ed., Academic Press, San Diego, 1999.
- [15] DELLER, J.R., HANSEN, J.H.L., PROAKIS, J.G: *Discrete-Time Processing of Speech Signals (IEEE Press Classic Reissue)*. Wiley-IEEE Press; Reprint edition, 1999. ISBN 0780353862

- [16] DONOHO, D. L.: Nonlinear Wavelet Methods for Recovery of Signals, Images, and Densities from noisy and incomplete data. In *Different Perspectives on Wavelets*, I. Daubechies, editor. AMS, Providence, 1993.
- [17] DONOHO, D.L., JOHNSTONE, I.M.: Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, **81**(3), pp. 425–455, 1994.
- [18] DONOHO, D.L., JOHNSTONE, I.M.: Ideal denoising in an orthonormal basis chosen from a library of bases. *C. R. Acad. Sci. Paris Ser. A*, **319**, pp. 1317–1322, 1994.
- [19] DUTILLEUX, P.: An implementation of the “algorithme à trous” to compute the wavelet transform. In *Wavelets: Time-Frequency Methods and Phase Space*, Inverse Problems and Theoretical Imaging, editors J.-M. Combes, A. Grossman, P. Tchamitchian. pp. 298–304, Springer-Verlag, Berlin, 1989.
- [20] HUBBARD, B.B.: *The World According to Wavelets*. 2nd ed., A K Peters, Wellesley, MA, USA, 1998. ISBN 1-56881-072-5
- [21] JAN, J.: *Číslicová filtrace, analýza a restaurace signálů*. Druhé vydání, Vysoké učení technické v Brně, 2002.
- [22] JOHNSTONE, I.M., SILVERMAN, B.W.: *Wavelet threshold estimators for data with correlated noise*. Research report, Stanford University and Bristol University, 1994.
- [23] GAO, H-Y.: Wavelet shrinkage denoising using the non-negative garrote. *J. Comput. Graph. Statist.*, **7** (4), pp. 469–488, 1998.
- [24] GAO, H-Y., BRUCE, A.G.: WaveShrink with firm shrinkage. *Research Report 39*, StatSci Division of MathSoft, Inc., 1996.
- [25] HERMANSKY, H., MORGAN, N.: RASTA Processing of Speech. *IEEE Trans. Speech and Audio Processing*, Vol. 2, No. 4, pp. 578–589, 1994.
- [26] HYVARINEN, A., KARHUNEN, J., OJA, E.: *Independent Component Analysis*. John Wiley & Sons, 2001. ISBN 0-471-40540-X
- [27] MALLAT, S.: A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7), s. 674–693, 1989.
- [28] MALLAT, S.: *A Wavelet Tour of Signal Processing*. 2nd edition, Academic Press, 1999. ISBN : 0-12-466606-X
- [29] NAJZAR, K.: Wavelety a řešení okrajových úloh. In *Sborník kursu Programy a algoritmy numerické matematiky 7*, str. 146–154, Matematický ústav AVČR, Praha, 1994.
- [30] NASON, G.P., SILVERMAN, B.W.: The stationery wavelet transform and some statistical applications. In *Wavelets and Statistics*, volume 103 of *Lecture Notes in Statistics*, editors A. Antoniadis, G. Oppenheim, pp. 281–300, Springer-Verlag, New York, 1995.

- [31] NEŠETŘIL, J.: *Teorie grafů*. SNTL, Praha, 1979.
- [32] MOULIN, P.: Wavelet Thresholding Techniques for Power Spectrum Estimation. *IEEE Transactions on Signal Processing*, vol. 42, no. 11, November 1994.
- [33] PERCIVAL, D.B., WALDEN, A.T.: *Wavelet methods for Time Series Analysis* (Cambridge Series in Statistical and Probabilistic Mathematics, 4). Cambridge University Press, 2000. ISBN 0521640687
- [34] PORUBA, J.: Speech Enhancement Based on Nonlinear Spectral Subtraction. In *Proc. 4th Int. Caracas Conference on Devices, Circuits and Systems*, Aruba, Dutch Caribbean, 2002, pp. T031-1 – T031-4. ISBN 0-7803-7381-2
- [35] PTÁČEK, L., VRBA, K., SMÉKAL, Z., a kol.: Metoda mapování spektrogramu. Výzkumná zpráva projektu MPO ČR FD-K/125, 2002.
- [36] QIU, L., YANG, H., KOH, S.N.: A Fundamental Frequency Detector of Speech Signals Based on Short Time Fourier Transform. In *Proc. Tencon '94*, vol. 1, pp. 526–530, IEEE 1994.
- [37] RUDIN, W.: *Analýza v reálném a komplexním oboru*. 2. přepracované vydání, Academia, Praha, 2003.
- [38] SOVKA, P., POLLAK, P.: The study of speech/pause detectors for speech enhancement methods. In *Proc. of the 4th European Conference on Speech Communication and Technology*, pp. 1575–1578, Madrid, Spain, 1995.
- [39] STRANG, G., NGUYEN, T.: *Wavelets and Filter Banks*. Wellesley Cambridge Press, 1996.
- [40] SWELDENS, W.: The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets. *Applied and Computational Harmonic Analysis*, 3, 186–200, 1996.
- [41] SYSEL, P.: Speech segmentation into phonemes by wavelet transform. In *Proc. Telecommunications and Signal Processing – TSP'2000*, pp. 126–129, Brno University of Technology, Brno, 2000.
- [42] SYSEL, P.: Modified Spectral Subtraction Method Utilizing Wavelet Transform. In *Proceedings of Research in Telecommunication Technology*. Research in Telecommunication Technology RTT 2001. Brno University of Technology, pp. 123–126, 2001. ISBN 80-214-1938-5
- [43] UNSER, M., BLU, T.: Mathematical Properties of the JPEG2000 Wavelet Filters. *IEEE Trans. Image Processing*, vol. 12, no. 9, pp. 1080–1090, September 2003.
- [44] VAIDYANATHAN, P. P.: *Multirate Systems and Filter Banks*. Prentice Hall PTR, 1992. ISBN 0136057187
- [45] VESELÝ, V.: Wavelets a jejich použití při filtrace dat. *Proceedings of the Summer School ROBUST'96* (J. Antoch and G. Dohnal, eds.), pp. 241–272, JČMF, 1997.
- [46] VESELÝ, V.: Kernel frame smoothing operators. In *Proceedings of the summer school ROBUST'2000*, Nečtiny near Plzeň, September 2000 (J. Antoch and G. Dohnal, eds.), JČMF (Society of Czech Mathematicians and Physicists), 2001, pp. 308–323.

- [47] VESELÝ, V.: Hilbert-space Techniques For Spectral Representation in Terms of Overcomplete Bases. In *Proceedings of the Summer school DATASTAT 01*, Czech Republic. Masarykova univerzita v Brně, 2002.
- [48] VIDAKOVIC, B.: *Statistical Modeling by Wavelets (Wiley Series in Probability and Statistics)*. John Wiley & Sons, New York, 1999.
- [49] VIDAKOVIC, B.: Nonlinear wavelet shrinkage with Bayes rules and Bayes factors, *J. Amer. Statist. Assoc.*, vol. 93, pp. 173–179, 1998.
- [50] VIDAKOVIC, B., BIELZA LOZOYA, C.: Time-adaptive wavelet denoising. *IEEE Trans. Signal Processing*, 46, pp. 2549–2554, 1998.
- [51] VÍCH, R., SMÉKAL, Z.: *Číslicové filtry*. Academia, 2000. ISBN 80-200-0761-X
- [52] VIRAG, N.: Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System. *IEEE Transaction on Speech and Audio Processing*, vol. 7, no. 2, Piscataway, NJ, 1999.
- [53] WALTER, G.G., SHEN, X.: *Wavelets and Other Orthogonal Systems*. 2nd ed., Chapman & Hall, 2001. ISBN 1-58488-227-1
- [54] WEYRICH, N., WARHOLA, G.T.: De-noising using wavelets and cross validation. In S. P. Singh, editor, *Approximation Theory, Wavelets and Applications*, vol. 454 of NATO ASI Series C, pp.523–532, 1995.
- [55] WICKERHAUSER, M. V.: *Adapted Wavelet Analysis from Theory to Software*. IEEE Press, A K Peters, Ltd., 1994.
- [56] ZHANG, J.-K., DAVIDSON, T.N., WONG, K.M.: Efficient Design of Orthonormal Wavelet Bases for Signal Representation. *IEEE Transactions on Signal Processing*, 52(7), pp. 1983-1996, July 2004.

Seznam vlastních prací

- [57] SYSEL, P., RAJMIC, P.: Application of Wavelet Transform to Audio Signal Compression. In *Proceedings of Telecommunications and Signal Processing TSP 2001*, pp. 143–147, VUT Brno, 2001. ISBN 80-214-1962-8.
- [58] RAJMIC, P.: Wavelet Spectrum Thresholding Rules. In *Proceedings of the Fourth International Conference on Recent Advantages in Soft Computing*, Nottingham Trent University, Nottingham, 2002. ISBN 1-84233-0764.
- [59] RAJMIC, P., SYSEL, P.: Wavelet Spectrum Thresholding Rules. In *Proceedings of the International Conference Research in Telecommunication Technology*, Žilina University, Žilina, 2002. ISBN 80-7100-991-1.

- [60] RAJMIC, P.: Statistical Approach to Wavelet Spectrum Thresholding. In *Proceedings of the Telecommunications and Signal Processing 2002 International Conference*, Brno University of Technology, Brno, 2002.
- [61] SMÉKAL, Z., HERMAN, I., KÉPESI, M., NAGY, Z., PLŠEK, M., RAJMIC, P., SYSEL, P., ŠKORPIL, V., VONDRA, M., VRBA, K.: *RTD technologie hláskové separace zamaskované v šumu*. Závěrečná výzkumná zpráva projektu GAČR č. 102/00/1084, FEKT VUT v Brně, Brno, 2002.
- [62] VRBA, K., SMÉKAL, Z., PORUBA, J., RAJMIC, P.: *Vývoj alternativních metod separace řeči*. Dlouhá výzkumná zpráva projektu MPO č. FD-K/125, FEKT VUT v Brně, 2002.
- [63] VRBA, K., SMÉKAL, Z., PORUBA, J., RAJMIC, P., PLŠEK, M., ŠTEFÍČEK, R.: *Možnosti alternativních metod separace řeči*. Dlouhá výzkumná zpráva projektu MPO č. FD-K/125, FEKT VUT v Brně, 2003.
- [64] RAJMIC, P.: *Využití matematické statistiky při konstrukci efektivních waveletových bazí*. Závěrečná zpráva projektu FRVŠ č. 2206/2003, FEKT VUT v Brně, 2003.
- [65] RAJMIC, P.: Statistical Properties of Wavelet Spectrum Thresholding Rules. In *Proceedings of 48. Internationales Wissenschaftliches Kolloquium*, Technische Universität Ilmenau, 2003. ISSN 1619-4098
- [66] RAJMIC, P.: Exact Risk Analysis of Wavelet Spectrum Thresholding Rules. In *Proceedings of 10th IEEE International Conference on Electronics, Circuits and Systems*, Sharjah, United Arab Emirates. University of Sharjah, United Arab Emirates, 2003.
- [67] RAJMIC, P., KLIMEK, J.: Odstranění „praskání“ LP desky pomocí waveletové analýzy. In *Proceedings of the 4th conference of Czech Student AES Section on Audio Technologies and Processing, ATP 2003*. s. 126–137, VUT v Brně, 2003. ISBN 80-214-2391-9
- [68] RAJMIC, P., KLIMEK, J.: Removing Crackle from an LP Record via Wavelet Analysis. In *Proc. of the 7th Conference on Digital Audio Effects (DAFx'04)*, Naples, Italy, Federico II University of Napoli, 2004. In print.
- [69] RAJMIC, P.: A Method for Real-Time Wavelet-Type Signal Processing. In *Proceedings of the RTT2004 Conference*, ČVUT, Praha, 2004. In print.
- [70] RAJMIC, P.: Statistical analysis of wavelet spectrum thresholding rules in order to suppress noise in signal. In *Proceedings of the Summer school DATASTAT 03*, Svatka, Czech Republic. Masarykova univerzita v Brně, 2004. In print.