



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

NEURONOVÉ SÍTĚ V ALGORITMECH VODOZNAČENÍ AUDIO SIGNÁLŮ

NEURAL NETWORKS IN AUDIO SIGNAL WATERMARKING ALGORITHMS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. ONDŘEJ KAŇA

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. RADEK ZEZULA, Ph.D.

BRNO 2009



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Diplomová práce

magisterský navazující studijní obor
Telekomunikační a informační technika

Student: Bc. Ondřej Kaňa

ID: 88463

Ročník: 2

Akademický rok: 2008/2009

NÁZEV TÉMATU:

Neuronové sítě v algoritmech vodoznačení audio signálů

POKYNY PRO VYPRACOVÁNÍ:

Nastudujte a teoreticky rozeberte základy umělých neuronových sítí a možnosti jejich využití v aplikacích digitálního vodoznačení audio signálů. Vámi zvolenou neuronovou síť implementujte společně s algoritmy vkládání a extrakce vodoznaku v prostředí Matlab. Otestujte funkčnost implementované audio vodoznačící metody a ověřte její robustnost při extrakci vodoznaku.

DOPORUČENÁ LITERATURA:

- [1] TUČKOVÁ, J.: Úvod do teorie a aplikací umělých neuronových sítí. Skripta FEL ČVUT v Praze, vydavatelství ČVUT, 2005, ISBN 80-01-02800-3.
- [2] Jirsík, V., Hráček, P.: Neuronové sítě, expertní systémy a rozpoznávání řeči. Skripta FEKT VUT v Brně.
- [3] Wu, G., Zhou, X.: A Fast Audio Digital Watermark Method Based on Counter-propagation Neural Networks. Computer Science and Software Engineering, Volume: 3, 2008, ISBN: 978-0-7695-3336-0.
- [4] Jun, W, Xiaofeng, L., Zhang, Y.: Advances in Neural Networks - ISNN 2005, Second International Symposium on Neural Networks, Chongqing, China, May 30 - June 1, 2005, Proceedings Part II, Vol. 3497, ISBN: 978-3-540-25913-8.

Termín zadání: 9.2.2009

Termín odevzdání: 26.5.2009

Vedoucí práce: Ing. Radek Zezula, Ph.D.

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

Abstrakt:

Digitální vodoznačení dat je jednou z technik, která se zabývá ochranou digitálních dat. Při návrhu systému vodoznačení je kladen důraz hlavně na robustnost a neslyšitelnost vodoznaku. Tato diplomová práce se zabývá problematikou vodoznačení audio signálů s využitím umělých neuronových sítí. Je zde popsána metoda vodoznačení ve frekvenční oblasti. Metoda je založená na psychoakustických principech lidského sluchu a technikách neuronových sítí.

Klíčová slova: vodoznačení, vodoznak, neuronová síť, Back-propagation, psychoakustický model, dct

Abstract:

Digital watermarking is a technique for digital multimedia copyright protection. The robustness and the imperceptibility are the main requirements of the watermark. This thesis deals with watermarking audio signals using artificial neural networks. There is described audio watermarking method in the DCT domain. Method is based on human psychoacoustic model and the techniques of neural networks.

Keywords: watermarking, watermark, neural network, Back-propagation, psychoacoustic model, dct

KAŇA, O. *Neuronové sítě v algoritmech vodoznačení audio signálů*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 48 s. Vedoucí diplomové práce Ing. Radek Zezula, Ph.D.

Prohlášení

Prohlašuji, že svou diplomovou práci na téma „Neuronové sítě v algoritmech vodoznačení audio signálů“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

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

V Brně dne

.....
podpis autora

Poděkování

Děkuji vedoucímu diplomové práce Ing. Radku Zezulovi, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne

.....
podpis autora

OBSAH

1	ÚVOD	10
2	UMĚLÉ NEURONOVÉ SÍŤE	11
2.1	Analogie biologické a umělé neuronové sítě	11
2.2	Model neuronu	12
2.2.1	Agregační funkce	14
2.2.2	Aktivační funkce	14
2.3	Principy učení umělých neuronových sítí	15
2.3.1	Hebbův zákon učení	15
2.3.2	Učení s učitelem	16
2.3.3	Učení bez učitele	16
2.4	Topologie neuronových sítí	16
2.5	Vícevrstvá síť typu Perceptron	18
2.6	Vícevrstvá síť Back-propagation	19
2.6.1	Back-propagation algoritmus s proměnným krokem učení	23
3	VODOZNAČENÍ AUDIO SIGNÁLŮ	24
3.1	Základní princip	24
3.2	Metody vodoznačení	25
4	DFMT algoritmus	26
4.1	Frekvenční maskování	26
4.1.1	Kritická pásma	27
4.2	Výpočet výkonového spektra	28
4.3	Nalezení tónových a netónových složek	29
4.3.1	Určení tónových složek	29
4.3.2	Určení netónových složek	30
4.4	Absolutní práh slyšení	31
4.5	Decimace tónových a netónových složek	31
4.6	Výpočet individuálního maskovacího prahu	32
4.7	Výpočet globálního maskovacího prahu	34
5	VLOŽENÍ VODOZNAKU	35
5.1	Učení neuronové sítě	38
6	EXTRAHOVÁNÍ VODOZNAKU	39
7	ROBUSTNOST ALGORITMU	41
8	ZÁVĚR	46
	LITERATURA	47
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK	48

SEZNAM OBRÁZKŮ

Obr. 2.1: Biologický neuron.....	12
Obr. 2.2: Schéma umělého neuronu	13
Obr. 2.3: Průběhy aktivačních funkcí.....	15
Obr. 2.4: Jednovrstvá Hopfieldova síť	17
Obr. 2.5: Vrstevnatá neuronová síť	17
Obr. 2.6: Neuron výstupní vrstvy	20
Obr. 2.7: Vývojový diagram algoritmu Back-propagation s proměnným krokem učení.....	23
Obr. 3.1: Blok vložení vodoznaku	24
Obr. 3.2: Blok extrakce vodoznaku.....	24
Obr. 4.1: Frekvenční maskování	26
Obr. 4.2: Spektrum signálu	29
Obr. 4.3: Nalezené tónové a netónové složky	30
Obr. 4.4: Absolutní maskovací práh.....	31
Obr. 4.5: Tónové a netónové složky po decimaci	32
Obr. 4.6: Maskovací křivky tónových složek	33
Obr. 4.7: Maskovací křivky netónových složek.....	33
Obr. 4.8: Globální maskovací práh	34
Obr. 5.1: Proces vkládání vodoznaku ve frekvenční oblasti	35
Obr. 5.2: Binární známka	36
Obr. 5.3: Topologie neuronové sítě využitě při procesu vkládání vodoznaku.....	38
Obr. 6.1: Proces extrakce vodoznaku ve frekvenční oblasti	39
Obr. 6.2: Vstupy a výstup natrénované neuronové sítě při extrakci vodoznaku.....	40
Obr. 7.1: Vytažený vodoznak s hodnotou NCC=0,818.....	41
Obr. 7.2: Vytažený vodoznak z nepozměněné nahrávky a)rock, b)pop, c)klasika	42
Obr. 7.3: Vytažený vodoznak po filtraci DP a)rock, b)pop, c)klasika	43
Obr. 7.4: Vytažený vodoznak po filtraci HP a)rock, b)pop, c)klasika	43
Obr. 7.5: Vytažený vodoznak po přidání šumu a)rock, b)pop, c)klasika.....	43
Obr. 7.6: Vytažený vodoznak po převzorkování a)rock, b)pop, c)klasika.....	44
Obr. 7.7: Vytažený vodoznak po mediánové filtraci a)rock, b)pop, c)klasika	44
Obr. 7.8: Vytažený vodoznak po kompresi ogg Pop a)120kbps, b)96kbps, c)60kbps	44
Obr. 7.9: Vytažený vodoznak po kompresi ogg Rock a)120kbps, b)96kbps, c)60kbps	45
Obr. 7.10: Vytažený vodoznak po kompresi ogg Klasika a)120kbps, b)96kbps, c)60kbps	45

SEZNAM TABULEK

Tabulka 2.1: Analogie mezi biologickými a umělými neuronovými sítěmi.....	12
Tabulka 2.2: Aktivační funkce	14
Tabulka 4.1: Rozdělení kritických pásem [10]	27
Tabulka 7.1: Hodnoty vodoznačených nahrávek bez změny obsahu.....	42
Tabulka 7.2: Hodnoty testu robustnosti po filtraci DP.....	43
Tabulka 7.3: Hodnoty testu robustnosti po filtraci HP.....	43
Tabulka 7.4: Hodnoty testu robustnosti po přidání šumu	43
Tabulka 7.5: Hodnoty testu robustnosti po převzorkování	44
Tabulka 7.6: Hodnoty testu robustnosti po mediánové filtraci	44
Tabulka 7.7: Hodnoty testu robustnosti po kompresi ogg	44
Tabulka 7.8: Hodnoty testu robustnosti po kompresi ogg	45
Tabulka 7.9: Hodnoty testu robustnosti po kompresi ogg	45

1 ÚVOD

V důsledku rychlého rozvoje internetu se rychle šíří i veškerá digitální data, ke kterým řadíme např. audio signály, obrazové nebo video signály. Rozvoj s sebou přinesl jak výhody, které vidíme ve snadném sdílení dat a jejich dostupnosti, tak i nevýhody. Jedním z problémů je nelegální šíření digitálních dat a porušování autorských práv.

Jednou z technik, která se zabývá ochranou digitálních dat je vodoznačení, vkládání vodoznaku do digitálních dat. Ve své diplomové práci se zabývám problematikou vodoznačení audio signálů s využitím neuronových sítí. Práce je rozdělena na dvě základní části. V první části práce je přehledný úvod do problematiky neuronových sítí. Je popsán základní model neuronu společně s nejčastěji se vyskytujícími sítěmi. Hluběji jsou pak rozebrány vícevrstvé neuronové sítě se zpětným šířením chyby. Samotná problematika neuronových sítí je poměrně nová oblast vědeckého zaměření. Její rozvoj mohl nastat až souběžně s rozvojem výkonné výpočetní techniky. Využití umělých neuronových sítí je široké, je však nutno k nim přistupovat realisticky a vědět, že řadu problémů je vhodnější řešit jinými metodami.

Druhá část práce se zabývá vodoznačením audio signálů. Je zde popsána metoda vodoznačení ve frekvenční oblasti. Metoda je založená na psychoakustických principech lidského sluchu, což zvyšuje nepostřehnutelnost vkládaného vodoznaku, a technikách neuronových sítí. Neuronová síť slouží k uchování vztahů mezi originálním audio signálem a odpovídajícím vodoznačeným audio signálem. Aplikace na vložení vodoznaků a jejich detekci byly zrealizovány v prostředí MATLAB.

2 UMĚLÉ NEURONOVÉ SÍTĚ

Umělé neuronové sítě představují jednu z mnoha podmnožin oblasti IT zvané strojové učení. Základy této nové vědní disciplíny byly položeny roku 1943, kdy neurobiolog Warren S. McCulloch společně s matematikem Walthrem Pittsem publikovali článek, ve kterém ukázali, že správně navržená neuronová síť může teoreticky vypočítat jakoukoliv spočitatelnou funkci [1].

Umělou neuronovou síť lze definovat jako paralelní distribuovaný systém tvořený mnoha jednoduchými procesory, které jsou uspořádány tak, aby byl systém schopen požadovaného zpracování informací [8].

Umělé neuronové sítě slouží především k sestavování statistických modelů ve značně komplexních systémech, kde nelze dostatečně spolehlivě odhalit charakter vazeb uvnitř systému, nebo v případech, kdy je exaktní matematický model systému tak složitý, že případná algoritmicizace úlohy je takřka nemožná.

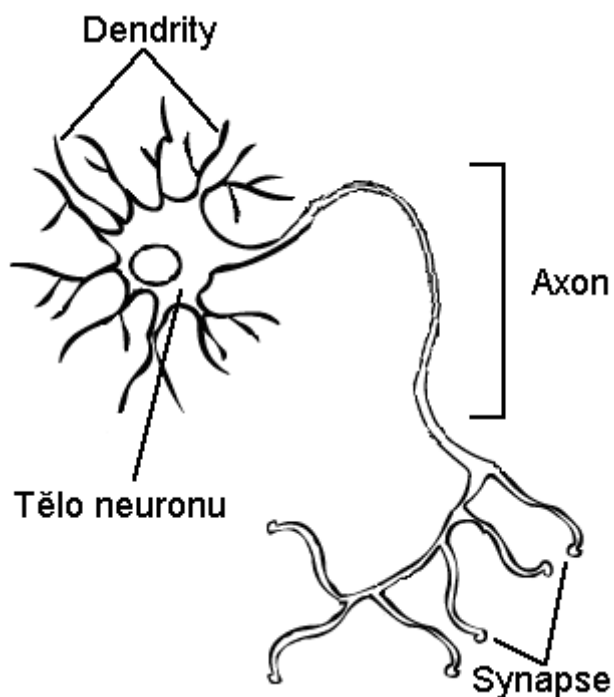
Jedním z nejvýznamnějších rysů umělých neuronových sítí od klasického pojetí výpočtů a programování je jejich schopnost učit se, adaptovat se. To znamená získávat znalosti učením pomocí množiny vzorů (známých hodnot vstupních a výstupních parametrů) bez nutnosti explicitní znalosti algoritmu řešení. Každá správně navržená a natrénovaná neuronová síť má pak také schopnost generalizace, tedy schopnost vhodně zpracovat i ty signály, které neobsahovala tréninková množina dat. Díky těmto vlastnostem našly umělé neuronové sítě uplatnění v mnoha vědních disciplínách a problémech. Mezi hlavní oblasti možných aplikací sítí patří především: predikce, transformace signálů, rozpoznávání obrazu, komprese a expanze signálu, řízení, analýza signálů, redukce šumu, oprava poškozených dat a další.

2.1 Analogie biologické a umělé neuronové sítě

Umělé neuronové sítě principiálně vycházejí z funkce skutečných neuronů v biologických nervových systémech, které tvoří základ informačních systémů všech živých organismů. Jejich podstatou je modelování struktury a činnosti biologických neuronových sítí. Základním strukturálním i funkčním stavebním elementem biologické nervové sítě je nervová buňka - neuron. Je specializovaná na zpracování, uchování a přenos informací. Zjednodušené schéma biologického neuronu je uvedeno na obr. 2.1.

Biologický neuron, jako většina buněk živých organismů, má svoji buněčnou membránu, která ho odděluje od zevního prostředí a chrání před vnějšími vlivy. Membrána obklopuje cytoplazmu s jádrem. Buněčné jádro se nachází v těle neuronu nazývaném soma. Z těla neuronu vystupují četné nervové výběžky – dendrity a axony. Dendrity jsou krátké a tenké, bývá jich větší počet, tvoří vstupy do neuronu. Axon je poměrně dlouhý a silný, bývá pro každý neuron pouze jediný a tvoří výstup. Konec axonu se větví a každá větev je ukončena tzv. synapsí. Synapse zajišťuje přenos informací do okolních neuronů, vyznačují se velkou plasticitou, jež souvisí s průchodností synapsí, která se mění při procesu učení [1].

Spojením neuronů pomocí synapsí vznikne již zmiňovaná biologická neuronová síť. Neuronová síť živých organismů je často velmi složitý útvar čítající miliardy neuronů, navzájem propojených, vzájemně na sebe působících a proměnlivých v čase. Struktura i vztahy mezi neurony se neustále vyvíjí.



Obr. 2.1 Biologický neuron

V tabulce 2.1 je uvedena základní analogie mezi biologickými a umělými neuronovými sítěmi.

Tabulka 2.1: Analogie mezi biologickými a umělými neuronovými sítěmi

Biologická neuronová síť	Umělá neuronová síť
tělo neuronu (soma)	matematický procesor
dendrit	vstup (n-rozměrný vektor hodnot)
axon	výstup (skalární hodnota)
synapse	vektor vah spojení

Další text se bude týkat pouze umělých neuronů, proto výrazem neuron bude dále označen umělý neuron.

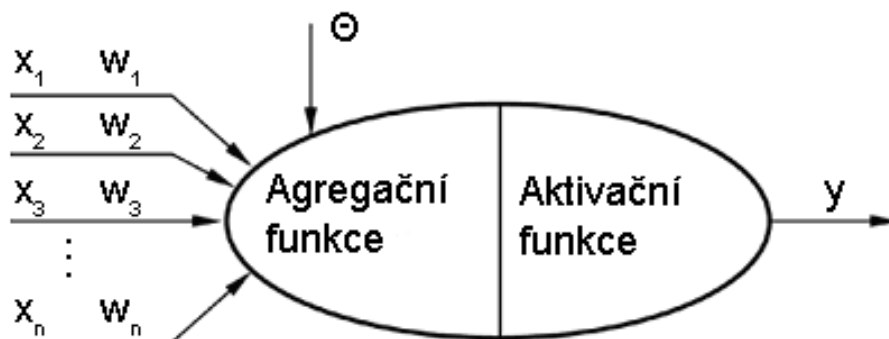
2.2 Model neuronu

V umělém neuronu dochází oproti originálu k náhradě biologických funkcí funkcemi matematickými. Existuje mnoho různých typů modelů umělých neuronů. Liší se navzájem vlastní topologií a typy matematických funkcí, které popisují jejich chování.

Umělý neuron si můžeme představit jako matematický procesor, který má konečný počet vstupů a pouze jediný výstup, jež může předávat libovolný počet shodných informací o svém stavu. Každý model neuronu má dvě části, agregační funkci a aktivační funkci, pomocí kterých transformuje vstupní hodnoty na výstupní.

Na umělou neuronovou síť, respektive na aktuální stav neuronu, je třeba se dívat jako na dynamický systém, poněvadž jeho vlastnosti se v závislosti na čase mění.

Nejrozšířenějším modelem je tzv. formální neuron, jež je často nazýván podle svých autorů McCulloch – Pittsův neuron. Schéma formálního neuronu je uvedeno na obr. 2.2.



Obr. 2.2 Schéma umělého neuronu

V tomto modelu představuje x elementy vstupního vektoru. V závislosti na konkrétní poloze neuronu v síti, mohou vstupy tvořit výstupy z jiných, předcházejících neuronů, nebo podněty z vnějšího okolí. Vstupní vektor může nabývat hodnot, které mohou být v „kvalitativní“ (obvykle nabývají booleovských hodnot) nebo „kvantitativní“ (vstupní veličina je vyjádřena reálným číslem) formě.

W jsou váhy spojení neuronu, které ovlivňují jednotlivé vstupy do neuronů (každý vstup je ohodnocen určitou hodnotou příslušné váhy spojení). Hodnoty vah vypovídají o průchodnosti, důležitosti daného spojení (vyjadřují uložení zkušeností do neuronu). Při učení se hledají optimální hodnoty vah, tak aby došlo ke shodě mezi výstupy zkoumaného procesu a výstupy neuronové sítě. Výpočty aktuálních hodnot vah a jejich postupné ladění jsou základem učících algoritmů neuronových sítí.

Θ se nazývá práh neuronu, ten představuje bariéru, kterou musí vstupní signál neuronu překonat, aby se mohl dál šířit neuronovou sítí. Hodnota prahu určuje, kdy je neuron aktivní (větší než práh) resp. pasivní (menší než práh).

Výstup y neuronu nakresleného na obr 2.2 své vstupní informace zpracovává a předává dále podle vztahu [2]:

$$y = S \left(\sum_{i=1}^n w_i x_i + \Theta \right), \quad (2.1)$$

kde

x_i - jsou vstupy neuronu a počet těchto neuronů je n ,

w_i - jsou synaptické váhy,

Θ - je práh daného neuronu,

$S(x)$ - je přenosová funkce neuronu (aktivační funkce).

Výraz v rovnici (2.1) bývá někdy označován jako vnitřní potenciál neuronu.

2.2.1 Agregáčn funkce

Agregační funkce neuronu určuje, jakým způsobem budou vstupní parametry kombinovány uvnitř neuronu. Při agregaci dochází k zobrazení vektoru vstupních signálů na skalární signál. Nejčastěji používané agregační funkce jsou:

- Lineární basicke funkce (LBF)
Mezi LBF patří vážená lineární kombinace vstupů (používá se pro McCulloch – Pittsův neuron popsaná vztahem (2.2), respektive (2.3).

$$u(t) = \sum_{i=1}^n w_i(t)x_i(t) \quad (2.2)$$

$$u(t) = \prod_{i=1}^n w_i(t)x_i(t) \quad (2.3)$$

- Radiální basicke funkce (RBF)
Tato funkce rozděluje vstupní prostor pomocí určitého počtu hyperkoulí [1]. Je popsána vztahem (2.4).

$$u(t) = \sqrt{\sum_{i=1}^n (x_i(t) - w_i(t))^2} \quad (2.4)$$

2.2.2 Aktivační funkce

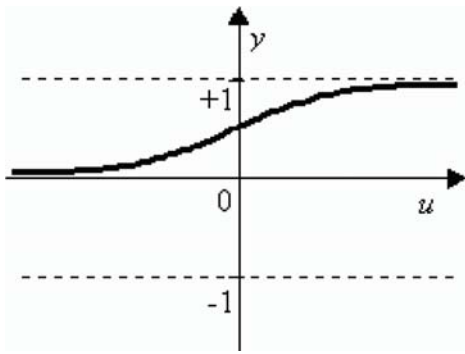
Aktivační funkce můžeme přirovnat k přenosovým funkcím, převádí hodnotu vstupního potenciálu na výstupní hodnotu z neuronu. Obecně lze tento vztah vyjádřit rovnicí (2.5). Podle zvoleného typu aktivační funkce se odvíjí náročnost výpočtů a čas trénování. Tvary přenosových funkcí bývají velmi různorodé.

$$y = f(u) \quad (2.5)$$

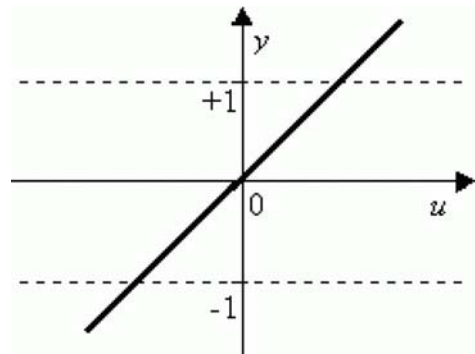
Funkční předpisy vybraných aktivačních funkcí jsou uvedeny v tabulce 2.2 („ γ “ je strmost funkce), jejich průběhy jsou znázorněny na obr 2.3.

Tabulka 2.2: Aktivační funkce

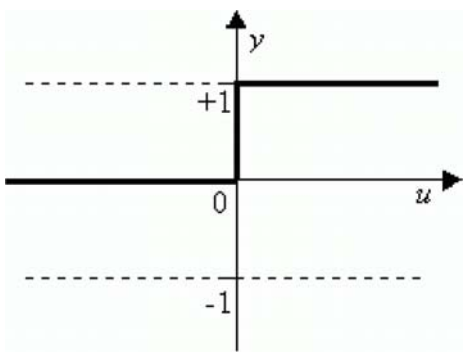
1 Skoková funkce	$f(u) = \begin{cases} 1 & \text{pro } u \geq 0 \\ 0 & \text{pro } u < 0 \end{cases}$
2 Lineární funkce	$f(u) = \gamma \cdot u + b$
3 Sigmoidální funkce	$f(u) = \frac{1}{1 + e^{-\gamma \cdot u}}$
4 Hyperbolicko-tangenciální funkce	$f(u) = \tanh(\gamma \cdot u)$



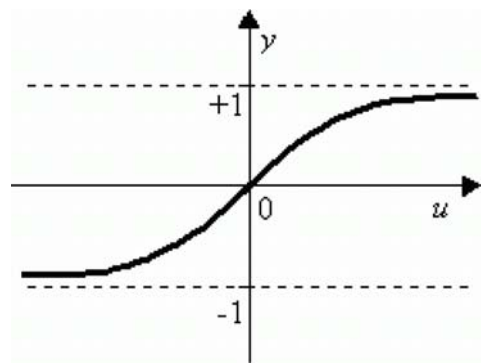
a) Sigmoidální funkce



b) Lineární funkce



c) Skoková funkce



d) Hyperbolicko-tangenciální funkce

Obr. 2.3 Průběhy aktivačních funkcí

2.3 Principy učení umělých neuronových sítí

Proces učení umělých neuronových sítí představuje dynamický proces, při kterém dochází především k modifikaci synaptických vah a prahů neuronů tak, aby mezi výstupy z modelované soustavy a výstupy z neuronové sítě byla co nejmenší odchylka. Tréninkovou množinou dat rozumíme takový výběr dat ze základního prostoru vstupů do neuronové sítě, který tento prostor dobře reprezentuje. Tato množina dat se pak použije k učicímu procesu neuronové sítě.

Učení lze realizovat induktivní nebo deduktivní metodou. Induktivní metoda vyvozujeme všeobecně platné závěry na základě pozorování množiny jevů (syntetický přístup). Při učení deduktivní metodou vystačíme s pozorováním jediného jevu a s jeho analýzou (analytický přístup) [9]. Při učení je nejčastěji využívána induktivní metoda a to buď s učitelem nebo bez učitele.

2.3.1 Hebbův zákon učení

Hebbovo učicí pravidlo představuje základ všech současných algoritmů učení. Jeho postuláty lze shrnout do dvou hlavních pravidel:

- pokud jsou dva spolu propojené neurony ve stejném okamžiku aktivní, pak jejich vazby jsou zesíleny a naopak
- pokud je aktivní jen jeden z neuronů, vazba mezi nimi je oslabena či zaniká

Každý neuron má pouze dva stavy, aktivní (1) a neaktivní (0). Tento zákon lze zapsat vyjádřit matematickým vztahem (2.6)

$$w_{ji}(t+1) = w_{ji}(t) + \eta \cdot y_i(t) \cdot x_j(t), \quad (2.6)$$

kde w_{ji} je váha spojení neuronů i a j , x_j je presynaptický stav neuronu j , y_i je postsynaptický stav i -tého neuronu, t je čas a η je konstanta rychlosti učení.

2.3.2 Učení s učitelem

Představuje nejběžnější způsob učení. Při učení s učitelem jsou známy požadované výsledky odpovídající tréninkové množině dat. Tyto výsledky se během učení porovnávají s výstupem ze sítě. Síť pak během učení nastavuje své parametry (prahy a váhy spojení) tak, aby se snížil rozdíl mezi skutečnou a požadovanou hodnotou na výstupu neuronové sítě.

2.3.3 Učení bez učitele

Učení bez učitele nebo také samoorganizace nemá k dispozici množinu vzorů. Využívá se schopnosti neuronových sítí rozeznat ve vstupech blízké vlastnosti a tak třídit vstupní data do shluků.

Samoorganizace se používá všude tam, kde neznáme učící množinu. Výpočet vzdáleností mezi vzory a aktuálními hodnotami tvoří kritérium učení. Jsou hledány extrémy (minimální vzdálenost vzoru a obrazu nebo maximální výstupní hodnota).

2.4 Topologie neuronových sítí

Topologií neuronové sítě rozumíme uspořádání neuronů a jejich vzájemné propojení. V dnešní době existuje řada typů neuronových sítí, každý typ má podle své struktury předurčeno vlastní využití.

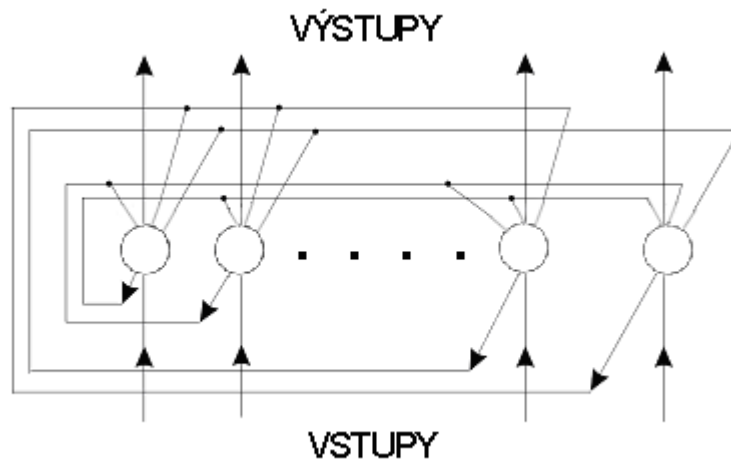
Z hlediska průchodu informací neuronovou sítí se rozlišují dva základní typy:

- Dopředné sítě:
Propojení neuronů existuje jenom mezi neurony přilehlých vrstev a je jednosměrné (tok informací prochází přímo od vstupu k výstupu sítě).
- Rekurentní sítě:
Díky zpětným vazbám je umožněn tok informací jak v přímém směru, tak i ve zpětné vazbě z jakéhokoliv výstupního bodu zpět na vstup sítě nebo určité vrstvy.

Z hlediska počtu vrstev lze neuronové sítě dělit na:

- **Jednovrstvé sítě:**

Jednovrstvé sítě neobsahují skryté vrstvy. Příkladem takové sítě může být Hopfieldova síť znázorněná na obr. 2.4, kde jsou neurony úplně propojeny.



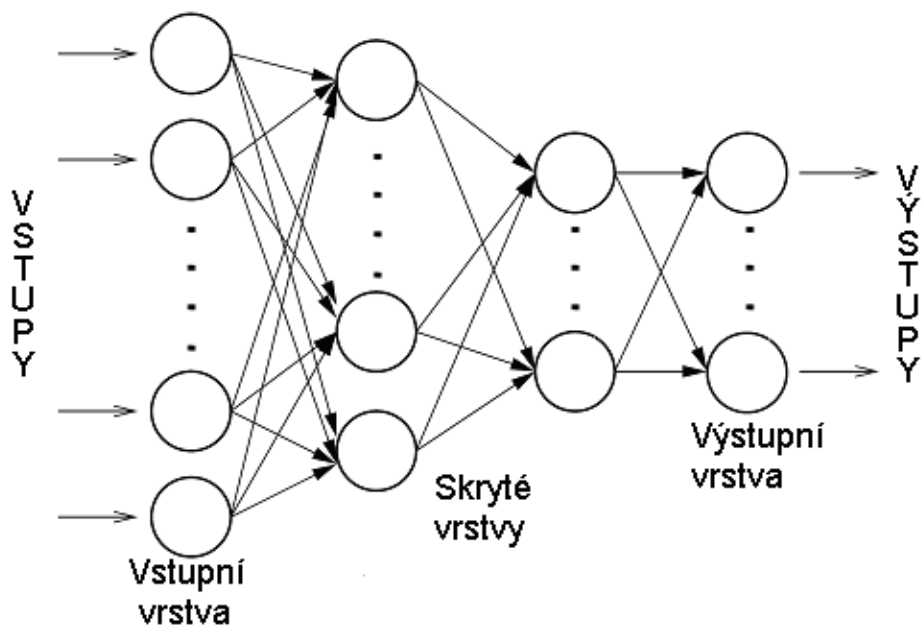
Obr. 2.4 Jednovrstvá Hopfieldova síť

- **Vícevrstvé (vrstevnaté) sítě:**

Je tvořena několika vrstvami neuronů. Každá vrstevnatá síť je složena z tří typů vrstev:

- jedna vstupní vrstva
- jedna nebo více skrytých vrstev
- jedna výstupní vrstva

Na obr. 2.5 je uveden příklad takovéto sítě.



Obr. 2.5 Vrstevnatá neuronová síť

Z hlediska hodnot parametrů v neuronových sítích rozlišujeme neuronové sítě na:

- Statické sítě:
V procesu trénování jsou nastaveny parametry sítě, dále při využívání neuronové sítě k řešení daného problému jsou tyto parametry konstantní.
- Dynamické sítě:
Parametry sítě jsou při procesu trénování přednastaveny a dále určitými metodami jsou adaptovány přímo při jejím využívání k řešení daného problému.

2.5 Vícevrstvá síť typu Perceptron

Základním výkonným prvkem perceptronových sítí je model neuronu s lineárně váženou agregační funkcí a aktivační funkcí, která je skoková [1]. Jde o zvláštní případ základního McCulloch-Pittsova modelu neuronu. Pomocí jednoho perceptronu je možné klasifikovat nanejvýš do dvou tříd. Učení perceptronu probíhá s učitelem.

Perceptrony se používají pro řešení jednoduchých klasifikačních úloh. Výstupem perceptronu je skalární veličina nabývající binárních hodnot, která se počítá podle rovnic (2.7) a (2.8). Důležitou podmínkou pro funkci perceptronu je, že vstupní data musí být lineárně separabilní.

$$u(x) = \sum_{i=1}^n w_i x_i + \Theta \quad (2.7)$$

$$y(x) = \begin{cases} 1 & \text{pro } u(x) \geq 0 \\ 0 & \text{pro } u(x) < 0 \end{cases} \quad (2.8)$$

K modifikaci vah spojení během procesu učení se využívá tzv. učícího zákona perceptronu. Pro adaptaci vah a prahů platí rovnice

$$w_{ij}(t+1) = w_{ij}(t) + e_j x_i, \quad (2.9)$$

$$\Theta_j(t+1) = \Theta_j(t) + e_j(t), \quad (2.10)$$

$$e = y - d, \quad (2.11)$$

kde t označuje čas, e je chybová hodnota aktuálního výstupu (rozdíl mezi skutečnou hodnotou y a požadovanou výstupní hodnotou d). Je-li výstup správný ($y=d$), váhy se nezmění. Má-li být vstup roven 1, ale je 0, k váhám se přičte vstupní hodnota. Má-li být vstup roven 0, ale je 1, od vah se odečte vstupní hodnota.

2.6 Vícevrstvá síť Back-propagation

Algoritmus Back-propagation (se zpětným šířením chyby) je jeden z nejužívanějších algoritmů používaných pro trénování a učení neuronových sítí. Byl vyvinut pro neuronové sítě obsahující alespoň jednu skrytou vrstvu. Příklad topologie sítě je na obr. 2.5. Nevýhodou této metody je velká citlivost na správnost tréninkových dat a na inicializaci parametrů sítě.

Algoritmus je založen na minimalizaci energie neuronové sítě. Energií se myslí míra naučenosti neuronové sítě (odchylka mezi skutečnou a získanou hodnotou z výstupu sítě pro danou trénovací množinu) [2]. K minimalizaci této funkce se využívá gradientní metoda. V principu se Back-propagation algoritmus aplikuje následujícím způsobem, na vstup se přivede matice vstupních parametrů, po průchodu sítí se porovnají data na výstupu s požadovanou hodnotou. Spočítá se hodnota chybové funkce, která se zpětně přepočítá do předchozích vrstev a opraví se váhy spojení. Jedná se o iterativní proces.

V následujícím textu bude popsána základní varianta algoritmu Back-propagation. Použité vztahy jsou přebrány z [2]. V základní variantě tohoto chybového učení jsou váhy modifikovány ve směru negativního gradientu chybové funkce. Algoritmus učení je popsán rovnicí (2.12)

$$w(t+1) = w(t) - \eta g(t), \quad (2.12)$$

kde w je množina všech vah a prahů v určitém časovém okamžiku, η je rychlost učení a g je gradient.

Pro chybu na výstupu z neuronu j , který je ve výstupní vrstvě, platí vztah (2.13)

$$e_j = y_j - d_j, \quad (2.13)$$

kde

y_j – j -tý výstup,

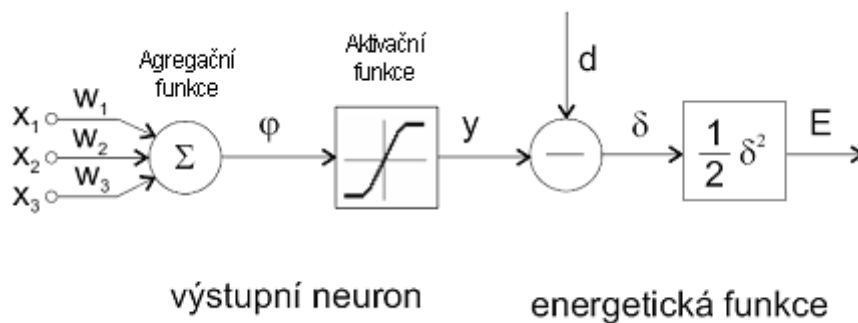
d_j – j -tý požadovaný výstup.

Chybová energie neuronu je rovna střední kvadratické chybě. Celková chybová energie je pak obdržena sumací chybových energií všech neuronů n ve výstupní vrstvě -vztah (2.14).

$$E = \frac{1}{2} \sum_{j=1}^n (y_j - d_j)^2 \quad (2.14)$$

Učící algoritmus představuje jakousi optimalizační metodu, která je pro danou neuronovou síť a trénovací množinu schopna nalézt váhové koeficienty a prahy.

Předpokládáme, že neuronová síť je tvořena neurony, jejichž chování je popsáno rovnicí 2.1. Blokové schéma libovolného neuronu i ve výstupní vrstvě je znázorněno na obr. 2.6.



Obr. 2.6 Neuron výstupní vrstvy

Pro dané hodnoty na vstupech neuronu bude neuron vykazovat určitou chybovou energii. Cílem je tuto hodnotu minimalizovat změnou váhových koeficientů, popřípadě prahů či vlastností aktivační funkce. To je možné vypočítat pomocí parciální derivace $\frac{\partial E}{\partial w_{ji}}$. Tuto složenou derivaci lze rozepsat jako součin dílčích derivací – vztah (2.15).

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial u_j} \frac{\partial u_j}{\partial w_{ji}} \quad (2.15)$$

$\frac{\partial E}{\partial y_j}$ je derivací energie $E = \frac{1}{2} \sum_{j=1}^n (y_j - d_j)^2$. Tato derivace se vypočte dle vztahu (2.16).

$$\frac{\partial E}{\partial y_j} = (y_j - d_j) \quad (2.16)$$

$\frac{\partial y_j}{\partial \varphi_j}$ je derivace aktivační funkce. Předpokládejme běžně používanou přenosovou funkci sigmoidu uvedenou v tabulce 2.2. Tuto derivaci můžeme vyjádřit vztahem jako funkci hodnoty y – vztah (2.17).

$$\frac{\partial y_j}{\partial u_j} = y_j(1 - y_j) \quad (2.17)$$

$\frac{\partial u_j}{\partial w_{ji}}$ je derivace agregační funkce $u_j = \sum_{i=1}^n w_{ji} \cdot y_i + \Theta_j$. Tato derivace se vypočte dle vztahu (2.18).

$$\frac{\partial u_j}{\partial w_{ji}} = y_i \quad (2.18)$$

Složením těchto dílčích výsledků dostáváme vztah (2.19).

$$\frac{\partial E}{\partial w_{ji}} = (y_j - d_j) \gamma_j (1 - y_j) y_i \quad (2.19)$$

Tímto výpočtem jsme získali gradient energetického prostoru. Pomocí vztahu (2.12) vypočteme vektor přírůstků vah Δw , jež platí pro všechny váhy v neuronové síti. Tedy vektor přírůstků vah spojení mezi neurony i a j je roven η násobkem součinu lokálního gradientu neuronu i a výstupu z neuronu j . Změna znaménka je nezbytná, poněvadž je orientován směrem k vrcholu (maximu).

Pro výstupní vrstvu platí:

$$\Delta w_{ji}(t) = \eta \delta_j(t) y_i(t) + \alpha \Delta w_{ji}(t-1), \quad (2.20)$$

$$\Delta \Theta_j(t) = \eta \delta_j(t) + \alpha \Delta \Theta_j(t-1). \quad (2.21)$$

Parametr α udává momentum tedy formu setrvačnosti, která umožňuje přeskočit lokální minima. Momentum se vypočte jako rozdíl dvou po sobě následujících změn váhy.

Váhy a prahy se upraví podle vztahů (2.22) a (2.23).

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t) \quad (2.22)$$

$$\Theta_i(t+1) = \Theta_j(t) + \Delta \Theta_j(t) \quad (2.23)$$

Pro výpočet vah a prahů ve skryté vrstvě se používají stejné úvahy jako při jejich výpočtu ve výstupní vrstvě. V dalším textu bude neuron j považován za neuron obsažený ve skryté vrstvě. Neuron l pak bude neuron výstupní vrstvy přímo napojený na neuron j . Pro jednoduchost se zaměříme na skrytou vrstvu, která přímo předchází vrstvu výstupní.

Pro každou váhu skryté vrstvy spočteme gradient energetického prostoru $\frac{\partial E}{\partial w_{ji}}$ dle vztahu:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_l} \frac{\partial y_l}{\partial u_l} \frac{\partial u_l}{\partial y_j} \frac{\partial y_j}{\partial u_j} \frac{\partial u_j}{\partial w_{ji}}. \quad (2.24)$$

Doposud jsme uvažovali pouze s jedním neuronem ve výstupní vrstvě, nyní bude ve výstupní vrstvě více neuronů. Tím pádem se výstup skrytého neuronu bude větvit. Hodnota δ_j je pak dána součtem příspěvků $w_{lj} \delta_l$ od všech neuronů výstupní vrstvy. Obecný zápis je popsán rovnicí (2.25).

$$\delta_j = \gamma_j (1 - y_j) \sum_{l=1}^n w_{lj} \delta_l \quad (2.25)$$

Hodnoty Δw_{ji} , $\Delta \Theta_j$, w_{ji} , Θ_j vypočteme podle vztahů:

$$\Delta w_{ji}(t) = \eta \delta_j(t) x_i(t) + \alpha \Delta w_{ji}(t-1), \quad (2.26)$$

$$\Delta \Theta_j(t) = \eta \delta_j(t) + \alpha \Delta \Theta_j(t-1), \quad (2.27)$$

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t), \quad (2.28)$$

$$\Theta_j(t+1) = \Theta_j(t) + \Delta \Theta_j(t), \quad (2.29)$$

kde x_i je i -tý vstup neuronové sítě nebo i -tý výstup předcházející skryté vrstvy. Výraz (2.25) pro výpočet δ_j lze zobecnit pro celou síť. Neuron j pak bude obsažen ve skryté vrstvě, neuron l bude obecně o jednu vrstvu blíže výstupu.

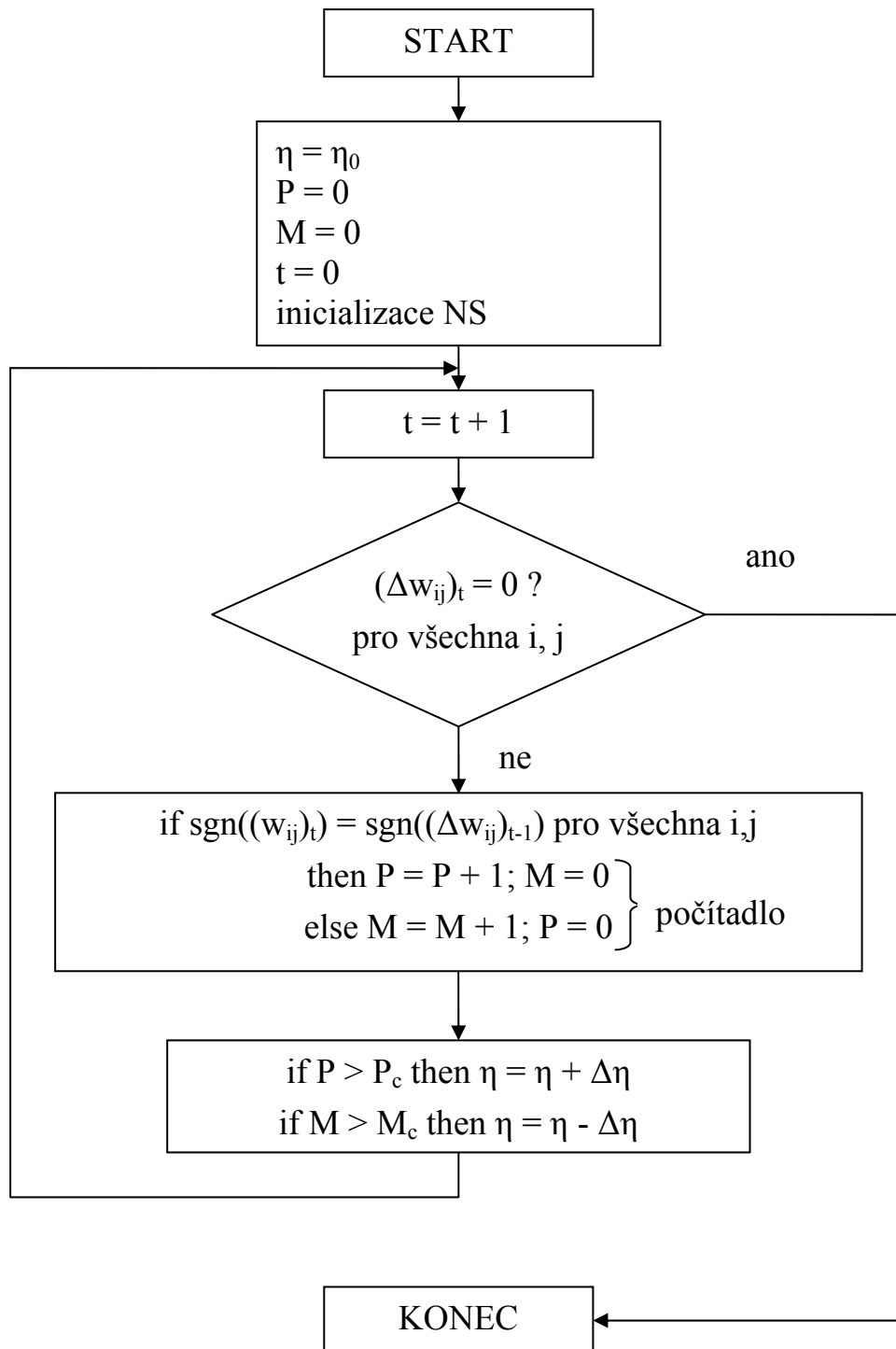
Celkový postup Back-propagation algoritmu:

- 1) Inicializace – náhodné nastavení vah spojení a prahů v určitém rozsahu
- 2) Volba vhodné tréninkové množiny dat
- 3) Dopředný výpočet – vybere se vzor z trénovací množiny, přiloží na vstupy sítě a vrstvu po vrstvě se nechá projít celou sítí. Pomocí vztahu (2.1) se vypočítá vektor výstupů sítě. Dále se pomocí vztahu (2.13) vypočte chyba pro výstupní vrstvu.
- 4) Zpětný výpočet – pomocí vztahů (2.19) pro neuron ve výstupní vrstvě a (2.25) pro neuron ve skryté vrstvě se vypočtou lokální gradienty sítě. Dále se vypočtou nové váhy spojení podle vztahu (2.26). Parametry η a α zvolíme tak, aby chybová funkce co nejrychleji konvergovala ke svému minimu a neuvízla v lokálním minimu.
- 5) Iterační výpočet – dopředný a zpětný výpočet se opakují tak dlouho, dokud se na výstupu nedosáhne požadovaných hodnot.

Algoritmus Back-propagation způsobil přelom ve vývoji neuronových sítí. Od svého vzniku byl mnohokrát modifikován. To bylo podmíněno hlavně jeho pomalým učením. Modifikované algoritmy jsou schopny konvergovat 10-krát až 100-krát rychleji. V následujícím kapitole je uveden příklad Back-propagation algoritmu s proměnným krokem učení.

2.6.1 Back-propagation algoritmus s proměnným krokem učení

Na obr. 2.7 jsou popsány jednotlivé kroky algoritmu Back-propagation s proměnným krokem učení, kde η je rychlost učení, P_c a M_c jsou konstanty (řídí vlastní zvětšení, zmenšení hodnoty η), t je čas. Vývojový diagram je převzat z [8].



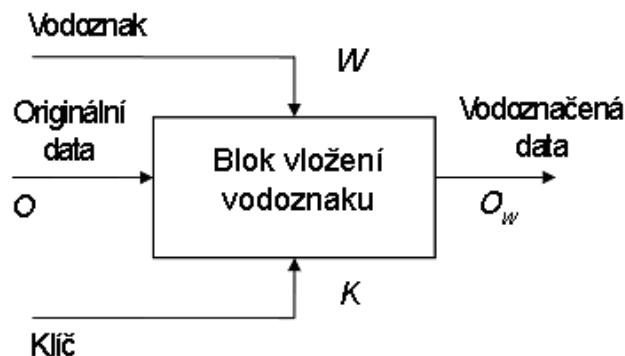
Obr. 2.7 Vývojový diagram algoritmu Back-propagation s proměnným krokem učení

3 VODOZNAČENÍ AUDIO SIGNÁLŮ

Vodoznačení audio signálu znamená přidání určité informace do užitečného signálu takovým způsobem, že je pro lidské ucho nepostřehnutelná, ale zároveň je možné ji pomocí počítačových algoritmů snadno detekovat, např. za účelem ověření pravosti zdrojových dat [4]. Při návrhu systému vodoznačení je kladen důraz hlavně na robustnost a neslyšitelnost vodoznaku.

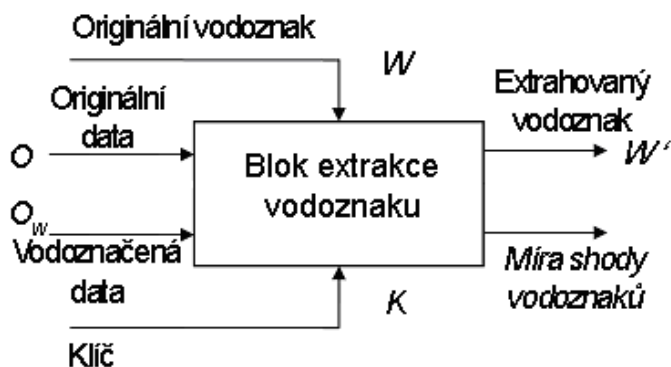
3.1 Základní princip

Technika vodoznačení využívá, stejně jako kryptografie, k vložení a vytažení vodoznaku stejného klíče. Klíč představuje danou posloupnost čísel, která definuje, kde se ve vodoznačených datech vodoznak vyskytuje. Každý systém vodoznačení se skládá ze dvou hlavních bloků: blok vkládání a blok detekce vodoznaku. Tento systém lze popsat množinou $(O, W, K, E_k, D_k, C_\tau)$, kde O jsou originální data, W značí vodoznak, K je šifrovací klíč, E_k značí proces vložení vodoznaku, D_k proces detekce vodoznaku, C_τ je porovnávací funkce (porovnává vytažený vodoznak s původním) [4]. Na obr. 3.1 je znázorněn proces vkládání vodoznaku.



Obr. 3.1 Blok vložení vodoznaku

Proces extrakce vodoznaku je znázorněn na obr. 3.2. Jedná se o inverzní operaci k procesu vkládání. Vstupní parametry jsou obvykle originální data, originální vodoznak, vodoznačená data a klíč. Počet vstupních parametrů je závislý na použité metodě. Na výstupu je pak extrahovaný vodoznak a míra shody extrahovaného vodoznaku s původním.



Obr. 3.2 Blok extrakce vodoznaku

3.2 Metody vodoznačení

K digitálnímu vodoznačení audio signálů se využívá mnoho technik, ty lze v zásadě rozdělit do dvou kategorií podle oblasti vkládání vodoznaku. Jde o metody v časové oblasti a frekvenční oblasti. Spousta těchto metod je založena na určitých druzích transformací. Nejčastěji jde o DFT (diskrétní Fourierova transformace), DCT (diskrétní kosinová transformace) a DWT (diskrétní vlnková transformace). Dále jsou uvedeny některé metody používané k vodoznačení audio signálů. Podrobný popis metod je uveden v [4].

- Metoda LSB

Tato metoda využívá nejméně významného bitu každého vzorku signálu k vložení vodoznaku. Šum obsažený v digitálních datech je použit jako nosný signál.

- Metoda Echo Hiding

Metoda vkládá vodoznak do originálního signálu jako ozvěnu. Využívá faktu, že lidské ucho není schopno zachytit signály kratší jak 10-20ms.

- Metoda Spread-Spectrum

Úzkopásmový signál vkládaného vodoznaku je modulován širokopásmovým nosným signálem, tím dojde k roztažení spektra signálu. Vodoznak je pak vkládán jako pseudonáhodný šum ve frekvenční oblasti.

4 DFMT algoritmus

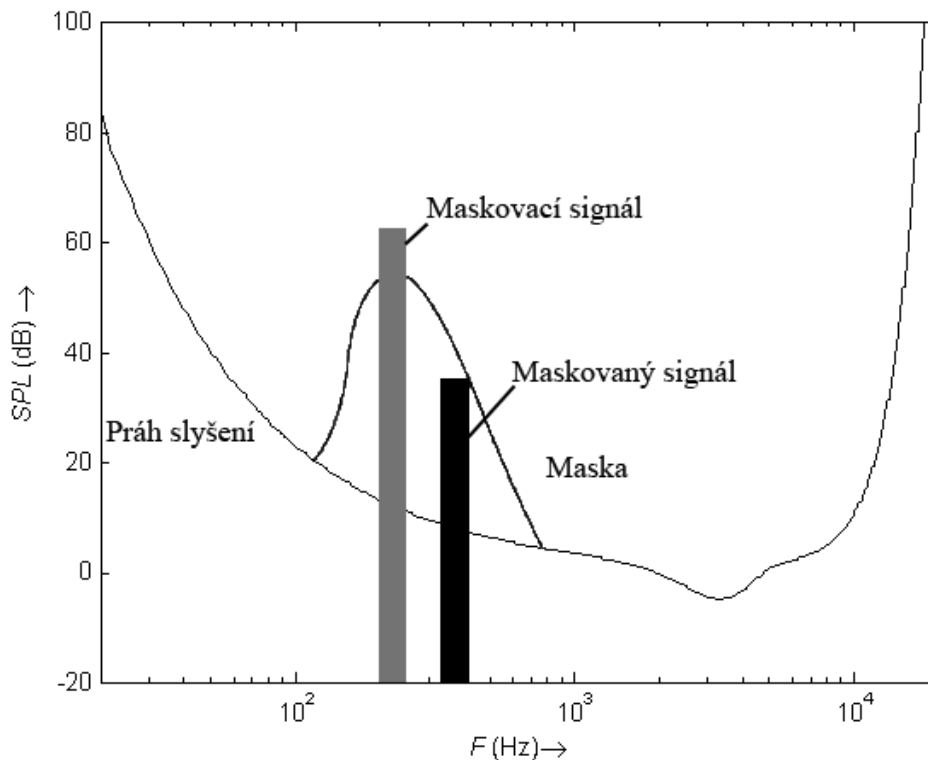
DFMT (determining frequency masking threshold) algoritmus pro výpočet úrovně frekvenčního maskování vychází z psychoakustického modelu pro MPEG layer 1 [5], [10]. Psychoakustický model analyzuje vstupní signál z hlediska jeho vnímání lidským sluchem. Výstupem je prahová (maskovací) křivka. Zvuky ležící pod touto křivkou nejsou lidským sluchem vnímány a jsou považovány za akusticky redundantní. Toho se využívá při ztrátové kompresi audio dat.

U aplikací vodoznačení psychoakustický model slouží k zajištění neslyšitelnosti vkládaného vodoznaku. Výpočet DFMT algoritmu se skládá z následujících kroků:

- Výpočet výkonového spektra
- Nalezení tónových a netónových složek
- Absolutní práh slyšení
- Decimace tónových a netónových složek
- Výpočet individuálních maskovacích prahů
- Výpočet globálního maskovacího prahu

4.1 Frekvenční maskování

Schopnost rozlišit dva zvukové tóny s malým rozdílem frekvencí je ovlivněna frekvenčním maskováním. Princip maskování v kmitočtové oblasti je znázorněn na obr. 4.1.



Obr. 4.1 Frekvenční maskování

Pokud znějí dva tóny současně, jeden z nich může potlačit slyšitelnost druhého. Lidské ucho totiž vnímá pouze tón o vyšší intenzitě a tón o nižší intenzitě je zamaskován. Tato neschopnost slyšet oba současné tóny se nazývá frekvenční maskování. Maskovací úroveň je závislá na frekvenci maskujícího tónu, hladině akustického tlaku (SPL) a tónové nebo netónové charakteristice maskovaného i maskujícího signálu. Vnímání tónů s blízkými frekvencemi je ovlivněno šířkou kritického pásma.

4.1.1 Kritická pásma

Kritické pásmo pro určitou frekvenci označuje úzké pásmo frekvencí okolo ní, které aktivuje shodnou část bazilární membrány v lidském uchu. Bazilární membrána představuje jakýsi frekvenční filtr – různé oblasti membrány zachytávají různé frekvence, zároveň se však částečně překrývají. Těto vlastnosti lidského ucha se využívá při procesu maskování frekvencí, kdy v rámci jednoho kritického pásma silné frekvence do značné míry maskují slabší části. Šířka kritického pásma dané frekvence je funkcí této frekvence, její hodnota je frekvenci přímo úměrná. V tabulce 4.1 je uvedeno rozdělení kritických pásem.

Tabulka 4.1: Rozdělení kritických pásem [10]

Číslo pásma	Střední frekvence (Hz)	Šířka pásma (Hz)
1	50	-100
2	150	100-200
3	250	200-300
4	350	300-400
5	450	400-510
6	570	510-630
7	700	630-770
8	840	770-920
9	1000	920-1080
11	1370	1270-1480
12	1600	1480-1720
13	1850	1720-2000
14	2150	2000-2320
15	2500	2320-2700
16	2900	2700-3150
17	3400	3150-3700
18	4000	3700-4400
19	4800	4400-5300
20	5800	5300-6400
21	7000	6400-7700
22	8500	7700-9500
23	10,500	9500-12000
24	13,500	12000-15500
25	19,500	15500-

Na základě konceptu kritického pásma byla vytvořena barková stupnice. Stupnice byla získána pomocí maskování signálního tónu šumem. Šířka pásma šumu, při kterém již tón nebyl slyšet (byl maskován) odpovídá kritickému pásmu a 1 barku. Frekvence v Hz může být přepočtena na Bark pomocí vztahu (4.1).

$$Bark = 13 \arctan(0,00076f) + 3,5 \arctan\left(\frac{f}{7500}\right)^2 \quad (4.1)$$

4.2 Výpočet výkonového spektra

Při výpočtu odhadu výkonu spektra (PSD – power spectral density), jsou bloky vstupního signálu převedeny z časové oblasti do frekvenční. Vstupní signál $x(n)$ je tedy segmentován na bloky o délce 512 vzorků. Délka bloků vychází ze specifikace psychoakustického modelu pro MPEG Layer 1. Kvůli zajištění minimálního prosakování spektra jsou bloky násobeny Hannovým oknem. Hannovo okno je definováno vztahem:

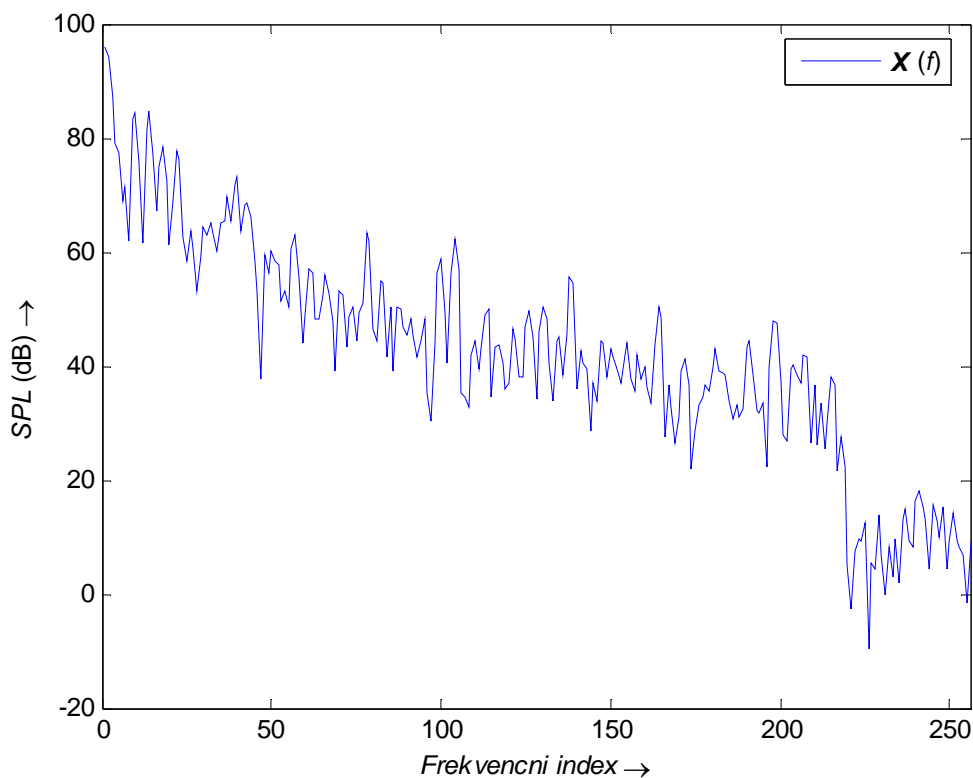
$$h(n) = \frac{\sqrt{8/3}}{2} \left[1 - \cos\left(2\pi \frac{n}{N}\right) \right], \quad (4.2)$$

kde $n=0\dots N-1$.

Samotný výpočet výkonového spektra $X(k)$ vstupního signálu $x(n)$ je proveden pomocí FFT podle vztahu (4.3). Výpočtem jsou získány hodnoty výkonu spektra v decibelech pro 256 frekvenčních indexů.

$$X(k) = 10 \log_{10} \left[\frac{1}{N} \left\| \sum_{n=0}^{N-1} x(n) h(n) \exp\left(-j2\pi \frac{nk}{N}\right) \right\|^2 \right] [dB] \quad \text{pro } 0 \leq k \leq \frac{N}{2} \quad (4.3)$$

Maximální hodnota výkonového spektra $X(k)$ je normalizována k referenční hodnotě SPL (Sound Pressure Level) 96dB. Na obr. 4.2 je uveden příklad výkonového spektra jednoho bloku audio signálu s vzorkovací frekvencí 44100Hz.



Obr. 4.2 Spektrum signálu

4.3 Nalezení tónových a netónových složek

Tónovost maskovacích složek má vliv na maskovací práh. Rozlišujeme tedy složky na tónové (harmonický signál) a netónové (šum). K výpočtu globálního maskovacího prahu je nezbytné určit tyto složky z FFT spektra. Na obr. 4.3 jsou znázorněny nalezené tónové a netónové složky pro daný blok audio signálu.

4.3.1 Určení tónových složek

Prvním krokem k určení tónových složek je nalezení lokálních maxim z výkonového spektra pomocí vztahu (4.4).

$$X(k) > X(k \pm 1) \quad (4.4)$$

Dalším krokem je určení tónových složek. To je provedeno prozkoumáním spektrálních složek výkonového spektra v okolí lokálních maxim uvnitř daného frekvenčního pásma podle následujícího kritéria:

$$X(k) > X(k \pm \Delta k) + 7dB, \quad (4.5)$$

kde

$$\Delta k \in \begin{cases} 2 & 2 < k < 63 & 0,17 - 5,5 \text{kHz} \\ (2,3) & 63 \leq k < 127 & 5,5 - 11 \text{kHz} \\ (2, \dots, 6) & 127 \leq k < 256 & 11 - 20 \text{kHz} \end{cases} \quad (4.6)$$

Úroveň akustického tlaku tónové masky je vypočtena podle vztahu (4.7).

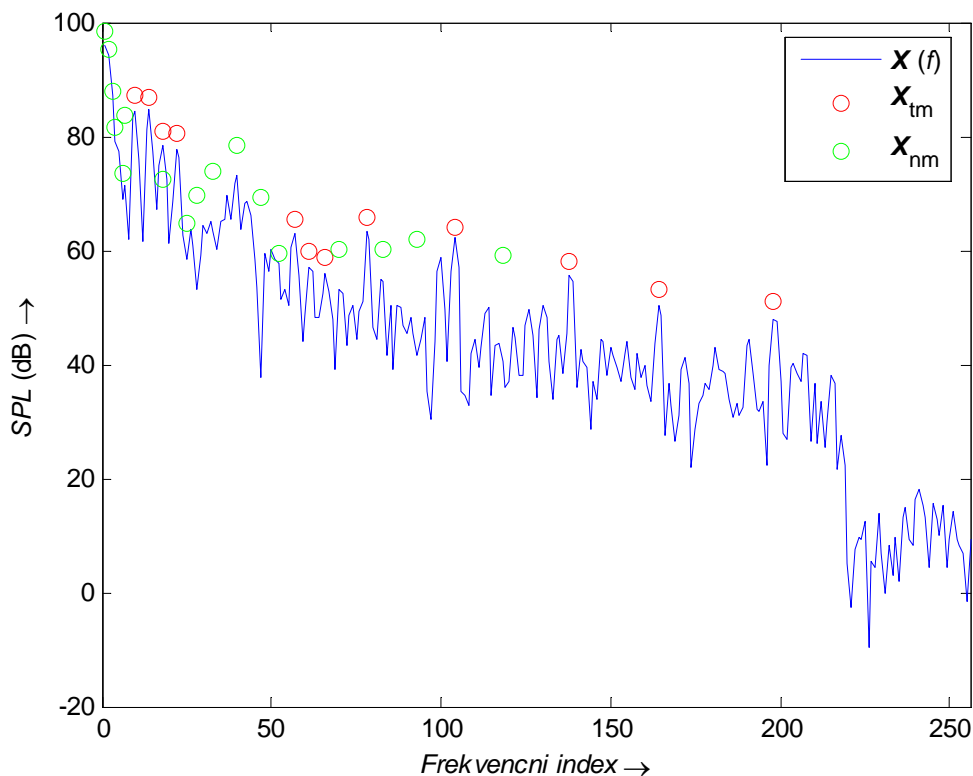
$$X_{tm}(k) = 10 \log_{10} \left(10^{\frac{X(k-1)}{10}} + 10^{\frac{X(k)}{10}} + 10^{\frac{X(k+1)}{10}} \right) [dB] \quad (4.7)$$

4.3.2 Určení netónových složek

Netónová složka se počítá pro každé kritické pásmo sumací zbylých spektrálních složek výkonového spektra, které nejsou v okolí X_{tm} , podle následujícího vztahu:

$$X_{nm}(\bar{k}) = 10 \log_{10} \sum_j 10^{\frac{X(j)}{10}} [dB] \quad (4.8)$$

pro všechny $X(j)$ jež nejsou členy $X_{tm}(k, k \pm 1, k \pm \Delta k)$, kde $\bar{k} = \left(\prod_{j=l}^u j \right)^{\frac{1}{u-l+1}}$, l je spodní a u horní spektrální složka hranice kritického pásma.

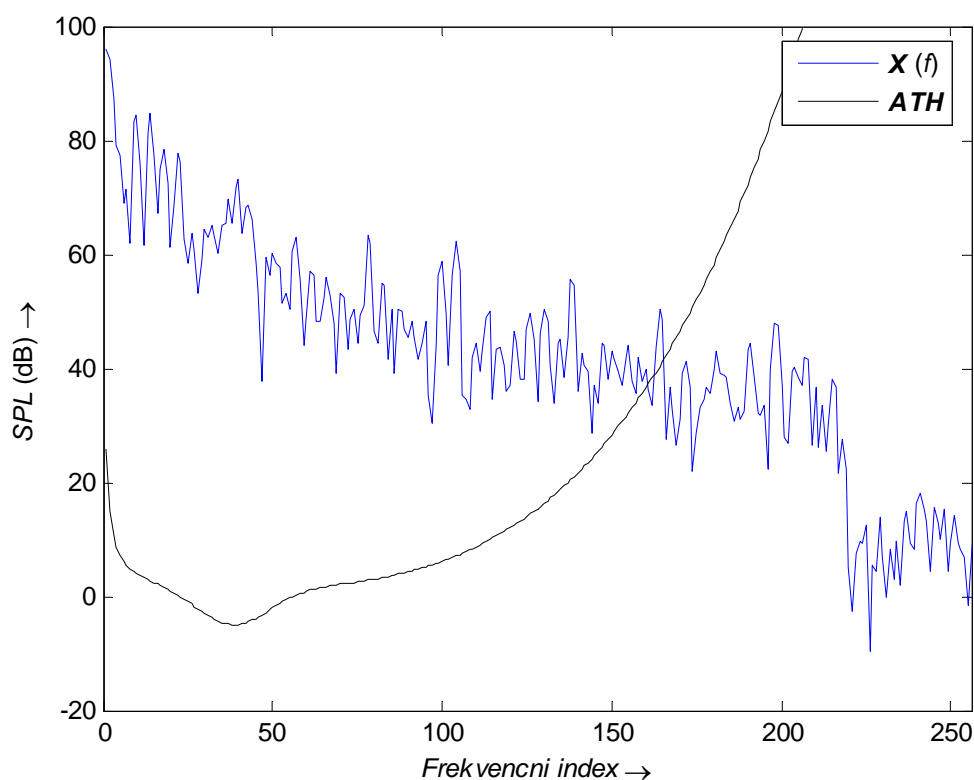


Obr. 4.3 Nalezené tónové a netónové složky

4.4 Absolutní práh slyšení

Absolutní práh slyšení (ATH) zobrazuje množství zvukové energie, které je nutné, aby byl tón dané frekvence detekovatelný lidským sluchem v nehlukném prostředí. Hladinu ATH lze aproximovat nelineární funkcí popsanou vztahem (4.9). Průběh této funkce je patrný z obr. 4.4.

$$ATH(f) = 3,64\left(\frac{f}{1000}\right)^{-0,8} - 6,5e^{-0,6\left(\frac{f}{1000}-3,3\right)^2} + 10^{-3}\left(\frac{f}{1000}\right)^4 \text{ [dB]} \quad (4.9)$$



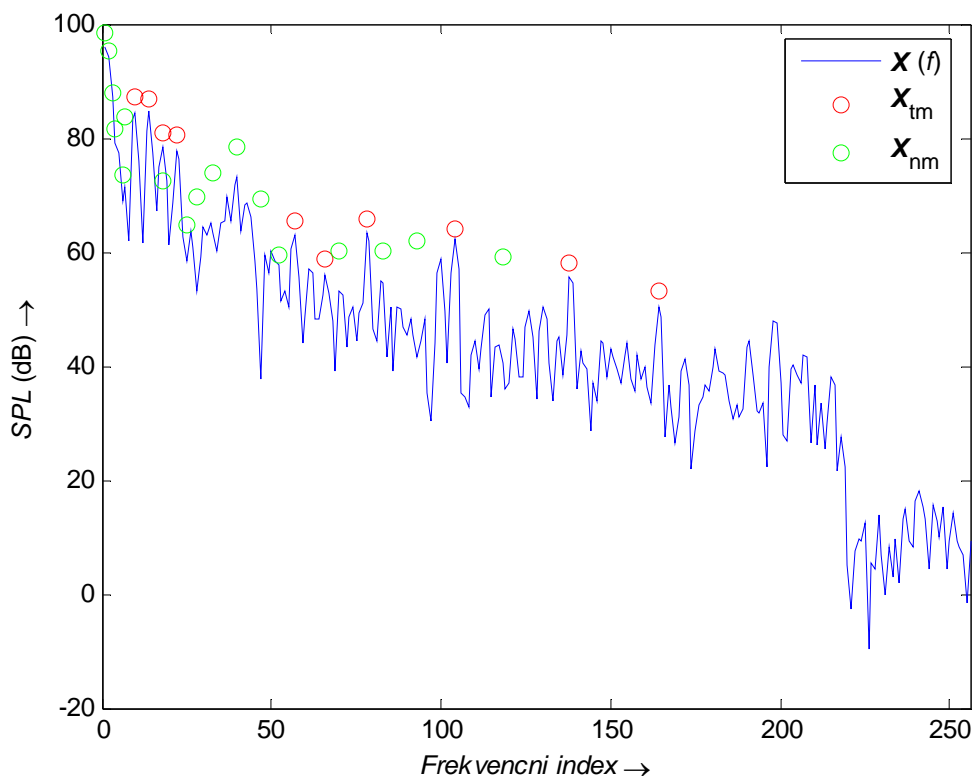
Obr. 4.4 Absolutní maskovací práh

4.5 Decimace tónových a netónových složek

Decimace se využívá k redukci počtu masek určených k výpočtu globálního maskovacího prahu. Provádí se podle dvou kritérií. Nejprve jsou odstraněny spektrální složky s nižší energií než hodnota absolutního maskovacího prahu.

$$X_{tm}(k) \geq ATH(k), X_{nm}(k) \geq ATH(k) \quad (4.10)$$

Posléze je provedena dodatečná decimace tónových složek. Pokud jsou tónové složky od sebe vzdálené méně než 0.5 Bark, je ponechána složka s vyšším výkonem. Na obr. 4.5 jsou znázorněny tónové a netónové složky po decimaci.



Obr. 4.5 Tónové a netónové složky po decimaci

4.6 Výpočet individuálního maskovacího prahu

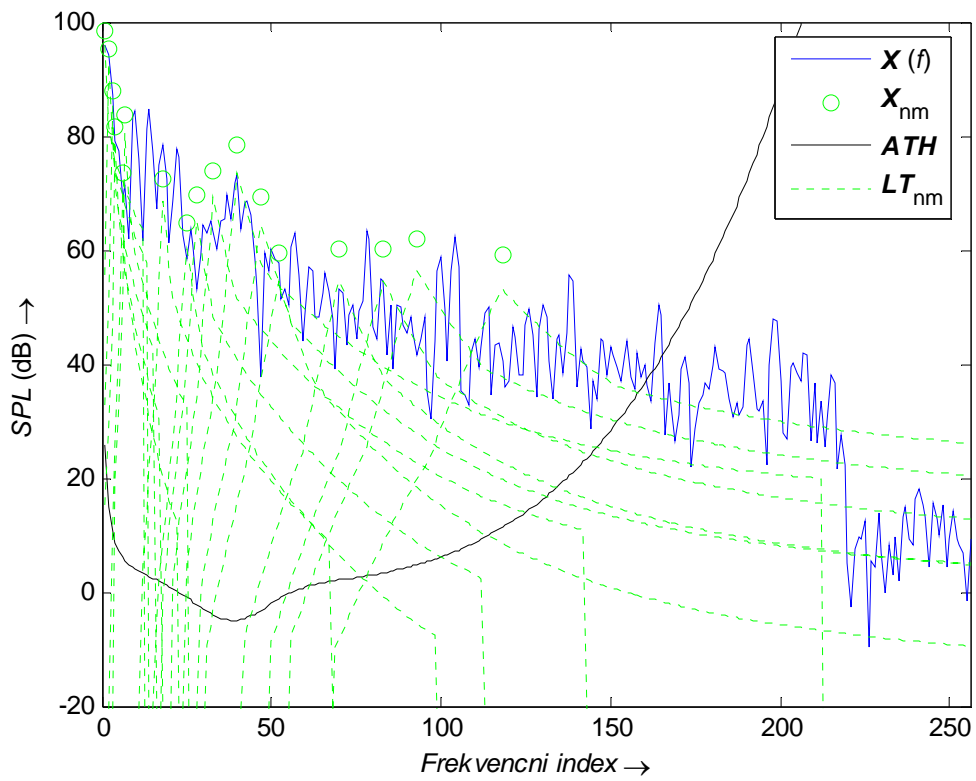
Tónové a netónové masky které jsme určily, neovlivňují jen frekvence uvnitř kritického pásma, ale také jeho okolí. Studie ukazují, že rozptyl maskování má strmost přibližně 25dB/Bark před maskou a -10dB/Bark po masce [11].

Rozptyl maskování můžeme popsat funkcí $SF(i,j)$, jež závisí na konkrétním frekvenčním indexu maskovací lokace i , pozici masky j , výkonovém spektru $X_{tm,nm}(j)$ a na vzdálenosti mezi maskou a maskovací lokací v Barcích $\Delta z = z(i) - z(j)$. Tato funkce určuje minimální hladinu blízkých frekvencí, kterou je lidský sluch schopen detekovat.

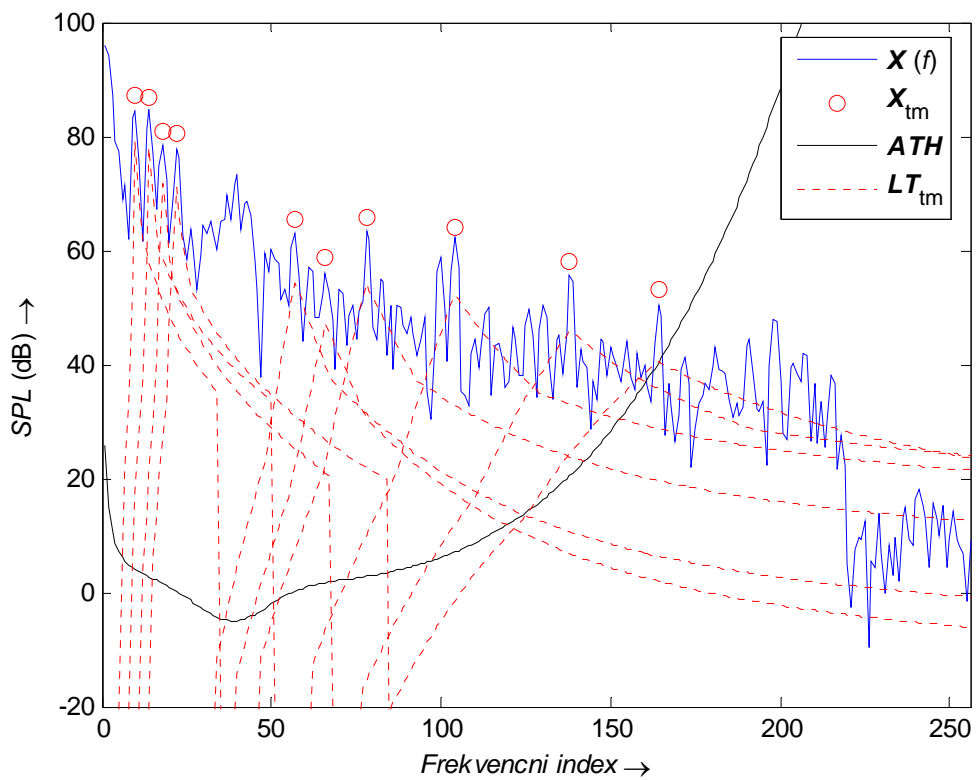
$$SF(i, j) = \begin{cases} 17\Delta z - 0,4X_{tm,nm}(j) + 11 & -3 \leq \Delta z < -1 \\ (0,4X_{tm,nm}(j) + 6)\Delta z & -1 \leq \Delta z < 0 \\ -17\Delta z & 0 \leq \Delta z < 1 \\ (0,15X_{tm,nm}(j) - 17)\Delta z - 0,15X_{tm,nm}(j) & 1 \leq \Delta z < 8 \end{cases} \quad (4.11)$$

Maskovací práh pro tónové LT_{tm} a netónové LT_{nm} masky je počítán podle vztahu 4.12 a jejich průběh je znázorněn na obr. 4.6 a obr. 4.7.

$$\begin{aligned} LT_{tm}(i, j) &= X_{tm}(j) - 0,25z(j) + SF(i, j) - 6,025 \text{ [dB]} \\ LT_{nm}(i, j) &= X_{nm}(j) - 0,175z(j) + SF(i, j) - 2,025 \text{ [dB]} \end{aligned} \quad (4.12)$$



Obr. 4.6 Maskovací křivky tónových složek



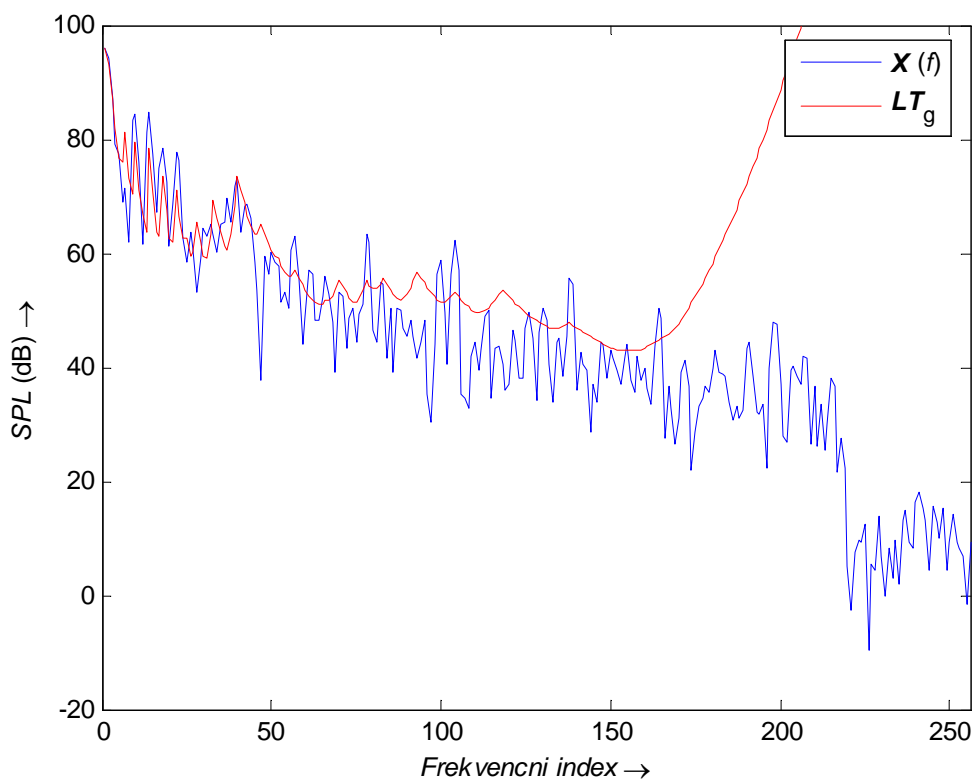
Obr. 4.7 Maskovací křivky netónových složek

4.7 Výpočet globálního maskovacího prahu

Globální maskovací práh $LT_g(i)$ pro frekvenční index i je součtem absolutního maskovacího prahu, individuálního tónového a netónového prahu dle následujícího vztahu:

$$LT_g(i) = 10 \log_{10} \left(10^{\frac{ATH(i)}{10}} + \sum_{l=1}^L 10^{\frac{LT_{tm}(i,l)}{10}} + \sum_{m=1}^M 10^{\frac{LT_{nm}(i,m)}{10}} \right) [dB], \quad (4.13)$$

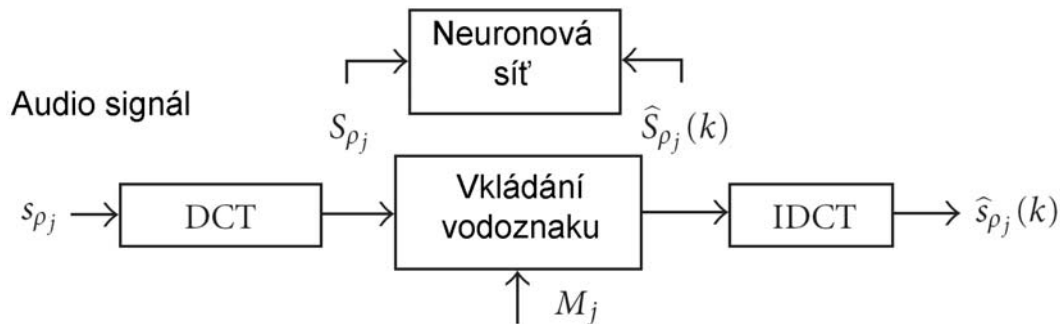
kde L , M udává počet tónových a netónových složek. Na obr. 4.8 je znázorněn průběh globálního maskovacího prahu pro daný blok audio signálu.



Obr. 4.8 Globální maskovací práh

5 VLOŽENÍ VODOZNAKU

Tato kapitola popisuje realizaci metody popsanou v literatuře [5], z dané literatury jsou přebrány vzorce i značení. Metoda využívá k vložení vodoznaku ve frekvenční oblasti DCT transformaci. Vodoznak tvoří binární známka o rozměrech $p \times q$. Proces vložení vodoznaku je schématicky znázorněn na obr. 5.1.



Obr. 5.1 Proces vkládání vodoznaku ve frekvenční oblasti

Vstupní audio signál X ve tvaru $X=(x_1, \dots, x_n)$ s N PCM (pulsní kódová modulace) vzorky je rozdělen do ϕ bloků. Každý blok obsahuje 256 vzorků. Množinu těchto bloků Ψ můžeme zapsat jako

$$\Psi = (s_1, \dots, s_i, \dots, s_\phi), \quad (5.1)$$

kde $s_i = (s_i(0), \dots, s_i(k), \dots, s_i(255))$ a $s_i(k)$ představuje k -tý vzorek i -tého bloku.

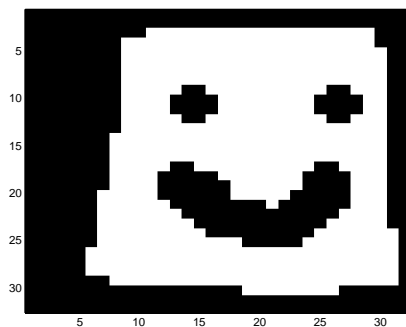
K zabezpečení informace související s vodoznakem proti různým druhům útoků, je využito pseudonáhodného číselného generátoru (PRNG). Ten generuje náhodné číslo r , na základě kterého se zajistí výběr bloků φ z množiny bloků Ψ . φ pak lze definovat jako

$$\varphi = \{s_{\rho_j} \mid j = 1, \dots, p \times q \text{ a } \rho_j \in \{0, \dots, \phi - 1\}\}, \quad (5.2)$$

kde $p \times q$ udává počet vybraných bloků. ρ_j je vyjádřeno vztahem (5.3).

$$\rho_j = r \bmod \phi \quad (5.3)$$

Jak bylo popsáno výše, vodoznak je představován binární známkou o rozměrech $p \times q$, jež můžeme popsat vztahem (5.4). V této práci byl použit vodoznak o rozměrech 32×32 pixelů. Tento vodoznak je znázorněn na obr. 5.2.



Obr. 5.2 Binární známka

$$H_{p,q} = (\sigma_{11}, \dots, \sigma_{1q}, \sigma_{21}, \dots, \sigma_{2q}, \dots, \sigma_{ik}, \dots, \sigma_{p1}, \dots, \sigma_{pq}) \quad (5.4)$$

$H_{p,q}$ je $(p \times q)$ -bitová binární posloupnost, kde σ_{ik} nabývá hodnot $\{0, 1\}$, $1 \leq i \leq p$ a $1 \leq k \leq q$. σ_{ik} představuje pixel binární známky na pozici (i, k) . Pro lepší práci se posloupnost $H_{p,q}$ převádí na vektorový tvar $w = (w_1, w_2, \dots, w_{pq})$ o $p \times q$ hodnotách, kde $w_j = 2\sigma_{ik} - 1$, $j = (i-1) \times q + k$, a $1 \leq j \leq p \times q$. w_j bude nabývat hodnot $\{-1, 1\}$ pro každé j , kde -1 představuje černý pixel ($\sigma_{ik} = 0$) a 1 bílý pixel ($\sigma_{ik} = 1$).

V dalším kroku se pomocí DCT transformace bloky $s_{\rho j}$ transformují do frekvenční oblasti. Transformace je popsána vztahem:

$$S_{\rho j}(l) = \sum_{n=1}^{256} \omega(n) s_{\rho j}(n) \cos \frac{\pi(2n-1)(l-1)}{512}, \quad (5.5)$$

kde $1 \leq l \leq 256$, $s_{\rho j}(n)$ představuje n -tý vzorek bloku $s_{\rho j}$ v časové oblasti, $S_{\rho j}(l)$ je l -tý DCT koeficient (frekvenční oblast) bloku $S_{\rho j}$ a $\omega(n)$ nabývá hodnot podle vztahu (5.6).

$$\omega(n) = \begin{cases} \frac{1}{256} & \text{pokud } n=1, \\ \sqrt{\frac{2}{256}} & \text{pokud } 2 \leq n \leq 256. \end{cases} \quad (5.6)$$

Transformací bloků φ na základě uvedených vztahů (5.5) a (5.6), získáme množinu bloků Φ . Ta je analogicky podle (5.1) popsána vztahem:

$$\Phi = \{S_{\rho j} \mid j = 1, \dots, p \times q \text{ a } \rho_j \in \{0, \dots, \phi - 1\}\}. \quad (5.7)$$

Během vodoznačícího procesu se vodoznak w vkládá do transformovaných bloků Φ . Jednotlivé bity vodoznaku w_j jsou pak ukryty v $S_{\rho j}(j_0)$ pro každé j , kde j_0 je konkrétní index každého transformovaného bloku. Vodoznak se vkládá do středního frekvenčního pásma, tím

pádem index $j_0 \in \{100, \dots, 200\}$. Při výpočtu indexu j_0 se využívá DFMT algoritmu popsaném v 4. kapitole. Výpočet indexu j_0 se skládá z následujících kroků:

krok 1: Pro každé $s_i \in \Psi$, se pomocí DFMT algoritmu vypočítá S_i a globální maskovací práh LTg_i , kde $i=1,2, \dots, \phi$.

krok 2: Pro všechna $S_i(j)$, kde $j=100, \dots, 200$ vypočítáme $LTg_i(j) - S_i(j) - \alpha$. α je konstanta.

krok 3: Určíme $j_0. j_0 = \arg \max\{LTg_i(j) - S_i(j) - \alpha\}$.

Vztah pro modifikaci $S_{\rho_j}(j_0)$ můžeme vyjádřit následovně:

$$\hat{S}_{\rho_j}(j_0) = S_{\rho_j}(j_0) + M_j, \quad (5.8)$$

kde $w_j \in \{-1, 1\}$, $M_j = w_j \times \alpha$ a α je konstanta. Hodnota α ovlivňuje vztah mezi nevnímátností a robustností vodoznaku. Nízká hodnota α zvyšuje nevnímátnost, nicméně se snižuje robustnost vodoznaku vůči útokům a různým manipulacím se signálem. Vložením vodoznaku do množiny transformovaných bloků Φ podle vztahu (5.8), získáme množinu $\hat{\Phi}$, jež je popsána vztahem (5.9).

$$\hat{\Phi} = \{\hat{S}_{\rho_j} \mid j = 1, \dots, p \times q \text{ a } \rho_j \in \{0, \dots, \phi - 1\}\} \quad (5.9)$$

Zpětnou inverzní kosinovou transformací (IDCT) \hat{S}_{ρ_j} , získáme vodoznačený audio blok \hat{s}_{ρ_j} . Množina těchto bloků $\hat{\Phi}$ je popsána vztahem:

$$\hat{\Phi} = \{\hat{s}_{\rho_j} \mid j = 1, \dots, p \times q \text{ a } \rho_j \in \{0, \dots, \phi - 1\}\}. \quad (5.10)$$

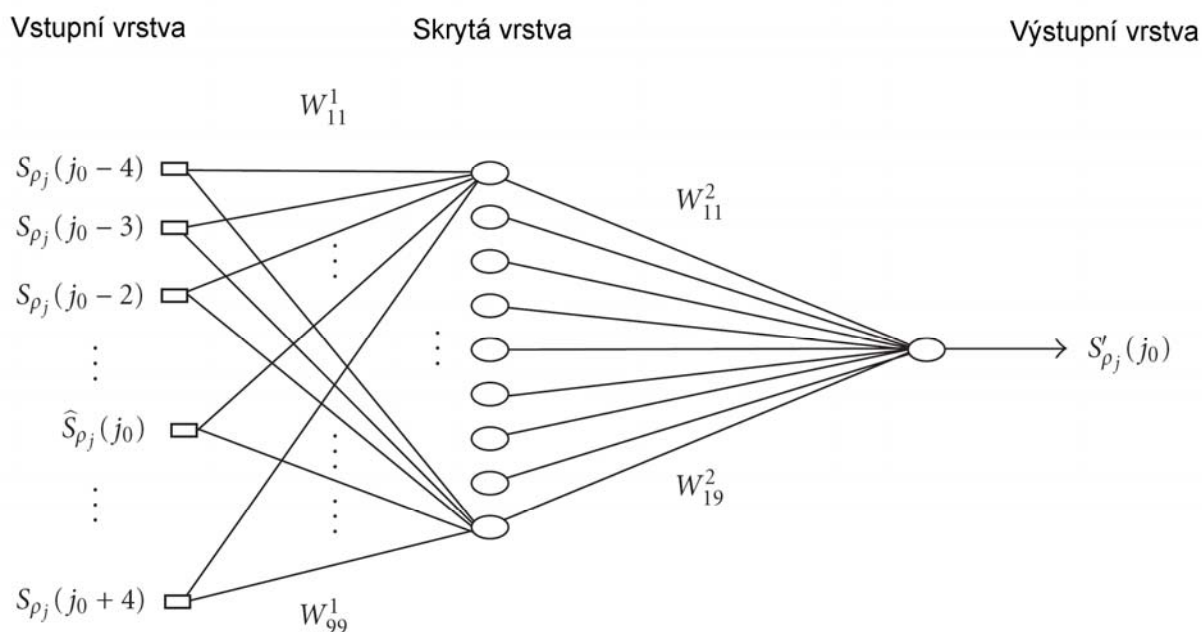
Vodoznačený audio signál je pak ve tvaru (5.11) respektive (5.12).

$$\hat{\Psi} = \{\hat{s}_1, \dots, \hat{s}_i, \dots, \hat{s}_\phi\} \quad (5.11)$$

$$\hat{X} = \{\hat{x}_1, \dots, \hat{x}_i, \dots, \hat{x}_\phi\} \quad (5.12)$$

5.1 Učení neuronové sítě

Při procesu vkládání se využívá neuronová síť 9-9-1 (neuronová síť má 9 vstupních neuronů, 9 neuronů ve skryté vrstvě a jeden výstupní neuron). Na obr. 5.3 je graficky znázorněna topologie této sítě [5].



Obr. 5.3 Topologie neuronové sítě využitě při procesu vkládání vodoznaku

K učení sítě je použit Back-propagation algoritmus s proměnným krokem učení. Podrobný popis algoritmu je uveden části 2.6.1. Trénovací množina Γ je vyjádřena vztahem

$$\Gamma = \{(A_j, B_j) | j = 1, 2, \dots, p \times q\}, \quad (5.13)$$

kde $|\Gamma|$ je $p \times q$. Vstupní vektor A_j neuronové sítě je

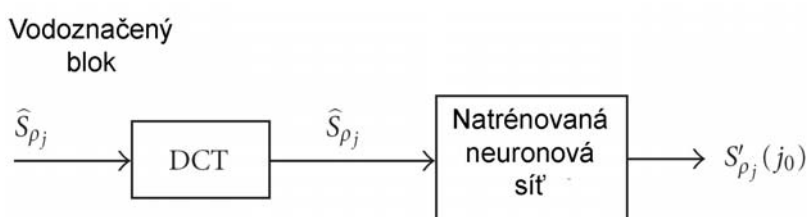
$$A_j = (S_{\rho_j}(j_0-4), \dots, S_{\rho_j}(j_0-1), \hat{S}_{\rho_j}(j_0), S_{\rho_j}(j_0+1), \dots, S_{\rho_j}(j_0+4)), \quad (5.14)$$

a požadovaný výstup B_j je $S_{\rho_j}(j_0)$.

6 EXTRAHOVÁNÍ VODOZNAKU

Proces extrakce vodoznaku je znázorněn na obr. 6.1. Při procesu vkládání vodoznaku se neuronová síť učí. Natrénovaná neuronová síť pak uchovává vztahy mezi originálním audio signálem a odpovídajícím audio signálem s vodoznakem. Při extrahování vodoznaku se tak nevyžaduje přístup k originálnímu audio signálu, ale jsou požadovány následující parametry:

- hodnoty synaptických vah
- inicializační vektor (seed) pro PRNG
- index j_0 každého bloku
- počet bitů $p \times q$ vodoznaku w



Obr. 6.1 Proces extrakce vodoznaku ve frekvenční oblasti

Nejprve pomocí vztahu (5.3) vybereme vodoznačené bloky audio signál $\hat{\phi}$ z množiny $\hat{\Psi}$. Každý vodoznačený audio blok \hat{s}_{ρ_j} v množině $\hat{\phi}$ je transformován pomocí vztahu (6.1), čímž získáme vodoznačený transformovaný audio blok \hat{S}_{ρ_j} ,

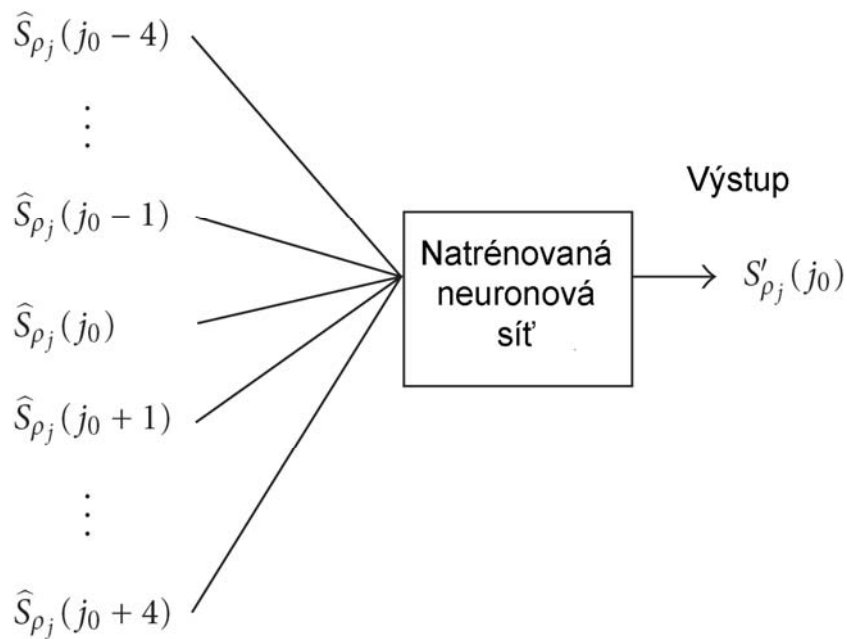
$$\hat{S}_{\rho_j}(l) = \sum_{n=1}^{256} \omega(n) \hat{s}_{\rho_j}(n) \cos \frac{\pi(2n-1)(l-1)}{512}, \quad (6.1)$$

kde $\hat{s}_{\rho_j}(n)$ představuje n -tý vzorek vodoznačeného audio bloku \hat{s}_{ρ_j} a $1 \leq l \leq 256$. $\hat{\Phi}$ tak představuje množinu vodoznačených transformovaných audio bloků.

Během procesu extrahování vodoznaku je natrénovaná neuronová síť využita k odhadu originálního audio signálu. Vstupní vektor pro trénovanou neuronovou síť je

$$(\hat{S}_{\rho_j}(j_0-4), \dots, \hat{S}_{\rho_j}(j_0-1), \hat{S}_{\rho_j}(j_0), \hat{S}_{\rho_j}(j_0+1), \dots, \hat{S}_{\rho_j}(j_0+4)). \quad (6.2)$$

Vektor je výběrem z \hat{S}_{ρ_j} množiny $\hat{\Phi}$, která může být také zkreslená díky útokům nebo různým manipulacím se signálem. Obr. 6.2 znázorňuje vstupní hodnoty a odpovídající výstup trénované neuronové sítě [5].



Obr. 6.2 Vstupy a výstup natrénované neuronové sítě při extrakci vodoznaku

Pomocí vztahů (6.3) a (6.4) je možné vodoznak w' rekonstruovat. $w_j' = -1$ představuje odpovídající černý pixel binární známky (vodoznaku), $w_j' = 1$ pak pixel bílý.

$$w' = (w_1', \dots, w_j', \dots, w_{pq}') \quad (6.3)$$

$$w_j' = \begin{cases} 1, & \text{pokud } [\hat{S}_{\rho_j}(j_0) - S'_{\rho_j}(j_0)] > 0, \\ -1, & \text{jinak} \end{cases} \quad (6.4)$$

7 ROBUSTNOST ALGORITMU

Při návrhu systému vodoznačení je kladen důraz hlavně na robustnost a neslyšitelnost vodoznaku. Zde popsaná metoda vodoznačení využívá frekvenčního maskování vycházejícího z psychoakustického modelu pro MPEG layer 1. Vložený vodoznak by tak měl být lidským sluchem nepostřehnutelný. Pro účely testování transparentnosti byl použit výpočet odstupe signálu od šumu SNR (signal to noise ratio) dle následujícího vztahu:

$$\text{SNR} = -10 \log_{10} \left(\frac{\sum_{n=1}^N [x(n) - x'(n)]^2}{\sum_{n=1}^N [x(n)]^2} \right) [dB], \quad (7.1)$$

kde $x(n)$ originální audio signál a $x'(n)$ je vodoznačený audio signál. Tento objektivní způsob hodnocení kvality můžeme brát jako správný, ale jeho vypovídací hodnota není zcela směrodatná.

Pro posouzení shody mezi originálním a extrahovaným vodoznakem je použito normalizované vzájemné korelace NCC (normalized cross-correlation):

$$\text{NCC} = \left(\frac{\sum_{i=0, j=0}^{I-1, J-1} w_{ij} w'_{ij}}{\sum_{i=0, j=0}^{I-1, J-1} (w_{ij})^2} \right), \quad (7.2)$$

kde I a J jsou horizontální a vertikální rozměr vodoznaku. w je originální vodoznak, w' vytažený vodoznak. Výstup funkce je v rozmezí 0 až 1. K stanovení hranice rozeznatelnosti vytaženého vodoznaku se svým originálem, jsem uvažil podle svého vizuálního subjektivního dojmu (obr. 7.1). Hranice rozeznatelnosti vodoznaku odpovídá hodnotě $\text{NCC}=0,81$.



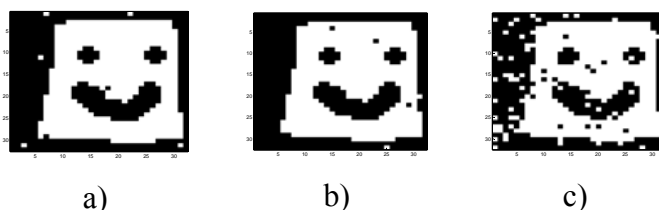
Obr. 7.1 Vytažený vodoznak s hodnotou $\text{NCC}=0,818$

Testy byly provedeny pro tři zvukové nahrávky různých hudebních stylů (rock, pop, klasika) formátu *.wav s vzorkovací frekvencí $F_{vz}=44100\text{Hz}$. Do nahrávek byl vložen

vodoznak o rozměrech 32×32 pixelů znázorněný na obr. 5.2. Hodnoty NCC a SNR jsou uvedeny v tabulce 7.1, na obr. 7.2 jsou pak znázorněny extrahované vodoznaky.

Tabulka 7.1: Hodnoty vodoznačených nahrávek bez změny obsahu

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock	0,998	34,7
Pop	0,994	37,1
Klasika	0,971	64,8



Obr. 7.2 Vytažený vodoznak z nepozměněné nahrávky a)rock, b)pop, c)klasika

U těchto nahrávek byla posléze provedena následující změna jejich obsahu:

- **Filtrace DP** – K filtraci byl použit filtr typu dolní propust s Butterworthovou aproximací 10-řádu s mezním kmitočtem 500Hz. Výsledné hodnoty NCC a SNR jsou uvedeny v tabulce 7.2, na obr. 7.3 jsou znázorněny extrahované vodoznaky.
- **Filtrace HP** – K filtraci byl použit filtr typu horní propust s Butterworthovou aproximací 10-řádu s mezním kmitočtem 18kHz. Výsledné hodnoty NCC a SNR jsou uvedeny v tabulce 7.3, na obr. 7.4 jsou znázorněny extrahované vodoznaky.
- **Přidání šumu do signálu** – Do nahrávek byl přidán aditivní šum (20%). Výsledné hodnoty NCC a SNR jsou uvedeny v tabulce 7.4, na obr. 7.5 jsou znázorněny extrahované vodoznaky.
- **Převzorkování** – Byla provedena decimace a následovná interpolace nahrávek činitelem $M=2$. Výsledné hodnoty NCC a SNR jsou uvedeny v tabulce 7.5, na obr. 7.6 jsou znázorněny extrahované vodoznaky.
- **Mediánová filtrace** – Byla provedena filtrace mediánovým filtrem o délce 3. Výsledné hodnoty NCC a SNR jsou uvedeny v tabulce 7.6, na obr. 7.7 jsou znázorněny extrahované vodoznaky.
- **Kompresi signálu** – Ke kompresi nahrávek byl využit formát ogg pro bitové toky 60kbps, 98kbps, 120kbps. Nahrávky byly posléze převedeny zpět do původního formátu *.wav. Výsledné hodnoty NCC a SNR jsou uvedeny v tabulkách 7.7-7.9, na obr. 7.8-7.10 jsou znázorněny extrahované vodoznaky.

Tabulka 7.2: Hodnoty testu robustnosti po filtraci DP

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock	0,992	34,6
Pop	0,990	25,4
Klasika	0,973	64,7



a) b) c)

Obr. 7.3 Vytažený vodoznak po filtraci DP a)rock, b)pop, c)klasika

Tabulka 7.3: Hodnoty testu robustnosti po filtraci HP

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock	0,993	1,8
Pop	0,987	1,6
Klasika	0,921	2,1



a) b) c)

Obr. 7.4 Vytažený vodoznak po filtraci HP a)rock, b)pop, c)klasika

Tabulka 7.4: Hodnoty testu robustnosti po přidání šumu

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock	0,935	17,7
Pop	0,889	13,2
Klasika	0,588	11,0



a) b) c)

Obr. 7.5 Vytažený vodoznak po přidání šumu a)rock, b)pop, c)klasika

Tabulka 7.5: Hodnoty testu robustnosti po převzorkování

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock	0,596	3,0
Pop	0,596	3,0
Klasika	0,552	3,0



a)



b)



c)

Obr. 7.6 Vytážený vodoznak po převzorkování a)rock, b)pop, c)klasika

Tabulka 7.6: Hodnoty testu robustnosti po mediánové filtraci

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock	0,649	23,2
Pop	0,661	24,8
Klasika	0,611	35,9



a)



b)



c)

Obr. 7.7 Vytážený vodoznak po mediánové filtraci a)rock, b)pop, c)klasika

Tabulka 7.7: Hodnoty testu robustnosti po kompresi ogg

Testovaná nahrávka	NCC [-]	SNR [dB]
Pop (120kbps)	0,835	25,4
Pop (96kbps)	0,814	22,1
Pop (60kbps)	0,550	17,6



a)



b)



c)

Obr. 7.8 Vytážený vodoznak po kompresi ogg Pop a)120kbps, b)96kbps, c)60kbps

Tabulka 7.8: Hodnoty testu robustnosti po kompresi ogg

Testovaná nahrávka	NCC [-]	SNR [dB]
Rock (120kbps)	0,856	23,8
Rock (96kbps)	0,824	19,8
Rock (60kbps)	0,598	15,8



a)



b)



c)

Obr. 7.9 Vytážený vodoznak po kompresi ogg Rock a)120kbps, b)96kbps, c)60kbps

Tabulka 7.9: Hodnoty testu robustnosti po kompresi ogg

Testovaná nahrávka	NCC [-]	SNR [dB]
Klasika (120kbps)	0,776	28,4
Klasika (96kbps)	0,728	25,3
Klasika (60kbps)	0,559	19,7



a)



b)



c)

Obr. 7.10 Vytážený vodoznak po kompresi ogg Klasika a)120kbps, b)96kbps, c)60kbps

8 ZÁVĚR

Cílem této práce bylo vytvoření algoritmu pro digitální vodoznačení audio signálů s využitím umělých neuronových sítí.

V první části práce byly teoreticky rozebrány základy umělých neuronových sítí. Byla popsána jejich analogie s biologickými neuronovými sítěmi, základní model neuronu a jeho funkce, principy učení neuronových sítí a jejich základní topologie. Podrobně byly rozebrány některé typy neuronových sítí. Konkrétně neuronová síť typu perceptron, síť se zpětným šířením chyby Back-propagation a její modifikovaná varianta s proměnným krokem učení.

V druhé části práce je uveden základní princip vodoznačení audio signálů. Je zde popsána metoda vodoznačení, která využívá k vložení vodoznaku ve frekvenční oblasti DCT transformaci. Tato metoda je založena na psychoakustickém modelu, který je používán v MPEG layer 1, a neuronové síti. Z důvodu rychlejší konvergence chybové funkce ke svému minimu, byla použita neuronová síť se zpětným šířením chyby Back-propagation s proměnným krokem učení. Při procesu vkládání vodoznaku se neuronová síť učí. Natrénovaná neuronová síť pak uchovává vztahy mezi originálním a vodoznačeným audio signálem. Díky tomu se při extrahování vodoznaku nevyžaduje přístup k originálnímu audio signálu.

Na závěr byl proveden test robustnosti dané metody. Metoda byla vystavena různým druhům útoků (filtrace, komprese, převzorkování, zašumění audio signálu). Efektivita detekce vodoznaku byla hlavně závislá na správném natrénování neuronové sítě. Vyšší hodnoty SNR byly do jisté míry ovlivněny délkou testovaných nahrávek a velikostí vkládaného vodoznaku. Díky využití maskovacích křivek psychoakustického modelu při vkládání vodoznaku, nebyla vložená informace ve výsledném vodoznačeném signálu téměř slyšet.

Z testování dané metody lze vyvodit, že využitím neuronových sítí v oblasti vodoznačení může dojít k zvýšení robustnosti metod proti běžným útokům. To je dáno hlavně schopností generalizace neuronové sítě.

LITERATURA

- [1] TUČKOVÁ, J.: *Úvod do teorie a aplikací umělých neuronových sítí*, Skripta FEL ČVUT v Praze, vydavatelství ČVUT, 2005, ISBN 80-01-02800-3.
- [2] JIRSÍK, V., HRÁČEK, P.: *Neuronové sítě, expertní systémy a rozpoznávání řeči*, Skripta FEKT VUT v Brně.
- [3] QUANG, X., ZHANG, H.: *Statistical audio watermarking algorithm based on perceptual analysis*, Proceedings of the 5th ACM workshop on Digital rights management 2005, Alexandria, VA, USA. November 07-07, 2005.
- [4] ARNOLD, M., SCHMUCKER, M., WOLTHUSEN, D. S.: *Techniques and Applications of Digital Watermarking and Content Protection*, Boston: Artech House Publishers, 2003, ISBN 1-58053-111-3.
- [5] TSAI, H., CHENG, J., YU, P.: *Audio Watermarking Based on HAS and Neural Networks in DCT Domain*, EURASIP Journal on Applied Signal Processing, pp. 252-263, 2003.
- [6] CHANG, CH., SHEN, W.: *Using Counter-propagation Neural Network for Digital Audio Watermarking*, Thesis, National Yunlin University of Science & Technology, Yunlin, 2005.
- [7] KETCHAM, M., VONGPRADHIP, S.: *Intelligent Audio Watermarking using Genetic Algorithm in DWT Domain*, International Journal Of Intelligent Technology Vol. 2, No. 2, 2007 pp. 135-140.
- [8] JIRSÍK, V.: *Umělá inteligence*, (přednáška) Brno: VUT Brno, 15.10.2007.
- [9] DRÁBEK, O., TAUFER, I., SEIDL, P.: *Umělé neuronové sítě – základy teorie a aplikace (4)*. CHEMagazín 2 (XVI), 2006, s. 33–36. ISSN 1210-7409.
- [10] TAKAYA, K.: *Multimedia Signals and Systems: MP3 – Mpeg 1, 2 layer 1, 2, 3 Audio Encoding*, University of Saskatchewan, March 26, 2008. Dostupné z: <<http://www.engr.usask.ca/classes/CME/462/notes/mycme462-MP3.pdf>>
- [11] W.A.W.S. Compression – Background on the Psychoacoustic Model. Dostupné z: <<http://is.rice.edu/~welsh/elec431/psychoAcoustic.html>>

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

<i>ATH</i>	absolute threshold
<i>DCT</i>	diskrétní kosinová transformace
<i>DFT</i>	diskrétní Fourierova transformace
<i>DFMT</i>	determining frequency masking threshold
<i>DWT</i>	diskrétní vlnková transformace
<i>FFT</i>	rychlá Fourierova transformace
<i>IDCT</i>	inversní diskretní kosinová transformace
<i>LBF</i>	lineární basická funkce
<i>LSB</i>	last significant bit
<i>LT_g</i>	globální maskovací práh
<i>LT_{nm}</i>	maskovací křivky pro netónové složky
<i>LT_{tm}</i>	maskovací křivky pro tónové složky
<i>NCC</i>	normalized cross-correlation
<i>PCM</i>	pulse code modulation
<i>PRNG</i>	pseudorandom number generator
<i>PSD</i>	power spectral density
<i>RBF</i>	radiální basická funkce
<i>SNR</i>	signal to noise ratio
<i>SPL</i>	sound pressure level [dB]
<i>X_{nm}</i>	spektrum netónových složek
<i>X_{tm}</i>	spektrum tónových složek