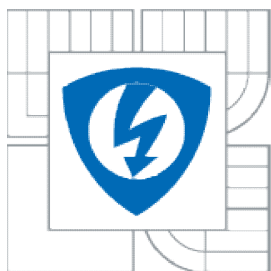




VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATION

PERCEPČNÍ KÓDOVÁNÍ ZVUKOVÝCH SIGNÁLŮ

PERCEPTUAL AUDIO CODING

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

BC. VLADIMÍR NOVÁK

VEDOUCÍ PRÁCE
SUPERVISOR
BRNO 2011

ING. JIŘÍ SCHIMMEL, PH.D.

ABSTRACT

This thesis describes Perceptual Audio Coding of MPEG1 Layer 3 format (ISO/IEC 11172-3), principles and algorithms of psychoacoustic model. MATLAB application for modeling of Psychoacoustic model 2 of this audio format is realized.

KEYWORDS

MPEG-1 Layer 3, MP3, perceptual coding, psychoacoustic model, spreading function

ABSTRAKT

Tato diplomová práce se zabývá percepčním kódováním zvukových signálů u formátu MPEG1 Layer 3 (ISO/IEC 11172-3). Popisuje principy a algoritmy psychoakustického modelu. Výstupem práce je program v prostředí MATLAB simulující Psychoakustický model 2 tohoto formátu.

KLÍČOVÁ SLOVA

MPEG-1 Layer 3, MP3, percepční kódování, psychoakustický model, maskovací křivka

NOVÁK, V. *Percepční kódování zvukových signálů*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2011. 57 s. Vedoucí diplomové práce Ing. Jiří Schimmel, Ph.D.

Prohlašuji, že svou diplomovou práci na téma Percepční kódování zvukových signálů jsem vypracoval samostatně pod vedením vedoucího diplomové práce s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny uvedeny v seznamu literatury na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení paragrafu 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení paragrafu 152 trestního zákona č. 140/1961 Sb.

V Brně dne

.....
podpis autora

Děkuji vedoucímu diplomové práce Ing. Jiřímu Schimmelovi, Ph. D. za velmi užitečnou metodickou pomoc a cenné rady pro zpracování práce. Dále bych chtěl poděkovat své rodinně za klid, který mi na práci poskytli.

V Brně dne

.....
podpis autora

OBSAH

1	Percepční kódování zvukového signálu	6
1.1	Lidské vnímání zvuku	6
1.2	Kritická pásma slyšení	7
1.3	Úrovně akustického tlaku	9
1.4	Maskování	10
1.4.1	Frekvenční maskování	10
1.4.2	Dočasné maskování	11
1.4.3	Funkce pro aplikaci maskovacích křivek	12
1.4.4	Sčítání maskovacích křivek	14
2	Standard MPEG-1	15
2.1	Princip MPEG	15
2.2	Vrstvový model MPEG-1	15
3	MPEG-1 Layer 3	17
3.1	Historie	17
3.2	Kodér	17
3.3	Dekodér	19
4	Psychoakustický model	21
4.1	Psychoakustický model 1	21
4.2	Psychoakustický model 2	22
4.2.1	Vstup signálu	24
4.2.2	Nepředvídatelnost	25
4.2.3	Výpočet energie a nepředvídatelnosti ve výpočetních pásmech	26
4.2.4	Konvoluce energie a nepředvídatelnosti s maskovací křivkou	27
4.2.5	Index tonality	27
4.2.6	Odstup signálu od šumu	28
4.2.7	Spektrum slyšitelného zvuku	28

4.2.8	Perceptuální entropie	29
4.2.9	Odstup signálu od masky.....	30
5	Simulace Psychoakustického modelu MPEG-1 Layer 3 v programu Matlab.....	32
5.1	Volba programu - MATLAB.....	32
5.2	Realizace programu	32
5.2.1	Ovládání programu	33
5.2.2	Realizace Psychoakustického modelu 2	34
5.2.3	Výběr vstupního signálu.....	34
5.2.4	Rychlá Fourierova transformace	35
5.2.5	Nepředvídatelnost.....	36
5.2.6	Zobrazení energie a nepředvídatelnosti v konvoluci s maskovacími křivkami.....	37
5.2.7	Index tonality	39
5.2.8	Odstup signálu od šumu	40
5.2.9	Spektrum slyšitelného zvuku.....	40
5.2.10	Odstup signálu od masky	42
6	Simulace banky filtrů a bloků mdct v mpeg 1 Layer 3 v programu matlab ..	43
6.1	Realizace programu	43
6.2	Banka filtrů.....	43
6.3	Blok MDCT	43
7	Závěr.....	45

SEZNAM OBRÁZKŮ

Obr. 1.1: Struktura vnitřního ucha - hlemýžďe s vyznačenými frekvencemi. [8].....	7
Obr. 1.2: Barková stupnice odvozená z tab 1.1.	8
Obr. 1.3: Práh slyšitelnosti	9
Obr. 1.4: Frekvenční maskování harmonickým signálem [4].....	11
Obr. 1.5: Dočasné maskování [4].....	12
Obr. 1.6: Funkce pro aplikaci maskovacích křivek.....	13
Obr. 3.1: Kodér MPEG-1 Layer 3 [7].....	18
Obr. 3.2: Dekodér MPEG-1 Layer 3 [7]	20
Obr. 4.1: Schéma psychoakustického modelu 1	21
Obr. 4.2: Blokové schéma psychoakustického modelu MPEG-1 Layer 3.....	23
Obr. 4.3: Okénkové funkce - Hannovo okno (modrá křivka), Blackmanovo okno (červená křivka), Hammingovo okno (zelená křivka).....	25
Obr. 4.4: Přepínací diagram oken [7].....	30
Obr. 5.1: Ovládací panel programu.....	33
Obr. 5.2: Časový průběh signálu 3	34
Obr. 5.3: Frekvenční spektrum signálu 3	35
Obr. 5.4: Spektrum fáze signálu 3.....	35
Obr. 5.5: Nepředvídatelnost signálu 3	36
Obr. 5.6: Nepředvídatelnost neperiodického signálu (signál 1).....	37
Obr. 5.7: Energie (červená barva) a maskovací křivka (modrá barva) pro signál 3 ..	37
Obr. 5.8: Energie po konvoluci s maskovací křivkou pro signál 3	38
Obr. 5.9: Energie (červená barva) a maskovací křivka (modrá barva) pro signál 2 ..	38
Obr. 5.10: Energie po konvoluci s maskovací křivkou pro signál 2	39
Obr. 5.11: Index tonality u signálu 3	39
Obr. 5.12: Odstup signálu od šumu (SNR) u signálu 3.....	40

Obr. 5.14: Práh energie (modrá barva) a práh slyšitelnosti (červená barva) u signálu 3	41
Obr. 5.15 Slyšitelné spektrum signálu 3	41
Obr. 5.16: Odstup signálu od šumu - energie signálu (zelená barva), práh energie (červená barva) vyjádřené v měřítkových koeficientech.....	42
Obr. 5.17: Odstup signálu od masky.....	42
Obr. 6.1: Výstup z MDCT bloku	44

ÚVOD

Pojem MP3 dnes zná téměř každý člověk a to nemusí být žádný expert na výpočetní techniku. Stal se nejpoužívanějším formátem pro přenos a reprodukci hudby a mluveného slova. Odborný název MPEG-1 Layer 3 již každý člověk nezná a přitom se pod ním skrývá právě tato zkratka.

Otázka zní proč a za jakým účelem tento formát na uchovávání hudby a její reprodukci vnikl. Dříve se hudba uchovávala na gramofonových deskách, později na magnetofonových páscích a CD. Všechna tato jmenovaná média dokázala uchovat jen omezenou část nahrávek (na jedno médium se vešlo průměrně 90 minut nahrávky). S lepší finanční dostupností počítačů postupně začaly vnikat kompresní algoritmy, které dokázaly redukovat velikost datových souborů. Tyto techniky se nejprve využívaly pro přenos informací a také pro archivaci dat. Bylo jen otázkou času, kdy se tyto algoritmy začnou používat i pro kompresy hudby.

V roce 1993 byl normalizován MPEG-1 Layer 3 a velmi rychle se rozšířil. Hlavním důvodem byl fakt, že soubory vytvořené touto kompresní technikou jsou velmi malé (průměrná písnička o délce 3 minut je velká cca 3 Mb) a kvalita reprodukce je také velmi dobrá.

Tato diplomová práce si klade několik cílů, které čtenáři přiblíží princip MP3. Je vysvětlen princip, jak člověk vnímá zvuk a jak se dají využít "nedokonalosti" lidského sluchu pro datovou kompresi audiosignálu. Poté jsou objasněny principy tzv. psychoakustické komprese, kterou využívá formát MPEG-1 Layer 3.

Dále je čtenář seznámen s pracovní skupinou MPEG a samotným formátem MPEG-1 Layer 3, který je specifikován v normě ISO/IEC 11172-3.

Hlavním cílem je popis a tvorba simulačního programu Psychoakustického modelu 2 vytvořeném v prostředí MATLAB, který má sloužit pro výuku studentů. Program musí názorně zobrazovat jednotlivé funkční bloky tohoto modelu, které objasní jak daný kódér MP3 funguje.

1 PERCEPČNÍ KÓDOVÁNÍ ZVUKOVÉHO SIGNÁLU

U percepčního kódování se při kompresi zvukového signálu využívá nedostatků lidského vnímání zvuku. Díky tomuto kódování můžeme přenášet a uchovávat zvukové záznamy v relativně malých datových souborech s přijatelnou kvalitou.

1.1 Lidské vnímání zvuku

Člověk zvuk vnímá smyslem nazývaným sluch. Podnětem pro sluch jsou zvukové vlny, což je podélné kmitání molekul vzduchu. Člověk je sluchem schopen rozeznat zvuky, tóny, intenzitu, výšku, zbarvení zvuku a směr, odkud zvuk přichází. Lidské ucho dokáže vnímat zvuky o frekvencích od 16 Hz do 20 kHz [1].

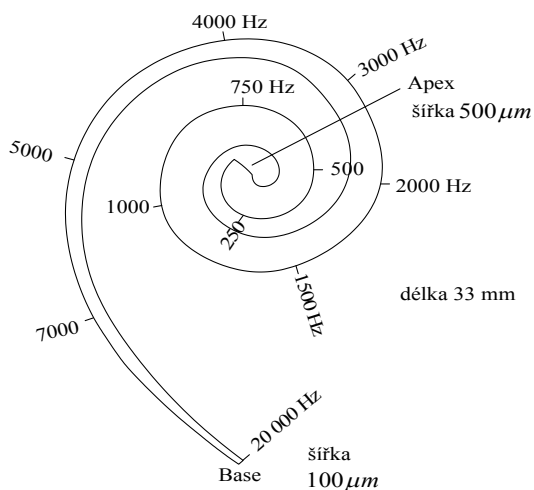
Lidské ucho je ústrojí sloužící k vnímání zvuku a je také centrem rovnováhy. Skládá se ze tří částí: zevní, střední a vnitřní ucho.

Zevní ucho má za úkol zachytit zvukové vlny a dovést je k bubínku. Skládá se z ušního boltce a ušní trubice (dlouhé asi 25 mm) zakončené bubínkem, což je pružná blána chránící střední ucho.

Střední ucho převádí zvukové vjemy z plynného prostředí zevního ucha na kapalně prostředí vnitřního ucha. Skládá se ze tří důležitých kůstek: kladívko, třmínek a kovadlinka. Střední ucho dokáže částečně sluch chránit proti hlasitým zvukům.

Vnitřní ucho začíná oválným a kulatým otvorem vyplněným blánou, do kterého je usazen třmínek. Na ně již navazuje hlemýžď, což je stočená trubice do dvou a půl otáčky vyplněná Cortiho orgánem, který převádí pomocí vláskových buněk (sluchových receptorů) energii zvukových vln na elektrický signál [1].

Maďarský vědec Georg von Békésy zjistil, že zvuk s vyššími frekvencemi člověk vnímá na základně hlemýžďe, a čím více se frekvence snižují, tím je vnímá dále uvnitř tohoto orgánu, viz obr. 1.1. [2]



Obr. 1.1: Struktura vnitřního ucha - hlemýžďe s vyznačenými frekvencemi. [8]

1.2 Kritická pásma slyšení

Na základě experimentů bylo zjištěno, že lidské ucho rozděluje zvuk do 24 frekvenčních pásem, které odpovídají určitým polohám na basilární membráně v hlemýždi ve vnitřním uchu. Tyto pásma nazýváme kritická pásma slyšení. Lidské ucho nedokáže v jednotlivých kritických pásmech rozlišit více frekvencí a tím vzniká tzv. frekvenční maskování. Jednotlivá pásma nejsou stejně veliká, pásma s nižšími frekvencemi mají menší frekvenční rozsah než pásma s frekvencemi vyššími.

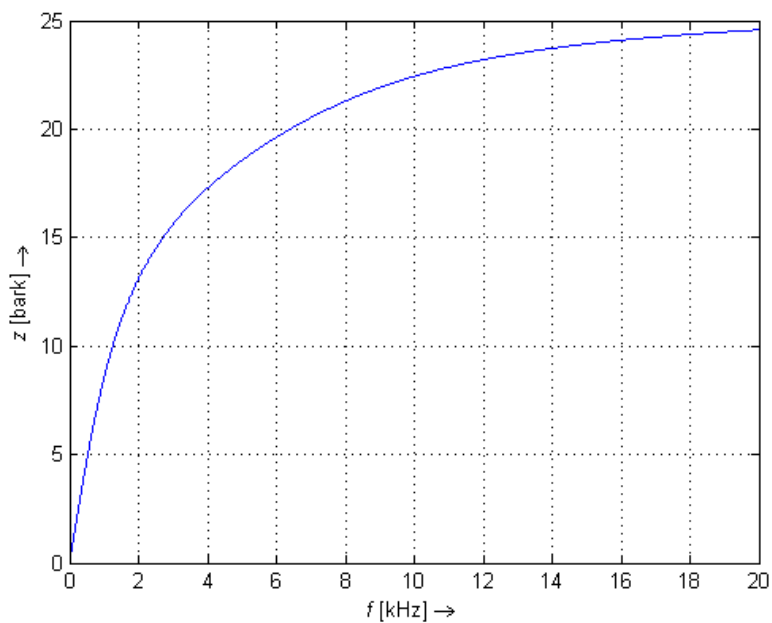
Pro psychoakustiku byla zavedena tzv. barková stupnice s jednotkami Bark (pojmenovaná po německém vědci Heinrichu Barkhausenovi), která vychází z fyzických dispozic lidského ucha, viz. obr. 1.2 a tab. 1.1. Jeden Bark se rovná jednomu segmentu o stejné délce v hlemýždi lidského ucha. Jednotlivé segmenty odpovídají 24 frekvenčním pásmům, ve kterých člověk vnímá frekvence stejně, viz obr. 1.1. Pro převod jednotek Hz na Barky se používá rovnice [2]

$$z(f) = 13 \cdot \arctan\left(\frac{0,76 \cdot f}{1000}\right) + 3,5 \cdot \arctan\left(\left(\frac{f}{7500}\right)^2\right) [Bark]. \quad (1.1)$$

Barkova stupnice se používá ke zjištění např. maskovacího efektu u signálu v psychoakustickém modelu.

Tab. 1.1: Kritická pásma s vyjádřením rozsahu frekvencí [2].

z (Bark)	f_{\min}(Hz)	f_{\max} (Hz)	Δf (Hz)
0	0	100	100
1	100	200	100
2	200	300	100
3	300	400	100
4	400	510	110
5	510	630	120
6	630	770	140
7	770	920	150
8	920	1080	160
9	1080	1270	190
10	1270	1480	200
11	1480	1720	240
12	1720	2000	280
13	2000	2320	320
14	2320	2700	380
15	2700	3150	450
16	3150	3700	550
17	3700	4400	700
18	4400	5300	900
19	5300	6400	1100
20	6400	7700	1300
21	7700	9500	1800
22	9500	12000	2500
23	12000	15500	3500
24	15500		



Obr. 1.2: Barková stupnice odvozená z tab 1.1.

1.3 Úrovně akustického tlaku

Lidské vnímání akustického tlaku zvukového signálu má logaritmický průběh. Dynamickým rozsahem rozumíme poměr nejhlasitějšího a nejtiššího vnímatelného zvuku. Dynamický rozsah lidského sluchu je přibližně od 10^{-5} do 10^2 Pa, a je frekvenčně závislý.

Hladina akustického tlaku (Sound Pressure Level - SPL) je veličina daná vztahem

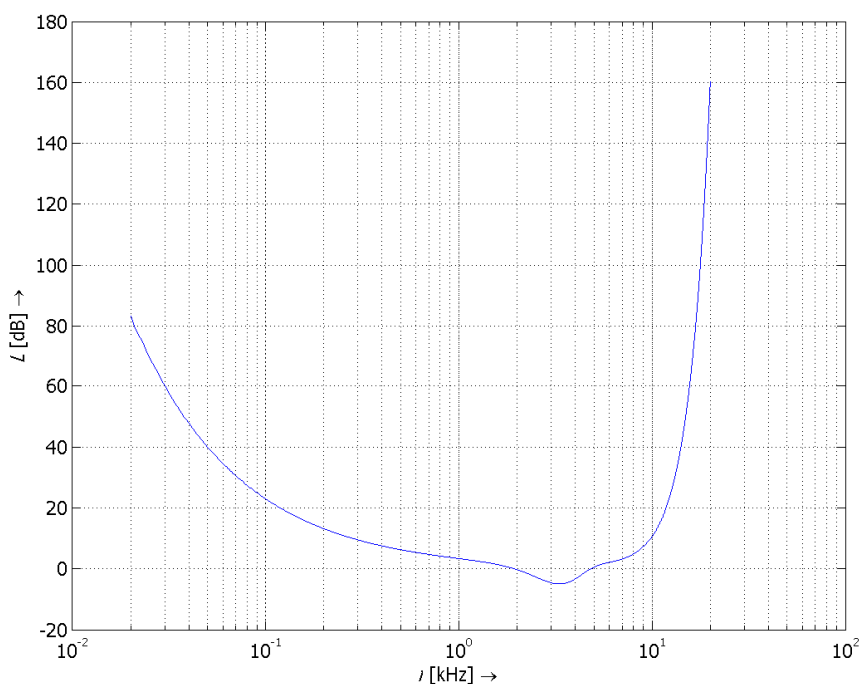
$$SPL = 20 \cdot \log\left(\frac{p}{p_0}\right) [dB], \quad (1.2)$$

kde p je daný akustický tlak a p_0 je tzv. referenční (prahový) akustický tlak, který je roven $20 \mu\text{Pa}$. Tato hodnota je platná pro kmitočet 1 kHz a pro běžné atmosférické podmínky [3].

Pro člověka se signál se stejnou úrovní akustického tlaku na různých frekvencích jeví jinak hlasitý.

Dalším důležitým pojmem je práh slyšitelnosti (threshold in quiet). Jedná se o závislost minimálně postřehnutelného akustického tlaku, který vyvolá u člověka s průměrným sluchem zvukový vjem. Člověk vnímá zvuk se stejnou intenzitou na různých frekvencích jinak, což je dáno křivkou na obr. 1.3 a vztahem

$$A(f) = 3,64 \cdot \left(\frac{f}{1000}\right)^{-0,8} - 6,5e^{-0,6 \cdot \left(\frac{f}{1000} - 3,3\right)^2} + 10^{-3} \left(\frac{f}{1000}\right)^4 [dB]. \quad (1.3)$$



Obr. 1.3: Práh slyšitelnosti

Největší citlivost lidského sluchu je v oblasti 3 až 4 kHz. Za dolní sluchovou mez je považována frekvence 16 Hz, za horní mez 20 kHz. Horní sluchová mez je proměnná, ovlivňuje jí například stáří posluchače [3].

Všechny frekvence, které jsou pod křivkou prahu slyšitelnosti, jsou pro lidské ucho neslyšitelné, a tudíž se mohou odebrat bez možného slyšitelného zkreslení.

1.4 Maskování

Maskování je definováno jako snížená schopnost vnímání zvuku z důvodu přítomnosti jiného zvuku. Kmitočtově blízké zvuky s nižší hladinou akustického tlaku nemusí být při současném znění zvuku s vyšší hladinou slyšitelné. Maskování souvisí s činností sluchových receptorů, které se při podráždění ocitnou po dobu několika milisekund ve stavu, kdy nejsou schopny reagovat na další vzruch [3].

Zvuku s nižší hladinou akustického tlaku říkáme maskovaný zvuk, zvuku s vyšší hladinou maskující zvuk. Zvuk s nižší hladinou akustického tlaku může překročit tzv. maskovací práh a tím začne být slyšitelný.

Maskování můžeme rozdělit na frekvenční (frequency masking) a dočasné maskování (temporal masking).

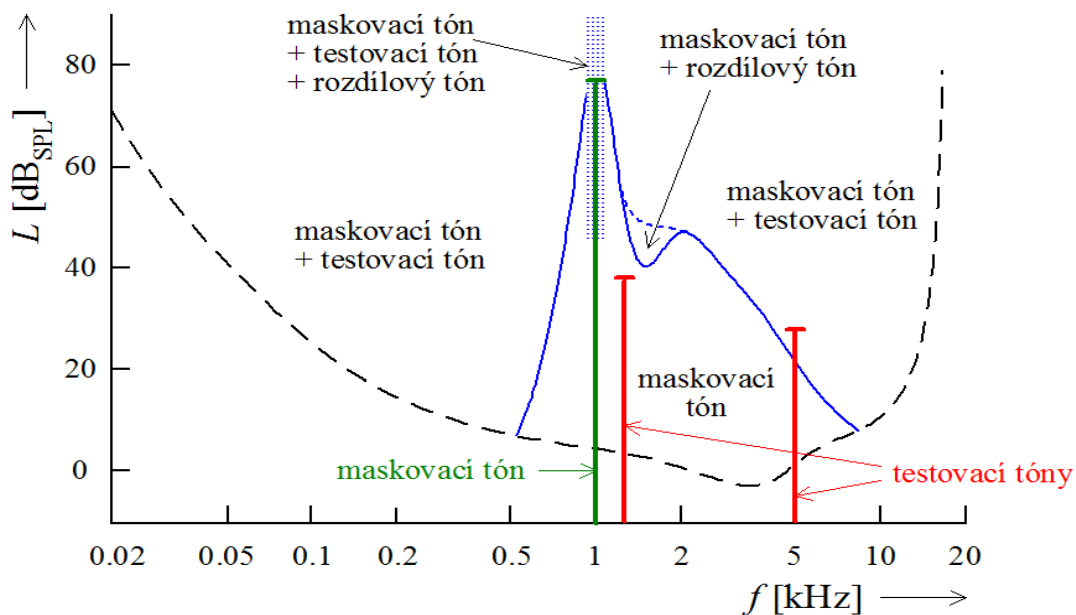
1.4.1 Frekvenční maskování

U frekvenčního maskování překrývá tón se silnější intenzitou jiný tón se slabší intenzitou. Oba tóny se vyskytují ve stejný časový okamžik. Maskování můžeme rozdělit na signály maskované harmonickým signálem nebo maskované šumem.

Na obr. 1.4 je zjednodušeně zobrazeno frekvenční maskování harmonického signálu. Práh slyšitelnosti je zobrazen čárkovaně. Zelenou barvou je zobrazen maskovací tón, který maskuje dva tóny maskované. Modrou křivkou je zobrazeno spektrum slyšitelného zvuku.

Maskování šumem je mnohem efektivnější, co se týče šířky pásma, než maskování signálem. Na rozdíl od signálu, ale nedosahuje takové intenzity.

Důležitým prvkem u psychoakustické komprese je správné rozlišení, zda se maskuje šumem či maskovacím tónem. Využívá se k tomu technika detekce lokálních maxim, kde maskovací signál bude maximum [4].



Obr. 1.4: Frekvenční maskování harmonickým signálem [4]

1.4.2 Dočasné maskování

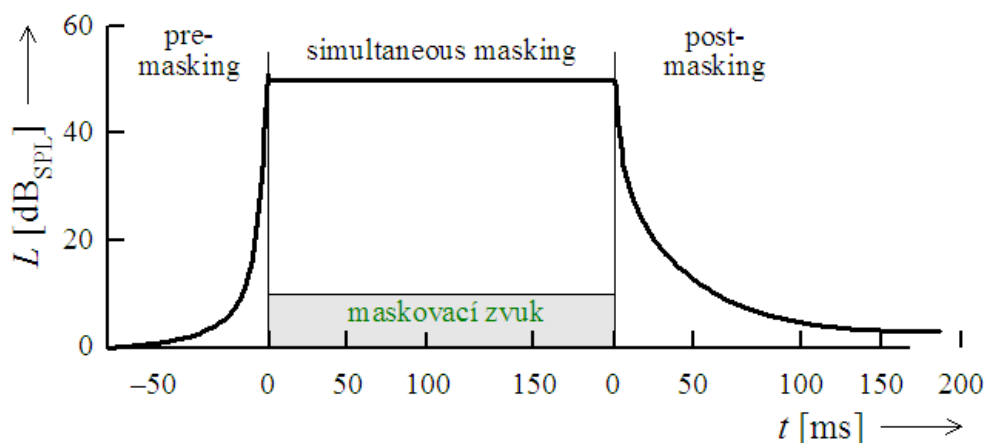
O dočasném maskování hovoříme, pokud se maskující a maskovací signál nevyskytují ve stejném čase, viz obr 1.5.

Zvuk s nižší hladinou akustického tlaku může být maskován při současném znění maskujícího signálu, potom hovoříme o frekvenčním maskování viz kap. 1.4.1.

Pre-masking je zpětné maskování, které vzniká v lidském uchu jako reakce na nastupující maskující signál. Zpětné maskování se využívá u psychoakustické komprese při kvantizaci vzorků. Při nástupu silnějšího signálu může dojít vinou kvantizace k tzv. pre-echu – zvuk slyšíme dopředu. Tento efekt se dá odstranit použitím krátkého okna (viz kap. 4.2.8).

Dopředné maskování (post-masking) je důsledek integračního chování lidského sluchu. Díky této vlastnosti sluchu dochází k zamaskování zvuku s nižší hladinou akustického tlaku, který přichází bezprostředně po signálu s hladinou vyšší [4].

Na obr. 1.5 je znázorněn časový průběh se zobrazením jednotlivých maskovacích efektů. Z obrázku je patrné, že zpětné maskování trvá mnohem kratší časový okamžik než maskování dopředné.



Obr. 1.5: Dočasné maskování [4]

1.4.3 Funkce pro aplikaci maskovacích křivek

Maskovací křivky ukazují závislost maskovacího prahu na kmitočtu. Jsou transformované do barkové stupnice, díky čemuž jsou snadněji aproximovatelné. Aplikací této funkce rozprostřeme bodové hodnoty do obou sousedních oddílů.

Nejjednodušší aproximací reprezentuje tzv. trojúhelníková maskovací funkce, která je daná vztahem

$$F(dz, L_M) = \begin{cases} 27dz & (\text{pro } dz < 0) \\ (-27 + 0,37 \text{ MAX}\{L_M - 40, 0\}) & (\text{pro } dz > 0) \end{cases}, \quad (1.4)$$

kde L_M je maskovací hladina akustického tlaku (SPL). Graficky je funkce zobrazena na obrázku 1.6 světlou barvou pro různé úrovně [5].

Mezi další aproximace patří tzv. Schroederova maskovací funkce, ze které se později vyvinula funkce použitá v normě MPEG 1 Layer 3. Je dána vztahem

$$F(dz, L_M, f) = (15,81 - I(L_M, f)) + 7,5(dz + 0,474) - (17,5 - I(L_M, f))\sqrt{1 + (dz + 0,474)^2}, \quad (1.5)$$

kde L_M je maskovací úroveň akustického tlaku (SPL) a I je funkce pro přizpůsobení úrovně, která je definována vztahem

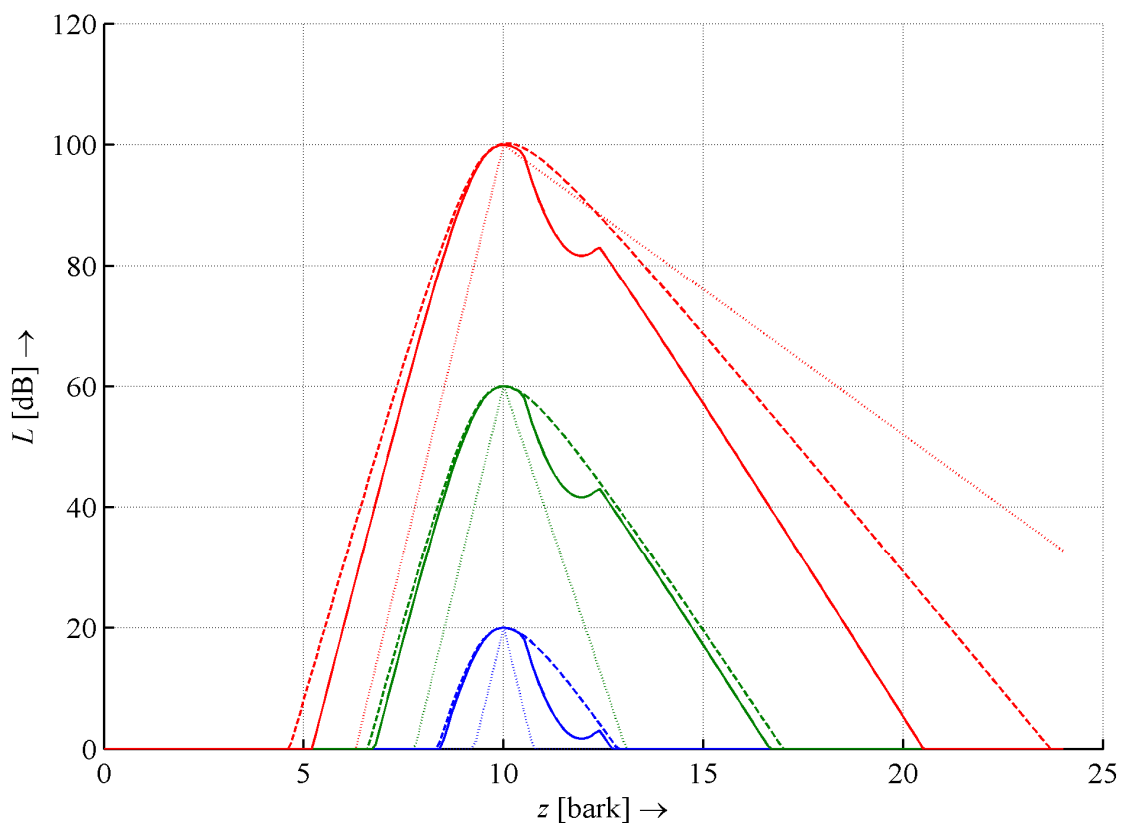
$$I(L_M, f) = \text{MIN} \left\{ \frac{5.10^{\frac{(L_M-96)}{10}} df}{dz}, 2 \right\}, \quad (1.6)$$

kteřá má frekvenční závislost danou rozptylem kritického pásma df/dz s frekvencí. Schroederova funkce je zobrazena na obrázku 1.6 přerušovanou čarou [5].

Norma ISO/IEC MPEG 1 layer 3 využívá ve svém psychoakustickém modelu 2 maskovací funkce, která vychází z Schroederovy aproximace. Je popsána vtahem

$$F(dz) = 15,8111389 + 7,5 \cdot (1,05 \cdot dz + 0,474) - 17,5 \cdot \sqrt{1 + (1,05 \cdot dz + 0,474)^2} + 8 \cdot \text{MIN}(0, (1,05 \cdot dz - 0,5)^2 - 2 \cdot (1,05 \cdot dz - 0,5)) \quad (1.7)$$

a zobrazena na obrázku 1.6 plnými barvami [5].



Obr. 1.6: Funkce pro aplikaci maskovacích křivek.

1.4.4 Sčítání maskovacích křivek

Pokud signál obsahuje více spektrálních složek, které se navzájem nemaskují, tak je výsledný maskovací práh dán součtem maskovacích křivek jednotlivých maskovacích složek dle vztahu

$$I_N = \left(\sum_{n=0}^{N-1} I_n^\alpha \right)^{\frac{1}{\alpha}}, \quad (1.8)$$

kde α je koeficient ovlivňující tvar křivky a je definován pro $1 \leq \alpha \leq \infty$. Při hodnotě $\alpha = 1$ se jednotlivé křivky sčítají, při hodnotě $\alpha = \infty$ se uvažuje nejvyšší práh v daném pásmu [5].

2 STANDARD MPEG-1

MPEG (Moving Picture Experts Group) je název pracovní skupiny, která se zabývá vývojem standardů pro audio a video kompresi. Skupina pracuje pod patronátem ISO (International Organization for Standardization) a IEC (International Electro-Technical commission).

Skupina MPEG standardizovala několik kompresních formátů, zde jsou stručně popsány ty nejdůležitější:

MPEG-1: Tento formát vznikl již v roce 1993. Je popsán v normě ISO/IEC 11172. Slouží pro kódování pohyblivého obrazu a zvuku s rychlostí až 1,4 Mb/s. Rychlost byla odvozena od možné rychlosti CD (Compact Disc). MPEG-1 zahrnuje i formát MP3 - zjednodušený název pro MPEG-1 layer 3, o kterém bude pojednáno v dalších kapitolách.

MPEG-2: Formát z roku 1995 sloužící ke kódování broadcastového vysílání a HDTV. Je popsán v normě ISO/IEC 13818. Je zpětně kompatibilní s předchozím formátem MPEG-1.

MPEG-4: Jedná se o formát pro kódování audiovizuálních objektů. Dokáže pracovat jak s reálným videem, tak s uměle vytvořeným - syntetickým. Využívá ještě vyšší komprese než MPEG-2. Formát vznikl v roce 1998 a je popsán v normě ISO/IEC 14496 [6].

2.1 Princip MPEG

Formát MPEG pro komprimaci videa nebo zvuku využívá ztrátovou kompresi. Datový vstupní tok se nejprve rozdělí do segmentů, které se transformují do frekvenční oblasti. Frekvenční vzorky se dále kvantizují a entropicky kódují.

Ve standardech (normách) pro každý typ MPEG je popsán jen formát bitového proudu a dekodér. Kodér ve standardu není popsán, jsou však k němu k dispozici pro členy skupiny MPEG referenční implementace, podle kterých se dá vytvořit platný bitový proud. V praxi potom dokáže jakýkoliv dekodér daného formátu MPEG dekodovat libovolný soubor stejného formátu MPEG bez ohledu na to, jakým kóděrem byl vytvořen [6].

2.2 Vrstvový model MPEG-1

Každý z výše popsaných standardů se skládá z tzv. vrstev (layers), které popisují způsob kódování obrazu, zvuku a jejich formáty. Formát MPEG-1 audio definuje tři vrstvy s rostoucí složitostí a efektivitou komprese.

Layer 1 využívá schéma založené na algoritmu MUSICAM (Masking pattern adapted Universal Subband Integrated Coding And Multiplexing). Vstupní signál se rozdělí po 12 vzorcích a bankou filtrů PQMF na 32 frekvenčních pásem aproximujících odezvu v kritických pásmech slyšení. Tím vznikne v každém úseku celkem 384 vzorků. Každý tento úsek se analyzuje a porovnává s psychoakustickým modelem. Po porovnání je provedena kvantizace a další zpracování signálu [6].

Layer 2 také využívá schéma založené na algoritmu MUSICAM. Pro úsporu bitů v hlavičce rozdělí signál na delší časové úseky (3x12 vzorků). Tím vznikne v každém úseku celkem 1152 vzorků [6].

Layer 3 využívá mnohem vyspělejší algoritmus ASPEC (Adaptive Spectral Perceptual Entropy Coding). Vstupní signál je rozdělen do 32 subpásem pomocí banky filtrů a prochází tzv. MDCT bloky, kde je každé pásmo transformováno do 18 spektrálních koeficientů. U MPEG 1 Layer 3 je také vylepšen psychoakustický model a využito bezztrátové Huffmanovo kódování [6].

3 MPEG-1 LAYER 3

MPEG-1 Layer 3 je jeden z nejrozšířenějších zvukových komprimovaných formátů. Používá se dnes nejen na přenos hudby, mluveného slova přes internet, ale je rozšířen i do hardwarových přehrávačů (např. malé přehrávače do kapsy, autorádia, hi-fi přehrávače).

3.1 Historie

Historie komprimace zvukových signálů se začala psát v roce 1979, kdy profesor Dieter Seitzer se svým týmem z University of Erlangen vytvořil první signálový procesor schopný zvukové komprese.

Dalším milníkem bylo rozšíření zvukové komprese o poznatky z lidského vnímání zvuku (o psychoakustiku), na kterém má největší zásluhu žák profesora Seitzera Karlheinz Brandenburg.

V roce 1987 začal spolupracovat Fraunhofer IIS (Institut Integrierte Schaltungen) s University of Erlangen na projektu pro vytvoření ztrátových kompresních algoritmů pro praktické využití v digitálních rozhlasových a telekomunikačních službách (DAP).

V roce 1989 popsal Karlheinz Brandenburg ve své disertační práci algoritmus OFC, o kterém můžeme mluvit jako o předchůdci formátu MP3. Tento algoritmus již obsahoval frekvenční rozdělení filtrů, nerovnoměrné kvantování, Huffmanovo kódování. OFC bylo později vylepšeno a rozšířeno a tím vznikl nový algoritmus ASPEC (Adaptive Spectral Perceptual Entropy Coding), který byl vytvořen pro vznikající skupinu MPEG. ASPEC se spolu s algoritmem MUSICAM staly součástí standardu MPEG.

V roce 1993 byla dokončena standardizace MPEG-1 Layer 3. V roce 1995 byla dokončena i standardizace MPEG-2, do které je upravený ASPEC také zahrnut [6].

3.2 Kodér

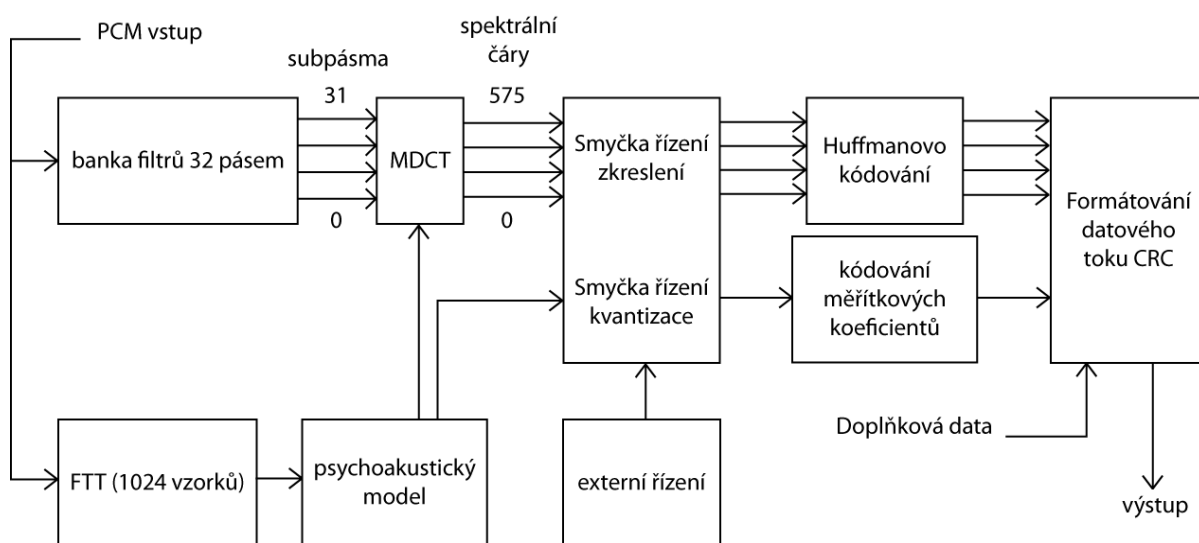
Kodér MPEG-1 Layer 3 slouží k převodu PCM (Pulzně kódová transformace) signálu na komprimovaný soubor s příponou mp3. Na rozdíl od dekodéru, proces kódování probíhá jen jednou a to při samotném převodu. U kódování není důležitá doba zpracování signálu, daleko více se musí brát na zřetel kvalita kódování. Signál nemusí být zpracován v reálném čase jako při dekódování, které slouží k samotnému poslechu nahrávky. Blokované schéma je zobrazeno na obrázku 3.1.

Pro vstup do kodéru se využívá PCM s vzorkovací frekvencí 32 kHz, 44,1 kHz nebo 48 kHz. Vzorkovací frekvence se musí volit tak vysoká, aby nedocházelo k tzv. aliasingu. Ke stanovení minimální vzorkovací frekvence se používá Shannon-Kotělníkův

teorém, který nám říká, že vzorkovací frekvence musí být rovna nejméně dvojnásobku nejvyšší frekvence spektra vzorkovacího signálu:

$$f_{vz} = 2 \cdot f_{max} [Hz] , \quad (3.1)$$

kde f_{vz} je vzorkovací frekvence a f_{max} nejvyšší frekvence spektra vzorkovacího signálu.



Obr. 3.1: Kodér MPEG-1 Layer 3 [7]

Kodér využívá tzv. polyfázové kvadraturní banky filtrů - PQMF [7], která signál rozdělí do 32 frekvenčních pásem se stejným počtem 512 vzorků. Na výstupu každého subpásma je vložena paměť, která dokáže pojmout až 36 hodnot. Dohromady tedy na výstupu je $32 \times 36 = 1152$ šestnáctibitových hodnot.

Dalším důležitým parametrem je velikost (délka) okna. Existují dva druhy oken: standardní o délce 36 hodnot a krátké o délce 12 hodnot. Krátké okno bývá použito třikrát, aby násobek byl stejný jako velikost standardního okna. Velikost okna se určuje na základě perceptuální entropie. Krátké okno se používá kvůli zabránění vzniku pre-echa při tzv. tranzientech – přechodných dějích. Při použití krátkého okna se třikrát zmenší délka trvání signálu v časové oblasti a tím se tento efekt neprojeví. Volbu velikosti okna určuje psychoakustický model [7].

Na výstup každého subpásma vytvořeného bankou filtrů je připojen blok MDCT, který převádí 36 vzorků z časové oblasti do frekvenční. Využívá se k tomu modifikovaná

diskrétní kosinova transformace. Před samotnou transformací se celý obsah vyrovnávací paměti vynásobí sinovým oknem, které je voleno dle použité velikosti okna.

Po transformaci se jednotlivé po sobě jdoucí vzorky překrývají z 50%. Z 36 frekvenčních vzorků je využito z první transformace posledních 18 a z další transformace prvních 18. Tím se docílí toho, že výstupních vzorků (spektrálních čar) je dohromady 576, což je polovina ze vstupních 1152 vzorků, viz. kap. 6 [7].

FFT transformace a psychoakustický model jsou podrobně popsány v kapitole 4.

Kvantizace probíhá ve třech smyčkách, které nazýváme hlavní, vnější a vnitřní smyčka. Hlavní smyčka předává svá data smyčce vnější a vnější smyčka je předává vnitřní.

V hlavní smyčce se nejprve provede vynulování proměnných. V případě, že nejsou vstupní data vynulovaná, se volá smyčka vnější. Hlavní smyčka také vypočítává počet nevyužitých bitů.

Vnější smyčka je zodpovědná za výpočet zkreslení pro všechny skupiny měřítkových koeficientů. Vypočítané hodnoty se porovnávají s povoleným zkreslením, což je výstup z psychoakustického modelu, viz kapitola 5. Měřítkové koeficienty jsou ve vnější smyčce ukládány pro další použití.

Vnitřní smyčka provádí samotnou kvantizaci, která je kvůli rozprostření šumu nelineární [7].

U MPEG-1 Layer 3 se používá bezeztrátové Huffmanovo kódování. Jedná se o kód s variabilní délkou kódového slova, který je vytvořen na základě četnosti výskytu jednotlivých znaků v souboru dat. Nejčastěji se vyskytující znak je zakódován nejkratším binárním řetězcem, naopak nejméně se vyskytující znak má delší kódový řetězec [7].

U MP3 se používá kódování v součinnosti s kvantizací. Kdy po kvantizaci již známe velikost výsledného souboru.

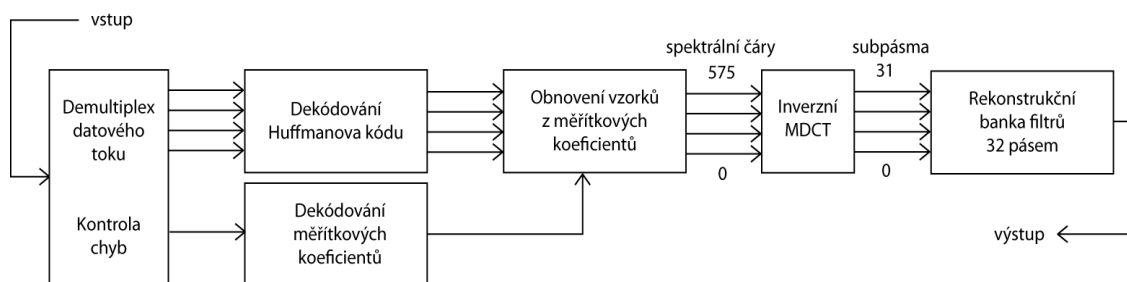
Ve formátu MP3 se využívá cyklický redundantní součet CRC-16 (16 bitů zabezpečení). Tento kontrolní součet je uložen přímo v bitovém toku za hlavičkou a je porovnáván s kontrolním součtem s dekodéru. Pokud není výsledek porovnání stejný, tak se daný rámeček utlumí a přehraje se předchozí rámeček. Tato metoda je pro lidský sluch příjemnější než přehrávání nekompletních (vadných) dat [7].

3.3 Dekodér

U dekodéru mp3 hraje velmi velkou roli i čas zpracování signálu. Pro posluchače je důležité, aby nahrávku slyšel v reálném čase, bez přerušování nebo dlouhých pauz na začátku. Dalším důležitým aspektem je jeho jednoduchost. Dekodér musí být jednoduše

realizovatelný nejen v počítačích pomocí softwarového přehrávače, ale i v miniaturních kapesních mp3 přehrávačích. Nesmí mít velké hardwarové nároky.

Standard ISO/IEC 1172-3 definuje požadavky na dekodér, kvůli kompatibilitě a možnosti přehrávání tohoto formátu na všech přehrávačích. Kodér není normou [7] přímo definován. Schéma dekodéru je na obrázku 3.2.



Obr. 3.2: Dekodér MPEG-1 Layer 3 [7]

V první fázi se data rozdělí a připraví pro zpětné dekodování Huffmanovým kódem. Proveďte se kontrola chyb a zvláště se dekodují měřítkové koeficienty. Po dekodování nám vznikne posloupnost hodnot kvantovaných frekvenčních koeficientů, které zpracujeme v bloku inverzní MDCT, kde jsou vzorky převedeny z frekvenční oblasti do časové. Výstupem je 36 vzorků v každém subpásmu. Po doplnění všech vzorků 31 nulami nám vznikne v každém subpásmu 1152 vzorků. Po aditivním sečtení všech vzorků se stejným časovým indexem dostáváme zpět rekonstruovaný PCM signál [7].

4 PSYCHOAKUSTICKÝ MODEL

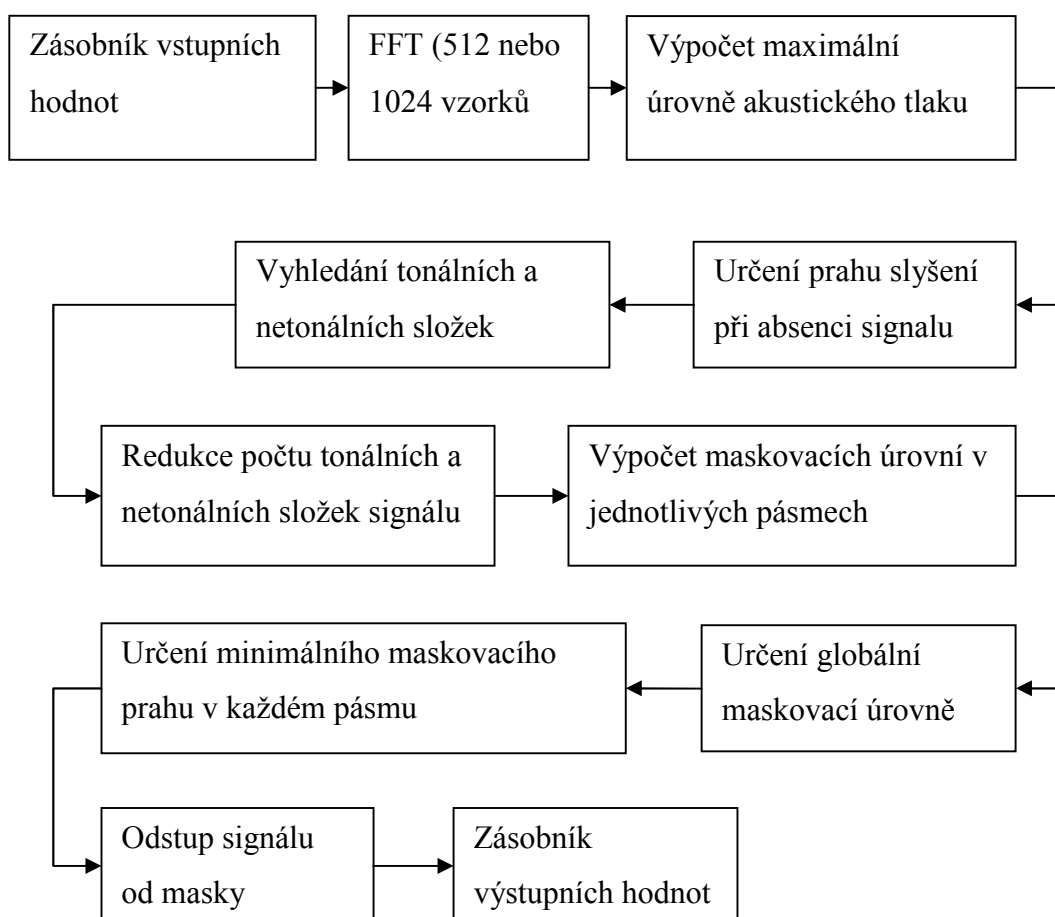
Psychoakustický model u standardu MPEG 1 slouží k aplikaci maskovacích křivek na vstupní signál. Model svými výstupními parametry ovlivňuje celý proces kódování signálu a kvantizaci a tím i velikost a kvalitu výstupního signálu.

Jak již bylo zmíněno, existují ve standardu ISO 11172-3 dva typy psychoakustického modelu: psychoakustický model 1 používaný u standardu MPEG 1 Layer 1 a 2 a psychoakustický model 2 používaný u MPEG 1 Layer 3.

V této diplomové práci je více popsán psychoakustický model 2, který je nasimulován v příloženém programu.

4.1 Psychoakustický model 1

Psychoakustický model 1 je blokově zobrazen na obr. 4.1.



Obr. 4.1: Schéma psychoakustického modelu 1

Vstupem je PCM signál, který se pomocí FFT převede do frekvenční oblasti. U MPEG 1 layer 1 je převeden na 512 vzorků u MPEG 1 layer 2 na 1024 vzorků. K přizpůsobení signálu se používá Hannovo okno. Více je o převodu do frekvenční oblasti použití oken popsáno v kapitole 4.2.1 [7].

Dalším krokem je určení maximální hodnoty hladiny akustického tlaku v každém pásmu. Dále se určí práh slyšení při absenci signálu. Následně se provede identifikace tonálních (harmonických) a netonálních (nestacionárních) složek signálu. Provede se redukce počtu maskovacích složek, zanechají se pouze významné. Dalším krokem je aplikování maskovacích křivek. Provede se součet maskovacích křivek v jednotlivých pásmech s prahem slyšení, tím vznikne globální maskovací práh. Z globálního maskovacího prahu se určí minimální maskovací práh v každém pásmu. Z těchto hodnot vypočítáme odstup signálu od masky SMR, což je výstupem psychoakustického modelu 1.

4.2 Psychoakustický model 2

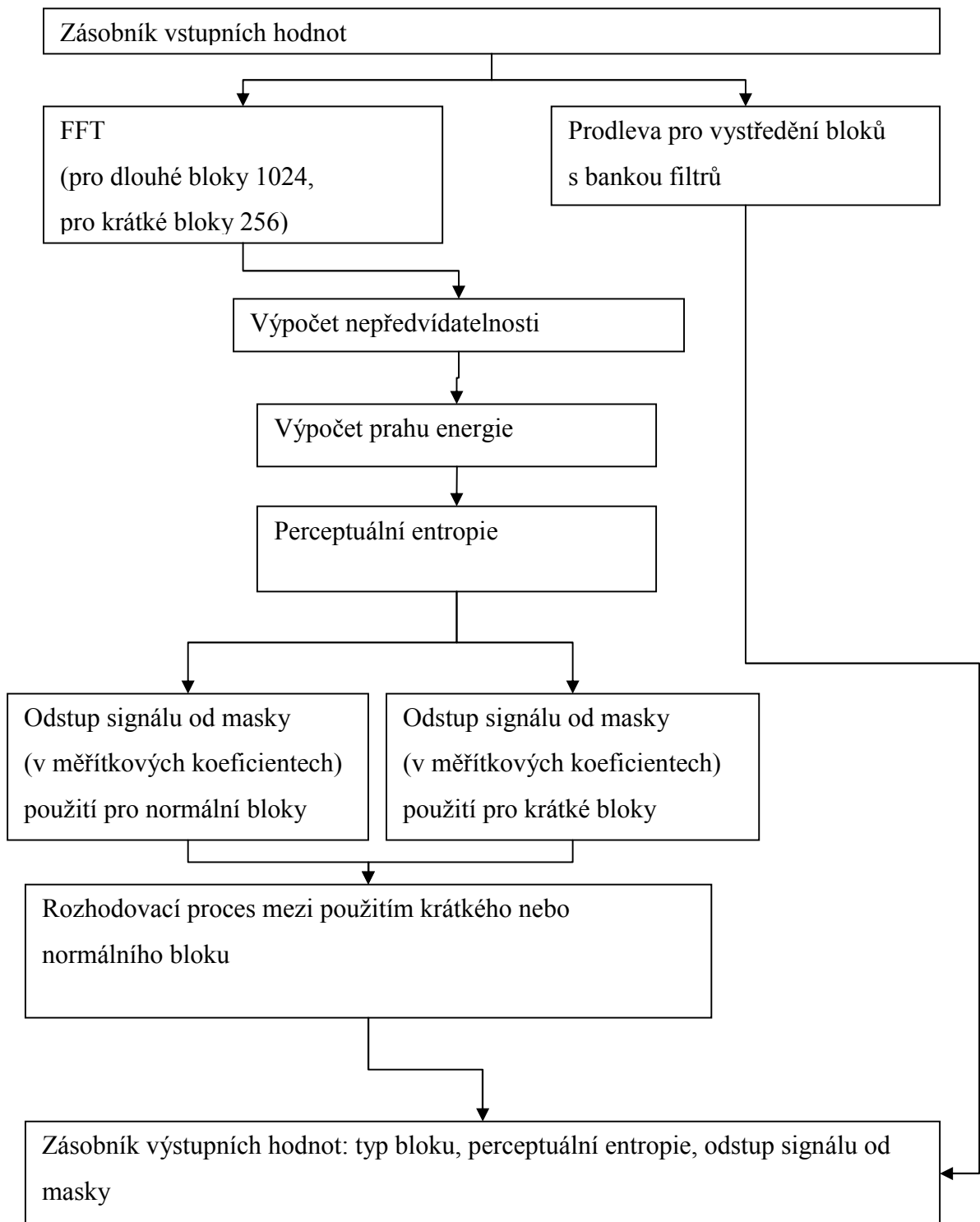
Schéma psychoakustického modelu 2 je znázorněno na obr. 4.2.

Vstupem do modelu je PCM signál, který je také vstupem banky filtrů. Banka filtrů s blokem MDCT způsobují časové zpoždění 256 vzorků a proto musí být provedena časová synchronizace [7].

Dalším důležitým faktorem je délka okna. Psychoakustický model využívá dvou typů oken standardního (délka 1024 vzorků) a krátkého (délka 256 bloků, které se používá 3x). O přepínání těchto oken je rozhodnuto pomocí percentuální entropie (viz kap. 4.2.8). Tyto okna je třeba správně vystředit na 1152 časových vzorků (pro standardní okna), respektive 384 časových vzorků (pro krátké okna), jimž odpovídá vstupní paměť MDCT bloku o délce 36 vzorků (12 vzorků pro krátké okna) [7].

Výstupy psychoakustického modelu 2 jsou:

- volba použití normálního či krátkého okna (viz kap. 4.2.8),
- percentuální entropie (viz kap. 4.2.8),
- odstup signálu od masky SMR (viz kap. 4.2.9).



Obr. 4.2: Blokové schéma psychoakustického modelu MPEG-1 Layer 3.

4.2.1 Vstup signálu

Vstupním vzorkem může být jakýkoliv PCM signál se vzorkovací frekvencí 32, 44,1 nebo 48 kHz.

Nejprve se vstupní signál rozdělí na 1024 vzorků pro dlouhé bloky a paralelně na 256 vzorků pro krátké bloky, které se samostatně převedou do frekvenční oblasti pomocí rychlé Fourierovy transformace (FFT), která je definovaná vztahem

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-2\pi i}{N}nk}, \text{ kde } k = 0, \dots, N - 1. \quad (4.1)$$

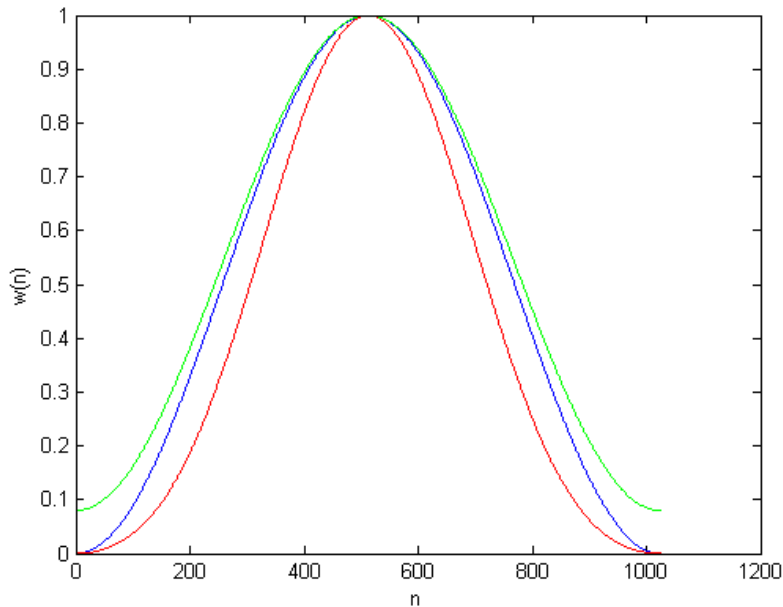
Před samotnou Fourierovou transformací je třeba vstupní signál upravit. Většinou se jedná o neperiodický signál o libovolné délce. Signál se musí rozdělit na dané periodické úseky, které na sebe spojitě navazují. Tato operace se provádí pomocí tzv. okénkových funkcí.

Nejjednodušší by bylo použít obdélníkové okno, které provede výřez jednotlivých posloupností bez úpravy amplitudy. V případě, že se jedná o signál, kde jednotlivé úseky nejsou rovné násobku periody signálu, bude výsledné periodické prodloužení vykazovat na mezích mezi jednotlivými úseky nespojitosti. Tento jev se nazývá prosakování spektra (spectral leakage). Řešením je použití okna, které svými parametry vhodně upravuje amplitudu signálu, aby k prosáknutí spektra nedošlo. Na obrázku 4.3 jsou zobrazeny nejčastěji používaná okna [8].

MPEG 1 Layer 3 používá tzv. Hannovo okno, které je definováno

$$w(n) = 0,5 \left(1 - \cos \left(2\pi \frac{n}{N} \right) \right), \text{ kde } 0 \leq n \leq N, \quad (4.2)$$

a plně svými parametry vyhovuje dané úpravě signálu.



Obr. 4.3: Okénkóvé funkce - Hannovo okno (modrá křívka), Blackmanovo okno (červená křívka), Hammingovo okno (zelená křívka).

Vstupní signál rozdělený po 1024 vzorcích pro dlouhé okno (nebo po 256 vzorcích pro krátké okno) je vynásoben Hannovým oknem a následně převeden do frekvenční oblasti v exponenciálním tvaru dle vztahu (4.1). Pro další použití je třeba si zvlášť uložit amplitudu r_ω a fázi φ_ω jednotlivých frekvenčních koeficientů ω , které nabývají hodnot od 0 do 511.

4.2.2 Nepředvídatelnost

Nepředvídatelnost se počítá pouze pro prvních 206 frekvenčních koeficientů. Pro ostatní frekvenční koeficienty je nepředvídatelnost nastavena na hodnotu 0,4. Tabulka 4.1 ukazuje způsob počítání jednotlivých frekvenčních koeficientů.

Tab. 4.1: Způsob rozdělení frekvenčních koeficientů pro počítání nepředvídatelnosti.

Koeficienty:	Způsob počítání:
$0 \leq \omega \leq 6$	Počítáno z normálních bloků FFT
$6 \leq \omega \leq 206$	Počítáno z druhého krátkého bloku FFT $(\omega + 2)/DIV4$
$\omega \geq 206$	0,4

Nepředvídatelností (unpredictability measure) zjistíme, jak je signál předvídatelný z předchozích dvou vypočtených hodnot [7]. Pro její výpočet potřebujeme vypočítat predikované hodnoty amplitudy a fáze ze vztahů

$$r_{\omega-pred} = 2r_{\omega}(t-1) - r_{\omega}(t-2), \quad (4.3)$$

$$\varphi_{\omega-pred} = 2\varphi_{\omega}(t-1) - \varphi_{\omega}(t-2), \quad (4.4)$$

kde $(t-1)$ reprezentuje minulý vzorek, $(t-2)$ předminulý vzorek, r_{ω} je amplituda po FFT a φ_{ω} je fáze po FFT.

Nyní už můžeme vypočítat samotnou nepředvídatelnost ze vztahu

$$c_{\omega} = \frac{\left((r \cdot \cos(f) - r_{pred} \cos(f_{pred}))^2 + (r \cdot \sin(f) - r_{pred} \sin(f_{pred}))^2 \right)^{0,5}}{r + |r_{pred}|}. \quad (4.5)$$

4.2.3 Výpočet energie a nepředvídatelnosti ve výpočetních pásmech

Další výpočty se provádí v tzv. výpočetních pásmech (calculation partition), což jsou frekvenční rozsahy dané šířkou koeficientů po převedení do frekvenční oblasti pomocí FFT. Rozsahy jsou dané normou ISO/IEC 11172-3 [7] v tabulkách D.3, např. pro vzorkovací frekvenci 44,1 kHz je počet pásem rovný hodnotě 57. V těchto tabulkách jsou již data přizpůsobena lidskému vnímání zvuku a neobsahují frekvence nad 20 kHz [7]. Výpočty, kde bude počítáno s výpočetními pásmy, jsou označeny indexem b .

Energie ve výpočetním pásmu je daná vztahem

$$e_b = \sum_{\omega_{low}}^{\omega_{high}} r_{\omega}^2, \quad (4.6)$$

a nepředvídatelnost je daná vztahem

$$c_b = \sum_{\omega_{low}}^{\omega_{high}} r_{\omega}^2 \cdot c_{\omega}, \quad (4.7)$$

kde ω_{high} je nejvyšší hodnota frekvence ve výpočtovém okně, ω_{low} je nejnižší hodnota frekvence výpočtového okna, které jsou dány hodnotami v tabulkách D.3 [7]. Nízké

frekvence jsou zastoupeny nižším rozlišením - jednou hodnotou. U vyšších frekvencí je toto rozlišení větší.

4.2.4 Konvoluce energie a nepředvídatelnosti s maskovací křivkou

Ve formátu MP3 je použita upravená modifikace Schroederovy aproximace a je dána vztahem (1.7). Tato funkce je pomocí konvoluce spojena s vypočítanou energií (4.6) a s nepředvídatelností (4.7), pomocí vztahů

$$ecb_b = \sum_{bb=1}^{bmax} e_{bb} \cdot f(bval_{bb}, bval_b), \quad (4.9)$$

$$ct_b = \sum_{bb=1}^{bmax} c_{bb} \cdot f(bval_{bb}, bval_b), \quad (4.10)$$

kde $f(bval_{bb}, bval_b)$ je maskovací křivka daná vztahem (1.7) [7].

Hodnoty se musí před dalším použitím normalizovat. Výsledná normovaná vázaná hodnota energie se vypočítá dle vztahu

$$en_b = ecb_b \cdot rnorm_b, \quad (4.11)$$

kde $rnorm_b$ je normalizační koeficient, který je dán vzorcem

$$rnorm_b = \frac{1}{\sum_{bb=1}^{bmax} sprdngf(bval_{bb}, bval_b)}. \quad (4.12)$$

Normalizována je i hodnota vázané nepředvídatelnosti s maskovací funkcí a to dle vztahu

$$cb_b = \frac{ct_b}{ecb_b} \quad (4.13)$$

kde ct_b (4.10) je nepředvídatelnost a ecb_b (4.9) je energie v konvoluci s maskovací křivkou [7].

4.2.5 Index tonality

Index tonality nabývá hodnot od 0 do 1, kde koeficienty blížíící se jedné jsou čistě tonální a koeficienty blížíící se k nule šumové. Rozdělení pomocí indexu tonality se

využívá pro frekvenční maskování. Tóny jsou maskovány méně než šum. Vzorec pro výpočet indexu tonality je

$$tb_b = -0,299 - 0,43 \log(cb_b), \quad (4.14)$$

kde cb_b je normovaná hodnota vázané nepředvídatelnosti s maskovací funkcí (4.13) [7].

4.2.6 Odstup signálu od šumu

Z výstupu indexu tonality vypočítáme odstup signálu od šumu SNR (Signal to Noise Ratio) pro jednotlivá výpočetní pásma. U formátu MP3 se používá maskování šumu nastavené pro celé výpočetní pásmo na 5,5 dB. U tónových složek je hodnota mnohem vyšší, pohybuje se až kolem 30 dB. Přesné hodnoty jsou dány opět normou [7] v tabulce D.3 položka TMN. Ve výpočtu SNR se porovnávají tabulkové hodnoty (norma [7], tab. D.3) položka *minval* (minimální práh) s vypočítanou hodnotou a vybírají se vyšší z obou hodnot, viz. vztah

$$SNR = \max(\minval_b, tb_b \cdot TMN + 5,5 \cdot (1 - tb_b)), \quad (4.15)$$

kde hodnoty \minval_b , TMN jsou tabulkové hodnoty [7] a tb_b je index tonality daný vztahem (4.14).

4.2.7 Spektrum slyšitelného zvuku

Z odstupu signálu od šumu již můžeme vypočítat vlastní práh energie pro jednotlivá výpočetní pásma dle vztahu

$$nb_b = en_b \cdot 10^{\frac{-SNR_b}{10}}, \quad (4.16)$$

kde en_b je normovaná hodnota vázané energie daná vztahem (4.11) a SNR je hodnota odstupu signálu od šumu daná vztahem (4.15) [7].

Nyní je provedena kontrola pre-echa, kdy jsou porovnány předchozí hodnoty prahů energie a z nich je zvoleno maximum:

$$thr_b = \text{MAX}(nb_b, (2 \cdot nb_b), (16 \cdot nb_b)), \quad (4.17)$$

kde nb_b je vlastní práh energie (4.16).

Po kontrole pre-echa jsou opět jednotlivá výpočetní pásma sloučena a počítá se s jednotlivými frekvenčními koeficienty. Převod je realizován vztahem

$$nb_\omega = \frac{nb_b}{\omega_{high_b} - \omega_{low_b} + 1}, \quad (4.18)$$

kde hodnoty ω_{high_b} a ω_{low_b} jsou dány normou v tabulce D.3 [7].

Dalším krokem je porovnání prahu energie a prahu slyšitelnosti, který je dán proměnnou $absthr_\omega$, který je definován hodnotami v tabulce D.4. obsažené v normě [7]. Výsledné slyšitelné spektrum signálu (práh energie) je dáno vztahem

$$thr_\omega = \max(nb_\omega, absthr_\omega), \quad (4.19)$$

kde nb_ω je převodní vztah (4.18).

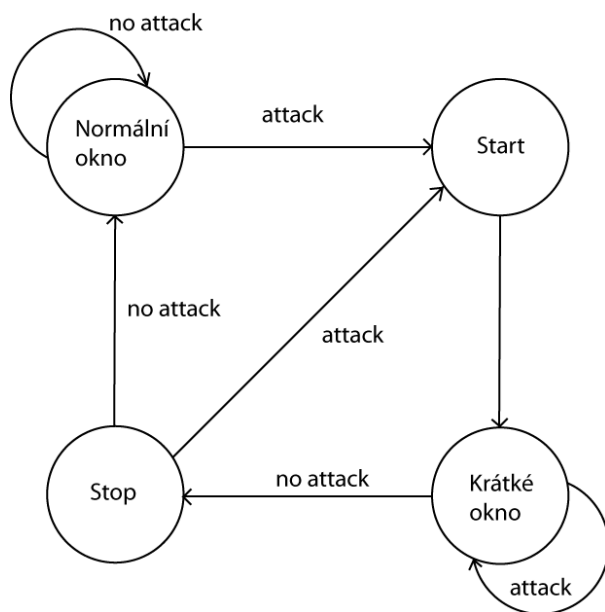
4.2.8 Perceptuální entropie

Perceptuální entropie udává průměrný minimální počet bitů vzorku spektrální čáry nutný k jejímu zakódování, aby nevznikl slyšitelný rozdíl od originálního signálu. Entropie je dána vztahem

$$pe = -\sum \left(cbwidth_k \cdot \log \left(\frac{tbr_k}{eb_k + 1} \right) \right), \quad (4.20)$$

kde index k značí jednotlivá výpočetní pásma, tbr_k je výsledné slyšitelné spektrum a eb_k je energie vstupního spektra signálu. Hodnota $cbwidth_k$ je počet frekvenčních koeficientů v jednotlivých výpočetních pásmech.

Perceptuální entropie také určuje, jaký typ okna bude použit v MDCT bloku. Přepínání oken je dáno přepínacím diagramem zobrazeným na obrázku 4.4. Přejít na krátké bloky je uskutečněno v případě, že perceptuální entropie je větší než 1800 bitů, na obrázku je tato situace znázorněna slovem "attack". Situaci, kdy entropie nepřekročila tuto hodnotu značíme slovem "no attack" [7].



Obr. 4.4: Přepínací diagram oken [7]

4.2.9 Odstup signálu od masky

Dalším výstupem psychoakustického modelu 2 je odstup signálu od šumu SMR (Signal to Mask Ratio). Tato hodnota se počítá zvlášť pro krátké i dlouhé bloky. Princip výpočtu je ale stejný. Výstupem jsou hodnoty v měřítkových koeficientech (scalefactor bands). SMR je dáno vztahem

$$SMR_n = 10 \log_{10} \left(\frac{epart_n}{npart_n} \right) , \quad (4.21)$$

kde $epart_n$ je energie vyjádřená v měřítkových koeficientech, která je dána vztahem

$$epart_n = \sum_{\omega=\omega_{low_n}}^{\omega_{high_n}} r_{\omega}^2 , \quad (4.22)$$

kde rozsahy ω_{high_n} a ω_{low_n} jsou specifikované v tabulce 4.2.

Proměnná $npart_n$ ve vzorci (4.21) je vypočítaný práh energie vyjádřený v měřítkových koeficientech, který se pro délku 1 (viz tabulka 4.2 položka width (n)) počítá pomocí vztahu

$$npart_n = \sum_{\omega=\omega_{low_n}}^{\omega_{high_n}} thr_{\omega} , \quad (4.23)$$

a pro délku rovnou 0 dle vztahu

$$npart_n = MIN(thr_{\omega low(n)}, \dots, thr_{\omega high(n)}) \cdot (\omega high_n - \omega low_n + 1), \quad (4.24)$$

kde rozsahy $\omega high_n$ a ωlow_n jsou specifikované v tabulce 4.2 [7].

Tab. 4.2: Rozdělení do měřítkových koeficientů [7]

Index	$\omega low (n+1)$ $\omega high (n)$	width (n)
0	1	0
1	17	0
2	33	0
3	49	0
4	65	0
5	81	0
6	97	0
7	113	0
8	129	0
9	145	0
10	161	0
11	177	0
12	193	0
13	209	1
14	225	1
15	257	1
16	273	1
17	289	1
18	305	1
19	321	1
20	337	1
21	353	1
22	369	1
23	385	1
24	401	1
25	417	1
26	433	1
27	499	1
28	465	1
29	481	1
30	497	1
31	513	1

5 SIMULACE PSYCHOAKUSTICKÉHO MODELU MPEG-1 LAYER 3 V PROGRAMU MATLAB

Hlavním cílem této diplomové práce je vytvoření grafického programu, který bude simulovat psychoakustický model. Program má sloužit pro studenty, kterým zde budou názorně pomocí grafů znázorněny jednotlivé úkony psychoakustického modelu.

Pro realizaci byl vybrán psychoakustický model 2, který se používá u standardu MPEG 1 Layer 3. Jedním z důvodů tohoto výběru je fakt, že se jedná o jeden z nejrozšířenějších hudebních kompresních formátů.

5.1 Volba programu - MATLAB

Celý program je realizován v matematickém programu Matlab, odladěn byl ve verzi Matlab R6.5.

Program Matlab byl zvolen z důvodu jednoduché implementace dat a jejich grafických výstupů, program má v sobě hodně funkcí již vytvořených. Mezi další výhody patří jeho zastoupení v laboratořích Vysokého učení v Brně, kde se používá a je studentům dobře znám.

Hlavní výhody implementace psychoakustického modelu 2 pomocí programu MATLAB:

- předpřipravené funkce,
- jednoduchá implementace různých typů grafů,
- program velmi dobře pracuje s maticovým počtem, který je v samotné práci využíván.

Mezi nevýhody tohoto prostředí patří:

- nepřenositelnost, lze spustit jen pomocí MATLABu - program není kompilovaný překlad, ale interpretovaný,
- rychlost zpracování proti profesionálním kódérům napsaných například v jazyce C.

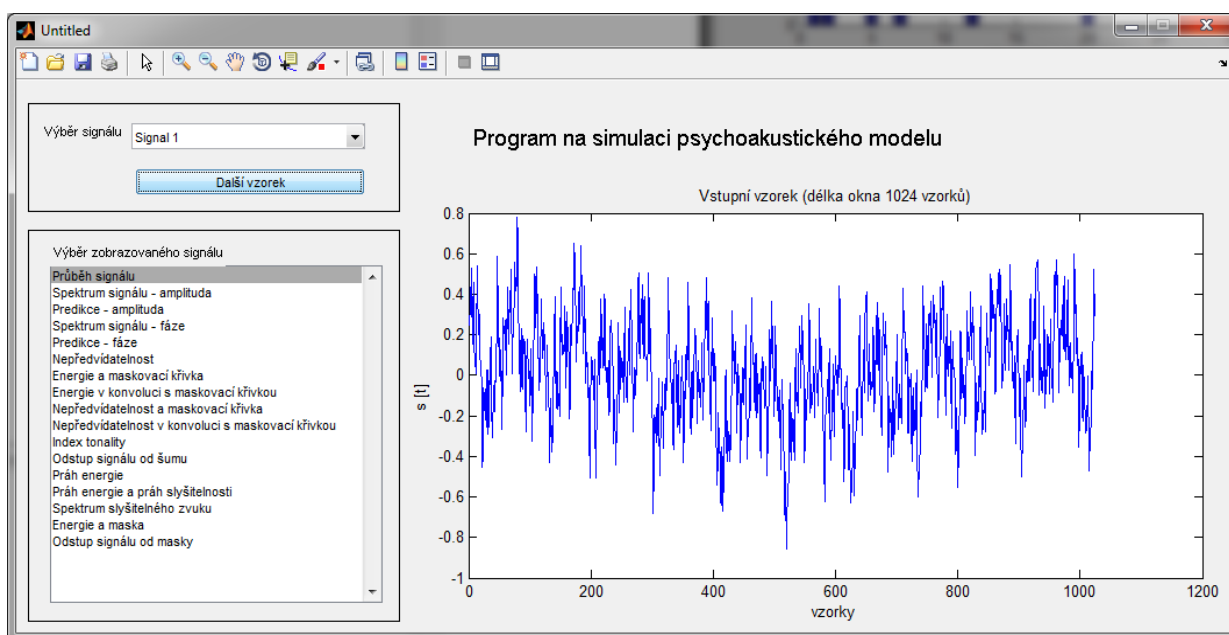
5.2 Realizace programu

Grafické prostředí v MATLABu je vytvořeno pomocí grafické nástavby HANDLE GRAPHICS. Základ programu byl vytvořen průvodcem prostředím GUIDE [10], do kterého byly implementovány jednotlivé funkce. Zobrazení grafů je realizováno v grafickém objektu AXES, který lze implementovat přímo do jednoho okna s ovládacími

prvky. Jednotlivé dílčí funkce jsou uloženy ve zvláštních souborech s příponou *.m a v hlavním programu je na ně odkazováno pomocí volání funkcí.

5.2.1 Ovládání programu

Program se spouští v programu MATLAB v Command window pomocí příkazu *model*. Je realizován v jednom okně, viz obrázek 5.1.



Obr. 5.1: Ovládací panel programu

Program je rozdělen do tří částí.

V první části volíme zdroj signálu. Je důležité, aby tato volba byla nastavena jako první. Také se zde nachází tlačítko "Další vzorek" na posunutí do dalšího okna.

Druhá část slouží pro výběr požadované zobrazené funkce. Zde je zapotřebí upozornit, že pro korektní zobrazení všech funkcí, které následují za zobrazením "Predikcí signálu", je třeba nastavit až nejméně třetí vzorek, pomocí tlačítka "Další". Důvodem je vypočet predikovaných hodnot z předešlých, které u prvních signálů nejsou.

Třetí část slouží k samotnému zobrazení výstupních grafů, které jsou realizovány dvourozměrným grafem typu *plot* pro zobrazení v časové a frekvenční oblasti. Pro zobrazení měřítkových koeficientů a výpočetních pásem je zvolen dvourozměrný graf typu *bar*, který je pro tyto funkce více názorný.

5.2.2 Realizace Psychoakustického modelu 2

Jednotlivé funkce programu jsou napsány dle normy ISO/IEC 11172-3 [7], nicméně některé funkce jsou z důvodu lepší grafické prezentace zjednodušeny. Model je vytvořen pouze pro práci s dlouhými okny. Perceptuální entropie, kontrola pre-echa není v programu aplikována.

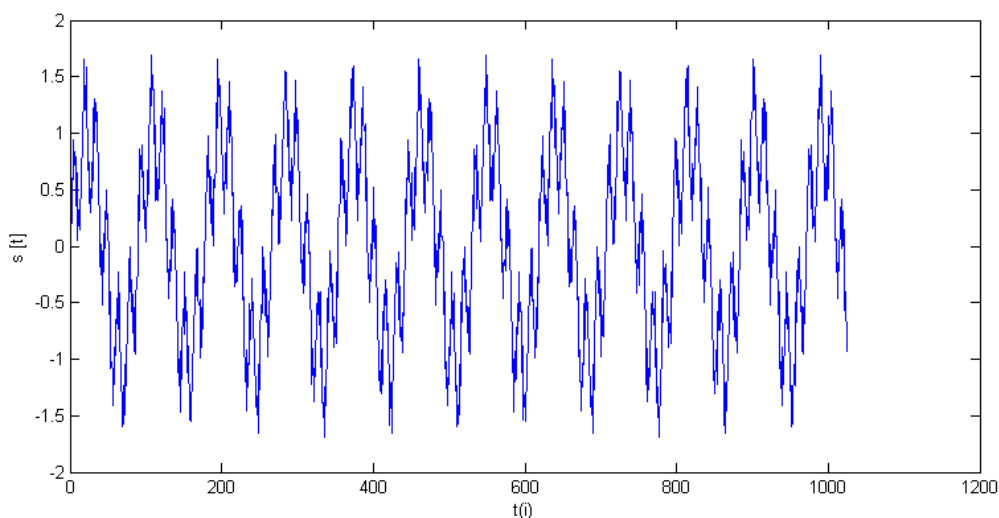
5.2.3 Výběr vstupního signálu

Vstupem do Psychoakustického modelu 2 je PCM signál, viz kapitola 4.2.1. V programu je vstup realizován dvěma způsoby.

První možností je využití vstupního signálu generovaného pomocí funkce přímo v programu MATLAB. V programu jsou zastoupeny tři uměle vytvořené signály. Signál 1 simuluje PCM signál, který je velmi podobný vstupním hudebním signálům. Signál 2 je též náhodný, ale má menší četnost spektrálních složek signálu. Signál 3 se skládá pouze ze tří spektrálních složek, jeho časový průběh je zobrazen na obr. 5.2. Z hlediska výuky studentů jsou poslední dva signály vhodné pro grafickou prezentaci průběhu jednotlivých funkcí psychoakustického modelu. Dále v textu budou použity grafické výstupy signálu 3, na kterém jsou nejvíce patrné jednotlivé úkony programu.

Druhou možností je nahrání souboru typu wav, což je realizováno pomocí funkce *waveread*.

Vstupní signál je ihned po vstupu rozdělen na jednotlivé bloky po 1024 hodnotách s posunovacím oknem *iblen*, které určuje délku posunutí viz kap. 4.2.

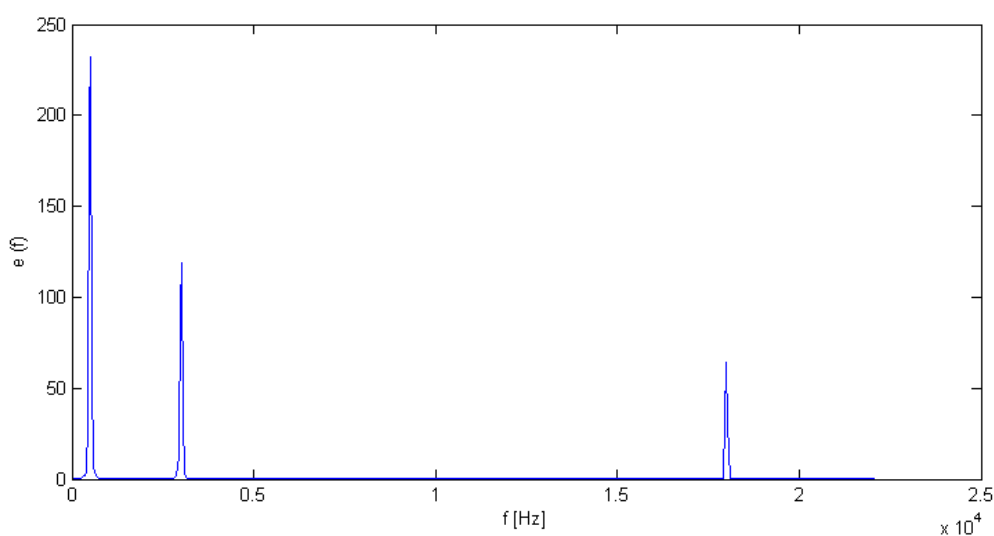


Obr. 5.2: Časový průběh signálu 3

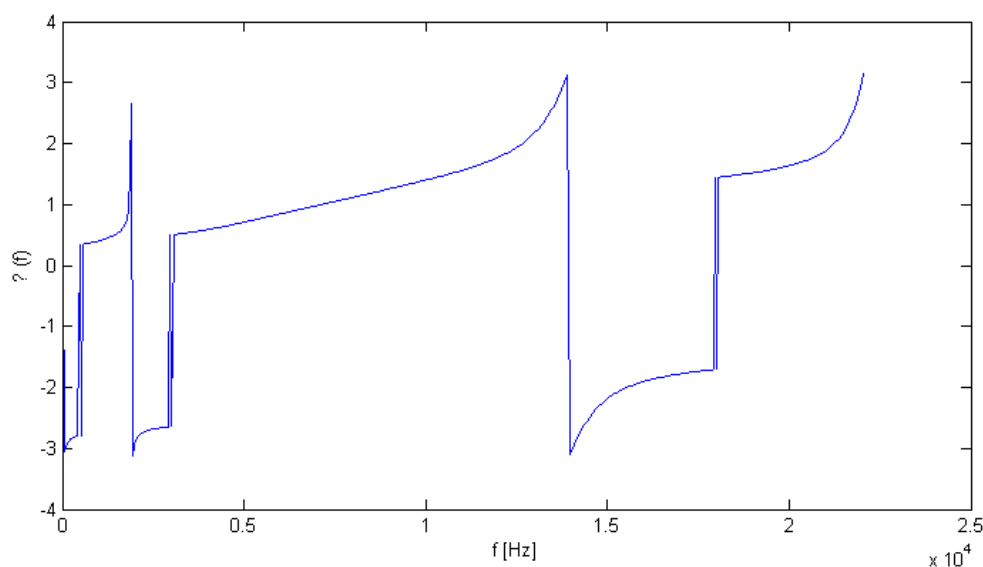
5.2.4 Rychlá Fourierova transformace

Pro převod PCM signálu do frekvenční oblasti se využívá rychlá Fourierova transformace (Fast Fourier Transform - FFT). Před transformací se signál ještě vynásobí Hannovým oknem (viz kap. 4.2.1). Na obrázku 5.3 je zobrazeno frekvenční spektrum signálu 3, kde jsou patrné již zmíněné (kap. 5.2) tři spektrální složky. Dalším výstupem je spektrum fáze daného signálu. Na obr. 5.4 je spektrum fáze signálu 3, kde jsou patrné jednotlivé fázové změny.

Rychlá Fourierova transformace a okénková funkce je řešena funkcí `fourtrans.m`.



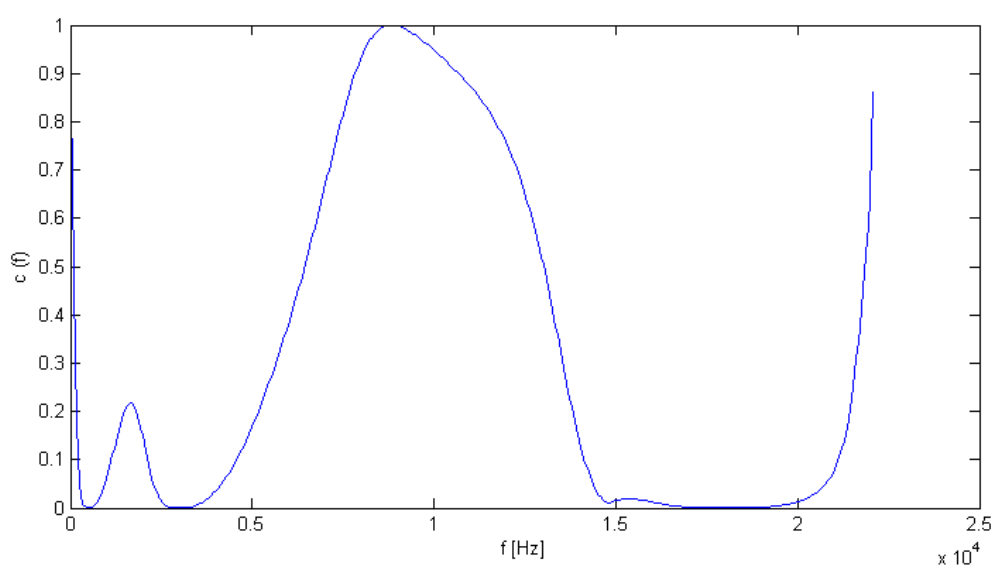
Obr. 5.3: Frekvenční spektrum signálu 3



Obr. 5.4: Spektrum fáze signálu 3

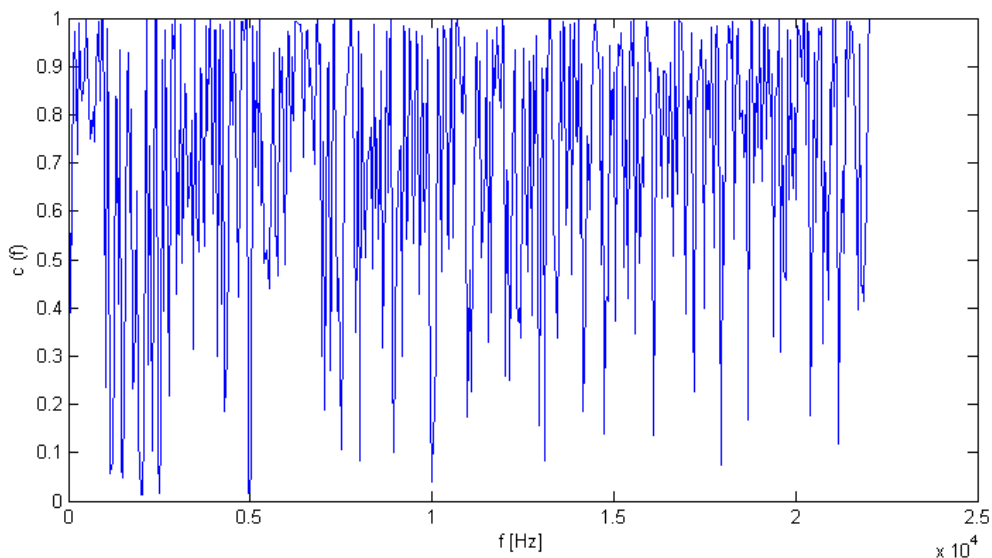
5.2.5 Nepředvídatelnost

K výpočtu nepředvídatelnosti je třeba dvou předchozích vypočtených hodnot (kap. 4.2.2). Předchozí vypočtené hodnoty frekvenčního spektra budou díky vstupnímu periodickému signálu 3 neměnné. U spektra fáze bude situace jiná. Fázové posuvy se mění s jinou frekvencí. Tento fakt ovlivní nepředvídatelnost, signály nebudou plně předvídatelné a některé hodnoty se budou blížit hodnotě 1 (viz obr. 5.5).



Obr. 5.5: Nepředvídatelnost signálu 3

U neperiodických signálů je nepředvídatelnost ovlivňována i frekvenčním spektrem. Signál je daleko méně předvídatelný z předešlých vypočtených hodnot. Na obr. 5.6 je pro znázornění této skutečnosti zobrazena nepředvídatelnost uměle vytvořeného náhodného signálu (signál 1).

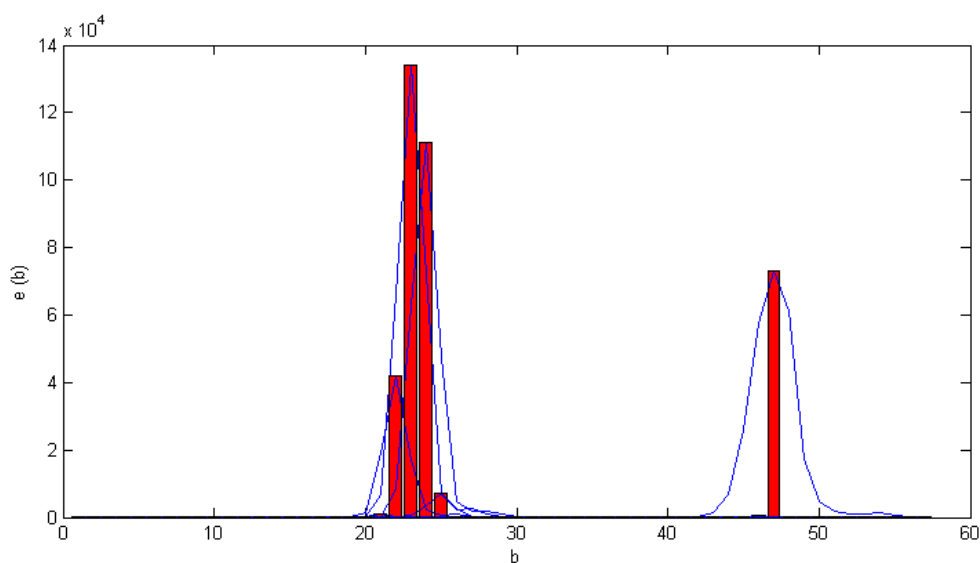


Obr. 5.6: Nepředvídatelnost neperiodického signálu (signál 1)

5.2.6 Zobrazení energie a nepředvídatelnosti v konvoluci s maskovacími křivkami

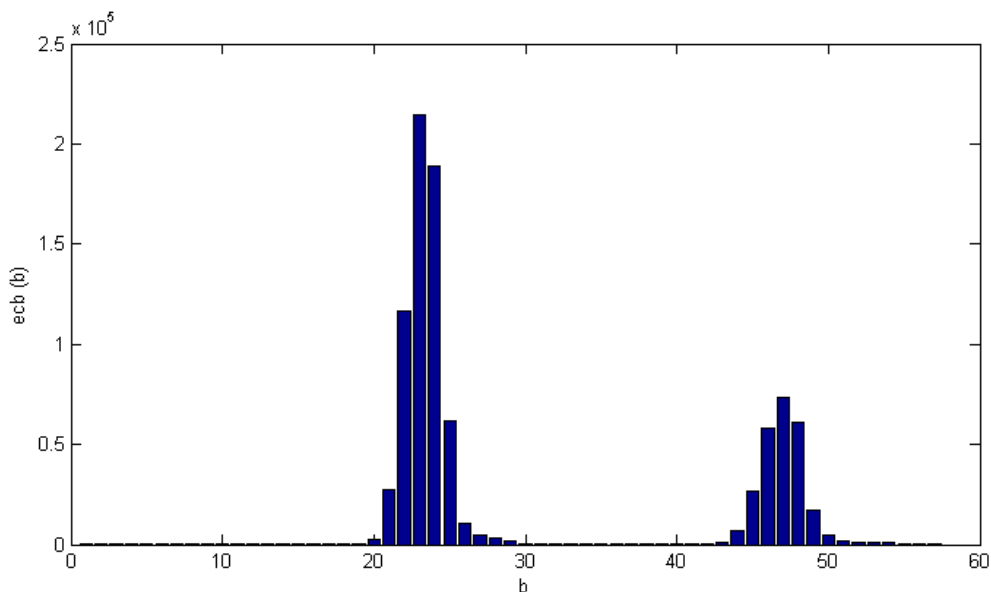
Další výpočty probíhají v tzv. výpočetních pásmech. Rozsahy hodnot jsou dané tabulkami D.3 v normě [7]. Pro vzorkovací frekvenci 44,1 kHz obsahuje 57 koeficientů (indexů) a je vypsána v příloze tab. A-1.

Na obr. 5.7 je zobrazena energie (viz kap. 4.2.3 a 4.2.4) převedená na jednotlivá výpočetní pásma b . Pro přehlednost je zobrazena červenou barvou pomocí vykreslovací funkce MATLABu *bar*. Modrou křivkou je zobrazena maskovací křivka (viz kap. 1.4.3), na které je patrné, jak se funkce rozprostírá do ostatních koeficientů.



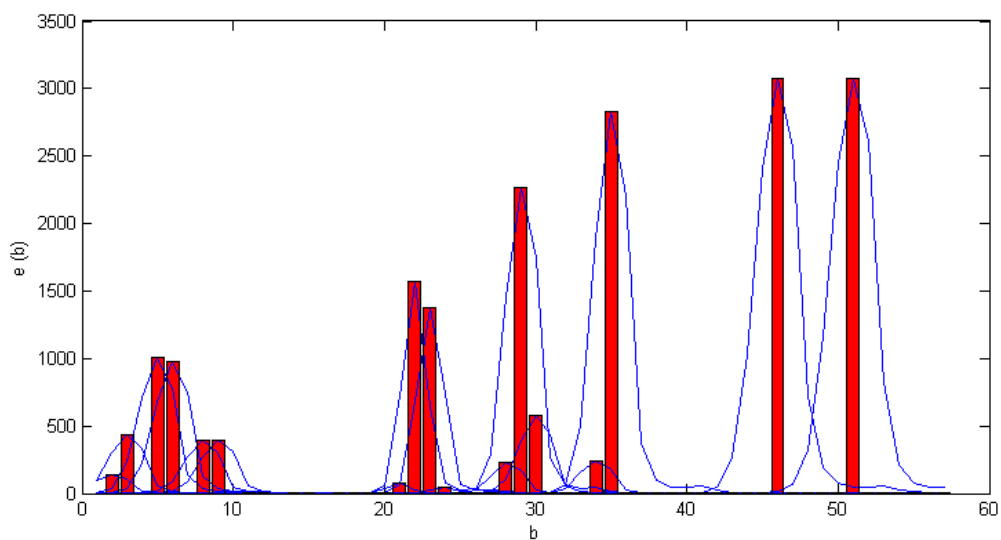
Obr. 5.7: Energie (červená barva) a maskovací křivka (modrá barva) pro signál 3

Na obr. 5.8 je zobrazena funkce z obr. 5.7 po konvoluci. Podobných grafických výsledků je dosaženo i při zobrazení nepředvídatelnosti s maskovací křivkou.

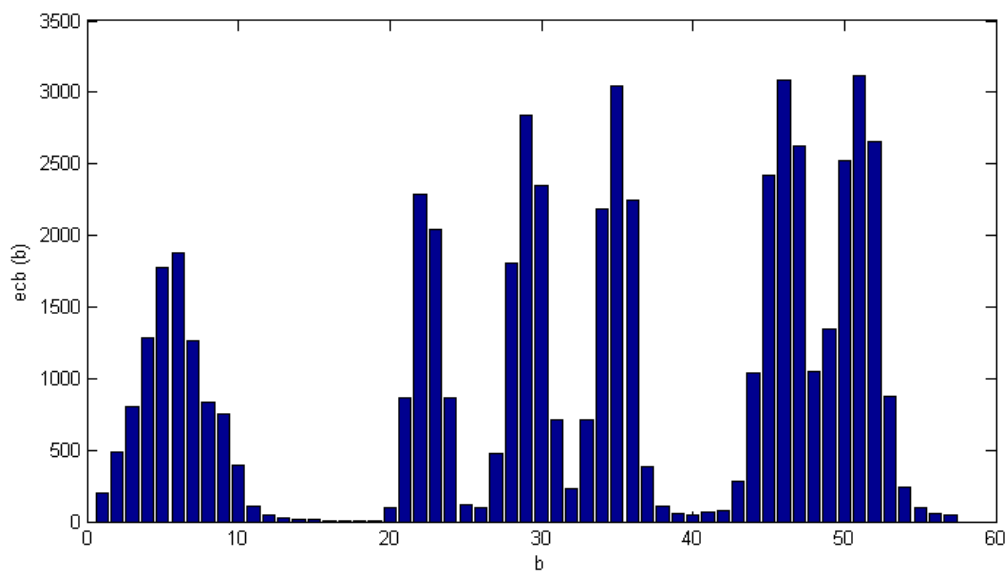


Obr. 5.8: Energie po konvoluci s maskovací křivkou pro signál 3

Na obrázku 5.9 a 5.10 je graficky znázorněno rozptřeni maskovací křivky do ostatních koeficientů (indexů) výpočetních oddílů u signálu 2. Z obrázků je lépe patrné, jak se jednotlivé koeficienty ovlivňují.



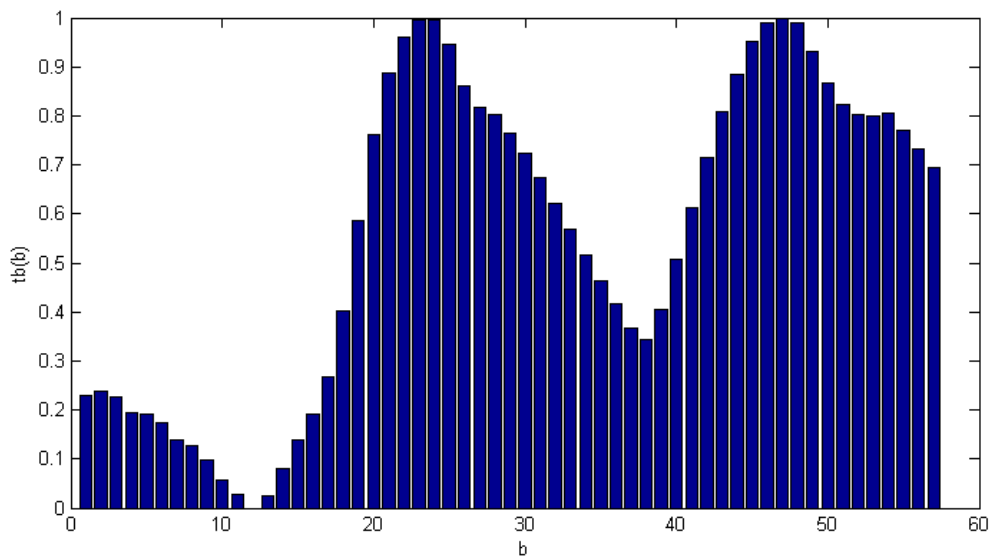
Obr. 5.9: Energie (červená barva) a maskovací křivka (modrá barva) pro signál 2



Obr. 5.10: Energie po konvoluci s maskovací křivkou pro signál 2

5.2.7 Index tonality

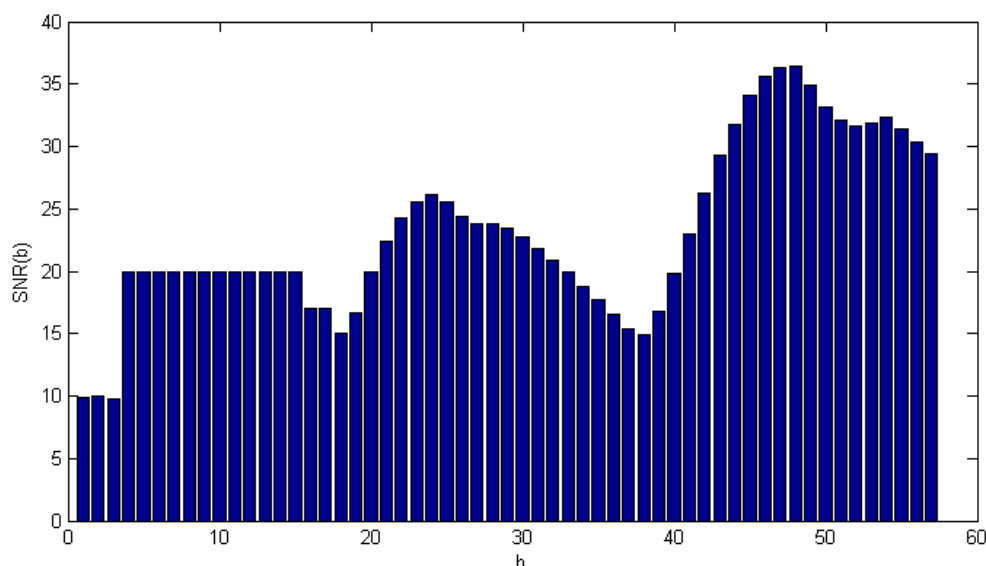
Index tonality určuje, zda jsou jednotlivé koeficienty tonální nebo šumové. Pro čistě tonální složku je hodnota koeficientu rovna 1, pro čistě šumovou složku je rovna 0 (viz kap. 4.2.5). Na obr. 5.11 je zobrazen index tonality pro signál 3.



Obr. 5.11: Index tonality u signálu 3

5.2.8 Odstup signálu od šumu

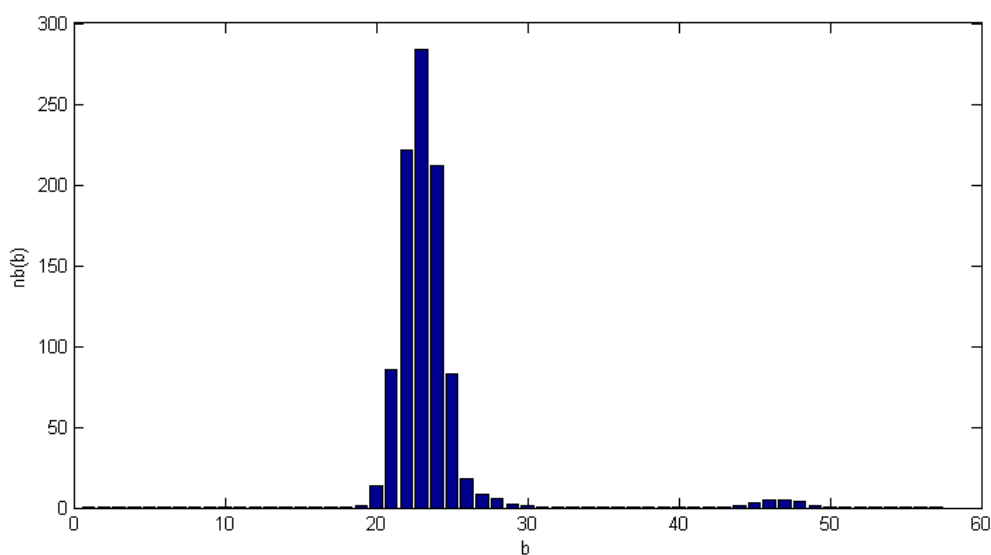
Odstup signálu od šumu je dán vztahem (4.15). Hodnoty jsou srovnávány s tabulkovými v normě [7], pro vzorkovací frekvenci 44,1 kHz jsou uvedené v příloze A v tabulce A-1. Na obrázku 5.12 jsou zobrazeny jednotlivé upravené koeficienty ve výpočtových pásmech.



Obr. 5.12: Odstup signálu od šumu (SNR) u signálu 3

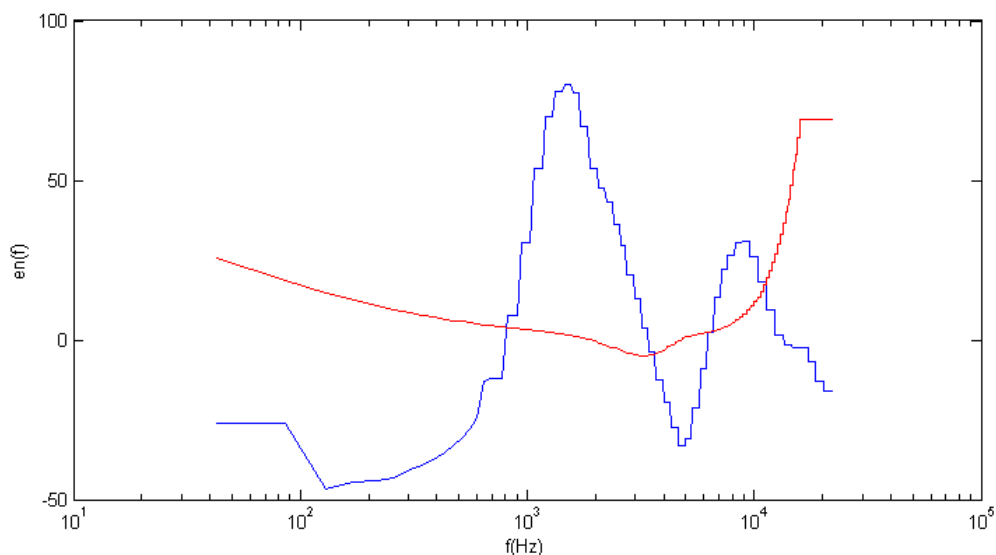
5.2.9 Spektrum slyšitelného zvuku

Na obr. 5.13 je zobrazen práh energie pro jednotlivá výpočtová pásma.



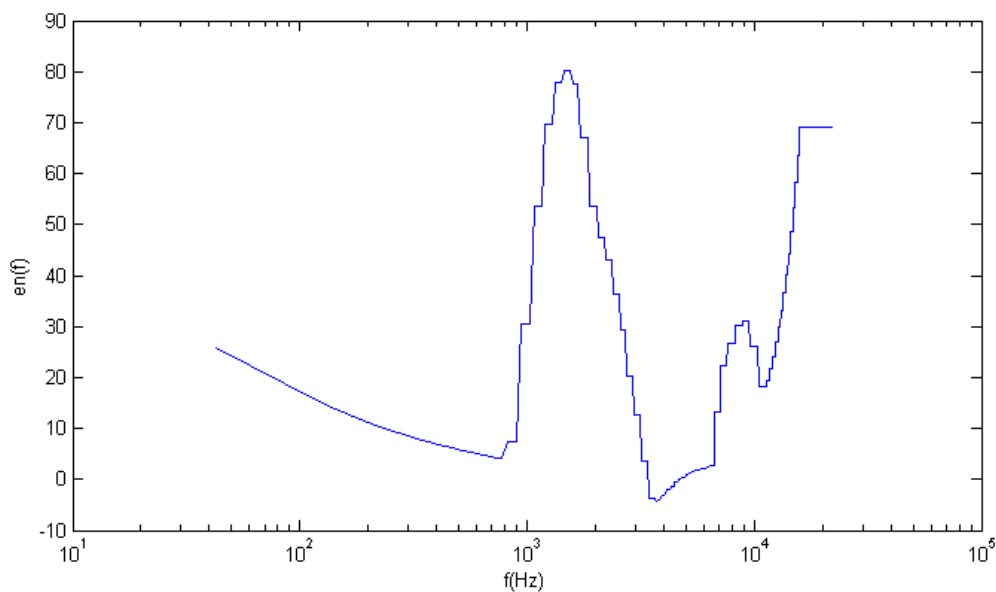
Obr. 5.13: Práh energie ve výpočtovém pásmu u signálu 3

Jednotlivá pásma jsou v dalším kroku sloučena (viz kap. 4.2.7) a převedena zpět na frekvenční koeficienty, které jsou pro názornost zobrazeny pomocí funkce *plot*. Na obr. 5.14 je zobrazen práh energie převedený do frekvenční oblasti pomocí vztahu (4.18) a práh slyšitelnosti, který je popsán v kap. 1.3.



Obr. 5.14: Práh energie (modrá barva) a práh slyšitelnosti (červená barva) u signálu 3

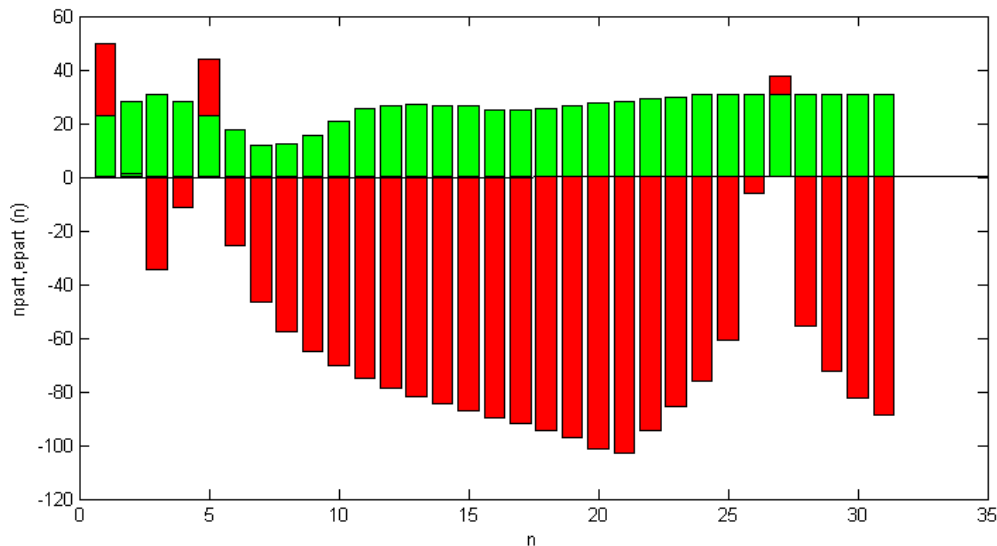
Dalším krokem je oříznutí prahu energie prahem slyšitelnosti, tím získáme slyšitelné frekvenční spektrum signálu, viz obr. 5.15.



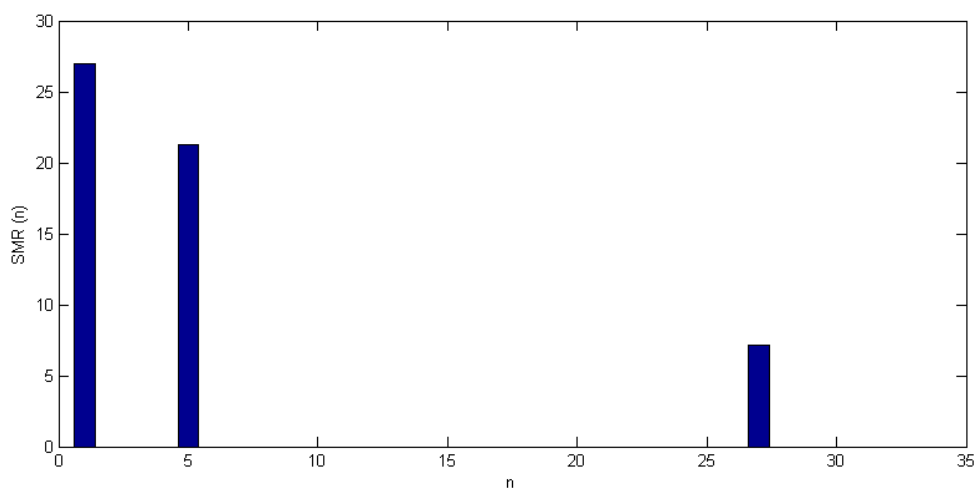
Obr. 5.15 Slyšitelné spektrum signálu 3

5.2.10 Odstup signálu od masky

Odstup signálu od masky (SMR) je popsán v kap. 4.2.9. Dle tabulky 4.2 [7] nabývá 32 hodnot - měřítkových koeficientů. Na obr. 5.16 je zobrazena vstupní energie signálu a práh energie. Z těchto hodnot se pomocí vztahu (4.21) vypočítá samotný odstup signálu od masky, který je zobrazen na obrázku 5.17.



Obr. 5.16: Odstup signálu od šumu - energie signálu (zelená barva), práh energie (červená barva) vyjádřené v měřítkových koeficientech



Obr. 5.17: Odstup signálu od masky

6 SIMULACE BANKY FILTRŮ A BLOKŮ MDCT V MPEG 1 LAYER 3 V PROGRAMU MATLAB

6.1 Realizace programu

Simulace banky filtrů a bloku MDCT je vytvořena v programu MATLAB pomocí funkce `graf_filter.m`. Řešení není provedeno pomocí grafického rozhraní GUIDE.

Vstupní parametry se musejí zadávat jako parametrické volání funkce (např. `graf_filter(12, 10)`), kde první hodnota (12) je počet pásem, na který byl signál rozdělen a druhá hodnota (10) uvádí výběr subpásma. Subpásem je dle normy ISO/IEC 11172-3 [7] 32.

Vstupní signál byl zvolen z vytvořených signálů pro simulaci psychoakustického modelu 2 (viz kap. 5). Jedná se o signál 3.

6.2 Banka filtrů

V kap. 3.2 je stručně popsána funkce banky filtrů. Signál je rozdělen na 32 subpásem po 512 vzorcích. Vzorky musejí být vynásobeny koeficienty C_i , které jsou dány normou [7] a jsou vypsány v příloze v tabulce A-2.

Dalším krokem je volba správného okna, které je určeno psychoakustickým modelem pomocí perceptuální entropie (viz kap. 4.2.8). Program je nastaven pro počítání s normálním oknem pomocí vztahu

$$z_i = x_i \sin\left(\frac{\pi}{36}\left(i + \frac{1}{2}\right)\right) \text{ pro } i = 0 \text{ až } 35, \quad (6.1)$$

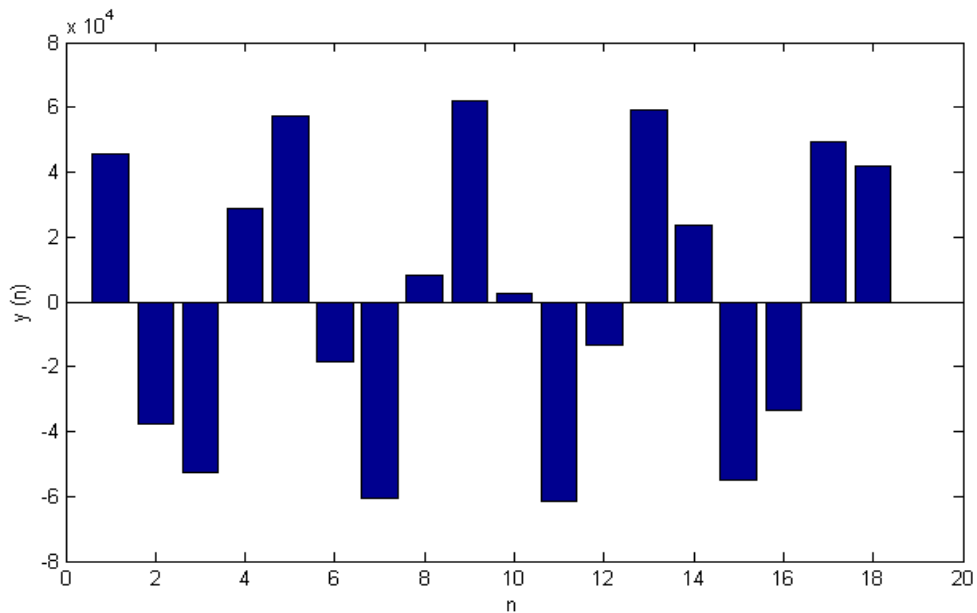
kde x_i je 36 vstupních vzorků z paměti (viz kap.3.2).

6.3 Blok MDCT

Vstupem bloku MDCT jsou přizpůsobené hodnoty pro normální okno vypočítané vztahem (6.1). V bloku MDCT se využívá díky překrývání jen 50 % vzorků (viz. kap. 3.2). Hodnoty jednotlivých spektrálních čar pro 17 vzorků jsou dány vztahem

$$y_i = \sum_{k=0}^{n-1} z_k \cos\left(\frac{\pi}{2n}\left(2k + 1 + \frac{n}{2}\right)(2i + 1)\right),$$

kde $k = 0$ až 2 , $i = 0$ až $\frac{n}{2} - 1$ a $n = 36$. Na obr. 6.1 je zobrazen výstup z MDCT bloku, který se dále zpracovává společně s výstupem psychoakustického modelu v bloku kvantizace.



Obr. 6.1: Výstup z MDCT bloku

7 ZÁVĚR

Hlavním cílem této diplomové práce bylo prostudování principů percepčního kódování zvukových signálu u standardu MPEG-1 Layer 3 a vytvoření programu v prostředí MATLAB pro simulaci jednotlivých kroků psychoakustického modelu s jejich grafickým znázorněním.

První kapitola pojednává o lidském vnímání zvuku, o jednotlivých metodách psychoakustiky, kterými dokážeme ovlivnit velikost zvukového souboru, aniž bychom ovlivnili zvukový vjem člověka.

Druhá kapitola se zabývá popisem jednotlivých standardů MPEG-1. Dává ucelený přehled, co tato pracovní skupina zaštiťuje. Na ní volně navazuje kapitola třetí, která popisuje už samotný formát MPEG-1 Layer 3, který je popsán v normě ISO/IEC 11172-3. Detailněji jsou zde popsány jednotlivé bloky kodéru a dekodéru a návaznosti mezi nimi. Čtenáři dá tato kapitola ucelený přehled o funkcionalitě celého formátu.

V kapitole čtvrté jsou popsány psychoakustické modely. Psychoakustický model 1, který se používá u standardu MPEG-1 Layer 1 a 2 je popsán velmi stručně. Psychoakustický model 2 je rozepsán hlouběji. Důvodem je jeho použití v simulačním programu. Popisují se zde jednotlivé vzorce, které jsou aplikovány v praktické simulaci.

V praktické části diplomové práce byl vytvořen program simulující psychoakustický model v prostředí MATLAB, který bude sloužit pro výuku studentů. Hlavním cílem bylo vytvořit názorné prostředí, ve kterém studenti uvidí jednotlivé funkční prvky a budou si moci uvědomit jejich návaznosti. Vytvořený program a jeho grafické výstupy jsou popsány v kapitole páté.

Kapitola šestá rozšiřuje kapitolu třetí o podrobnější popis banky filtrů a bloku MDCT. Tyto bloky zpracovávají paralelně vstupní signál s psychoakustickým modelem, který jejich funkčnost ovlivňuje. V prostředí MATLAB byl vytvořen jednoduchý program, který tyto bloky simuluje. Bohužel z důvodu jednoduchosti a více výukového pojetí těchto programů, není součástí této práce jejich vzájemné propojení.

ZKRATKY

ASPEC	Adaptive Spectral Perceptual Entropy Coding
FFT	Fast Fourier Transform
MDCT	Modified Discrete Cosine Transform
MPEG	Motion Picture Experts Group
PCM	Pulse-Code Modulation
PQMF	Pseudo-Quadrature Mirror Filter
SMR	Signal to Mask Ratio
SNR	Signal to Noise Ratio
SPL	Sound Pressure Level

LITERATURA

- [1] *Atlas lidského těla v obrazech*. Zpracoval kolektiv autorů. REBO, Česká republika, 2008, ISBN: 978-80-7234-896-1
- [2] <http://www.neuroreille.com/promenade/francais/cochlea/fcochlea.htm> [online], stav k 15.5.2011
- [3] SMÉKAL, Z., Číslicové zpracování řeči, Vysoké učení technické v Brně, Česká republika, 2009
- [4] ZWICKER E., FASTL H., *Psychoacoustics*, Springer, USA, 1999. ISBN 3-540-65063-6
- [5] BOSI M., GOLDBERG R.E. *Introduction to Digital Audio Coding and Standards*. Kluwer Academic Publishers Group, USA, 2003, ISBN: 1-4020-7357-7
- [6] MPEG home page [online], dostupné <http://mpeg.chiariglione.org>, stav k 10.12.2010
- [7] ČSN EN ISO/IEC 11172/3 *Informační technologie - Kódování pohyblivých obrazů včetně doprovodného zvuku pro číslicový záznam do rychlosti 1,5 Mbit/s, část 3: Zvuk (audio)*. Český normalizační institut, Česká republika, 1996.
- [8] SMÉKAL, Z., Číslicové zpracování signálu, Vysoké učení technické v Brně, Česká republika, 2010
- [9] <http://elektronika.kvalitne.cz/ATMEL/necoteorie/transformat/AVRFFT/AVRFFT.html> [online], stav k 24.4.2011
- [10] ZAPLATÍLEK K., DOŇÁR B., MATLAB - Tvorba uživatelských aplikací, Technická literatura BEN, Česká republika, Praha, 2004, ISBN: 80-73000-133-0

PŘÍLOHA A

Tab. A-1: Rozsahy výpočetních pásem pro vzorkovací frekvenci 44,1 kHz

index	ωlow	ωhigh	bval	minval	TMN
1	1	1	0.00	0.0	24.5
2	2	2	0.43	0.0	24.5
3	3	3	0.86	0.0	24.5
4	4	4	1.29	20.0	24.5
5	5	5	1.72	20.0	24.5
6	6	6	2.15	20.0	24.5
7	7	7	2.58	20.0	24.5
8	8	8	3.1	20.0	24.5
9	9	9	3.45	20.0	24.5
10	10	10	3.88	20.0	24.5
11	11	11	4.28	20.0	24.5
12	12	12	4.67	20.0	24.5
13	13	13	5.6	20.0	24.5
14	14	14	5.42	20.0	24.5
15	15	15	5.77	20.0	24.5
16	16	16	6.11	17.0	24.5
17	17	19	6.73	17.0	24.5
18	20	22	7.61	15.0	24.5
19	23	25	8.44	10.0	24.5
20	26	28	9.21	7.0	24.5
21	29	31	9.88	7.0	24.5
22	32	34	10.51	4.4	25.0
23	35	37	11.11	4.5	25.6
24	38	40	11.65	4.5	26.2
25	41	44	12.24	4.5	26.7
26	45	48	12.85	4.5	27.4
27	49	52	13.41	4.5	27.9
28	53	56	13.94	4.5	28.4
29	57	60	14.42	4.5	28.9
30	61	64	14.86	4.5	29.4
31	65	69	15.32	4.5	29.8
32	70	74	15.79	4.5	30.3
33	75	80	16.26	4.5	30.8
34	81	86	16.73	4.5	31.2
35	87	93	17.19	4.5	31.7
36	94	100	17.62	4.5	32.1
37	101	108	18.5	4.5	32.5
38	109	116	18.45	4.5	32.9
39	117	124	18.83	4.5	33.3
40	125	134	19.21	4.5	33.7
41	135	144	19.60	4.5	34.1

42	145	155	20.00	4.5	34.5
43	156	166	20.38	4.5	34.9
44	167	177	20.74	4.5	35.2
45	178	192	21.12	4.5	35.6
46	193	207	21.48	4.5	36.0
47	208	222	21.84	4.5	36.3
48	223	243	22.20	4.5	36.7
49	244	264	22.56	4.5	37.1
50	265	286	22.91	4.5	37.4
51	287	314	23.26	4.5	37.8
52	315	342	23.60	4.5	38.1
53	343	371	23.95	4.5	38.4
54	372	401	24.30	4.5	38.8
55	402	431	24.65	4.5	39.1
56	432	469	25.00	4.5	39.5
57	470	513	25.33	3.5	39.8

Tab. A-2: Keficienty C_i pro výpočet vektoru Z

i	koef		i	koef		i	koef		i	koef
1	0.000000000		129	0.000971317		257	0.035780907		385	0.000971317
2	-0.000000477		130	0.000953674		258	0.035758972		386	0.000983715
3	-0.000000477		131	0.000930786		259	0.035694122		387	0.000991821
4	-0.000000477		132	0.000902653		260	0.035586357		388	0.000995159
5	-0.000000477		133	0.000868797		261	0.035435200		389	0.000994205
6	-0.000000477		134	0.000829220		262	0.035242081		390	0.000898437
7	-0.000000477		135	0.000783920		263	0.035007000		391	0.000980854
8	-0.000000954		136	0.000731945		264	0.034730434		392	0.000968933
9	-0.000000954		137	0.000674248		265	0.034412861		393	0.000954151
10	-0.000000954		138	0.000610352		266	0.034055710		394	0.000935555
11	-0.000000954		139	0.000539303		267	0.033659935		395	0.000915051
12	-0.000001431		140	0.000462532		268	0.033225536		396	0.000891685
13	-0.000001431		141	0.000378609		269	0.032754898		397	0.000866413
14	-0.000001907		142	0.000288486		270	0.032248020		398	0.000838757
15	-0.000001907		143	0.000191689		271	0.031706810		399	0.000809669
16	-0.000002384		144	0.000088215		272	0.031132698		400	0.000779152
17	-0.000002384		145	-0.000021458		273	0.030526638		401	0.000747204
18	-0.000002861		146	-0.000137329		274	0.029890060		402	0.000714302
19	-0.000003338		147	-0.000259876		275	0.029224873		403	0.000680923
20	-0.000003338		148	-0.000388145		276	0.028532982		404	0.000646591
21	-0.000003815		149	-0.000522137		277	0.027815342		405	0.000611782
22	-0.000004292		150	-0.000661850		278	0.027073860		406	0.000576973
23	-0.000004768		151	-0.000806808		279	0.026310921		407	0.000542164
24	-0.000005245		152	-0.000956535		280	0.025527000		408	0.000507355

25	-0.000006199		153	-0.001111031		281	0.024725437		409	0.000472546
26	-0.000006676		154	-0.001269817		282	0.023907185		410	0.000438213
27	-0.000007629		155	-0.001432419		283	0.023074150		411	0.000404358
28	-0.000008106		156	-0.001597881		284	0.022228718		412	0.000371456
29	-0.000009060		157	-0.001766682		285	0.021372318		413	0.000339031
30	-0.000010014		158	-0.001937389		286	0.020506859		414	0.000307560
31	-0.000011444		159	-0.002110004		287	0.019634247		415	0.000277042
32	-0.000012398		160	-0.002283096		288	0.018756866		416	0.000247478
33	-0.000013828		161	-0.002457142		289	0.017876148		417	0.000218868
34	-0.000014782		162	-0.002630711		290	0.016994476		418	0.000191212
35	-0.000016689		163	-0.002803326		291	0.016112804		419	0.000165462
36	-0.000018120		164	-0.002974033		292	0.015233517		420	0.000140190
37	-0.000019550		165	-0.003141880		293	0.014358521		421	0.000116348
38	-0.000021458		166	-0.003306866		294	0.013489246		422	0.000093937
39	-0.000023365		167	-0.003467083		295	0.012627602		423	0.000072956
40	-0.000025272		168	-0.003622532		296	0.011775017		424	0.000052929
41	-0.000027657		169	-0.003771782		297	0.010933399		425	0.000034332
42	-0.000030041		170	-0.003914356		298	0.010103703		426	0.000017166
43	-0.000032425		171	-0.004048824		299	0.009287834		427	0.000000954
44	-0.000034809		172	-0.004174709		300	0.008487225		428	-0.000013828
45	-0.000037670		173	-0.004290581		301	0.007703304		429	-0.000027180
46	-0.000040531		174	-0.004395962		302	0.006937027		430	-0.000039577
47	-0.000043392		175	-0.004489899		303	0.006189346		431	-0.000050545
48	-0.000046253		176	-0.004570484		304	0.005462170		432	-0.000060558
49	-0.000049591		177	-0.004638195		305	0.004756451		433	-0.000069618
50	-0.000052929		178	-0.004691124		306	0.004072189		434	-0.000077724
51	-0.000055790		179	-0.004728317		307	0.003411293		435	-0.000084400
52	-0.000059605		180	-0.004748821		308	0.002774239		436	-0.000090122
53	-0.000062943		181	-0.004752159		309	0.002161503		437	-0.000095367
54	-0.000066280		182	-0.004737377		310	0.001573563		438	-0.000099182
55	-0.000070095		183	-0.004703045		311	0.001011848		439	-0.000102520
56	-0.000073433		184	-0.004649162		312	0.000475883		440	-0.000105381
57	-0.000076771		185	-0.004573822		313	-0.000033379		441	-0.000106812
58	-0.000080585		186	-0.004477024		314	-0.000515938		442	-0.000108242
59	-0.000083923		187	-0.004357815		315	-0.000971317		443	-0.000108719
60	-0.000087261		188	-0.004215240		316	-0.001399517		444	-0.000108719
61	-0.000090599		189	-0.004049301		317	-0.001800537		445	-0.000108242
62	-0.000093460		190	-0.003858566		318	-0.002174854		446	-0.000107288
63	-0.000096321		191	-0.003643036		319	-0.002521515		447	-0.000105858
64	-0.000099182		192	-0.003401756		320	-0.002841473		448	-0.000103951
65	0.000101566		193	0.003134727		321	0.003134727		449	0.000101566
66	0.000103951		194	0.002841473		322	0.003401756		450	0.000099182
67	0.000105858		195	0.002521515		323	0.003643036		451	0.000096321
68	0.000107288		196	0.002174854		324	0.003858566		452	0.000093460
69	0.000108242		197	0.001800537		325	0.004049301		453	0.000090599

70	0.000108719		198	0.001399517		326	0.004215240		454	0.000087261
71	0.000108719		199	0.000971317		327	0.004357815		455	0.000083923
72	0.000108242		200	0.000515938		328	0.004477024		456	0.000080585
73	0.000106812		201	0.000033379		329	0.004573822		457	0.000076771
74	0.000105381		202	-0.000475883		330	0.004649162		458	0.000073433
75	0.000102520		203	-0.001011848		331	0.004703045		459	0.000070095
76	0.000099182		204	-0.001573563		332	0.004737377		460	0.000066280
77	0.000095367		205	-0.002161503		333	0.004752159		461	0.000062943
78	0.000090122		206	-0.002774239		334	0.004748821		462	0.000059605
79	0.000084400		207	-0.003411293		335	0.004728317		463	0.000055790
80	0.000077724		208	-0.004072189		336	0.004691124		464	0.000052929
81	0.000069618		209	-0.004756451		337	0.004638195		465	0.000049591
82	0.000060558		210	-0.005462170		338	0.004570484		466	0.000046253
83	0.000050545		211	-0.006189346		339	0.004489899		467	0.000043392
84	0.000039577		212	-0.006937027		340	0.004395962		468	0.000040531
85	0.000027180		213	-0.007703304		341	0.004290581		469	0.000037670
86	0.000013828		214	-0.008487225		342	0.004174709		470	0.000034809
87	-0.000000954		215	-0.009287834		343	0.004048824		471	0.000032425
88	-0.000017166		216	-0.010103703		344	0.003914356		472	0.000030041
89	-0.000034332		217	-0.010933399		345	0.003771782		473	0.000027657
90	-0.000052929		218	-0.011775017		346	0.003622532		474	0.000025272
91	-0.000072956		219	-0.012627602		347	0.003467083		475	0.000023365
92	-0.000093937		220	-0.013489246		348	0.003306866		476	0.000021458
93	-0.000116348		221	-0.014358521		349	0.003141880		477	0.000019550
94	-0.000140190		222	-0.015233517		350	0.002974033		478	0.000018120
95	-0.000165462		223	-0.016112804		351	0.002803326		479	0.000016689
96	-0.000191212		224	-0.016994476		352	0.002630711		480	0.000014782
97	-0.000218868		225	-0.017876148		353	0.002457142		481	0.000013828
98	-0.000247478		226	-0.018756866		354	0.002283096		482	0.000012398
99	-0.000277042		227	-0.019634247		355	0.002110004		483	0.000011444
100	-0.000307560		228	-0.020506859		356	0.001937389		484	0.000010014
101	-0.000339031		229	-0.021372318		357	0.001766682		485	0.000009060
102	-0.000371456		230	-0.022228718		358	0.001597881		486	0.000008106
103	-0.000404358		231	-0.023074150		359	0.001432419		487	0.000007629
104	-0.000438213		232	-0.023907185		360	0.001269817		488	0.000006676
105	-0.000472546		233	-0.024725437		361	0.001111031		489	0.000006199
106	-0.000507355		234	-0.025527000		362	0.000956535		490	0.000005245
107	-0.000542164		235	-0.026310921		363	0.000806808		491	0.000004768
108	-0.000576973		236	-0.027073860		364	0.000661850		492	0.000004292
109	-0.000611782		237	-0.027815342		365	0.000522137		493	0.000003815
110	-0.000646491		238	-0.028532982		366	0.000388145		494	0.000003338
111	-0.000680923		239	-0.029224873		367	0.000259876		495	0.000003338
112	-0.000714302		240	-0.029890060		368	0.000137329		496	0.000002861
113	-0.000747204		241	-0.030526638		369	0.000021458		497	0.000002384
114	-0.000779152		242	-0.031132698		370	-0.000088215		498	0.000002384

115	-0.000809669		243	-0.031706810		371	-0.000191689		499	0.000001907
116	-0.000838757		244	-0.032248020		372	-0.000288486		500	0.000001907
117	-0.000866413		245	-0.032754898		373	-0.000378609		501	0.000001431
118	-0.000891685		246	-0.033225536		374	-0.000462532		502	0.000001431
119	-0.000915051		247	-0.033659935		375	-0.000539303		503	0.000000954
120	-0.000935555		248	-0.034055710		376	-0.000610352		504	0.000000954
121	-0.000954151		249	-0.034412861		377	-0.000674248		505	0.000000954
122	-0.000968933		250	-0.034730434		378	-0.000731945		506	0.000000954
123	-0.000980854		251	-0.035007000		379	-0.000783920		507	0.000000477
124	-0.000898437		252	-0.035242081		380	-0.000829220		508	0.000000477
125	-0.000994205		253	-0.035435200		381	-0.000868797		509	0.000000477
126	-0.000995159		254	-0.035586357		382	-0.000902653		510	0.000000477
127	-0.000991821		255	-0.035694122		383	-0.000930786		511	0.000000477
128	-0.000983715		256	-0.035758972		384	-0.000953674		512	0.000000477

PŘÍLOHA B

Obsah přiloženého CD:

Diplomová práce\

fb\
- filtr.m

- graf_filter.m
- kefcienty_C1.txt
- mdct.m
- window_L.m

pm2\
- bloky.m

- energie.m
- fourtrans.m
- konvoluce.m
- model.fig
- model.m (spouštěcí soubor)
- nepredvidatelnost.m
- predikce.m
- SMR.m
- SNR.m
- spreading.m
- tabulka.m
- tabulka2.m
- tonalita.m

diplomova_prace.docx

diplomova_prace.pdf