

**ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE**

---

**PROVOZNĚ EKONOMICKÁ FAKULTA**



**Katedra informačních technologií**

Diplomová práce

**Digitální zvuk**

**Jakub Glanc**

© ČZU v Praze 2011

## **Čestné prohlášení**

Prohlašuji, že jsem diplomovou práci na téma „Digitální zvuk“ vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

**V Českém Brodě dne 8. dubna 2011**

.....

**Jakub Glanc**

## **Poděkování**

Rád bych touto cestou poděkoval vedoucímu diplomové práce Ing. Miloši Ulmanovi, Ph.D. za odborné vedení a cenné rady při psaní práce.

*Digitální zvuk*

*Digital sound*

## **Souhrn**

Diplomová práce se zabývá problematikou digitálního zvuku a jeho formátů. Popisuje základní veličiny a důležité skutečnosti z oblasti akustiky. Dále je předmětem práce popis vybraných digitálních audio formátů a porovnání ztrátových audio formátů mezi sebou s následným vyhodnocením tohoto porovnání a vyvozením závěrů pro užití formátu pro konkrétní případy. Další součástí analytické části je vícekriteriální analýza variant, která řeší výběr ztrátového audio formátu pro ukládání audio souborů v domácích podmínkách na základě stanovených kritérií.

## **Klíčová slova**

Digitální, zvuk, formát, mp3, ogg, wma, frekvence, porovnání, vícekriteriální analýza

## **Summary**

This thesis deals with digital sound and its formats. It describes the fundamentals and important facts from the field of acoustics. Further work is subject to a description of selected digital audio formats and audio formats loss comparison between the subsequent evaluation of this comparison, and draw conclusions for a particular format for use cases. Another part of the analytical part of the multi-criteria analysis of options to address the choice of lossy audio format for storing audio files in the home under the conditions laid down criteria.

## **Key words**

Digital, audio, format, mp3, ogg, wma, frequency, comparison, multicriteria analysis

## OBSAH

<b>1</b>	<b>Úvod .....</b>	<b>12</b>
<b>2</b>	<b>Cíl práce a metodika.....</b>	<b>13</b>
<b>3</b>	<b>Digitální zvuk a zvukové formáty.....</b>	<b>14</b>
3.1	Zvuk.....	14
3.2	Sluchový systém.....	14
3.2.1	Binaurální slyšení.....	15
3.2.2	Hlasitost.....	16
3.2.3	Barva zvuku .....	17
3.2.4	Frekvenční rozsah .....	17
3.2.5	Maskování (frekvenční, temporal masking) .....	18
3.3	Digitální záznam zvuku.....	20
3.3.1	Rozdíl mezi analogovým a digitálním záznamem zvuku .....	20
3.3.2	Postup digitalizace (vzorkovací frekvence, kódování, kanály...)	22
3.3.3	Kvalita zvuku a objem dat.....	24
3.3.4	Komprese audio dat.....	25
3.3.5	Způsoby záznamu.....	29
3.4	Digitální audio formáty .....	32
3.4.1	Ztrátové .....	33
3.4.2	Bezztrátové.....	40
3.4.3	MIDI.....	43
<b>4</b>	<b>Porovnání digitálních zvukových formátů .....</b>	<b>44</b>
4.1	Bezztrátové formáty – 256 kbps.....	44
4.1.1	WAV .....	45
4.1.2	FLAC.....	46
4.1.3	APE .....	46
4.1.4	AIFF .....	47
4.2	Ztrátové formáty – 256 kbps .....	47
4.2.1	MP3 .....	48
4.2.2	WMA.....	49

4.2.3	OGG .....	50
4.2.4	AAC .....	50
4.2.5	MPC .....	51
4.3	Bezztrátové formáty – 128 kbps.....	53
4.4	Ztrátové formáty – 128 kbps .....	53
4.4.1	MP3 .....	54
4.4.2	WMA.....	55
4.4.3	OGG .....	55
4.4.4	AAC .....	56
4.4.5	MPC .....	57
4.5	Bezztrátové formáty – 96 kbps.....	59
4.6	Ztrátové formáty – 96 kbps .....	59
4.6.1	MP3 .....	60
4.6.2	WMA.....	61
4.6.3	OGG .....	61
4.6.4	AAC .....	62
4.6.5	MPC .....	62
4.7	Určení vhodného ztrátového formátu pro ukládání hudebních dat v domácích podmínkách na základě vícekritériální analýzy .....	65
4.7.1	Metoda TOPSIS .....	67
4.7.2	Metoda ORESTE .....	69
4.7.3	Metoda váženého součtu .....	70
<b>5</b>	<b>Závěr .....</b>	<b>73</b>
<b>6</b>	<b>Seznam použitých zdrojů .....</b>	<b>75</b>
<b>7</b>	<b>Přílohy.....</b>	<b>79</b>
7.1	Porovnání všech ztrátových formátů .....	79
7.2	Loga jednotlivých formátů .....	80

## Seznam tabulek

Tabulka 1 - Záznamové formáty .....	29
Tabulka 2 - Přehled ztrátových formátů.....	33
Tabulka 3 - Přehled bezztrátových formátů.....	41
Tabulka 4 - Bezztrátové formáty 256 kbps .....	44

Tabulka 5 - Ztrátové formáty 256 kbps.....	48
Tabulka 6- Bezztrátové formáty 128 kbps.....	53
Tabulka 7 - Ztrátové formáty 128 kbps.....	53
Tabulka 8 - Bezztrátové formáty 96 kbps.....	59
Tabulka 9 - Ztrátové formáty 96 kbps.....	60
Tabulka 10 - Kriteriaální matice.....	67
Tabulka 11 - Výsledky metody TOPSIS.....	68
Tabulka 12 - Výsledky metody ORESTE.....	70
Tabulka 13 - Výsledky metody váženého součtu.....	71

### Seznam obrázků

Obrázek 1 - Sluchový systém.....	15
Obrázek 2 - Křivky stejné hlasitosti.....	16
Obrázek 3 - Příklad barvy zvuku.....	17
Obrázek 4 - Spektrum zvuku.....	18
Obrázek 5 - Frekvenční maskování.....	19
Obrázek 6 - Temporal masking.....	19
Obrázek 7 - Digitální/analogový signál.....	20
Obrázek 8 - Zakódování digitálního signálu.....	21
Obrázek 9 - Vzorkování.....	22
Obrázek 10 - Kvantování.....	23
Obrázek 11 - Kódování.....	24
Obrázek 12 - Princip PCM.....	27
Obrázek 13 - Princip SBC.....	29
Obrázek 14 - Záznamová hlava magnetofonu.....	31
Obrázek 15 - Princip magneto-optického záznamu digitalizovaných dat na minidisk.....	32
Obrázek 16 - Princip MP3.....	37
Obrázek 17 - Princip AAC.....	39
Obrázek 18 - Spektrum WAV 256 kbps.....	45
Obrázek 19 - Spektrum FLAC 256 kbps.....	46
Obrázek 20 - Spektrum APE 256 kbps.....	46
Obrázek 21 - Spektrum AIFF 256 kbps.....	47
Obrázek 22 - Spektrum MP3 256 kbps.....	48
Obrázek 23 - Spektrum WMA 256 kbps.....	49
Obrázek 24 - Spektrum OGG nejvyšší kvalita.....	50
Obrázek 25 - Spektrum AAC 256 kbps.....	50
Obrázek 26 - Spektrum MPC 256 kbps.....	51
Obrázek 27 - Spektrum MP3 128 kbps.....	54
Obrázek 28 - Spektrum WMA 128 kbps.....	55
Obrázek 29 - Spektrum OGG střední kvalita.....	55
Obrázek 30 - Spektrum AAC 128 kbps.....	56



Obrázek 31 - Spektrum MPC 128 kbps .....	57
Obrázek 32 - Spektrum MP3 96 kbps .....	60
Obrázek 33 - Spektrum WMA 96 kbps.....	61
Obrázek 34 - Spektrum OGG nízká kvalita .....	61
Obrázek 35 - Spektrum AAC 96 kbps .....	62
Obrázek 36 - Spektrum MPC 96 kbps .....	62

### **Seznam grafů**

Graf 1 - Porovnání ztrátových formátů při 256 kbps .....	49
Graf 2 - Porovnání ztrátových formátů při 128 kbps .....	53
Graf 3 - Porovnání ztrátových formátů při 96 kbps .....	64
Graf 4 - Odpovědi respondentů.....	66
Graf 5 - Porovnání ztrátových formátů celkem.....	79

# 1 Úvod

První zmínky o digitálním zvuku se objevily již ve 30. letech minulého století. PCM neboli pulzně-kódová modulace, která vznikla na základě řady výzkumů a pokusů britského vědce Alexe Reevse dala základ dalšímu vývoji v této oblasti.

Průkopníkem v digitálním audiu je především německý Fraunhoferův ústav pro mediální komunikaci v Ilmenau, který mimo jiné postavil základy pro v současné době nejpoužívanější a nejpopulárnější ztrátový formát komprese audia – MP3.

Ztrátová komprese se stala významnou především na konci 20. století společně s rozvojem počítačové gramotnosti obyvatelstva a větší dostupností počítačové techniky. Zároveň se v této době začala prodávat velká škála přenosných přehrávačů podporujících především formát MP3.

V současné době již existuje velká řada ztrátových i bezztrátových formátů digitálního audia, které se liší svými vlastnostmi a samozřejmě také použitelností pro různé účely.

Dnes již není ani pro běžné uživatele problémem zkopírovat si svoje oblíbené CD na harddisk počítače nebo do některého z řady přenosných přehrávačů nebo i mobilních telefonů, které již běžně podporují větší množství jak ztrátových tak i bezztrátových formátů.

Řada formátů se snaží ohrozit dlouholeté výsadní postavení formátu MP3 na čele žebříčku oblíbenosti mezi širokým spektrem uživatelů. Některým se daří prosadit u určité uživatelské skupiny, ale postavení MP3 se dlouhodobě nemění a oblíbenost tohoto formátu je dlouhodobě stabilní.

## 2 Cíl práce a metodika

Cílem teoretické části diplomové práce je charakteristika digitálního zvuku a jeho formátů, popis forem záznamu digitálního zvuku a základních charakteristik vážících se k této problematice.

Cílem analytické části diplomové práce je porovnání těchto formátů na různých bit rates a jejich zhodnocení podle kritérií uživatele a použitelnosti v různých oblastech života. Formáty jsou porovnávány na základě autorem stanovených kritérií a také na základě frekvenční analýzy spekter každého z nich.

Dalším cílem analytické části diplomové práce je vybrat vhodný ztrátový formát digitálního audia pro ukládání rockové a populární hudby v domácích podmínkách. Tohoto cíle bude naplněno s pomocí tří metod vícekritériální analýzy variant, jejichž výsledky budou následně porovnány.

První kapitola obsahuje úvod, v kterém je práce stručně představena.

Druhá kapitola, Cíl práce a metodika, popisuje metody, s jejichž pomocí byla práce sepsána a cíl, ke kterému by měla celá práce směřovat.

Kapitola 3 je sepsána na základě literární rešerše, vysvětluje princip digitalizace zvuku a obecně popisuje jednotlivé zvukové formáty. Na začátku této kapitoly je krátká zmínka o sluchovém systému člověka a základních vlastnostech zvuku jako např.: intenzita, barva, frekvenční rozsah a podobně.

Kapitola 4 obsahuje analýzu a srovnávací test zvukových formátů, jejich zhodnocení a popis výhod a nevýhod jednotlivých formátů. Je zde doporučen vhodný ztrátový formát digitálního audia pro ukládání hudby v domácích podmínkách.

### **3 Digitální zvuk a zvukové formáty**

S postupem technologií ve všech oblastech lidského života se vyvíjí i oblast audia. Od mechanického záznamu analogového zvuku jsme dnes již postoupili a velký krok dál k digitálnímu zvuku. V této první kapitole budou objasněny některé základní pojmy nezbytné k dalšímu porozumění textu. Bude se jednat o základní akustické veličiny a pojmy úzce související s danou problematikou.

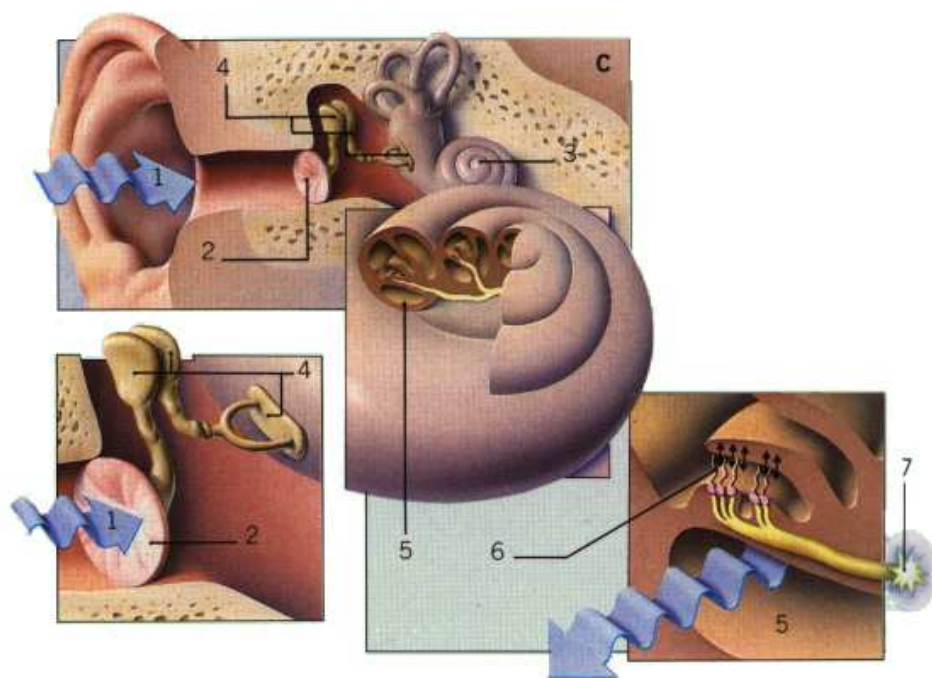
#### **3.1 Zvuk**

Různé druhy zvuku nás obklopují každý den. Hluk na ulici, rozhovor s přáteli, poslouchání hudební nahrávky. Ve všech těchto případech se jedná o mechanické vlnění, které je člověk schopen zachytit a poté zpracovat uchem. Frekvence takového vlnění se musí pohybovat v rozmezí přibližně 16 Hz – 20 khz, aby byl lidský sluchový systém schopen toto vlnění zachytit a zpracovat. Rychlost šíření těchto vln je v závislosti na teplotě kolem 332 m/s.

Člověk se již dlouhou dobu snaží o co nejefektivnější zaznamenávání zvuku tak, aby bylo možné daný zvuk nebo skladbu přehrávat znovu a znovu. Proto zvukový záznam prošel od jeho vzniku celou řadou podob od rytí stopy do voskového válečku až po dnešní digitalizaci analogového signálu a jeho následné uchovávání v digitální formě na celé řadě médií jako například na pořád ještě oblíbeném kompaktním disku (CD).

#### **3.2 Sluchový systém**

Sluch je jedním z pěti smyslů člověka. Díky faktu, že člověk disponuje dvěma přijímači (uši), může lokalizovat podle dopadu jednotlivých zvukových vln na jedno či druhé ucho s určitým posunem i směr, odkud zdroj zvuku vychází.



Obrázek 1 - Sluchový systém

Zdroj: <http://panwiki.panska.cz/index.php/Soubor:Ucho2.jpg>, 5.2. 2011

Zvukové vlny jsou zachyceny uchem (bubínkem) (2), poté vlny putují do hlemýžďe za pomoci sluchových kůstek (4). Hlemýžď (5) je vyplněn určitým druhem organické kapaliny, jejíž tlak se zde na základě přichozího vlnění z bubínku mění. Tyto změny jsou zachycovány pomocí stereocilií na vláskových buňkách (6) a převáděny na nervové impulsy (7), které nakonec postupují sluchovým nervem až do mozku. [20]

### 3.2.1 Binaurální slyšení

Pod tímto pojmem je třeba si představit vnímání zvuku dvěma ušima. Fakt, že člověk je vybaven dvěma ušima, nám umožňuje lokalizovat přesně zdroj zvuku, z kterého dané vlnění vychází. Pokud se zdroj zvuku nachází přímo před posluchačem, potom je stejný úhel i dráha, které musí vlnění urazit při cestě do levého i pravého ucha. Obě uši vnímají stejnou hlasitost, zbarvení i fázové zpoždění. V tomto případě sluchový orgán (viz. předchozí kapitola) vyhodnotí, že zdroj vlnění se nachází přímo před posluchačem.

Nachází-li se ale zdroj mimo přímou osu, jsou dráhy, úhly a fáze zvukového vlnění v každém uchu rozdílné. Sluchový systém dokáže reagovat i na velmi nepatrné odchylky.

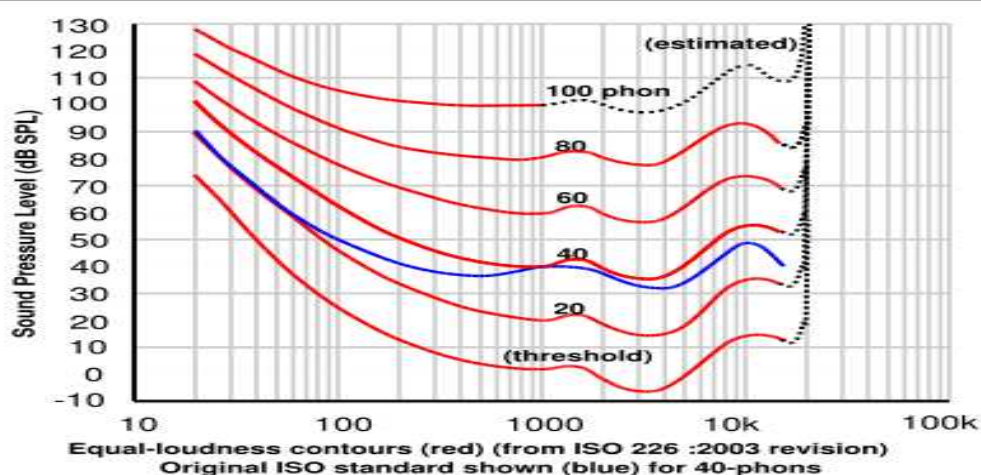
Z poměru hlasitostí, barvy zvuku a fázových rozdílů mezi vjemy pravého a levého ucha se vyhodnotí směr, ze kterého vlnění vychází. Tomuto se říká schopnost lokalizace zvuku. Citlivost vyhodnocování stoupá směrem v vyšším frekvencím. Sluchový systém dokáže lépe lokalizovat zdroj zvuku, který se nachází v rovině před posluchačem než zdroj, který se nachází nad, pod nebo za posluchačem. V případě poškození jednoho ucha posluchače se schopnost lokalizace podstatně zhoršuje. [21]

### 3.2.2 Hlasitost

Rozeznávají se dva druhy hlasitosti – subjektivní a objektivní.

Míra subjektivní hlasitosti je dána velikostí akustického tlaku, který vzniká při rezonanci zvukového zdroje a dopadá v podobě vln na uši člověka. Proto je hlasitost oproti intenzitě zvuku subjektivní, jelikož se nedá přesně změřit a každý člověk vnímá stejné vlnění jinak. Pro určení subjektivní hlasitosti při rozdílných frekvencích se užívá jednotka phon. Práh slyšitelnosti je touto jednotkou vyjádřen v hodnotě 0 – 4 Ph a práh bolestivosti 120 – 130 Ph. U tónů o hlasitosti pohybujících se v prahu bolestivosti je nebezpečí trvalého poškození sluchu.

Objektivní hlasitost se odborně nazývá intenzita zvuku a měří se v jednotkách decibel. Tato obdoba hlasitosti je oproti subjektivní hlasitosti přesně měřitelná.

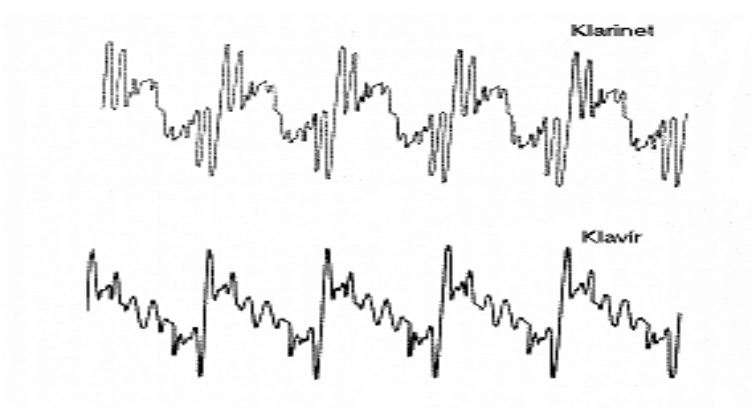


Obrázek 2 - Křivky stejné hlasitosti

Zdroj: <http://pctuning.tyden.cz>, 20.12.2010

### 3.2.3 Barva zvuku

Každý zvuk, který vydává ať už hudební nástroj nebo například lidské mluvicí ústrojí, obsahuje kromě základní frekvence i určitý podíl vyšších harmonických frekvencí. Tyto vyšší harmonické frekvence dávají každému hudebnímu nástroji jeho charakteristickou barvu. Proto například kytara zní jinak než klavír a podobně.



Obrázek 3 - Příklad barvy zvuku

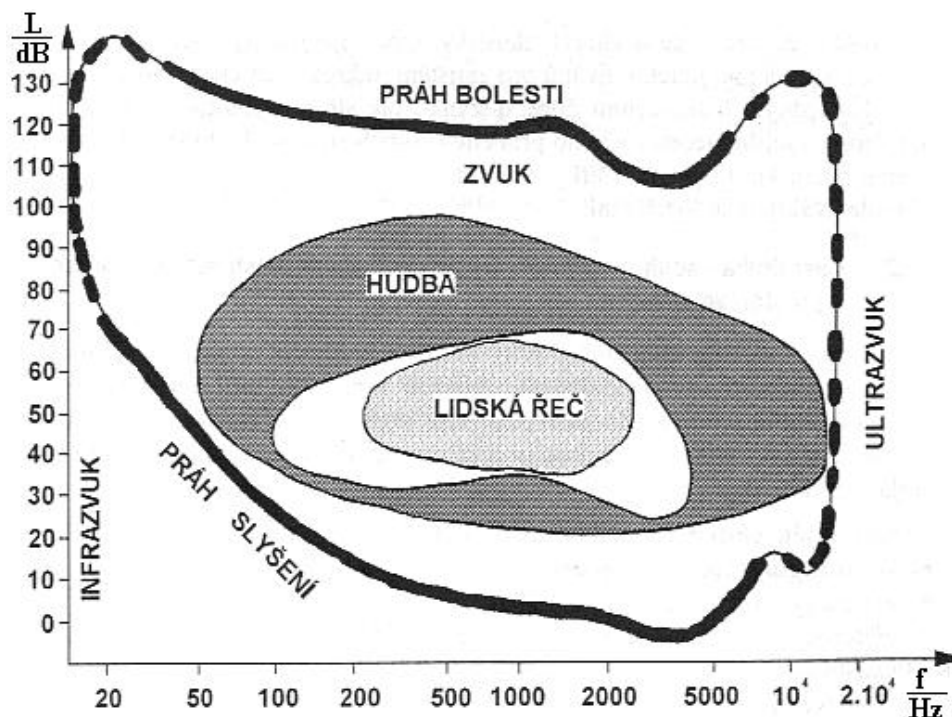
Zdroj: [www.quido.cz/fyzika/ images/Zvuk\\_graf.gif](http://www.quido.cz/fyzika/images/Zvuk_graf.gif), 1.1.2011

Samozřejmě to samé platí i o lidské řeči. Každá hláska obsahuje tzv. formanty, což jsou obdoby vyšších harmonických frekvencí u hudebních tónů. Například sykavky obsahují velmi vysoké frekvence a proto mají starší lidé problém s vnímáním těchto hlásek.

Mimo vyšších harmonických frekvencí ovlivňují barvu zvuku ještě způsob vzniku tónu (např. fouknutí do trubky, úder do bubnu...) a také doznívání (dokmit zdroje).

### 3.2.4 Frekvenční rozsah

Rozsah slyšitelných frekvencí je pro každého člověka jiný. V literatuře se udává, že frekvenční rozsah naprosto zdravého člověka se pohybuje v rozmezí 16 Hz – 20 khz (někdy se také udává 16 Hz – 16 khz nebo 20 Hz – 20 khz). S věkem člověka se snižuje schopnost sluchového ústrojí vnímat vyšší frekvence (tedy se snižuje horní hranice tohoto rozsahu). Pro běžnou lidskou komunikaci (řeč) jsou nejdůležitější frekvence v rozsahu 500 Hz – 4000 khz. V těchto úrovních se běžně pohybují frekvence lidské řeči. [26]



Obrázek 4 - Spektrum zvuku

Zdroj: <http://fyzika.jreichl.com/index.php?sekce=browse&page=210>, 20.3.2011

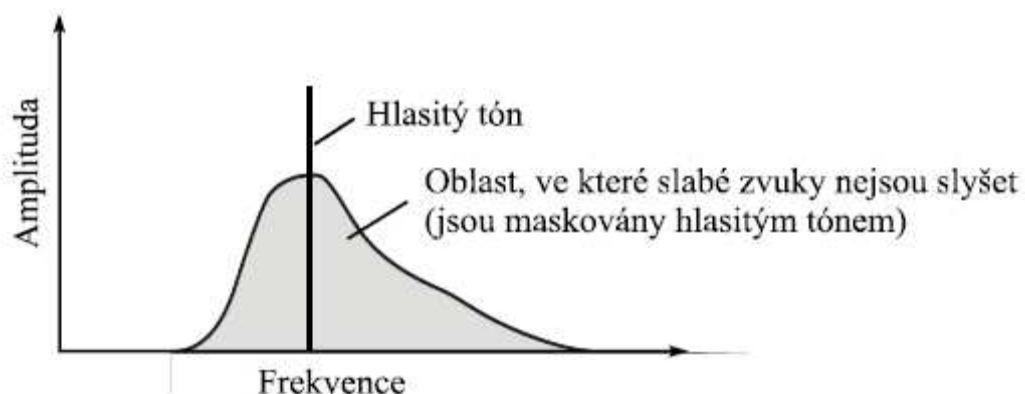
Všechny zvukové vlny vyskytující se pod minimální hranicí 16 Hz (20 Hz) se označují jako infrazvuk a nejsou běžným lidským uchem slyšitelné. Stejně je tomu s vlnami v pásmu nad 20 kHz. Zvuku, který frekvenčně spadá do tohoto pásma, se říká ultrazvuk. Ultrazvuku se využívá mimo jiné v medicíně.

### 3.2.5 Maskování (frekvenční, temporal masking)

Slabé zvuky lidský sluch vnímá pouze je-li naprosté nebo alespoň poměrné ticho. Když se objeví silnější zvuky, okamžitě přehluší zvuky slabší. Tomuto jevu se říká maskování a můžeme říci, že silnější zvuk maskuje slabší. Silnější zvuk maskuje především zvuky, které mají velmi blízkou frekvenci jako daný silnější zvuk (jsou mu frekvenčně blízké). Na následujícím obrázku je znázorněna křivka závislosti maskování zvuku na frekvenci v blízkosti hlasitého tónu. Z obrázku je patrné, že křivka je plošší směrem k vyšším frekvencím. Proto také při dělení spektra jsou u vyšších frekvencí větší šířky pásem a u nižších frekvencí jsou tyto pásma užší. Principu maskování se například



hojně využívá u standardu MPEG, kde jsou kódovány pouze ty zvuky, které jsou slyšitelné.

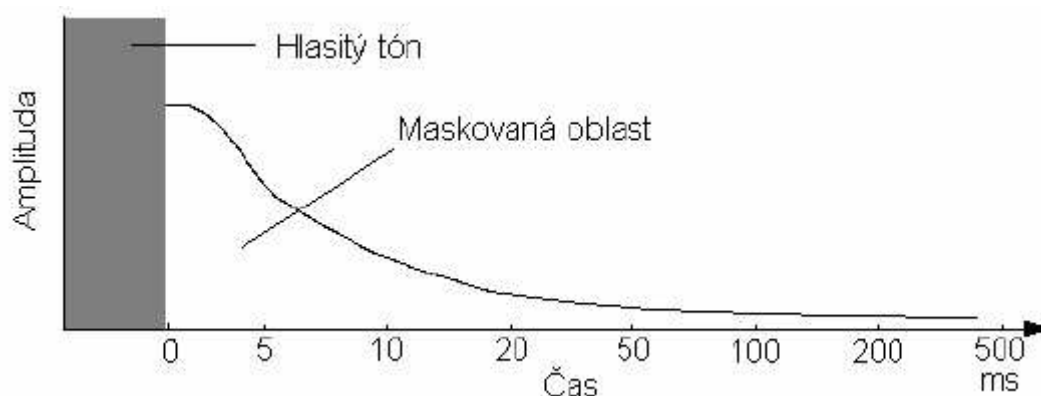


Obrázek 5 - Frekvenční maskování

Zdroj: <http://phoenix.inf.upol.cz/esf/ucebni/kompresse.pdf>, 13.3.2011

#### Maskování po silném zvuku (temporal masking)

Po skončení silného zvuku určitou dobu trvá, než je lidské ucho schopno vnímat zvuky slabší. Tato skutečnost je zobrazena na obrázku, který následuje za tímto odstavcem. Maskovací efekt vzniká i před silným zvukem a ne pouze po něm. Tato skutečnost ukazuje, že mozek potřebuje určitý čas pro zpracování sluchového vjemu. Efekt maskování se před silným zvukem pohybuje mezi 2 a 5 ms a po silném zvuku se tato doba prodlužuje až na 100 ms. [33]



Obrázek 6 - Temporal masking

Zdroj: <http://phoenix.inf.upol.cz/esf/ucebni/kompresse.pdf>, 13.3.2011

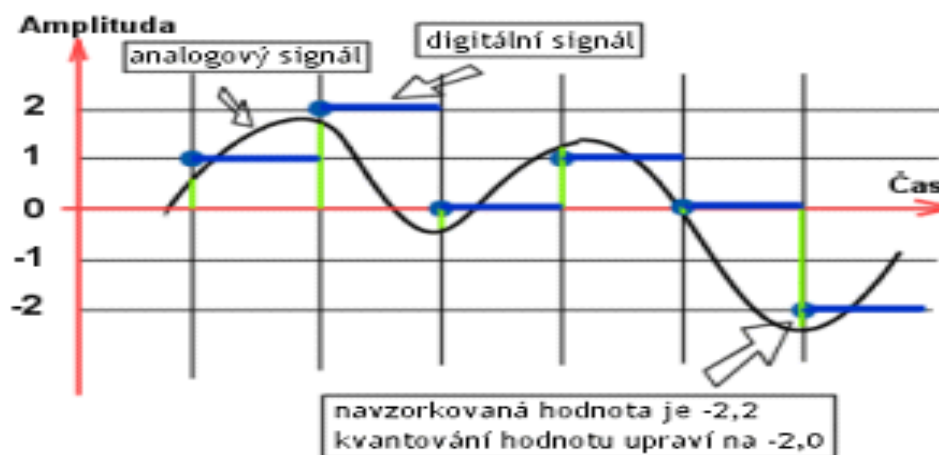
### 3.3 Digitální záznam zvuku

Digitální zvuk se objevil až dlouho po tom analogovém. První techniky, které se používaly v souvislosti se zpracováním a nahráváním analogového zvuku, zahrnovaly například mechanické válce a kotouče, magnetické dráty, pásky a optické filmy, které byly závislé na napětí a vyrovnavání stavů zajišťovaných vakuovou trubicí zesilovače.

V případě digitálního záznamu je informace v tzv. binární formě. Signál má pouze dva stavy (0 a 1), které se mění v předurčených momentech v závislosti na stavových hodinách. Pokud je binární signál znehodnocen šumem, je odmítnut přijímačem, jako signál je hodnocen pouze takový signál, který je nad nebo pod úrovní určitého „prahu“ (určitý stanovený bod). Nicméně signál bude přenášen s konečnou šířkou pásma a to bude omezovat rychlost, při kterém se napětí mění. Přehlušující šum může mít za důsledek posunutí bodu, ve kterém přijímač usoudí, že došlo ke změna stavu. Časová nestabilita má stejný efekt. [34]

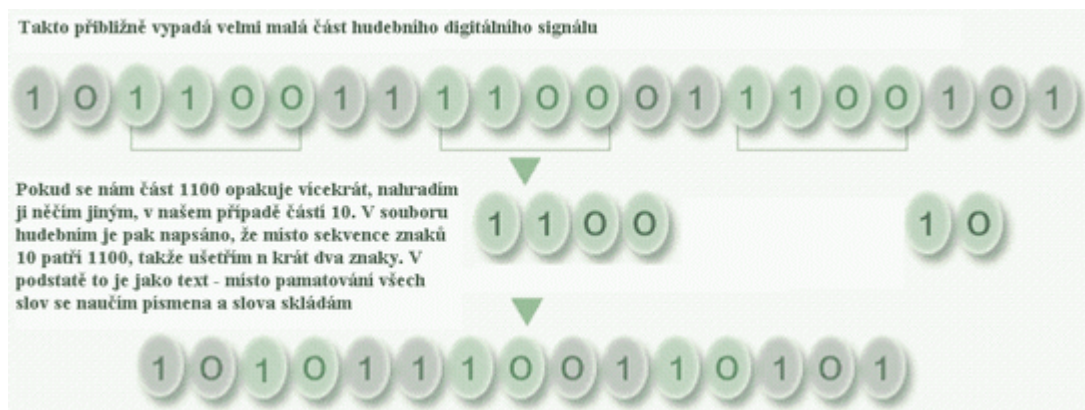
#### 3.3.1 Rozdíl mezi analogovým a digitálním záznamem zvuku

Rozdíl mezi analogovým a digitálním signálem se dá dobře vysvětlit na následujícím obrázku. Z něj je patrné, že digitální signál je hrubší než signál analogový. Tato „hrubost“ je dána především vzorkovací frekvencí a kvantováním. Čím vyšší vzorkovací frekvence a počet kvantizačních úrovní je zvoleno, tím více se digitální signál přibližuje analogovému.



Obrázek 7 - Digitální/analogový signál

Zdroj: [houmrovo.gvn.cz/kvinta/data/prezentace/zvuk/prezentace.pptx](http://houmrovo.gvn.cz/kvinta/data/prezentace/zvuk/prezentace.pptx), 3.1.2011



Obrázek 8 - Zakódování digitálního signálu

Zdroj: <http://mag.repro.cz/view.php?cisloclanku=2004011704> , 12.10.2010

### 3.3.1.1 Výhody digitálního audia

Kvalita digitálního zvuku je nezávislá na hlavě (snímací, záznamové) a médiu. Audio frekvenční odezva, linearita a šum jsou určeny pouze kvalitou provedeného konverzního procesu. Výjimečný dynamický rozsah a linearita jsou snadno dosažitelné kombinované s osvobozením od modulace hluku a přeslechů.

Digitální záznam, jak již bylo řečeno, není nic víc než série čísel a proto může být mnohokrát kopírován bez jakékoliv degradace kvality. Z předchozí věty tak vyplývá, že životnost digitálního záznamu je závislá pouze na životnosti média, na kterém je tento záznam právě uložen.

Používání technik pro opravu chyb eliminuje efekt odpadávání. V zákaznických produktech mohou být tyto techniky využívány jako výhody pro usnadnění požadavků na manipulaci.

Použití korekce časové základny při znovupřehrávání zamezuje kolísání frekvence a může být dále použita pro synchronizaci více strojů.

Používání digitálního záznamu a korekce chyb dovoluje, aby odstup signálu od šumu nebyl nijak velký. Při záznamu na pásek tak stopa může být úzká a proto je dosaženo úspory místa na pásce i přes větší šířku pásma.

Je možné sestavit mimořádně přesné a stabilní filtry a ekvalizéry s vlastními fázemi linearity. [34]

V moderní době se kromě těchto spíše technických výhod vyskytly i ostatní vyplývající z moderních trendů a neustálého pokroku. Mezi tyto výhody bezesporu patří možnost distribuce digitálních formátů přes internet, jejich snadná přenositelnost, možnost mít svoji oblíbenou hudbu kdykoli u sebe například v mobilním telefonu nebo nějakém audio či multimediálním kapesním přehrávači.

### 3.3.2 Postup digitalizace (vzorkovací frekvence, kódování, kanály...)

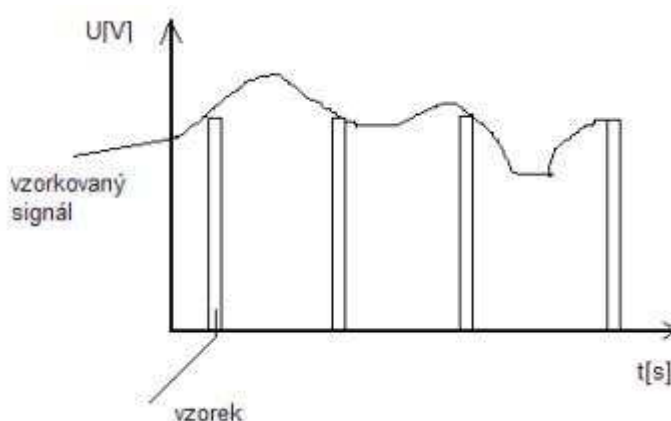
V tomto procesu jde o převedení spojitého analogového signálu na signál digitální (číslicová forma). Pro převod analogového signálu na digitální se využívá tzv. A/D převodník.

Celý proces se dá obecně shrnout do 3 na sebe navazujících kroků:

- 1) Vzorkování
- 2) Kvantování
- 3) Kódování

Jednotlivé fáze digitalizace budou dále rozebrány v následujících podkapitolách.

#### 3.3.2.1 Vzorkování

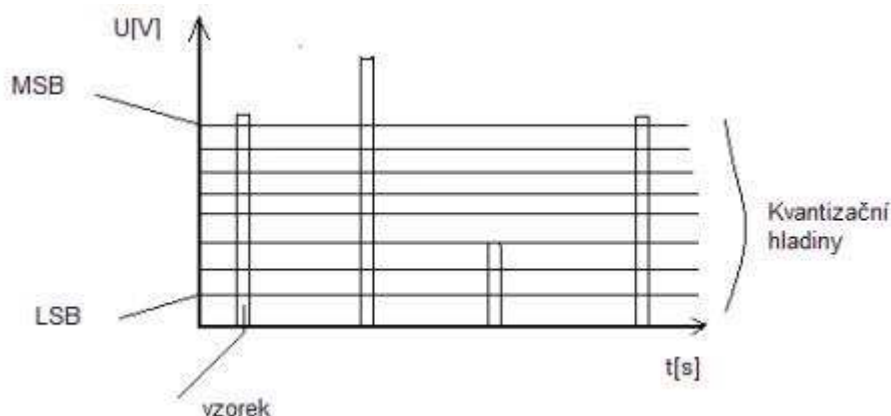


Obrázek 9 - Vzorkování

Zdroj:[http://www.jenseneboj.com/pracestudentu/data/Komprese%20multim%EDi%ED.htm#\\_Toc131170347](http://www.jenseneboj.com/pracestudentu/data/Komprese%20multim%EDi%ED.htm#_Toc131170347), 10.2.2011

Na úplném začátku procesu převodu analogového signálu do digitální podoby stojí vzorkování. V této fázi jsou z analogového signálu odebírány vzorky. Tyto vzorky nejsou nic jiného než informace o aktuální hodnotě amplitudy v daném okamžiku, která je vyjádřena určitým napětím. Pro vzorkování se používá minimální frekvence, jejíž výše je stanovena tak, aby nedocházelo k aliasingu (zkreslení signálu vlivem podvzorkování). Velikost této frekvence je podle Nyquistova kritéria dvojnásobkem maximální frekvence vzorkovaného signálu.

### 3.3.2.2 Kvantování

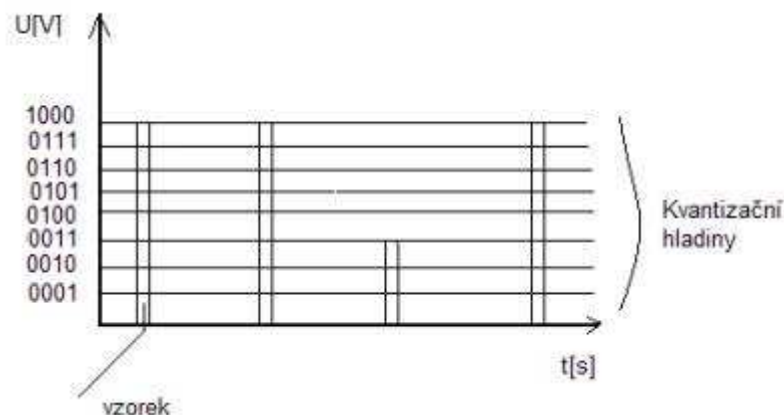


Obrázek 10 - Kvantování

Zdroj:[http://www.jenseneboj.com/pracestudentu/data/Kompresse%20multim%E9di%ED.htm#\\_Toc131170347](http://www.jenseneboj.com/pracestudentu/data/Kompresse%20multim%E9di%ED.htm#_Toc131170347), 10.2.2011

V této fázi dochází ke snižování počtu stavů z nekonečného množství na přesně definovaný počet. K tomuto snižování dochází přiřazováním kvantizačních hladin, které jdou jednoznačně vyjádřit binárním kódem, ke vzorkům signálů. Přiřazování ke kvantizačním úrovním se provádí na základě rozhodovacích úrovní, které se nachází mezi jednotlivými kvantizačními hladinami. V případě, že vzorek přesáhne rozhodovací hladinu, je přiřazen k vyšší kvantizační hladině. Pokud vzorek rozhodovací hladinu nedosáhne, je přiřazen k nižší kvantizační hladině. Počet kvantizačních hladin je dán výsledným počtem bitů v kódovém slově podle vztahu  $2^n$  (kde  $n$  je počet bitů kódového slova).

### 3.3.2.3 Kódování



Obrázek 11 - Kódování

Zdroj: [http://www.jenseneboj.com/pracestudentu/data/Komprese%20multim%ED9di%ED.htm#\\_Toc131170347](http://www.jenseneboj.com/pracestudentu/data/Komprese%20multim%ED9di%ED.htm#_Toc131170347), 10.2.2011

V závěrečné fázi celého procesu je každému nakvantovanému vzorku přiděleno vyjádření v binárním kódu. Po dokončení této fáze je proces digitalizace kompletní a je k dispozici hotový digitální signál. [13]

### 3.3.3 Kvalita zvuku a objem dat

Objem dat je stále velmi důležitým faktorem pro ukládání jakýchkoli souborů včetně audio souborů. Proto je například možné na jedno CD uložit v jiném formátu, než je klasický formát CD Audio, mnohonásobně více audio souborů než v tomto formátu. Ve formátu CD Audio (.cda) je možné na běžné CD umístit pouze 80 minut záznamu.

Právě kvůli tomuto byly vynalezeny ztrátové komprese audia. Ve formátu .cda jsou zaznamenávány i zvuky, které lidské ucho běžně neslyší a tudíž jsou tzv. „redundantní“. Pomocí ztrátových kompresí je možné tyto zvuky (frekvence) odfiltrovat a použít pouze frekvence, které jsou běžně slyšitelné a tím dosáhnout vysokého kompresního poměru například až 12:1. Tím je také dosaženo úspory datového prostoru.

### Důležité ukazatele, které ovlivňují kvalitu zvuku

**Bitová hloubka, velikost vzorku nebo rozlišení** – rozdílné podmínky pro vyjádření počtu bitů reprezentujících určitý vzorek audio signálu. Běžně používané bitové hloubky jsou 8, 16 a v některých profesionálních případech i 24 bitů. Čím více bitů, tím více „vzorků“

může být vyjádřeno a tím věrněji může digitální audio soubor reprezentovat původní analogový audio signál.

**Přenosová rychlost** – množství dat, které musí být přeneseno každou sekundu, aby bylo možné reprodukovat audio signál

**Decibel (db)** – logaritmická stupnice, která měří hlasitost zvuku. Když jsou dva zvuky od sebe 10 db, znamená to, že jeden zvuk je desetkrát hlasitější než druhý.

**Frekvence** – cykly za sekundu, také známé jako Hertz, vyjádření míry změny zvukové vlny. Čím vyšší počet cyklů za sekundu, tím vyšší frekvence tónu pro lidské ucho.

**Hertz (Hz)** – cykly za sekundu, jedná se o způsob vyjádření frekvence audia. Jeden Hertz je jeden kompletní cyklus zvukové vlny za sekundu. Frekvence se dá vyjádřit i v kilohertzích (khz, tisíc cyklů za sekundu)

**Vzorek** – samostatný snímek zvukové vlnové křivky v daném okamžiku.

**Vzorkovací frekvence** – počet vzorků za sekundu, které jsou vybrány pro sestavení digitálního audio souboru (vyjádřena v khz). Čím vyšší vzorkovací frekvence, tím vyšší kvalita zvuku.

**Vlnová délka** – délka jednoho cyklu audio signálu nebo vlnové křivky. Jedná se o inverzi k frekvenci. Delší vlnová délka reprezentuje nižší frekvence, kratší vlnová délka vyšší frekvence. [10]

### 3.3.4 Komprese audio dat

#### 3.3.4.1 PCM

Zkratka anglických slov Pulse Code Modulation (PCM) označuje modulační metodu pro převod analogového signálu na digitální signál. Tuto metody vytvořil v roce 1937 britský vědec Alec Reeves. PCM nebyla komerčně využívána až do 50. let 20. století, kdy byl zpopularizován vynález tranzistoru.

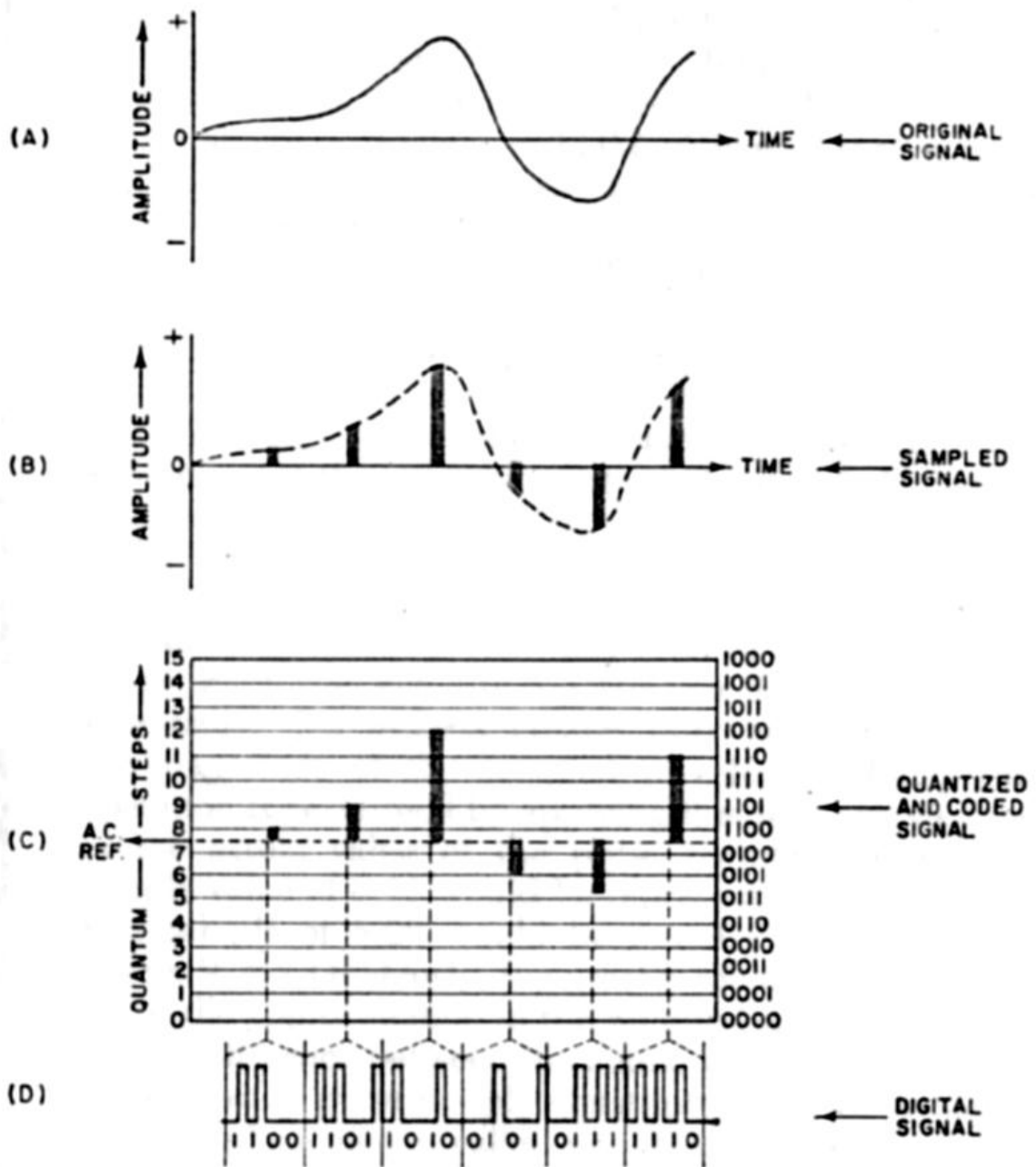
Principem této modulace je pravidelné odečítání hodnot signálu pomocí A/D převodníku a jejich následný záznam do binární podoby. Důležitými parametry jsou vzorkovací frekvence a jemnost rozlišení jednotlivých hodnot (úroveň kvantizace). Velikost vzorkovací frekvence je například u běžné telefonní linky 8 khz přes vzorkovací frekvenci audio CD, která činí 44,1 khz, až po ještě vyšší frekvence, které používají profesionální záznamová zařízení. D/A převodník se používá pro opačný převod, tedy digitálního zvuku na analogový.

Pokud je použita příliš nízká vzorkovací frekvence, nestačí zaznamenané hodnoty k věrné rekonstrukci signálu a tím dochází ke zhoršení kvality. Podle Nyquistovi vzorkovací věty (Harry Theodor Nyquist), má být vzorkovací frekvence minimálně dvojnásobná vzhledem k frekvenci zaznamenaného signálu. Když tato podmínka není splněna, vzniká jev zvaný aliasing (zfalšování frekvence).

Rozlišením hodnot (kvantováním) je dána přesnost hodnot v jednotlivých bodech. Obvykle je použito 8 nebo 16 bitů (256 nebo 65536 možných hodnot). Při každém kvantování je do signálu zaneseno určité množství kvantizačního šumu. Úroveň tohoto šumu s rostoucím rozlišením klesá. Pokud je použito nízké rozlišení, je kvantizační šum slyšitelný. [28]

Postup PCM je graficky znázorněn na následujícím obrázku:





**Fig. 2. Basic operating steps in pulse-code modulated transmission. (A) The original speech signal is (B) sampled to produce pulses corresponding to instantaneous voltage amplitude of audio waveform which is then (C) quantized into incremental levels and (D) encoded into equivalent digital signal.**

Obrázek 12 - Princip PCM

Zdroj: <http://homepage.mac.com/oldtownman/recording/images2/pcm01.jpg>

### **3.3.4.2 MDCT**

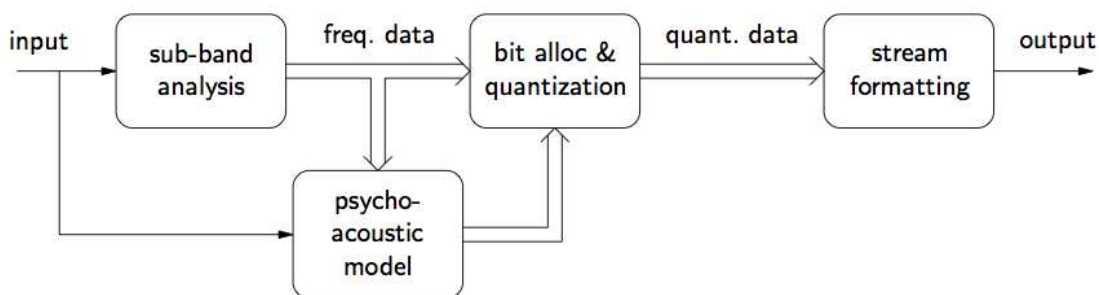
Modifikovaná diskretní kosinová transformace (MDCT) je příbuzná transformace k Fourierově transformaci založená na typu IV diskretní kosinové transformace (DCT-IV). Je vytvořena pro provedení na nepřetržitých blocích velkých datových sad, kde jsou další bloky překrývány tak, že druhá polovina jednoho bloku se shoduje s první polovinou následujícího bloku. Toto překrývání kromě energie stlačení DCT dělá MDCT atraktivní speciálně pro aplikace kompresí signálu, jelikož toto pomáhá se vyvarovat nesrovnalostem vyplývajícím z hranic bloků. V důsledku těchto výhod je MDCT dnes používána ve většině moderních ztrátových audio formátech jako je MP3, WMA, OGG a AAC, které jsou popsány v dalších kapitolách této práce.

MDCT byla navržena Princenem, Johnsonem a Bradleyem v roce 1987 podle dřívější práce Princena a Bradleyho z roku 1986, kdy se snažili rozvinout základní princip MDCT, princip rušení time-domain aliasingu (TDAC).

V případě MP3 není aplikována MDCT přímo ale spíše na výstup 32 pásmové kvadraturní banky filtrů (PQF). Taková kombinace s bankou filtrů se nazývá hybridní filtr nebo podpásmové MDCT. Podobně je tomu například u formátu AAC. [15]

### **3.3.4.3 SBC**

SBC je kódovací systém pro zvuk, který je zvláště navržen pro přenos audio signálu při středně velkém datovém toku s nízkou výpočetní složitostí. Funguje na principu podpásmového kódování, které rozděluje audio signál do několika frekvenčních pásem. Za pomoci tohoto rozdělení je možné jednotlivá pásma kódovat samostatně. Jednoduchá psychoakustická jednotka přerozdělí bitový tok mezi jednotlivá pásma podle jeho energie. Jde vlastně o maskování s využitím nedokonalosti lidského sluchu, který slyší široké spektrum frekvencí, ale dostatečně silný signál dokáže zamaskovat ty ostatní. Princip maskování je základem naprosté většiny moderních ztrátových kompresí, jako je například MP3. Takto upravený signál jednotlivých pásem je dále kódován v tomto případě metodou APCM. SBC podle specifikace používá 4 nebo 8 pásem, adaptivní algoritmus rozložení bitového toku a adaptivní PCM kvantizéry. Základní princip SBC je graficky znázorněn na následujícím obrázku. [24]



Obrázek 13 - Princip SBC

Zdroj: <http://cnx.org/content/m32106/latest/img015.png>, 11.2.2010

### 3.3.5 Způsoby záznamu

Záznamová technika prošla jako každá technika velkým vývojem. V této kapitole jsou popsána různá záznamová média a způsoby, jakými se na tato média audio signál zaznamenává. V následující tabulce jsou zmíněny některé starší formáty záznamu audio signálu. [17]

Rok	Záznamové médium	Popis
1877 – 1929	Voskové válce	2 - 4 minuty záznamu, vosk nebo směs vosku
1929 – 1960	Acetátové disky	vínyl na papírovém, skleněném nebo kovovém podkladu
1945 – 1955	Nahrávací drát	stočený drát, 15 - 30 minut, pouze jednostranný
1945 – současnost	Kotouč s páskou	známý cívkový magnetofon
1965 – současnost	Kompaktní kazeta	páska v uzavřeném obalu
1977 – současnost	Mini/mikrokazeta	malé 2 - 4 cm kazety
1984 – současnost	CD, MP3 a ostatní dig. záz.	signál nahrán přímo v dig. formě na různá média

Tabulka 1 - Záznamové formáty

Zdroj: <http://ucblibraries.colorado.edu/systems/digitalinitiatives/docs/digital-audio-bp.pdf>, 11.3.2011

#### 3.3.5.1 CD

Kompaktní disk neboli CD přinesla na trh společnost Philips z Nizozemska a firma Sony. V roce 1980 navrhly společně tento záznamový formát a postupně ho představily v Evropě a Japonsku (1982) a Spojených státech amerických (1983). Původní představený formát nesl název CD-Audio (CD-A, CD-DA nebo Red Book CD, jak je specifikován ve

standardu 908 Mezinárodní komise pro elektrotechniku (IEC). Další CD formáty byly postupně představovány světu: CD-ROM (1984), CD-i (1986), CD-WO (CD-R, 1988), Video-CD (1994) a CD-RW (1996).

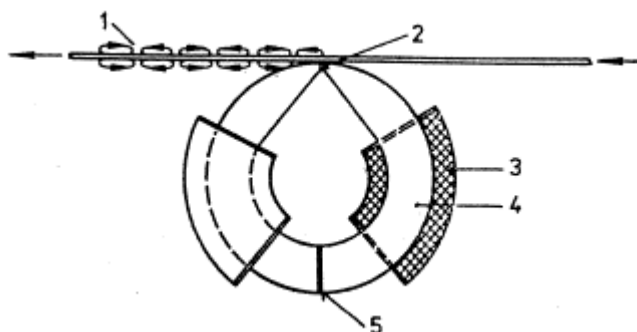
#### 3.3.5.1.1 Fyzické rozměry a specifikace

Kompaktní disk je vlastně kruhová placka o průměru 120 mm s otvorem uprostřed o průměru 15 mm a tloušťkou 1,2 mm. 7,5 mm od okraje otvoru začíná upínací oblast, která dosahuje do vzdálenosti až 23 mm od otvoru. Tato zčásti průhledná a zčásti pokovovaná oblast může obsahovat viditelný nápis s razítkem výrobce. Poté následuje 2 mm tzv. uváděcí oblast. Tato oblast obsahuje kontrolní údaje využívané přehrávači. Dostáváme se tedy již na poloměr 25 mm. Následuje 33 až 33,5 mm široká oblast tzv. programová oblast, která rozšíří průměr z 25 na maximálně 58 mm. Předposlední oblastí na disku je tzv. vyváděcí oblast a jako poslední je 1 mm oblast, která je nevyužita a pouze tak rozšiřuje okraj nosiče.

U komerčně lisovaných CD-A disků jsou šachty lisovány vstříknutím polykarbonátového substrátu s vysokým indexem lomu (1,55) do horního plastového povrchu. Tato datová vrstva je pak potažena 50 až 100 nm pokovenou vrstvou poskytující odrazivost, dále je přidán 1000 až 3000 nm plast nebo lak, který působí jako ochranná vrstva, a konečně je natištěn cca 5000 nm popisek. Šachty se vytvoří jako vypukliny od paprsku laseru. Šachty jsou cca 600 nm široké a typický disk může mít až 2 miliardy těchto „rýh“. Hloubka jednotlivých šachet je rovna přibližně čtvrtině vlnové délky světla v podkladu, tedy cca  $500 / 4 = 125$  nm (pohybuje se od cca 110 do 150 nm). Délky šachet (rýh) se liší od minima 830 nm - 970 nm (což představuje 3krát minimální dobu kanálového bitu) do maximální výše 3000 - 3600 nm (což odpovídá zhruba jedenáctinásobku), skutečná délka závisí na dráze lineární rychlosti, která se může pohybovat v intervalu 1,2 až 1,4 m / s. Šachty jsou umístěny podélně vedle sebe ve spirálovitém vzoru. Začínají na vnitřní straně průměru disku a postupně směřují po spirálovité dráze ven. Celková dráha těchto spirálovitých drah je zhruba 5800m a rozkládají se na 35 – 35,5 mm disku. K chybám jsou náchylnější stopy blíže k vnějšímu okraji disku a z důvodu, že ne všechny nahrávky zabírají celý disk, byl tento systém záznamu navržen tak, aby vždy začínal u vnitřního okraje disku. Sousední stopy jsou od sebe vzdáleny 1600 nm odděleně od středu ke středu. [16]

### 3.3.5.2 Magnetický pásek

Zmagnetováním magneticky tvrdého materiálu s velkou remanencí  $B_r$  před štěrbinou záznamové hlavy vzniká magnetický záznam zvuku. Nejčastěji je materiálem použitým pro záznam magnetofonový pásek, který je tvořen magneticky neaktivním a mechanicky dostatečně odolným a trvanlivým plastovým nosičem, na jehož povrchu je nanesena vrstva z magneticky tvrdých feritů. Pohonné zařízení táhne pásek s konstantní rychlostí, která může být od 2,4 cm/s přes 4,75 cm/s, 9,55 cm/s, 19,05 cm/s až po 38,1 cm/s kolem štěrbin záznamové hlavy viz. následující obrázek. [27]



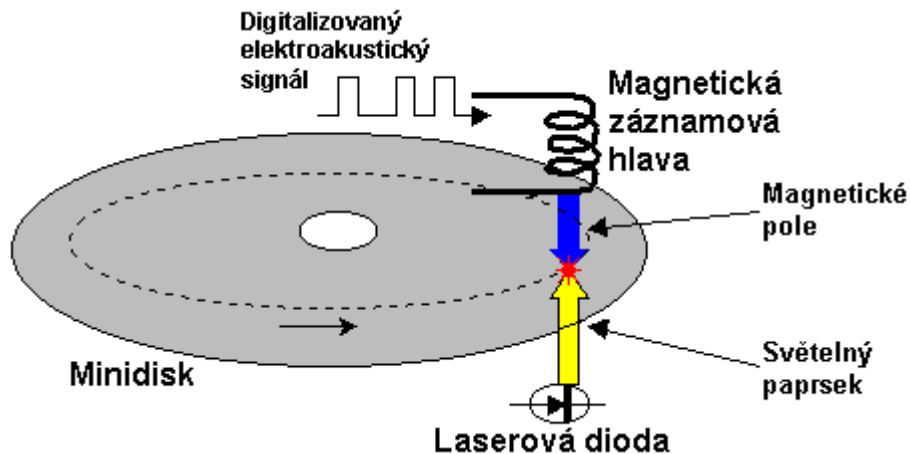
Obrázek 14 - Záznamová hlava magnetofonu

Zdroj: <http://slaboproud.sweb.cz/elt2/stranky1/elt042.htm>, 12.2.2010

Popis obrázku: magnetický tok prvků pásku (1), aktivní štěrbin (2), cívka (3), jádro hlavy (4), zadní štěrbin (5)

### 3.3.5.3 MiniDisc

MiniDisc využívá princip magneticko-optického záznamu digitalizovaných zvukových dat. Záznamové médium je polykarbonátový rotující kotouč, který má průměr 6,4 cm a je opatřen aktivní vrstvou se speciálními ferromagnetickými a optickými vlastnostmi. Povrch MD je třívrstvý. První vrstva (záznamová) je ze slitiny železa, kobaltu, terbia a chrómu. Druhá vrstva (reflexní) je vyrobena ze slitiny hliníku a titanu a třetí vrstvou je silikonová ochranná vrstva. Přesné vycentrování a snížení mechanického namáhání kotouče v místě nasazení na náhonovou hřídel přístroje zajišťuje kovová středovka nalisovaná na středovém otvoru minidisku. Vnější rozměry celého minidisku jsou 7 x 6,75 x 0,5 cm (v x š x h).



Obrázek 15 - Princip magneto-optického záznamu digitalizovaných dat na minidisk

Zdroj: <http://astronuklfyzika.cz/Minidisk.htm>

Na obrázku výše je znázorněn princip záznamu na minidisk. Optická a magnetická hlava jsou proti sobě a mezi nimi se pohybuje kotouč minidisku. Polovodičový laser vyzařuje při zaznamenávání nemodulovaný světelný paprsek o výkonu kolem 5 mW. Při dopadu paprsku na povrch disku je zahřáno místo dopadu aktivní vrstvy na tzv. Curieovu teplotu. Curieova teplota je taková teplota, při které daný materiál ztrácí feromagnetické vlastnosti a stává se z něj materiál paramagnetický (u minidisku je tato teplota zhruba 180 °C). Když paprsek projde je teplota okamžitě snížena pod Curieův bod a materiál se opět stává feromagnetickým. Pokud je do magnetické hlavy, která je umístěna naproti z druhé strany disku, přiveden signál, zůstane díky magnetickému poli v aktivní vrstvě stopa tohoto signálu jako změna magnetické orientace. Tato změna orientace vede v důsledku Faradayova jevu k změně polarizace odraženého světelného paprsku při přehrávání. [32]

### 3.4 Digitální audio formáty

Pokud se podíváme na oblíbenost digitálních audio formátů mezi uživateli P2P sítí, je podle článku zveřejněném na serveru [www.zdnet.com](http://www.zdnet.com) 65% všech zvukových souborů na základě objemu přenosu v MP3 formátu, ale překvapujících 12,3% je v open-source formát OGG souborů (téměř všechny jsou výlučně přenášeny na síti BitTorrent, zejména v Asii). [37]

### 3.4.1 Ztrátové

Ztrátové komprese doslovně znamená zmenšení objemu dat za pomoci odstranění méně potřebných dat (informací). Po provedení ztrátové komprese již není nikdy možné získat zpět původní data.

	Výrobce	Rok vzniku	Patentováno	Použitá komprese	Přenosová rychlost
AAC	ISO/IEC MPEG audio comitee	1997	Ano	MDCT + podpásmové kódování	8 - 529 kbit/s
MPC	Frank Klemm/MDT	1993	Ne	Podpásmové kódování	3 - 1300 kbit/s
MP3	ISO/IEC MPEG audio comitee	1997	Ano	MDCT + podpásmové kódování	8 - 384 kbit/s
OGG	Xiph.org foundation	2000	Ne	MDCT	Variabilní
WMA	Microsoft	1999	Ano	MDCT	4 - 768 kbit/s

Tabulka 2 - Přehled ztrátových formátů

Zdroj: [http://en.wikipedia.org/wiki/Comparison\\_of\\_audio\\_codecs](http://en.wikipedia.org/wiki/Comparison_of_audio_codecs), 12.12.2010

Ztrátové komprese zjednodušeně probíhá tím způsobem, že se originální audio stopa rozdělí na vzorky určitých kmitočtových pásem a odstraní se všechny vzorky z těch pásem, která jsou pro lidské uši neslyšitelná. Tyto hranice se však nedají stanovit přesně pro každého člověka, jelikož každý člověk má pásmo slyšitelnosti posunuto trochu jinak.

Kvalita zvuku a tím i velikost výsledného audio souboru se dá ovlivnit zvolením počtu vzorků vlnových signálu za sekundu. V anglickém jazyce se toto nazývá sampling rate (vzorkovací frekvence). Čím menší je vzorkovací frekvence, tím menší je výsledná velikost audio souboru. Pro účely testu této práce budou soubory komprimovány do různých bit rate (přenosová rychlost). Přenosová rychlost je vlastně počet bitů použitých za jednotku času k vyjádření nepřetržitého média po zdrojovém kódování (datové kompresi).

Kromě pozitivního vlivu (šetření datového prostoru) má však ztrátové komprese audia i svoje nevýhody. Nevýhodou je, že kvalita zvuku se snižuje rychleji než velikost souboru a tím se kvalita audio signálu vzdaluje od CD kvality. Pro většinu populace je ale i toto dostatečná kvalita. Pořád je tedy pro většinu lidí důležitá skutečnost, že velikost takto komprimovaného souboru může odpovídat například pouze 10% velikosti originálního nekomprimovaného audio souboru. [7]

Z této skutečnosti vyplývá, že ztrátové kompresní formáty jsou mnohem populárnější než ty bezztrátové a je jich mnohem větší množství.

V této práci je popsáno a otestováno pět v současné době nejpoužívanějších a nejkvalitnějších audio formátů pro ztrátovou kompresi.

#### 3.4.1.1 *MPEG-1 (layer I, layer II, layer III - .mp3)*

Z této skupiny algoritmů pro ztrátovou kompresi je nejrozšířenější a nejvíce používaný MPEG-1 layer III neboli algoritmus pro kompresi do zvukového formátu .mp3.

MP3 je vlastně zkratka slov Motion Picture Experts Group Layer 3. Jedná se o komprimovaný formát zvuku, který dokáže vytvořit soubory s velikostí 10 až 20krát menší, než jsou audio soubory na kompaktním disku CD. Jedna minuta hudby na CD potřebuje zhruba 10 MB datového prostoru, této velikosti skladby běžných délek ve formátu MP3 ani zdaleka nedosahují.

MP3 je formát, který patří do skupiny formátů pod souhrnným názvem MPEG. Původně byl MPEG používán k vysoké kompresi videa, ale jen málokdo ví, že je kromě tohoto používán i pro kvalitní kompresi zvuku. Samozřejmě, že s postupem času se objevila celá řada konkurentů.

Hlavním rysem MP3 souborů je komprese MPEG Audio Layer 3, která převádí stereo WAV soubor do MP3 souboru s parametry 44 khz, stereo, 16 bit, 128 kb/s. Tento soubor má kvalitu, která se blíží kvalitě Audio CD.

Největšími výhodami MP3 jsou velikost souboru (komprese při „kvalitě CD“ 12:1), skladba o délce 4 minuty zabere v nekomprimovaném WAVu zhruba 40 MB datového místa, v případě MP3 pouze něco kolem 4 MB. I ten nejlepší sluchový analytik má jisté problémy s rozeznáním výsledného komprimovaného zvuku od originálu.

MPEG se dělí na 3 vrstvy podle způsobu zpracování souboru:

**Layer 1** – filtr, který odděluje jednotlivá spektra, pracuje s poměrem hlasitostí a závislost citlivosti frekvencí je aplikována na jednotlivé framy a používá stejnou frekvenci. Psychoakustický model v tomto případě využívá pouze frekvenčního maskování. Jedná se o velmi nedokonalý formát, který je v praxi prakticky nepoužitelný (MP1).



**Layer 2** – v případě tohoto formátu pracuje dělicí filtr s 3 framy najednou a psychoakustický model využívá již jednoduchého temporal maskování. Tento formát byl zjednodušeným předchůdcem formátu MP3. V dnešní době se s ním uživatelé setkají jen zřídka kdy (MP2).

**Layer 3** – v tomto případě se již jedná o formát MP3, který používá mnohem lepší dělicí filtr s měnící se frekvencí. Psychoakustický model již plně využívá temporal maskování. Jedná se o v dnešní době nejrozšířenější formát pro zvukové soubory (MP3).

## Konkurence

Formát MP3 byl průkopníkem v oblasti digitálního audia. Následovaly stejně zaměřené formáty jako AAC, WMA a další. Velký důraz byl u těchto „nástupců“ kladen především na zabezpečení proti pirátskému kopírování skladeb. Samozřejmostí již byla snaha o lepší kvalitu zvuku a větší komprese. [4]

## Vývoj

V roce 1979 bylo poprvé navrženo psychoakustické kódování, které MP3 používá. O návrh tohoto kódování se postaral Manfred R. Schröder z AT & T Bell Labs v Murray Hill a M.A. Krasner.

Předchůdci finální verze MP3 byly kodeky Optimum Coding in the Frequency Domain (OFC) a Perceptual Transform Coding (PXFМ). Společně s block-switching od společnosti Thomson byly OFC a PXFM sloučeny do jednoho kodeku (ASPEC), který vyhrál výběrové řízení společnosti MPEG, ale byl omylem zamítnut. [12]

*„MP3 pak vychází přímo z OFC a PXFM kodeků a reprezentuje výsledek spolupráce Dr. Karlheinz Brandenbura, pracujícího jako postgraduál u AT&T Bell Labs, s Jamesem D. Johnstonem u téže firmy, ve spolupráci s Fraunhoferovým sdružením (Fraunhofer Society) pro integrované obvody v Erlangenu (Německo). Nadále vývoj a implementace formátu MP3 probíhá ve Fraunhoferově institutu v Německu. Pracovní skupina je vskutku mezinárodní: Leon van de Kerkhof (Nizozemsko), Gerhard Stoll (Německo), Leonardo Chiariglione (Itálie), Yves-François Dehery (Francie), Karlheinz Brandenburg (Německo) a James D. Johnston (USA). Tato skupina převzala řešení APSECu, použila řešení filtru Layer2 (MP2), přidala vlastní nápady a výsledkem byl*

*MPEG Layer3, tedy MP3. Hlavním úkolem této práce bylo dosáhnout lepší kvality zvuku při stejném datovém toku oproti MP2.“ [12]*

### ***Slabiny MP3 komprese***

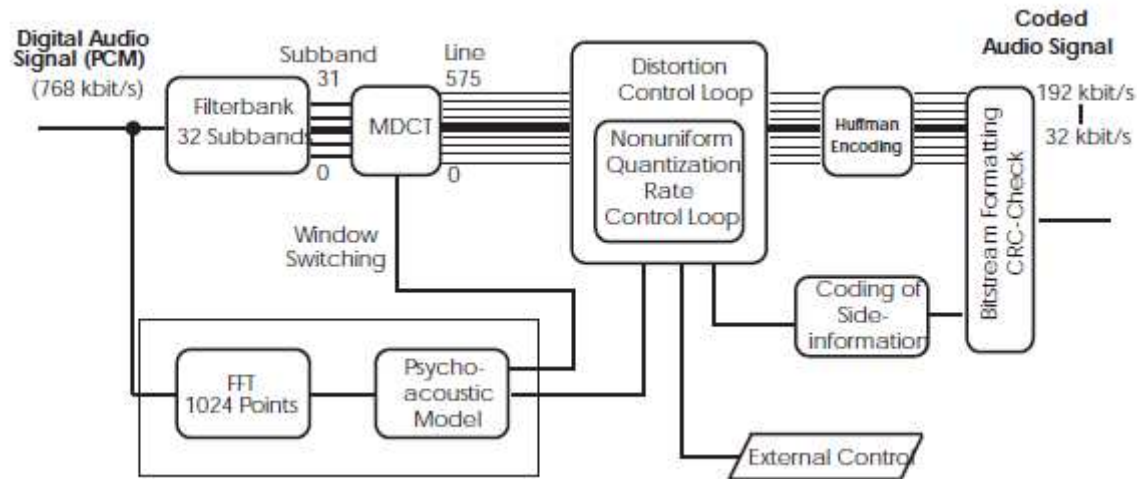
*„Při kompresi mluveného slova jsou výsledky výrazně horší. Popsané maskování a potlačování tónů způsobuje, že u mluveného slova může být ve slově potlačena počáteční nebo koncová slabika. Mohou být také zkracovány pauzy mezi jednotlivými slovy. To působí u mluveného slova značně rušivě. Pro kompresi hlasu jsou vhodné jiné metody např. AMR, G.729, OGG Vorbis, nebo Speex. Výsledná kvalita ovšem závisí na zvoleném datovém toku.“ [12]*

Formát MP3 vzhledem ke svým kvalitám především při nízkých přenosových rychlostech se doporučuje používat například novinářům pro reportážní příspěvky. Jak bylo řečeno v úvodu této kapitoly, formát MP3 je nejčastěji stahovaným formátem audia mezi uživateli P2P sítí. Díky svému silnému prosazování v době začínající éry digitálního audia a silné podpoře výrobců si bude tento formát i v budoucích letech udržovat svoji silnou pozici.

Zjednodušený popis postupu komprese zdrojového zvuku do formátu mp3 je následující:

Po vstupu signálu do kodéru mp3 je zvuk pomocí banky filtrů rozdělen do 32 frekvenčních pásem. Kvantizér a kodér poté v jednotlivých pásmech za pomoci psychoakustického modelu vyhodnocuje změnu prahu slyšitelnosti a vyřazuje ty zvuky, které leží úrovně pod ním. Ostatní signály, jejichž úroveň spadá nad tento práh slyšitelnosti, jsou dále kvantovány nelineárně. Nelineárně z toho důvodu, že pro silnější signály stačí menší přesnost než pro slabší, protože mají větší odolnost proti kvantizačnímu zkreslení. Kromě toho, že jsou signály dále kvantizovány nelineárně jsou také kvantizovány s proměnnou hloubkou (tohoto se využívá zejména pro neměnné tóny). Po provedení těchto výše popsaných kroků se v každém z 32 pásem nachází již jen signály nad prahem slyšitelnosti a každý je zakódován optimálně vzhledem ke svému charakteru. V tomto momentě jsou data statisticky upravována na základě Huffmanova kódování obdobně jako u počítačových souborů metodou bezztrátové komprese. Pro zpětné dekódování je nutné znát podmínky, za kterých tyto úpravy proběhly (celková amplituda,

počet bitů pro kvantizaci atd.). V poslední fázi se data uspořádají do tzv. rámců obsahujících vedle jednotlivých vzorků zkomprimovaného audiosignálu záhlaví, bity pro synchronizaci, zabezpečení proti chybám a informace nutné pro dekódování. [2] [25]



Obrázek 16 - Princip MP3

Zdroj: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.26.5956&rep=rep1&type=pdf>, 20.3.2011

### 3.4.1.2 Ogg Vorbis (.ogg)

Vývoj standardu OGG začal v roce 1993. OGG byl již od začátku vyvíjen jako open source a proto je bez jakýchkoli patentových vzorů. Byl vyvíjen jako substitut za MP3 a WMA a v současné době je možná i stejně známý a populární jako samotný formát MP3. Především algoritmus pro OGG je stále vylepšován, což něco vypovídá o své flexibilitě. Mimo jiné se kvalita zvuku pořád zlepšuje s každým dalším vývojem algoritmu, soubory jsou zpětně kompatibilní a mohou být přehrávány i ve starších přehrávačích.

Stejně jako MP3 OGG nabízí kódování na různých přenosových rychlostech. Pomocí této komprese jsou části písně zakódovány s vyšší kompresí než ostatní, což závisí na zdroji. Většinou je tato komprese doprovázena malým šumem nebo malými přerušováními. OGG je také jedním z mála formátů, který podporuje multikanálovou kompresi. Kanály okolí mohou být teoreticky komprimovány více než dvěma kanály. [1]

OGG je často přehlíženým formátem především u výrobců kapesních přehrávačů, ale svojí kvalitou se může rovnat nebo dokonce převyšovat MP3, jak bude dále porovnáno v testu, který je součástí této práce.

### **Nastavení kodeku a pravděpodobnostní model**

Vorbis je původně výzkumný kodek a jeho současný design odráží touhu dovolit během desetiletí neustále kodek zlepšovat. Z těchto důvodů se konfigurovatelné aspekty nastavení kodeku přiklání spíše k extrému budoucí adaptace.

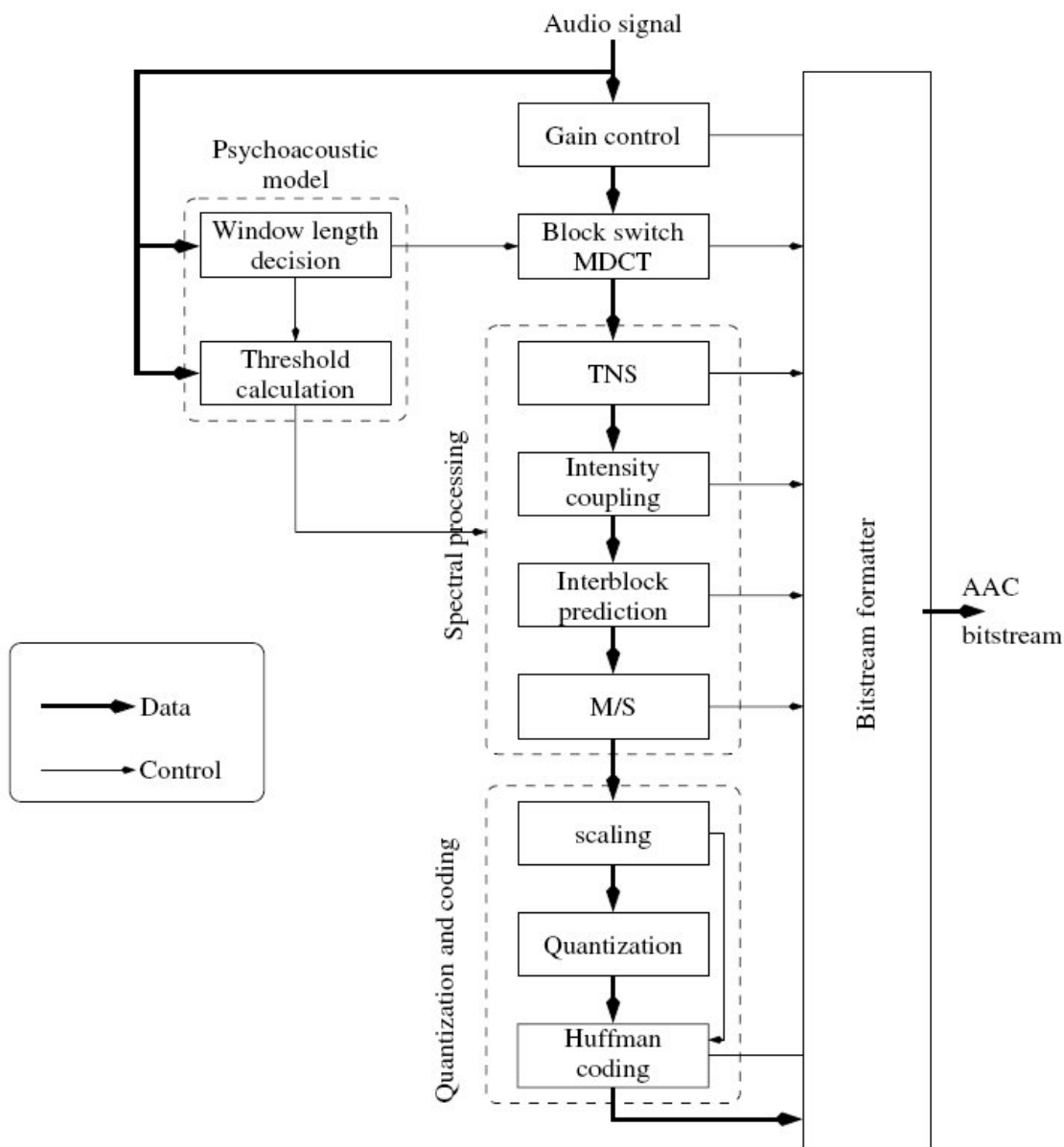
Jedním z nejkontroverznějších rozhodnutí ve Vorbisu (a také nejvíce neobvyklým pro vývojáře Vorbisu) je, že celý pravděpodobnostní model kodeku, Huffmanovy a VQ číselníky jsou zabaleny do hlavičky bitového toku kolem rozsáhlých nastavení parametrů kodeku (většinou se jedná o stovky polí). Toto dělá nemožným (stejně jako by tomu bylo u formátů MPEG) označení jednoho framu vlaječkou v každém audio paketu nebo začít dekódovat na jakémkoli framu v bitovém toku bez předchozího přenesení hlavičky nastavující nastavení kodeku.

Vorbis hlavičky jsou potřebné pro začátek dekódování. Velikost hlavičky je neomezená, ačkoli pro streamování platí pravidlo, které říká, že 4kB nebo méně je doporučená velikost hlavičky (a Xiph.Org 's Vorbis enkodér se tímto doporučením řídí).  
[36]

#### **3.4.1.3 *Advanced audio coding (.aac)***

Po velkém úspěchu MP3 především u internetových uživatelů byla brzy odhalena omezení tohoto nového formátu: při nízkých bit rates MP3 rychle ztrácí kvalitu, i když velikost souboru se pozoruhodně snížila. To bylo období, kdy se objevil formát AAC. Formát AAC byl také vyvinut ve Frahoferově institutu a původním plánem bylo, že se AAC stane nástupcem MP3. Cíle však nebylo dosaženo mimo jiné také proto, že při bit ratu 96 kbps jsou zaznamenány opravdu velké ztráty kvality. Možná právě tato skutečnost je důvodem, proč AAC není tak známým formátem, i když byl již představen v roce 1999. Především ale neexistují prakticky žádné softwarové přehrávače, které by tento formát podporovaly. Prodej licence pro jeho další používání zajišťuje firma Dolby Laboratories.  
[14]

Na následujícím obrázku je zjednodušeně vyobrazen kodér AAC a postupný tok audio signálu tímto prvkem.



Obrázek 17 - Princip AAC

Zdroj: [http://i.cmpnet.com/audiodesignline/2007/10/datacomp\\_fig16.13.jpg](http://i.cmpnet.com/audiodesignline/2007/10/datacomp_fig16.13.jpg), 2.3.2011

#### 3.4.1.4 Wma

Windows Media Audio (WMA) je stále vyvíjeným formátem společnosti Microsoft. Zpočátku byl integrován do PC přehrávače Windows Media Player jako alternativa pro patentovaný formát MP3. Po vydání prvních verzí se však tento formát neseťkal s příliš pozitivními ohlasy z řad veřejnosti.

Tento přístup uživatelů trval až do příchodu verze WMA9, který se rychle stal oblíbeným formátem a celá řada výrobců dnešních přenosných zařízení zařadila podporu tohoto formátu jako standard pro tato zařízení společně s MP3.

Nová verze WMA podporuje i vícekanálovou kompresi pro systémy 5.1 a bezztrátovou kompresi. Další užitečnou funkcí je podpora tzv. MultiBitrate, kdy je soubor umístěn například na serveru a obsahuje větší počet stop s různými bit rate. Uživatel si následně může vybrat například s ohledem na rychlost svého internetového připojení ten, který mu bude nejvíce vyhovovat.

Windows Media Audio (WMA) se postupem času stal opravdu kvalitním formátem pro digitální zvuk, který může směle konkurovat například i formátu AAC. Jediným oficiálním kóděrem pro tento formát je Windows Media Player, jelikož WMA je součástí skupiny formátů Windows Media. Z důvodu uzavřené specifikace formátu WMA nemá tento formát příliš velkou podporu pod jinými operačními systémy, než jsou systémy Windows od společnosti Microsoft.

WMA je praktickým formátem pro použití v internetových rádiích nebo pro poslech on-line hudby. [6] [22]

#### **3.4.1.5 Mpc**

Formát Musepack (MPC) má své základy stejně jako MP3 v formátu MPEG-1 layer II. Podle mnoha poslechových testů je při přenosové rychlosti kolem 200 kbps téměř nerozeznatelný od originálního nekomprimovaného audia. Největším problémem je ale velmi malá podpora výrobců přenosných přehrávačů a tím i prakticky skoro žádné povědomí o tomto formátu mezi širokou laickou veřejností.

V minulých letech to již vypadalo, že se vývoj mpc zastaví. Naštěstí se tak nestalo a formát mpc začíná získávat na popularitě. Bohužel za tímto formátem v současné době nestojí žádná velká firma, která by se postarala o jeho větší prosazení, tak jak je tomu například v případě konkurenčních formátů MP3 (Thomson), WMA (Microsoft) a AAC (Apple). [9]

#### **3.4.2 Bezztrátové**

Je-li pro člověka podstatnější věrnost hudby, kterou poslouchá, a není limitován datovým místem jistě se nebude zajímat o ztrátovou kompresi, ale zaměří se na formáty digitálního audia, které využívají bezztrátovou kompresi. Tato komprese není sice tak

účinná, ale zkomprimovaný soubor je možné pomocí dekomprese vrátit zpět do původní podoby.

Bezztrátová komprese audia se dá nejlépe přirovnat ke známým archivačním technikám ZIP nebo RAR. Jsou odstraněny nadbytečné bity, ale zůstává zaznamenáno, kde tyto bity chybí a díky tomu je možné poté audio signál pomocí dekomprese upravit do původního stavu. Při přehrávání souboru dochází k postupné dekompresi, takže se přehrává neporušená originální zvuková stopa, u které je pouze zmenšena její velikost v paměti. Tímto se dá dosáhnout úspory datového místa až kolem 50%. Většinou ale úspora místa bývá menší. Proto jsou soubory v bezztrátových formátech prakticky nevhodné pro běžné přenosné přehrávače, kde je málo paměťového místa. Avšak například pro uložení své oblíbené sbírky CD na disk PC jsou tyto formáty naopak velmi vhodné. Za použití téměř jakékoliv bezztrátové komprese je možné uložit přibližně 1000 CD na zhruba 300 GB paměti pevného disku PC. [7]

Nejdůležitější formáty využívající bezztrátovou kompresi jsou uvedeny v dalších částech této kapitoly.

	Výrobce	Rok vzniku	Patentováno	Vzorkovací frekvence
WAV	IBM a Microsoft	1998	Ne	1 Hz - 16.777216 Mhz
FLAC	Xiph.org foundation, Josh Coalson	2000	Ne	1 Hz - 655350 Hz
APE	Matthew T. Ashland	2000	Ne	8, 11.025, 12, 16, 22.05, 24, 32, 44.1, 48 khz

Tabulka 3 - Přehled bezztrátových formátů

Zdroj: [http://en.wikipedia.org/wiki/Comparison\\_of\\_audio\\_codecs](http://en.wikipedia.org/wiki/Comparison_of_audio_codecs), 12.12.2010

#### 3.4.2.1 Wav (.wav)

Wav je zkratka pro audio formát Waveform audio file format. Byl vytvořen společnostmi IBM a Microsoft hlavně pro ukládání zvukových dat do PC. Vychází z obecnějšího formátu RIFF. I přes to, že je možné ukládat do WAV souboru zvuk komprimovaně, například pomocí GSM komprese, ADPCM a dalších, většinou se používá nekomprimovaný zvuk v lineárně pulzně kódové modulaci. Stejně je uložen zvuk na Audio CD a je tedy velmi snadné provádět převod mezi těmito formáty.

Jelikož je PCM bezztrátový formát, využívá se formátu WAV hlavně při zpracování zvuku. Z důvodu úspory datového místa se dnes pro ukládání audio záznamů používají především komprimované formáty jako je MP3, WMA, OGG a další, kterými zabývají další kapitoly této práce. Pokud je z nějakého objektivního důvodu bezztrátový formát vyžadován, je lepší z hlediska úspory datového místa použít například FLAC, Monkey's audio (APE) nebo i další formáty. [35]

#### **3.4.2.2 Aiff**

Aiff znamená Audio Interchange File Format a jedná se o společný formát pro ukládání a přenos určitého vzorku zvuku. Formát byl vyvinut společností Apple Computer a jedná se o standardní audio formát pro počítače Macintosh. Je také hojně používán na počítačích Silicon Graphic Incorporated (SGI).

Audio soubory ve formátu AIFF mají příponu .aif nebo .ief. Tento formát je bezztrátový a nepodporuje tedy kompresi dat, z čehož vyplývá, že soubory AIFF mají velkou datovou velikost. Existuje však i formát AIFF-Compressed, který kompresi dat podporuje a to až 6:1. [8]

#### **3.4.2.3 Free lossless audio codec (.flac)**

FLAC je další z řady open-source a rozšířených zvukových kodeků, který je určen výhradně pro bezztrátovou kompresi. Pro snížení datového obsahu souboru využívá flac lineární predikci a celý princip fungování komprese by se dal přirovnat k programům jako je WinRAR nebo WinZip. Komprimovaný soubor je ve většině případů menší o 40 – 50% než originální nekomprimovaný soubor. [29]

#### **3.4.2.4 Monkey's audio (.ape)**

Formát Monkey's audio je v současné době nejpoblárnější formát bezztrátové komprese. Nespornou výhodou tohoto formátu je prakticky stoprocentní shoda s originálním souborem po zvukové stránce. Dalšími výhodami jsou rychlost komprese a dekomprese. Poměr komprese vzhledem k originálnímu .wav souboru je běžně ale pouze 1:2.

Velkou nevýhodou je podpora tohoto formátu pouze v operačním systému Windows a nemožnost streamování audia v tomto formátu přes internet. [11]



### 3.4.3 MIDI

MIDI je formát digitálního audia, který se hodí především pro uživatele, kteří hrají na nějaké klávesové nástroje. Nejedná se vlastně o formát ale o komunikační rozhraní sloužící k sériové komunikaci mezi nástroji, počítači a dalšími přístroji.

Musical Instrument Data Interface (MIDI) byl jako norma zaveden v roce 1983 a drtivá většina profesionálních a poloprofesionálních klávesových nástrojů toto rozhraní podporuje.

MIDI je vlastně jen sled příkazů, které říkají, kdy se má jaký tón zapnout, s jakou intenzitou a na jakou dobu. Zjednodušeně se dá říci, že se jedná o digitální notový zápis. Výsledný zvuk je tedy závislý na přístroji, na kterém je MIDI přehráván. Z tohoto vyplývá velká výhoda MIDI. Záznam je co do datového objemu velmi malý, je možné jej po nahrání snadno upravovat a libovolně měnit zvuk nahrané stopy.

MIDI se tedy využívá zejména tam, kde je potřeba dále s klávesovou stopou kreativně pracovat.

Standardně má soubor MIDI příponu .mid a především v 90. letech minulého století byl toto nejvyužívanější způsob pro ozvučení počítače, her a programů. [31]

## 4 Porovnání digitálních zvukových formátů

V analytické části práce byly porovnány především ztrátové audio formáty při různých přenosových rychlostech. V další části jsou interpretovány výsledky vícekriteriální analýzy variant pro výběr vhodného ztrátového audio formátu pro ukládání audio souborů v domácích podmínkách.

Pro test byla zvolena skladba Elevation od irské skupiny U2. Původní soubor ve formátu .cda byl bezztrátově převeden do formátu .wav, z kterého následnými kompresemi vznikaly soubory ostatních zvukových formátů použité v tomto testu. Byly použity soubory při bitovém toku 256, 128 a 96 kbps.

Audio soubor ve formátu .wav byl převeden pomocí jednotlivých komprimačních algoritmů do nejvíce užívaných ztrátových a bezztrátových audio formátů. Poté byly tyto soubory porovnány na základě porovnání spektrálních analýz (oříznutí frekvencí) jednotlivých souborů.

Pro test byly použity programy Switch sound file converter 2.03 a Audacity 1.3.11 Beta.

### 4.1 Bezztrátové formáty – 256 kbps

U bezztrátových formátů jsou v rámci práce hodnoceny pouze ukazatele doba převodu v jednotkách sekund a velikost výsledného souboru v megabytech.

Formát	Doba převodu	Velikost souboru (MB)
FLAC, 256 Kbps, 44 Khz, stereo	36 sec	24.4
AIFF, 256 Kbps, 44 Khz, stereo	4 sec	38.5
APE	18 sec	23.8

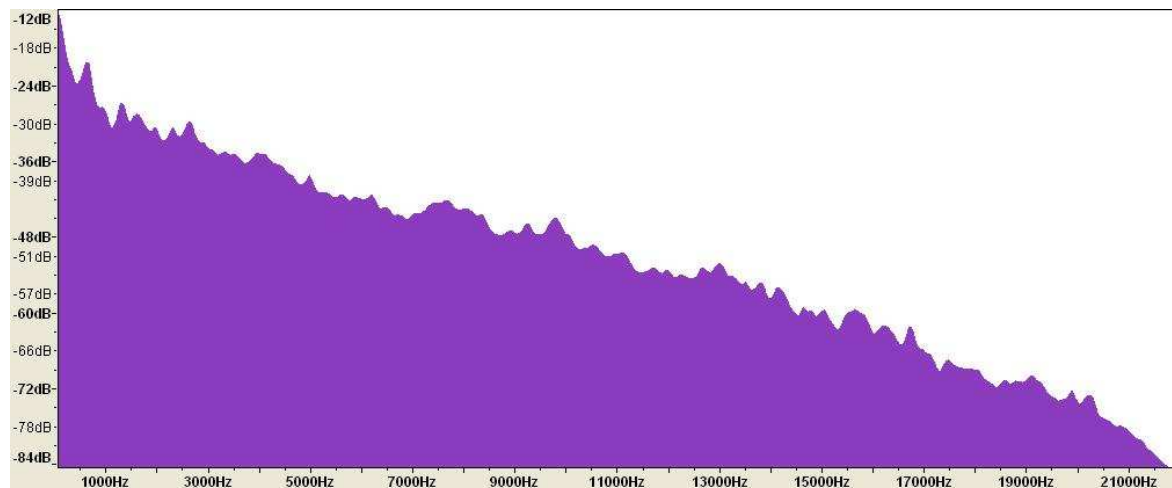
Tabulka 4 - Bezztrátové formáty 256 kbps

Zdroj: Autor

Při zvolení přenosové rychlosti 256 kbps jsou výsledky přehledně popsány v tabulce 3, která předchází tomuto textu. Nejlépe na základě kritéria doba převodu dopadl formát aiff následovaný monkey's audio (ape) a nejdéle se převáděl soubor do formátu flac. Při srovnání druhého kritéria (velikost souboru) nejlepších výsledků dosáhl formát

ape, následovaný formátem flac a až za nimi se umístil soubor ve formátu aiff. Pokud jsou uvažována pouze tato kritéria, je doporučen při přenosové rychlosti 256 kbps formát monkey's audio (ape).

#### 4.1.1 WAV



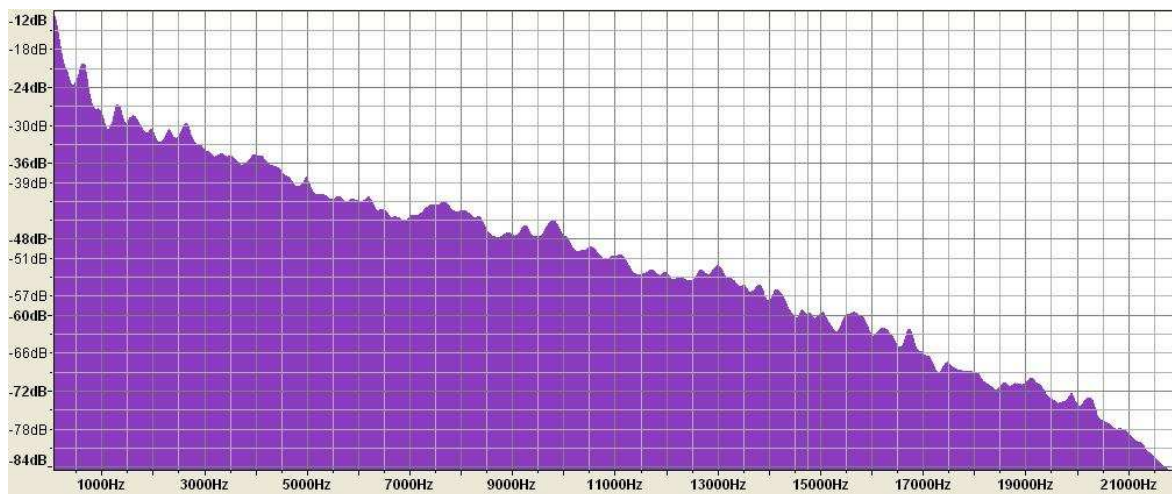
Obrázek 18 - Spektrum WAV 256 kbps

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza WAV souboru je zobrazena na obrázku 18 výše. Je patrné, že nekomprimovaný originální audio souboru obsahuje všechny frekvence, které jsou podle obecné definice slyšitelné pro lidské ucho. Tato skutečnost si ale vybírá svojí daň na velikosti souboru, který zabírá necelých 40 MB datového prostoru.

Poslechový prožitek by měl být v tomto případě nejkvalitnější, ovšem toto je značně individuální. Každý člověk má jiný frekvenční rozsah, který vnímá, a s věkem je postupně lidský sluchový systém méně citlivý na poslech vyšších frekvencí.

#### 4.1.2 FLAC

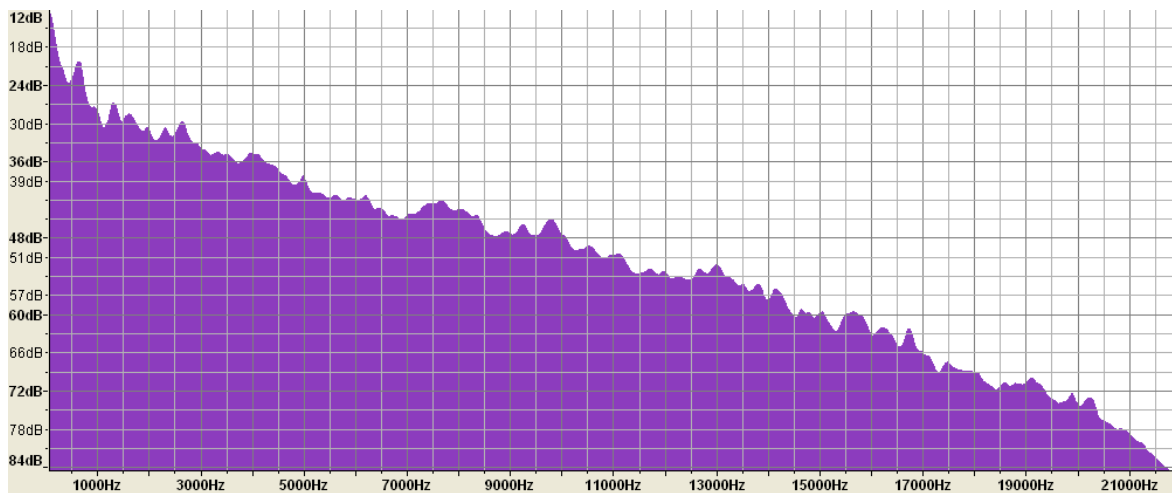


Obrázek 19 - Spektrum FLAC 256 kbps

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza flac souboru je znázorněna na obrázku 19 výše. Spektrum frekvencí je v tomto případě velmi podobné originálnímu wav souboru, což vyplývá i z toho, že se jedná o bezztrátovou kompresi. Při době převodu 36 sekund a výsledné velikosti souboru 24,4 MB jen odborník pozná rozdíl ve výsledné zvukové reprodukci.

#### 4.1.3 APE



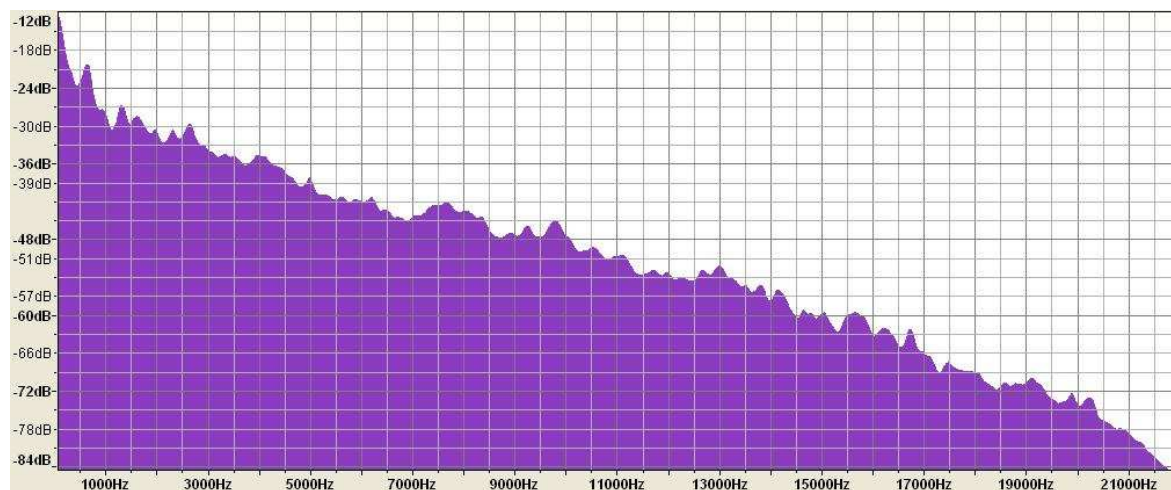
Obrázek 20 - Spektrum APE 256 kbps

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza souboru ve formátu monkey's audio (ape) je znázorněna výše na obrázku 20. Jak již bylo napsáno dříve, tento soubor byl vyhodnocen jako vítěz v tomto testu pro přenosovou rychlost 256 kbps.

Doba převodu 18 sekund a velikost souboru 23,8 MB je ideálním kompromisem mezi všemi zde uvažovanými bezztrátovými formáty při dané přenosové rychlosti.

#### 4.1.4 AIFF



Obrázek 21 - Spektrum AIFF 256 kbps

Zdroj: Vlastní zpracování v programu Audacity

Posledním testovaným bezztrátovým formátem je formát aiff, jehož spektrální analýza je vyobrazena na obrázku 21. Tento formát při přenosové rychlosti 256 kbps disponuje naprosto bezkonkurenční dobou převodu. Velikost 38,5 MB je ale podstatně vyšší než u ostatních konkurentů.

#### 4.2 Ztrátové formáty - 256 kbps

Ztrátové formáty byly na různých přenosových rychlostech hodnoceny podle stejných kritérií jako v případě bezztrátových formátů. Bylo ale přidáno ještě další kritérium, které je patrné ze spektrálních grafů jednotlivých souborů. Jedná se tedy o ořezávání určitých frekvencí, na kterých je ztrátová komprese založena. Hodnoty, při kterých jednotlivé kodeky ořezávají vysoké frekvence, nejsou v tabulce v rámci této kapitoly. V této kapitole se autor zaměřil na optické porovnání těchto spekter. Přesná čísla jsou použita ve vícekritériální analýze variant, která následuje za tímto porovnáním.

Formát	Doba převodu	Velikost souboru (MB)
MP3	55 sec	6.99
WMA 9.1	18 sec	7.02
OGG Vorbis	62 sec	6.68
AAC	65 sec	3.59
MPC	50 sec	5.62

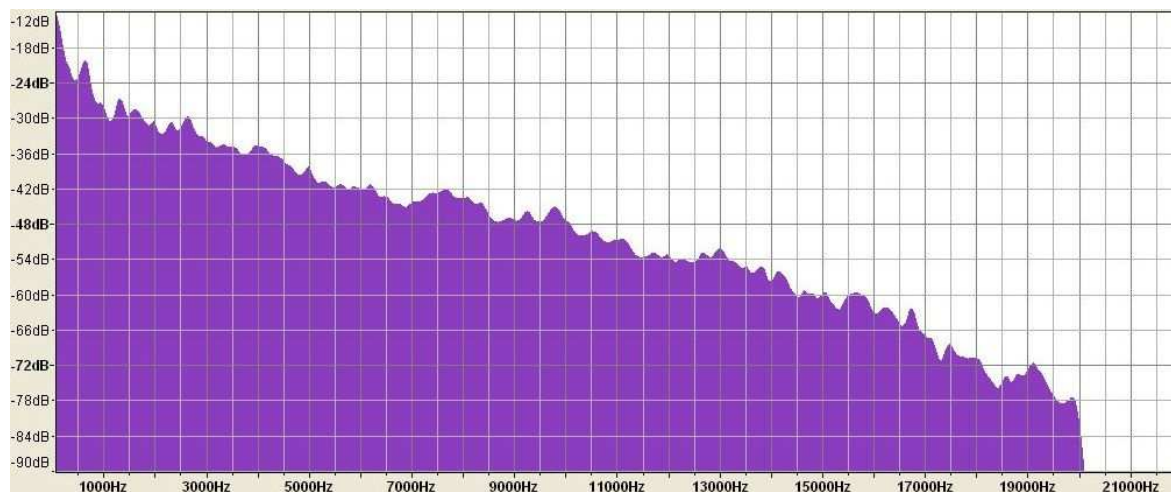
Tabulka 5 - Ztrátové formáty 256 kbps

Zdroj: Autor

Podle předcházející tabulky disponuje nejrychlejším převodem na této úrovni formát WMA, který je následovaný formátem MPC a MP3. Na posledních dvou místech se umístily formáty OGG a AAC. Formát AAC naopak obsadil první místo na základě velikosti souboru, kde dosáhl opravdu skvělý výsledek 3,59 MB. Na druhém místě se v tomto testu umístil MPC následovaný OGG, MP3 a WMA.

Na základě těchto výsledků lze doporučit na této úrovni přenosové rychlosti formát MPC. Dobré výsledky mají i formáty WMA a AAC. Naopak celkově nejhůře v tomto testu dopadly formáty OGG a MP3.

#### 4.2.1 MP3



Obrázek 22 - Spektrum MP3 256 kbps

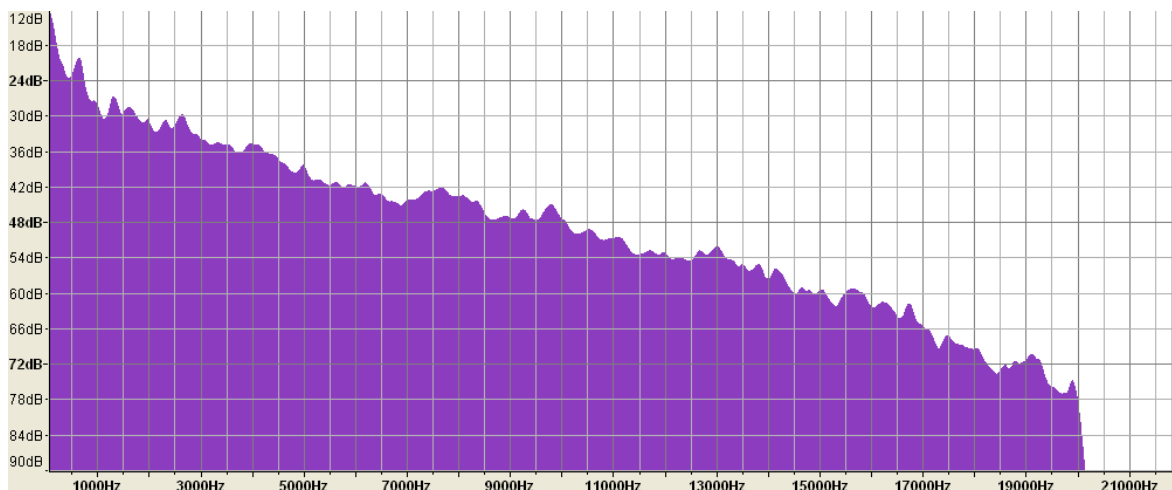
Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza souboru ve formátu MP3 je zobrazena na obrázku 22. Z obrázku je patrné, že již nyní se projevují prvky ztrátové komprese. Originální soubor obsahoval i frekvence přes 21000 Hz. Další věcí, která je z obrázku patrná, je způsob ořezávání

frekvencí tímto kodekem. Kodek eliminuje velmi náhle frekvence kolem 20 khz. Jedná se o prudké oříznutí těchto frekvencí, které se v originálním souboru vyskytovaly.

S dobou převodu 55 sekund a velikostí výsledného souboru skoro 7 MB se tento formát při přenosové rychlosti 256 kbps řadí mezi dva nejhorší.

#### 4.2.2 WMA



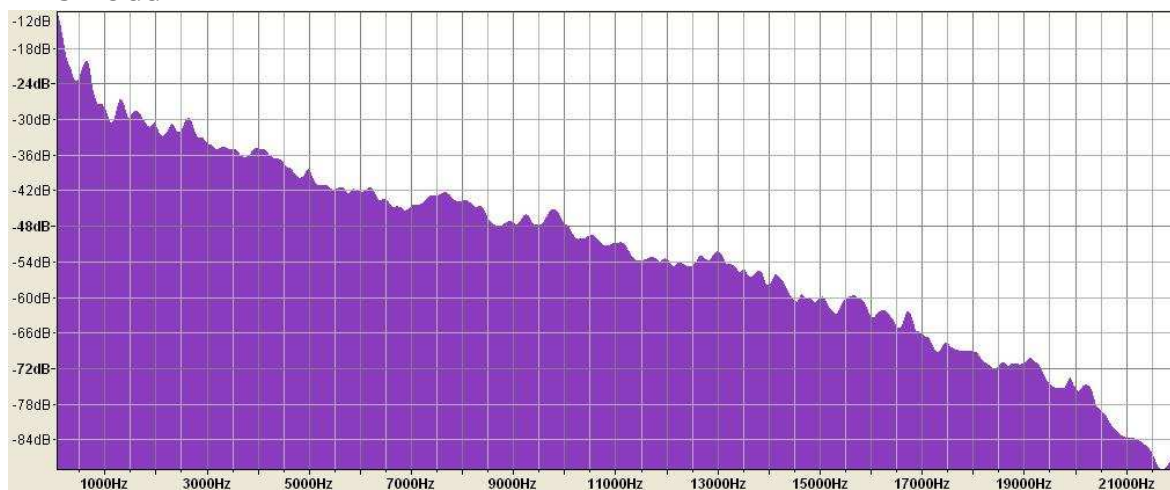
Obrázek 23 - Spektrum WMA 256 kbps

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza souboru ve formátu WMA je zobrazena na obrázku 23. Zobrazení frekvenčního spektra je v tomto případě velmi podobné předchozímu formátu MP3. Kodek také v tomto případě ořezává vyšší frekvence velmi prudce prakticky ve stejných oblastech jako kodek MP3.

WMA v tomto případě disponuje nejrychlejší dobou převodu 18 sec ale na druhou stranu největší velikostí výsledného souboru přes 7 MB.

### 4.2.3 OGG



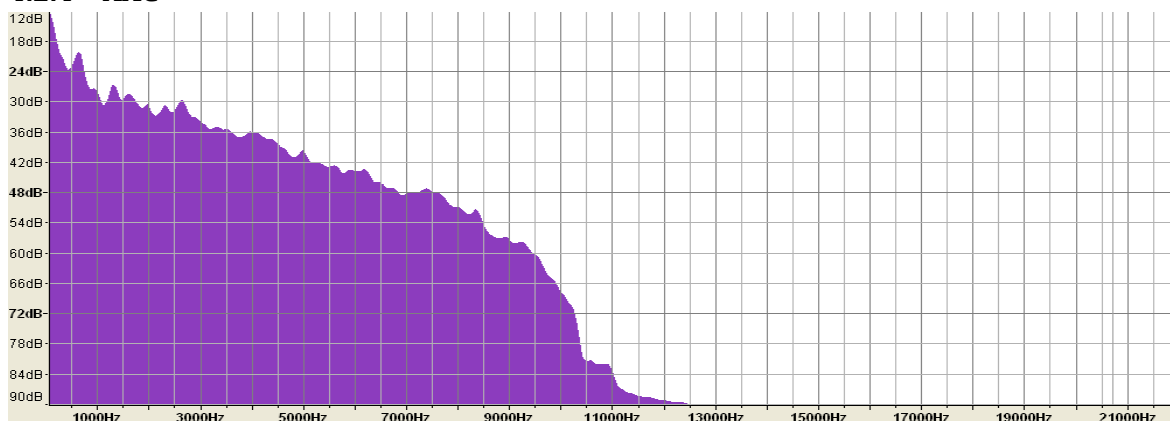
Obrázek 24 - Spektrum OGG nejvyšší kvalita

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza formátu OGG je zobrazena na obrázku 24. Již na první pohled je patrný rozdíl oproti předchozím dvěma formátům. Kodek ogg vorbis neořezává vyšší frekvence tak agresivním způsobem jako například WMA nebo MP3. Je vidět, že v komprimovaném souboru se nepatrně vyskytují i frekvence vyšší než 21 khz, tak jak tomu bylo v případě originálního WAV souboru.

Velkou nevýhodou nicméně je dlouhá doba převodu, která je druhou nejvyšší ze všech srovnávaných ztrátových formátů při této přenosové rychlosti. Velikost výsledného souboru 6,68 MB je nicméně lepší než u předchozích dvou porovnávaných souborů.

### 4.2.4 AAC



Obrázek 25 - Spektrum AAC 256 kbps

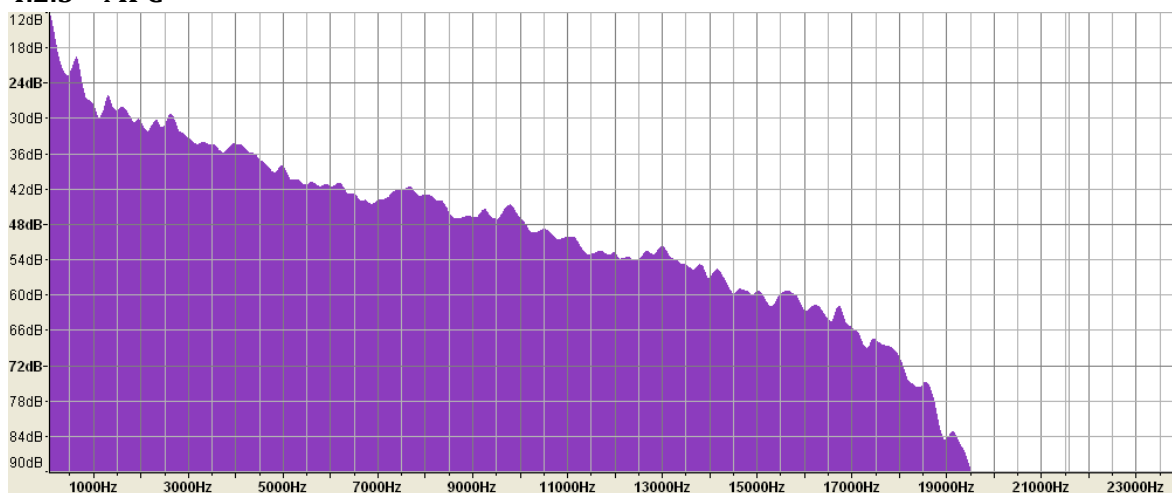
Zdroj: Vlastní zpracování v programu Audacity



Formát AAC se soustředí především na nižší frekvence, jak je patrné z předchozího obrázku 25. Tento kodek začíná s ořezáváním frekvencí nejdříve ze všech. Prakticky již někdy kolem 10 khz se tento formát hodně odlišuje od ostatních. Frekvence nad 13 khz se ve výsledném souboru již vůbec nevyskytují.

Další nevýhodou tohoto formátu je velmi dlouhá doba převodu. Nespornou výhodou je ovšem velikost výsledného souboru, která při dané přenosové rychlosti nemá mezi srovnávanými soubory konkurenci.

#### 4.2.5 MPC



Obrázek 26 - Spektrum MPC 256 kbps

Zdroj: Vlastní zpracování v programu Audacity

Formát MPC, který byl vyhodnocen jako vítěz v tomto testu při přenosové rychlosti 256 kbps, má svoje frekvenční spektrum vykresleno na obrázku 26. Z tohoto obrázku je patrné, že formát MPC neořezává vyšší frekvence tak agresivně jako MP3 nebo WMA ale spíše postupně. Pokud bereme v úvahu skutečnost, že lidský sluch slyší frekvence mezi 20 Hz a 20 khz, formát mpc pokrývá skoro celé toto spektrum.

Při době převodu pod 1 minutu a velikosti výsledného souboru něco málo přes 5,5 MB je tento formát vítězem prvního testu ztrátových formátů při přenosové rychlosti 256 kbps.



Graf 1 - Porovnání ztrátových formátů při 256 kbps

Po vyexportování číselných dat o jednotlivých spektrech byla tato data naimportována do programu Microsoft Excel a byl vytvořen graf 1, v kterém jsou syntetizovány a přehledně porovnány všechny spektrální průběhy jednotlivých ztrátových formátů při přenosové rychlosti 256 kbps.

Je patrné, že na nižších frekvencích jsou frekvenční průběhy dosti podobné kromě formátu AAC, který začíná ořezávat frekvence nejdříve, následovaný MPC, MP3, WMA a jako poslední začíná ořezávat formát OGG.

### 4.3 Bezztrátové formáty – 128 kbps

V následující tabulce jsou zaznamenány výsledky testu bezztrátových formátů při přenosové rychlosti 128 kbps.

Formát	Doba převodu	Velikost souboru (MB)
FLAC	7 sec	24.4
AIFF	2 sec	19.2
APE	13 sec	23.8

Tabulka 6- Bezztrátové formáty 128 kbps

Zdroj: Autor

Při porovnávání bezztrátových formátů při přenosové rychlosti 128 kbps na základě stejných kritérií jako v předchozím případě tentokrát nejlépe dopadl formát AIFF, který při nejrychlejší době převodu dosáhl i největší úspory datového místa. Ostatní dva formáty dopadly prakticky stejně. Celkově jsou mezi všemi ztrátovými formáty nevelké rozdíly.

### 4.4 Ztrátové formáty – 128 kbps

V následující tabulce jsou zaznamenány výsledky testu ztrátových formátů na úrovni přenosové rychlosti 128 kbps.

Formát	Doba převodu	Velikost souboru (MB)
MP3	46 sec	3.49
WMA 9.1	22 sec	3.52
OGG Vorbis	76 sec	4.87
AAC	44 sec	3.46
MPC	46 sec	4.76

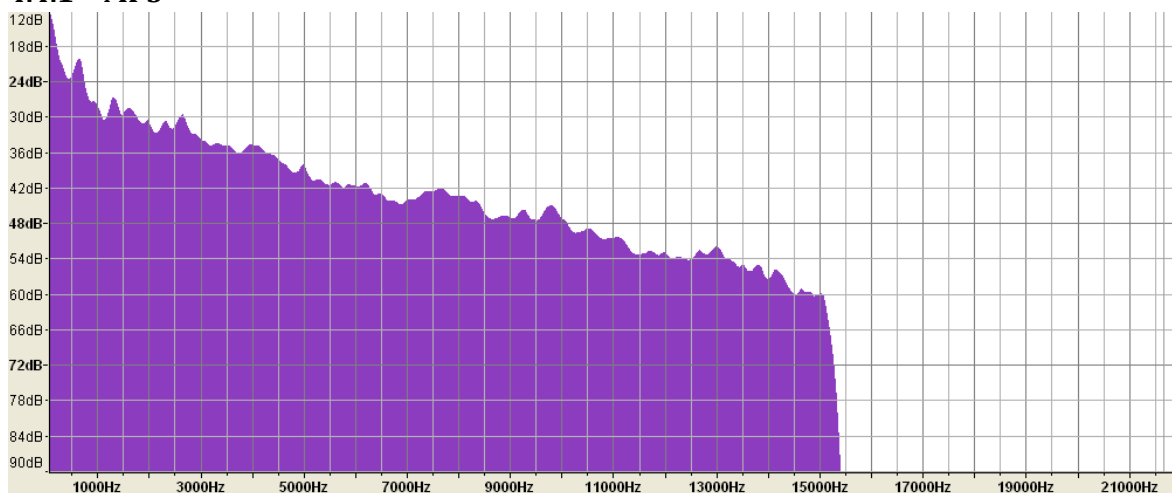
Tabulka 7 - Ztrátové formáty 128 kbps

Zdroj: Autor

Dle tabulky 7 má nejrychlejší kompresi formát WMA 9.1, stejně tomu bylo i na vyšší přenosové rychlosti. Následován je formátem AAC, dále formáty MP3 a MPC a nejhorší je v tomto směru formát OGG. Nejméně datového prostoru zabírá soubor ve formátu AAC, který je následován formátem MP3, WMA, MPC a nakonec i formátem OGG, který v testu na této úrovni pouze na základě těchto dvou kritérií dopadl nejhůře.

Nejlépe hodnoceným formátem na základě dvou kritérií z tabulky 7 je formát AAC těsně následovaný formátem WMA 9.1.

#### 4.4.1 MP3



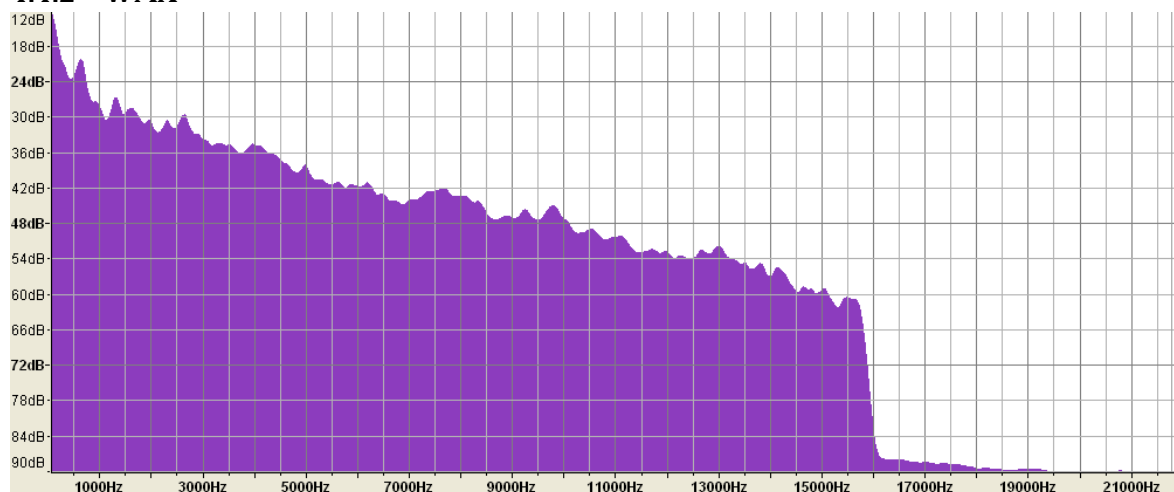
Obrázek 27 - Spektrum MP3 128 kbps

Zdroj: Vlastní zpracování v programu Audacity

Stejně jako tomu bylo s ořezáváním frekvencí v případě MP3 při přenosové rychlosti 256 kbps je tomu stejně i na obrázku 27 při rychlosti 128 kbps. Jediným rozdílem je místo, v kterém kodek začíná tyto vyšší frekvence ořezávat. V předchozím případě to bylo kolem 20 khz, nyní začíná kodek agresivně ořezávat již frekvence těsně za 15 khz hranicí.

Doba převodu 46 sekund není v porovnání s ostatními nejhorší a výsledná velikost souboru definuje tento formát na této přenosové rychlosti jako průměrný.

#### 4.4.2 WMA



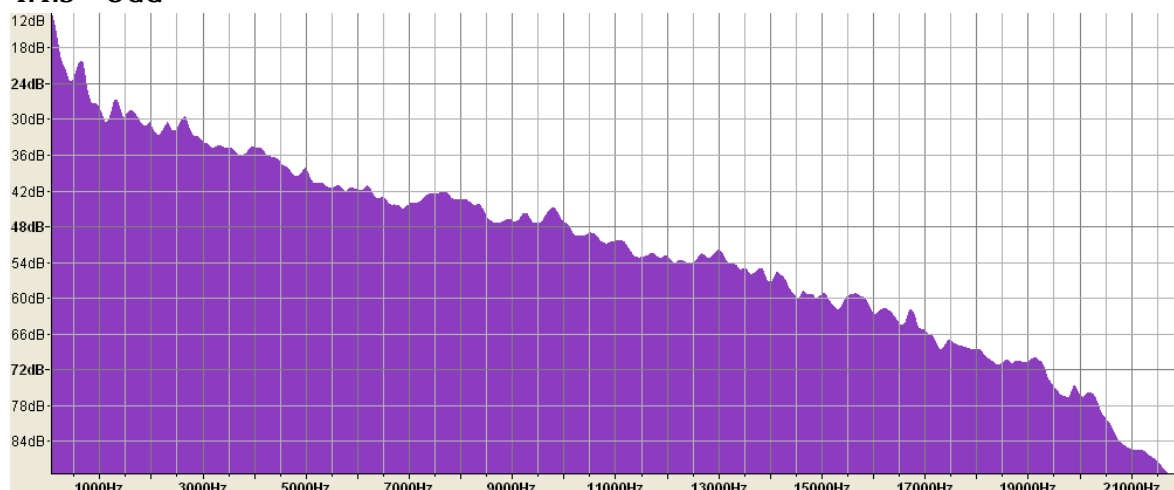
Obrázek 28 - Spektrum WMA 128 kbps

Zdroj: Vlastní zpracování v programu Audacity

Frekvenční analýza formátu WMA na úrovni přenosové rychlosti 128 kbps je opět velmi podobná spektru MP3. Nejvíce si jsou oba kodeky podobné v agresivitě ořezávání vyšších frekvencí, i když formát WMA posunul tuto hranici o pár stovek Hz doprava oproti MP3 k hranici 16 kHz.

Znovu nejrychlejší komprese a přijatelné velikost souboru pasuje tento formát opět na jedno ze dvou prvních míst i při této přenosové rychlosti.

#### 4.4.3 OGG



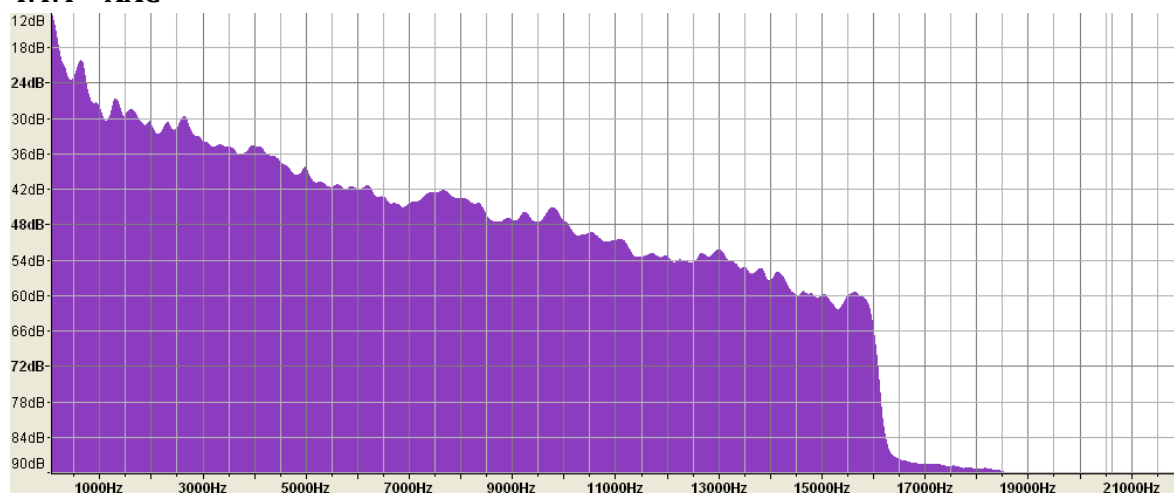
Obrázek 29 - Spektrum OGG střední kvalita

Zdroj: Vlastní zpracování v programu Audacity

Analýza spektra formátu OGG při přenosové rychlosti 128 kbps je znázorněna na obrázku 29. Oproti všem porovnávaným ztrátovým formátům tento formát ořezává nejméně frekvencí i při tomto bitovém toku a tudíž by se jeho zvuk měl nejvíce blížit originálnímu WAV souboru.

Sráží ho ovšem obrovsky dlouhá doba komprese a také výsledná velikost je nejhorší v této skupině porovnávaných formátů.

#### 4.4.4 AAC



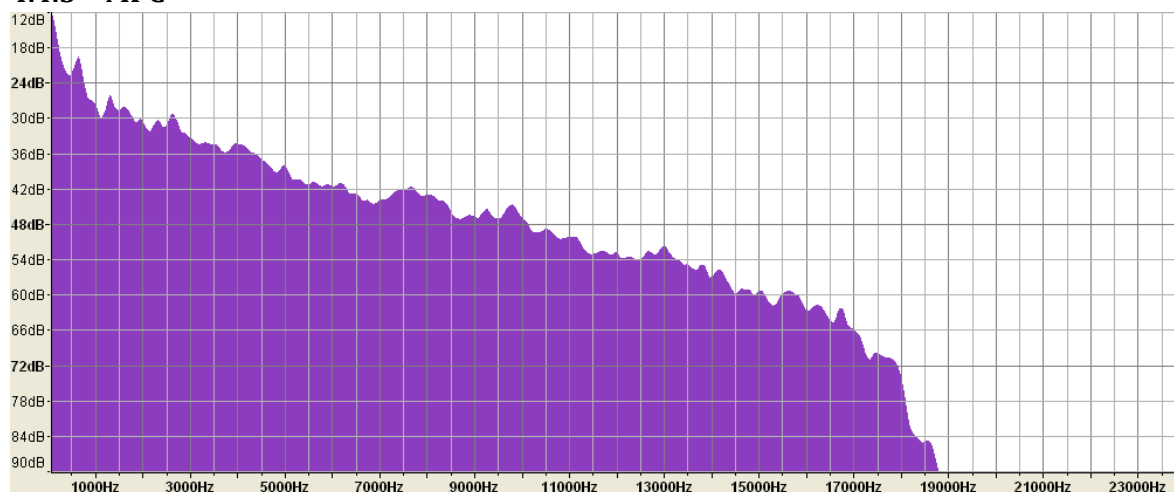
Obrázek 30 - Spektrum AAC 128 kbps

Zdroj: Vlastní zpracování v programu Audacity

Jak je vidět na spektrální analýze formátu AAC při přenosové rychlosti 128 kbps, bod, ve kterém začíná kodek ořezávat frekvence se posunul trochu více doprava, což je přesný opak toho, co se děje u ostatních formátů. Na této přenosové rychlosti AAC ořezává frekvence kolem 16 kHz.

Díky tomuto posunu a nízké výsledné velikosti se společně s WMA prosadil tento formát na přední místa testu na úrovni 128 kbps.

#### 4.4.5 MPC

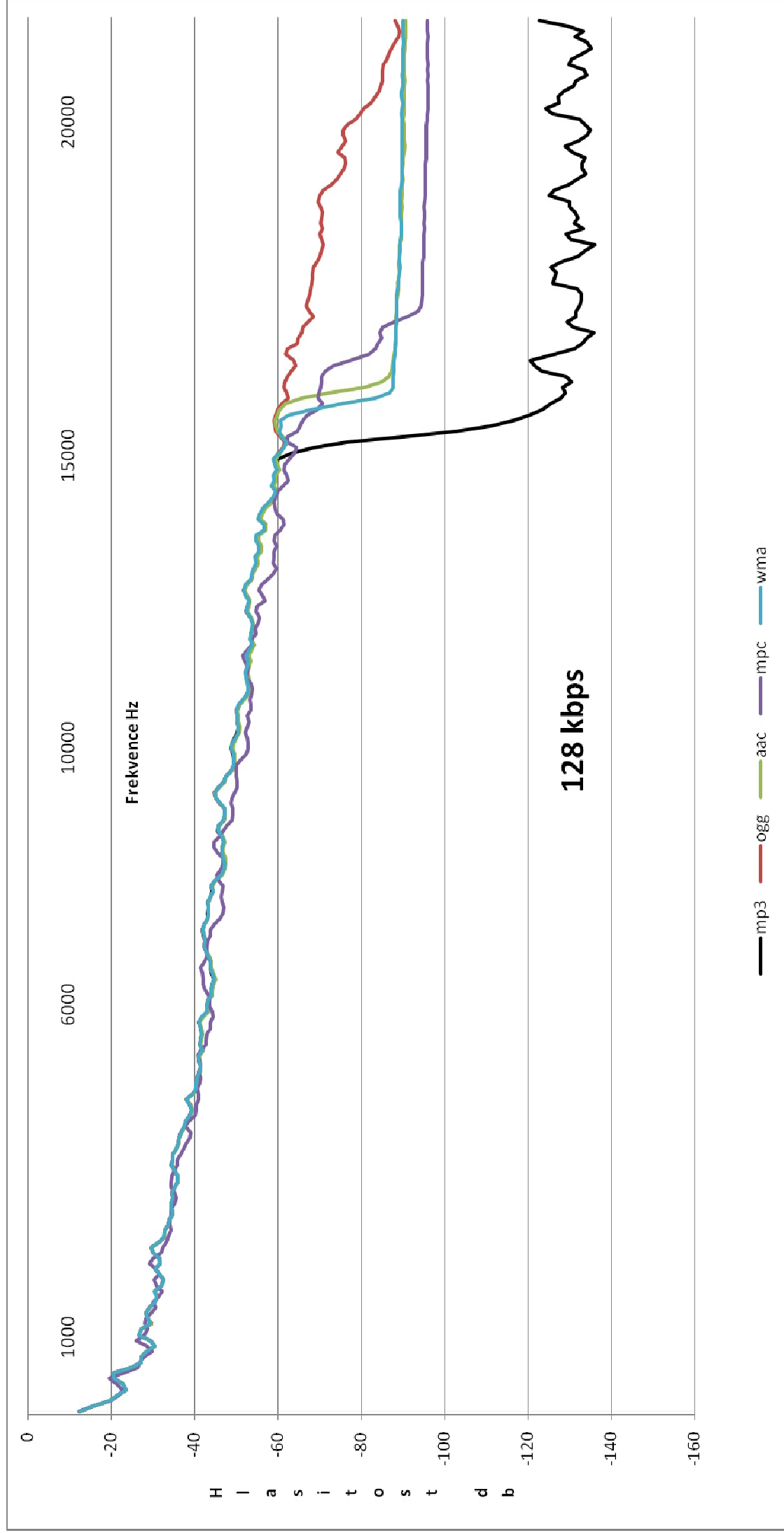


Obrázek 31 - Spektrum MPC 128 kbps

Zdroj: Vlastní zpracování v programu Audacity

Vítěz předchozího testu při 256 kbps se v tomto případě propadl o pár míst dozadu. Spektrální analýza tohoto souboru je na obrázku 31. Kodek opět ořezává frekvence plynuleji než formáty WMA nebo MP3.

Ovšem doba převodu 46 sekund a velikost výsledného souboru 4,76 MB sráží tento formát spíše do pole poražených v testu při 128 kbps.



Graf 2 - Porovnání ztrátových formátů při 128 kbps



Opět byla vyexportována data a vložena do programu Microsoft Excel. Na základě těchto dat vznikl graf 2, kde jsou jasně vidět rozdíly v jednotlivých spektrech porovnávaných ztrátových formátů.

Zhruba do 15 khz se formáty tolik neliší, za touto hranicí ovšem začínají již ořezávat vyšší frekvence. Jako první začíná ořezávat MP3 následovaný MPC. Dalším je poté formát WMA, AAC a jako poslední OGG.

Rozdíl je také ale v agresivitě tohoto ořezávání, kde například formáty OGG nebo MPC ořezávají více postupně než formát MP3.

#### 4.5 Bezztrátové formáty – 96 kbps

V následující tabulce číslo 8 jsou porovnány bezztrátové formáty podle kritérií doba převodu a velikost výsledného (komprimovaného) souboru.

Formát	Doba převodu	Velikost souboru (MB)
FLAC	5 sec	25.3
AIFF	6 sec	13.9
APE	22 sec	24.3

Tabulka 8 - Bezztrátové formáty 96 kbps

Zdroj: Autor

Nejkratší dobou převodu disponuje na této úrovni bitové rychlosti formát FLAC, který je následován AIFF a APE. Nejmenší datový prostor zabírá souboru ve formátu AIFF, který je následovaný formátem APE a FLAC.

Podle těchto kritérií je na této úrovni nejlepším bezztrátových formátem formát AIFF.

#### 4.6 Ztrátové formáty – 96 kbps

V následující tabulce 9 jsou zapsány výsledky testu ztrátových formátů na úrovni 96 kbps.

Formát	Doba převodu	Velikost souboru (MB)
MP3	52 sec	2.62
WMA 9.1	19 sec	2.64
OGG Vorbis	87 sec	3.27
AAC	44 sec	2.59
MPC	48 sec	3.67

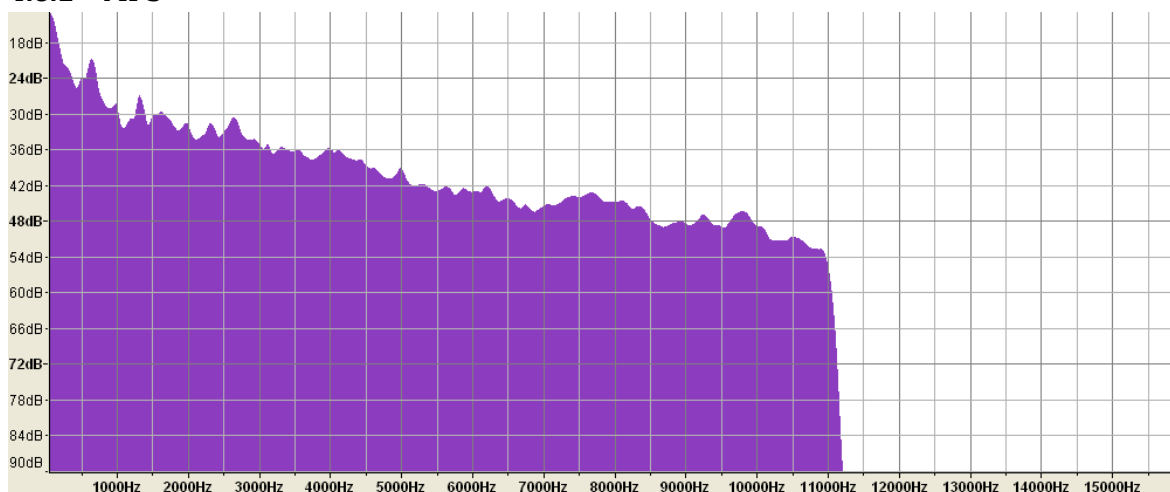
Tabulka 9 - Ztrátové formáty 96 kbps

Zdroj: Autor

Nejrychlejší převodem stejně jako na ostatních přenosových rychlostech disponuje formát WMA, který následují AAC, MPC, MP3 a OGG. Nejméně datového prostoru zabírá soubor ve formátu AAC následovaný MP3, WMA, OGG a MPC.

Podle těchto kritérií se dá tedy na úrovni 96 kbps doporučit opět formát WMA a formát AAC.

#### 4.6.1 MP3



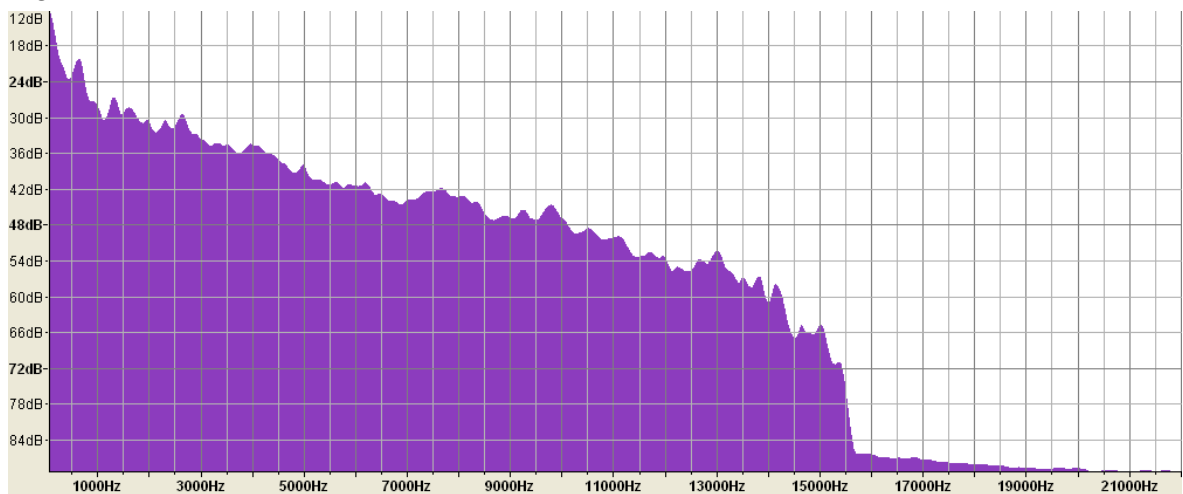
Obrázek 32 - Spektrum MP3 96 kbps

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza formátu MP3 při 96 kbps je zobrazena na obrázku 32. MP3 opět agresivně ořezává vyšší frekvence. Tentokrát se tento bod ale posunul již někam pod 11 khz.

S dobou převodu 52 sekund a velikostí souboru 2,62 MB se i v tomto případě formát MP3 umístil pouze jako průměrný ztrátový formát vzhledem k ostatním.

#### 4.6.2 WMA



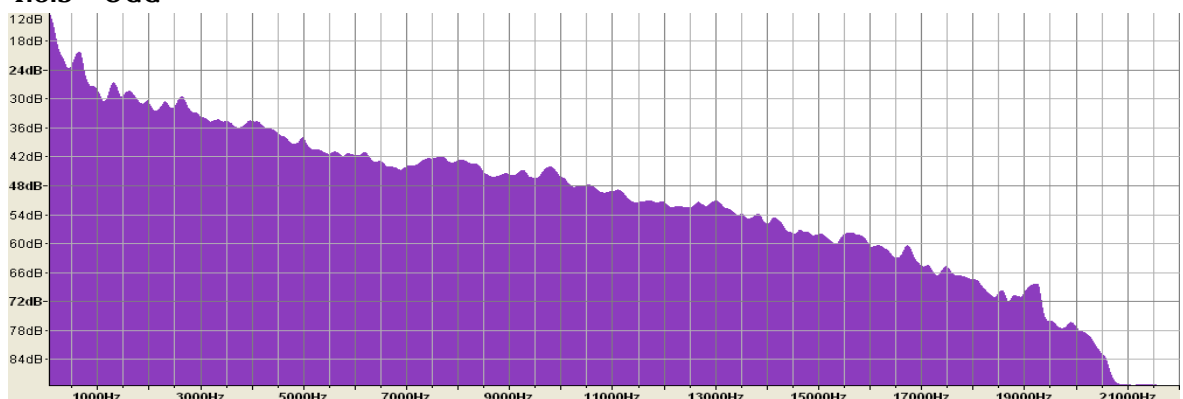
Obrázek 33 - Spektrum WMA 96 kbps

Zdroj: Vlastní zpracování v programu Audacity

Analýza spektra formátu WMA při 96 kbps je znázorněna na obrázku 33. Ořezávání frekvencí již není tak agresivní jako v předchozích případech, ale bod, v kterém ořezávat začíná se posunul opět více doleva, ale i tak soubor obsahuje množství frekvencí, které ostatní formáty na této úrovni neobsahují.

Navíc disponuje stále rychlou kompresí a velikostí souboru srovnatelnou s formátem MP3. Společně s formátem AAC je WMA na prvních dvou místech i tohoto srovnání.

#### 4.6.3 OGG



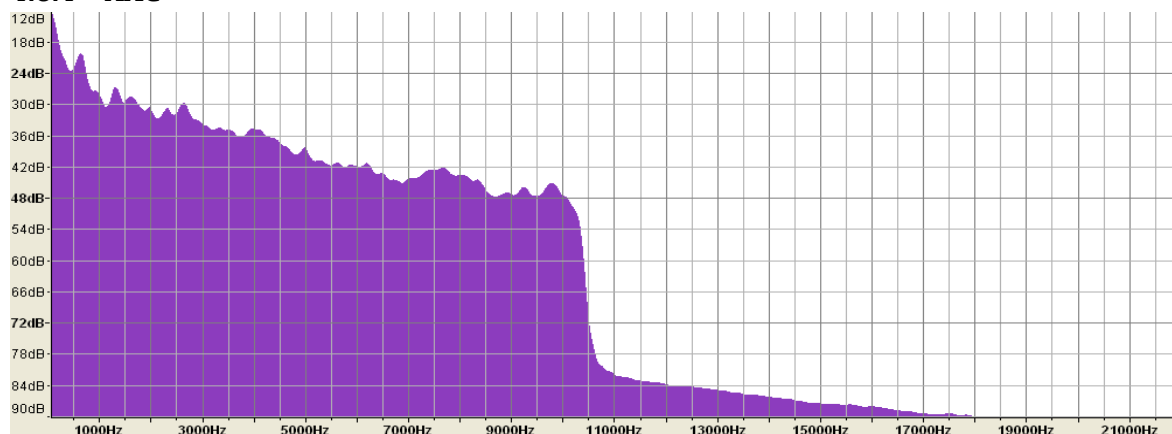
Obrázek 34 - Spektrum OGG nízká kvalita

Zdroj: Vlastní zpracování v programu Audacity

Spektrální analýza formátu OGG při 96 kbps je na obrázku 34. I v tomto případě se bod pro ořezávání frekvencí posunul více doleva, ale i tak je z tohoto pohledu OGG stále nedostihnutelný. Obsahuje pořád většinu frekvencí jako originální WAV.

Handicapem OGGu je stále ale dlouhá doba převodu a i velikost souboru nepatří právě k těm nejmenším.

#### 4.6.4 AAC

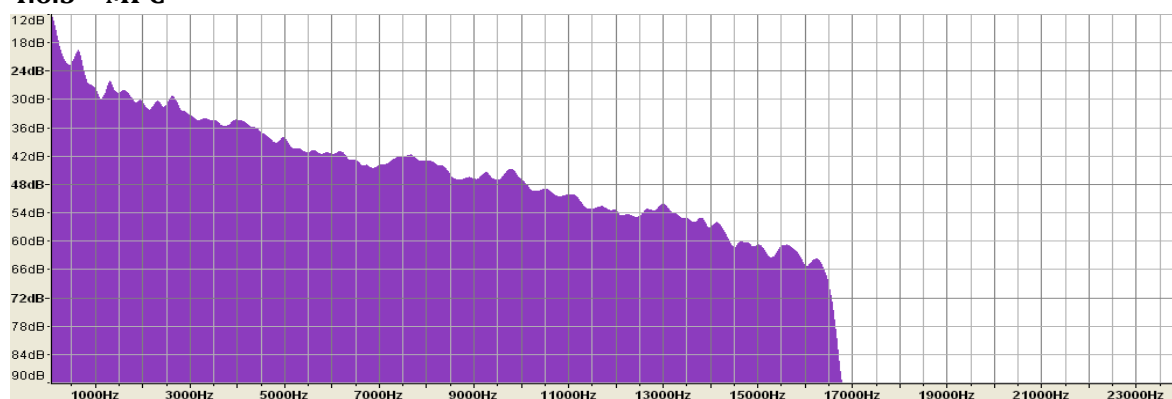


Obrázek 35 - Spektrum AAC 96 kbps

Zdroj: Vlastní zpracování v programu Audacity

Jako v případě přenosové rychlosti tak i v tomto případě se formát AAC umístil vysoko. Společně s formátem WMA se dá považovat za vítěze tohoto porovnání. Sice ořezává frekvence již kolem 10 khz, ale vše vynahrazuje velmi malou velikostí výsledného souboru.

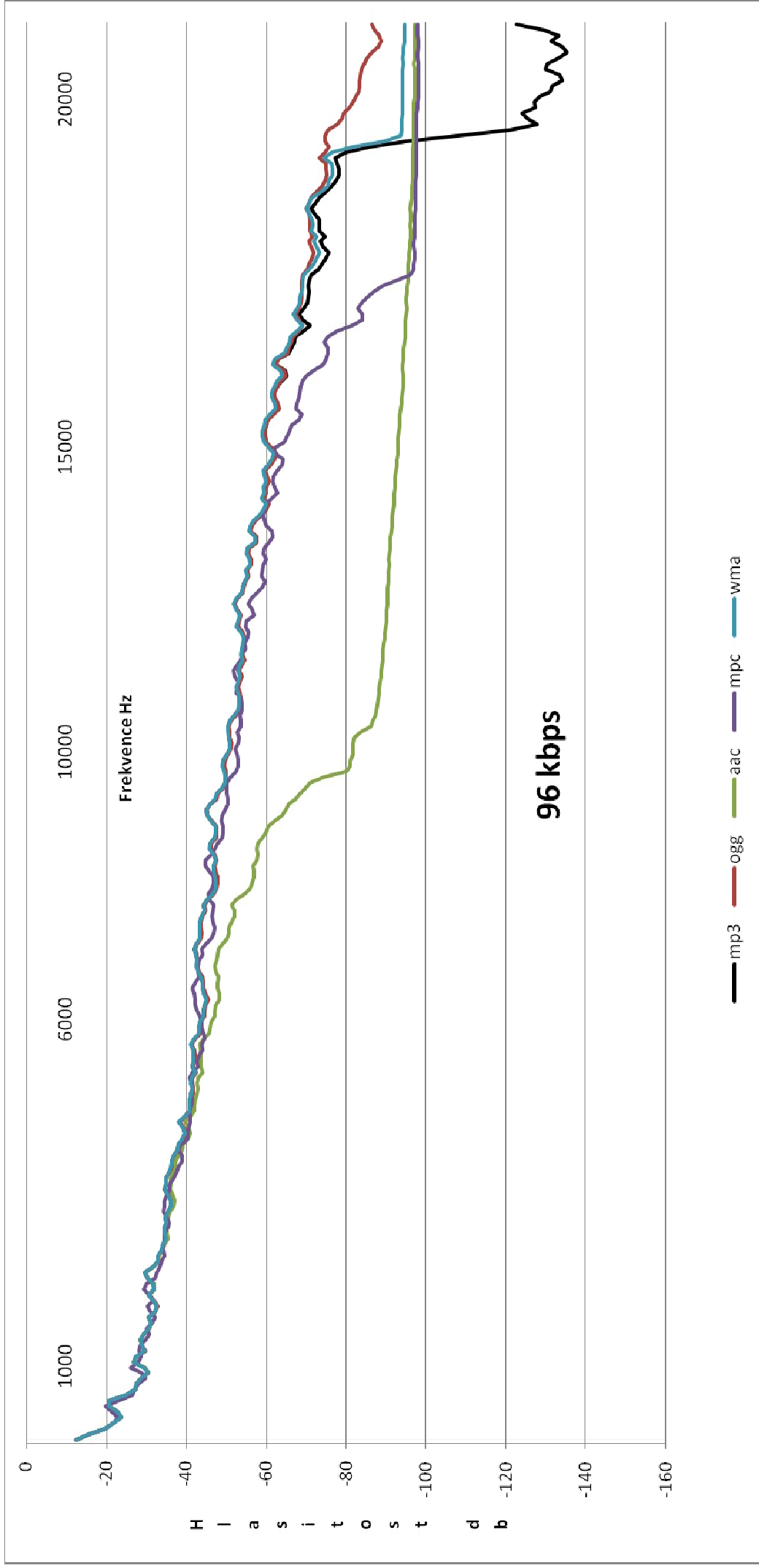
#### 4.6.5 MPC



Obrázek 36 - Spektrum MPC 96 kbps

Zdroj: Vlastní zpracování v programu Audacity

Jak je patrné ze spektrální analýzy formátu MPC při 96 kbps, tento formát začíná ořezávat více agresivněji než na předchozích přenosových rychlostech. Velikost souboru je však příliš velká a tak se tento formát umístí v porovnání s ostatními až na předposledním místě jen těsně před formátem OGG.



Graf 3 - Porovnání ztrátových formátů při 96 kbps

V grafu 3 jsou porovnány všechny ztrátové formáty na úrovni přenosové rychlosti 96 kbps. Je patrné, že nejdříve začíná frekvence ořezávat formát AAC, následovaný formáty MPC, MP3, WMA a OGG.

Na závěr těchto porovnání ztrátových a bezztrátových formátů na různých přenosových rychlostech je třeba vyvodit doporučení a zhodnotit předchozí výsledky.

Jak již bylo patrné přímo z výsledků testů je nejvhodnějším formátem podle zvolených kritérií formát WMA, který se hodí pro většinu hudebních souborů. Pokud by uživatel chtěl vhodný formát pro ukládání mluveného slova, bylo by vhodné použít formát AAC, který ořezává vyšší frekvence a ponechává výrazné především frekvenční pásmo, v kterém se běžně vyskytuje lidská řeč. Pokud by šlo uživateli především o věrnost zvuku z hlediska shody s originálem, byl by doporučen formát OGG, který sice celkově nedopadl dobře, ale z pohledu obsahu jednotlivých frekvencí a shody s originálem by dosáhl nejlepších výsledků. Naopak z výsledků je patrné, že velmi rozšířený formát MP3 spadá spíše do průměru mezi těmito formáty.

Co se týče bezztrátových formátů byl by doporučen jeden z dvojice formátů AIFF nebo APE.

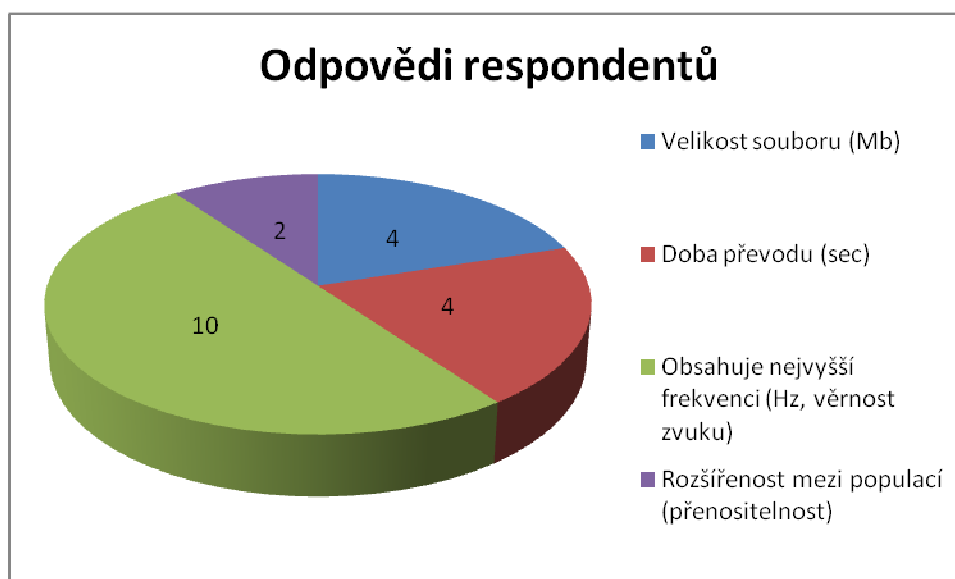
#### **4.7 Určení vhodného ztrátového formátu pro ukládání hudebních dat v domácích podmínkách na základě vícekriteriální analýzy**

K určení ztrátového formátu, který je nejvhodnější pro ukládání audio dat v domácích podmínkách, byly použity metody TOPSIS, Váženého součtu a ORESTE, které patří do oblasti metod vícekriteriální analýzy variant. Aby autor nevolil kritéria subjektivně, bylo dotázáno 20 respondentů, kteří si hudbu doma archivují, které z vybraných kritérií je pro ně při výběru formátu nejdůležitější. Tyto odpovědi jsou poté základem pro stanovení vah kritérií metodou alokace 100 bodů. K výpočtu byl použit program Microsoft Excel 2010 společně s přídatným modulem MCAKOSA vytvořeným na katedře operační a systémové analýzy PEF ČZU v Praze.

Bylo zvoleno kritérium doba převodu, která vyjadřuje dobu v sekundách potřebnou k převodu originálního WAV souboru do příslušného ztrátového formátu. Podle odpovědí vzorku 20 lidí je toto kritérium nejdůležitější pro 4 z dotazovaných. Dalším kritériem byla

velikost souboru v megabytech. Toto kritérium bylo opět nejdůležitější pro 4 z 20 dotazovaných. Kritérium pojmenované „obsahuje nejvyšší frekvenci“, které vyjadřuje, jak je zvuk věrný k originálu a také při jaké frekvenci kodér ořezává frekvence z daného souboru. Tato věrnost audia byla nejdůležitější pro celou polovinu z dotazovaných. Posledním kritériem byla zvolena rozšířenost mezi populací (podle [www.fileinfo.com](http://www.fileinfo.com)). Toto kritérium vyjadřuje, jak je který formát mezi lidmi rozšířen a používán. Rozšiřitelnost je důležitá z hlediska přenositelnosti a možnosti přehrávání na různých zařízeních.

Výsledky dotazování jsou souhrnně znázorněny na následujícím grafu.



Graf 4 - Odpovědi respondentů

Zdroj: Autor

V následující tabulce jsou přehledně sepsány potřebné hodnoty pro provedení vícekritériální analýzy variant. Dále jsou zde zapsány váhy jednotlivých kritérií společně s jejich povahou (min, max). Tato tabulka sloužila jako kritériální matice pro vyhodnocení vícekritériální analýzy variant metodami TOPSIS, Váženého součtu a ORESTE pomocí přídatného modulu MCAKOSA.



Formát	Doba převodu (sec)	Velikost souboru (MB)	Obsahuje nejvyšší frekvenci (Hz)	Rozšířenost mezi populací dle Fileinfo
MP3 256	55	6,99	19850	2
WMA 9.1 256	18	7,02	20013	2
OGG Vorbis 256	62	6,68	20379	4
AAC 256	65	3,59	10335	2
MPC 256	50	5,62	18678	5
MP3 128	46	3,49	15286	2
WMA 9.1 128	22	3,52	15938	2
OGG Vorbis 128	76	4,87	20339	4
AAC 128	44	3,46	16141	2
MPC 128	46	4,76	18012	5
MP3 96	52	2,62	11003	2
WMA 9.1 96	19	2,64	15489	2
OGG Vorbis 96	87	3,27	20053	4
AAC 96	44	2,59	10559	2
MPC 96	48	3,67	16593	5
povaha	min	min	max	min
váhy	0,2	0,2	0,5	0,1

Tabulka 10 - Kriteriační matice

Zdroj: Autor a [www.fileinfo.com](http://www.fileinfo.com)

Tato kriteriační matice byla použita jako podklad pro použití přídatného modulu MCAKOSA a vypočtení následujících metod vícekritériační analýzy variant.

#### 4.7.1 Metoda TOPSIS

„Tato metoda posuzuje varianty z hlediska jejich vzdálenosti od ideální a bazální varianty. Postupujeme takto:

- 1) Převědeme minimalizační kritéria na maximalizační.
- 2) Zkonstruujeme normalizovanou kriteriační matici  $R$  podle vzorce . Sloupce matice  $R$  jsou tak vektory jednotkové délky.
- 3) Vypočteme váženou kriteriační matici  $W$  dle vztahu  $w_{ij} = v_j r_{ij}$ .
- 4) Určíme ideální a bazální variantu vzhledem k hodnotám matice  $W$ , tj. , .
- 5) Vypočteme vzdálenosti variant od ideální varianty  $d_i^+$  a od bazální varianty  $d_i^-$ .

6) Nakonec spočteme relativní ukazatele vzdáleností variant od bazální varianty podle vzorce  $ci=di-/(dI++ di-)$ . Hodnoty těchto ukazatelů se pohybují mezi 0 a 1, přičemž 0 se nabývá pro bazální a 1 pro ideální variantu.

7) Varianty seřadíme sestupně podle hodnoty relativních ukazatelů vzdálenosti. [3]

	Metoda TOPSIS	
	Vzdálenost od bazální varianty	Pořadí
MP3 256	0,50728	13
WMA 9.1 256	0,612127	5
OGG Vorbis 256	0,482624	14
AAC 256	0,422761	15
MPC 256	0,527704	9
MP3 128	0,630706	4
WMA 9.1 128	0,748916	2
OGG Vorbis 128	0,508786	12
AAC 128	0,669846	3
MPC 128	0,579392	7
MP3 96	0,524122	11
WMA 9.1 96	0,765122	1
OGG Vorbis 96	0,527022	10
AAC 96	0,541954	8
MPC 96	0,594262	6

Tabulka 11 - Výsledky metody TOPSIS

Zdroj: Autor

Předcházející tabulka kompletuje výsledky metody TOPSIS. Jako nejlepší kompromisní varianta byla zvolena možnost ukládání audio souborů do formátu WMA 9.1 při bitové rychlosti 96 kbps. Tento výsledek je samozřejmě ovlivněn odpověďmi dotazovaných respondentů a stanovení kritérií na jejich základě. Na druhém místě se v této variantě umístil opět formát WMA 9.1 tentokrát však při přenosové rychlosti 128 kbps. Jako třetí skončil formát AAC při přenosové rychlosti 128 kbps. První z oblíbených formátů MP3 se umístil až na čtvrtém místě.

Dá se tedy říci, že na základě použití metody TOPSIS byl vyhodnocen nejlépe formát WMA 9.1.

#### 4.7.2 Metoda ORESTE

**Metoda popisuje následující algoritmus:**

„Nejprve sestavíme vektor pořadí kritérií. Indiferentní kritéria jsou ohodnocena průměrnými pořadovými čísly.

$$q = (q_1 \dots q_n) \text{ kde } q \text{ je pořadí } j\text{-tého kritéria}$$

Analogicky sestavíme matici uspořádání variant dle jednotlivých kritérií. Indiferentní kritéria jsou opět ohodnocena průměrnými pořadovými čísly.

$P = (p_{ij}), i = 1, \dots, p \text{ a } j = 1, \dots, k$  kde  $p_{ij}$  je pořadí varianty  $a_i$  podle  $j$ -tého kritéria.

Ze znalosti vektoru  $q$  a matice  $P$  je možné vypočítat vzdálenost  $D$  od fiktivního počátku  $D = (d_{ij})$  je matice vzdáleností od fiktivního počátku

Pro prvky této matice platí:

$$d_{ij} = [0.5(p_{ij})^r + 0.5(q_j)^r]^{1/r} \text{ kde } r \in \mathbb{R}$$

Obvykle se vzdálenost měří pomocí tzv. Dujmovičovy metriky, kdy  $r = 3$  a nazývá se Dujmovičův exponent. Hodnoty vzdáleností  $d_{ij}$  se dále uspořádají vzestupně a ohodnotí pořadovými čísly (příp. průměrnými pořadovými čísly v případě, že se některé hodnoty shodují).

Tím vznikla matice pořadových čísel  $R = (r_{ij})$ . Z matice pořadových čísel je možno určit její řádkové součty:

$$r_i = \sum_{j=1}^n r_{ij} \quad i = (1, \dots, n)$$

Vzestupným uspořádáním hodnot  $r_{ij}$  dostáváme uspořádání (kvaziuspořádání) jednotlivých variant. “ [18]

	Metoda ORESTE	
	Hodnota ri	Pořadí
MP3 256	129	9
WMA 9.1 256	93	4
OGG Vorbis 256	143	11
AAC 256	161	14
MPC 256	164	15
MP3 128	119,5	7
WMA 9.1 128	95,5	5
OGG Vorbis 128	141	10
AAC 128	81,5	2
MPC 128	151	12
MP3 96	113	6
WMA 9.1 96	74	1
OGG Vorbis 96	121,5	8
AAC 96	88	3
MPC 96	155	13

Tabulka 12 - Výsledky metody ORESTE

Zdroj: Autor

Při pohledu na tabulku číslo 12, která sumarizuje výsledky metody ORESTE je opět patrné, že jako nejlepší byl vyhodnocen formát WMA 9.1 při přenosové rychlosti 96 kbps. Na druhém místě však došlo ke změně, jelikož ho obsadil formát AAC při přenosové rychlosti 128 kbps a třetí je opět formát AAC při přenosové rychlosti 96 kbps.

Celkově lze ale však opět říci, že na základě metody ORESTE lze doporučit jednoznačně formát WMA 9.1, jehož výsledky jsou nejlepší (1., 4. a 5. Místo).

#### 4.7.3 Metoda váženého součtu

**„Postup metody váženého součtu je dán následujícími kroky:**

1. *Uřídíme ideální variantu  $H$  s ohodnocením ( $h_1, \dots, h_n$ ) a bazální variantu  $D$  s ohodnocením ( $d_1, \dots, d_n$ ).*
2. *Vytvoříme standardizovanou kritériální matici  $R$ , jejíž prvky získáme pomocí vzorce:*

$$r_{ij} = \frac{y_{ij} - d_j}{h_j - d_j}$$

Matice  $R$  již představuje matici hodnot funkce užitku z  $i$ -té varianty podle  $j$ -tého kritéria, protože prvky této matice jsou lineárně transformovanými kriteriálními hodnotami tak, že  $r_{ij} \in \{0,1\}$ . Potom bazální variantě odpovídá hodnota nula a ideální variantě hodnota jedna.

3. Pro jednotlivé varianty vypočteme agregovanou funkci užitku :

$$u(a_i) = \sum_{j=1}^n v_j r_{ij}$$

4. Varianty seřadíme sestupně podle hodnot  $u(a_i)$  a potřebný počet variant s nejvyššími hodnotami užitku považujeme za řešení problému.“ [5]

	Metoda váž. součtu	
	Užitek	Pořadí
MP3 256	0.6677	6
WMA 9.1 256	0.7817	1
OGG Vorbis 256	0.6211	9
AAC 256	0.3186	15
MPC 256	0.5857	11
MP3 128	0.6246	8
WMA 9.1 128	0.7253	3
OGG Vorbis 128	0.6602	7
AAC 128	0.6743	5
MPC 128	0.6030	10
MP3 96	0.4333	14
WMA 9.1 96	0.7514	2
OGG Vorbis 96	0.6864	4
AAC 96	0.4357	13
MPC 96	0.5758	12

Tabulka 13 - Výsledky metody váženého součtu

Zdroj: Autor

Metoda Váženého součtu jen potvrdila výsledky předchozích metod. Zde se formát WMA 9.1 prosadil dokonce na první tři místa mezi všemi srovnávanými formáty. Poprvé se zde více prosadil také formát OGG, který se v poslední době stává pořád více populárním mezi širokou veřejností.

Podle všech metod vícekriteriální analýzy variant, které byly v této práci použity, dopadl nejlépe formát WMA 9.1, který je tedy doporučen jako nejlepší ztrátový formát pro ukládání audio souborů v domácích podmínkách.

### **Komentář k výsledkům vícekriteriální analýzy:**

Výsledky vícekriteriální analýzy provedené autorem jsou podpořeny i výsledky nezávislého testu společnosti ZDLabs, která prováděla test mezi posluchači. Tito lidé měli za úkol označit, který ze ztrátových zvukových formátů se jim jeví více podobný originálnímu audiu. 97,6% posluchačů dalo přednost formátu WMA.

*„Formát Windows Media Audio (WMA) je součástí programového vybavení Windows Media firmy Microsoft. Podle prohlášení firmy Microsoft nabízí zvukovou kvalitu srovnatelnou s audio CD při datovém toku 64kbps, což představuje přibližně dvakrát účinnější kompresi oproti MP3. Tuto skutečnost potvrdil nezávislý výzkum ZDLabs (jako MP3 kodek byl použitý software firmy Xing, který nenabízí nejlepší kvalitu výsledných MP3 souborů). WMA je zabezpečený SDMI kompatibilní formát. Krátce po svém uvedení na trh v loňském roce byl však prolomen. Ve spojení s ASF - Advanced Streaming Format může být WMA streamován. Díky tomu se formát Windows Media stal přímým konkurentem zavedeného formátu RealAudio a velmi rychle získává na oblibě. V současné době WMA formát podporuje celá řada přenosných hardwarových i softwarových přehrávačů. Formát WMA je jednoznačně hlavním kandidátem na přední místo mezi kompresními audio formáty blízké budoucnosti.“ [30]*

## 5 Závěr

Výstupem literární rešerše je charakteristika problematiky digitálního zvuku, základních veličin, důležitých parametrů a fyzikálních skutečností spjatých s tímto tématem.

Dále jsou v teoretické části práce stručně charakterizovány nepoužívané i ztrátové i ztrátové formáty digitálního audia.

V praktické části diplomové práce se autor zaměřil na dosažení stanovených cílů z kapitoly 2 – Cíl práce a metodika. Jednalo se o porovnání ztrátových a bezztrátových formátů na základě vybraných kritérií na různých bitových rychlostech. Tato kritéria byla v případě ztrátových formátů ještě doplněna o porovnání spektrálních analýz jednotlivých zvukových souborů. Druhým cílem bylo vybrání vhodného ztrátového formátu pro ukládání především rockové a populární hudby v domácích podmínkách. Tohoto cíle bylo dosaženo pomocí tří metod vícekritériální analýzy variant, konkrétně metod TOPSIS, ORESTE a Váženého součtu.

Jako nejlepší bezztrátový formát byl na základě porovnání kritérií vybrán formát AIFF, kterému zdatně sekunduje formát APE.

V případě ztrátových formátů si nejlépe vedl formát WMA. Při bitové rychlosti 256 kbps se však na prvním místě umístil formát MPC.

Z výsledků vícekritériální analýzy variant nejlépe vyšel opět formát WMA tesně následovaný formátem AAC. Formát AAC zaujal především svojí malou velikostí, která je dána potlačením většiny vyšších frekvencí, které ale v případě hudby mohou být velmi podstatné. Proto by autor jednoznačně doporučil formát WMA, u kterého k ořezávání těchto vyšších frekvencí dochází daleko později. Formát AAC se ale může výborně hodit uživatelům pro ukládání běžné lidské řeči, jejíž frekvence se pohybuje právě v mezích, které tento formát nepotlačuje.

Výsledky porovnání na základě kritérií a spektrální analýzy a výsledky vícekritériální analýzy variant se prakticky shodují a proto by autor doporučil uživatelům používat pro ukládání a poslech digitálního audia formát WMA (konkrétně nejlepších výsledků dosáhl tento formát při bitové rychlosti 96 kbps).

Tento výsledek podporuje i výsledek nezávislého testu ZDLabs, který je citován v předcházející kapitole.

Budoucnost digitálního audia je otevřená a v tomto rychle se vyvíjejícím prostředí se jistě brzy dočkáme dalších novinek, které posunou hranice užití digitálního audia zase o něco dál. Nicméně se autor domnívá, že pozice formátu MP3 bude ještě v několika dalších letech neotřesitelná, i když, jak je patrné z celé této práce, tento formát rozhodně nepatří k tomu nejlepšímu, co se v oblasti digitálního audia nachází.



## 6 Seznam použitých zdrojů

1. All-streaming-media.com [online]. 2011 [cit. 2011-03-01]. Description of popular audio formats. Dostupné z WWW: <<http://all-streaming-media.com/streaming-media-faq/faq-Description-of-popular-audio-formats-MP3-WMA-OGG-AAC-WAV.htm>>.
2. BRANDENBURG, Karlheinz . MP3 AND AAC EXPLAINED. AES 17th International Conference on High Quality Audio Coding [online]. 2008, [cit. 2011-03-21]. Dostupný z WWW:  
<<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.26.5956&rep=rep1&type=pdf>>
3. BROŽOVÁ, Helena. Kvantitativní podpora rozhodování [online]. 2009 [cit. 2011-03-21]. Vícekriterální analýza variant. Dostupné z WWW:  
<<http://pef.czu.cz/~BROZOVA/CASESTUDY/index.html>>.
4. Ceskolipska.cz [online]. 2004 [cit. 2011-03-13]. Mp3. Dostupné z WWW:  
<<http://ceskolipska.cz/files/25/mp3.doc>>.
5. HOUŠKA, Milan. Pef.czu.cz/~houska [online]. 2009 [cit. 2011-03-21]. Metoda váženého součtu. Dostupné z WWW:  
<[http://pef.czu.cz/~houska/EMM/Materialy/Cviceni/VAV/Vazeny\\_soucet.htm](http://pef.czu.cz/~houska/EMM/Materialy/Cviceni/VAV/Vazeny_soucet.htm)>.
6. CHOLT, David. Formáty zvuku ve výpočetní technice [online]. Plzeň, 2010. 5 s. Semestrální práce. Západočeská univerzita v Plzni. Dostupné z WWW:  
<[http://www.kiv.zcu.cz/~lobaz/dtp1/semestralky2009-2010/brozura/Cholt\\_brozura.pdf](http://www.kiv.zcu.cz/~lobaz/dtp1/semestralky2009-2010/brozura/Cholt_brozura.pdf)>.
7. Ifanda.cz [online]. 2009 [cit. 2010-02-02]. Ztrátové a bezztrátové formáty zvuku na PC. Dostupné z WWW: <<http://ifanda.cz/clanky/pocitace/ztratove-bezztratove-formaty-zvuku-na-pc>>.
8. Infopedia.funsite.cz [online]. 2010 [cit. 2010-09-21]. Formát AIFF. Dostupné z WWW:  
<<http://infopedia.funsite.cz/index.php?text=1181-format-aiff>>.
9. Jech.webz.cz [online]. 2011 [cit. 2011-01-11]. Digitální kompresní formáty. Dostupné z WWW: <<http://jech.webz.cz/formaty.php>>.

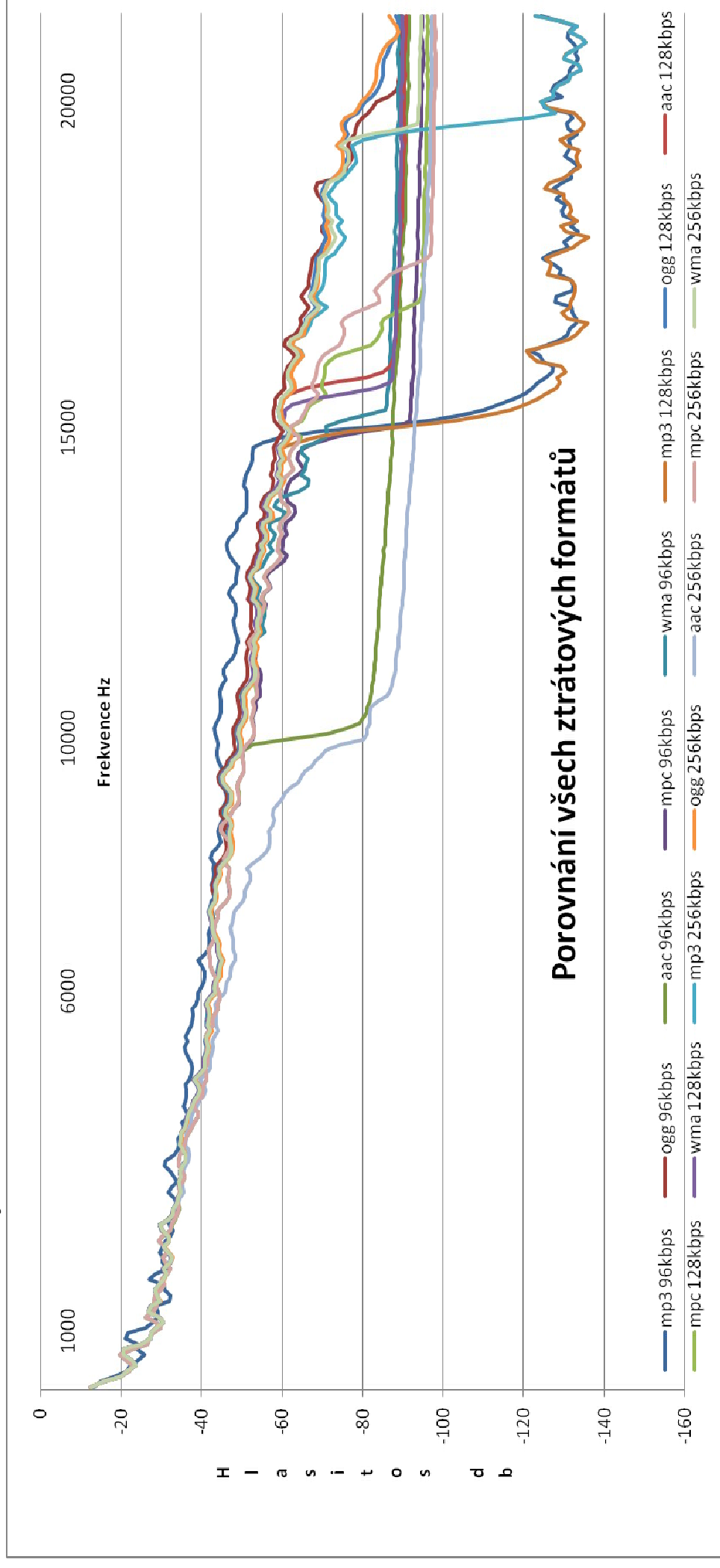
10. JOHNSON, Marc. Media matters [online]. Columbia University, 2000 [cit. 2010-12-13]. Audio and Video File Formats and File Sizes. Dostupné z WWW: <<http://www.media-matters.net/docs/resources/Digital%20Files/General/Columbia%20University.pdf>>.
11. KROČÁK, Pavel. Absolventi.gymcheb.cz [online]. 2007 [cit. 2011-01-21]. Zvukové kodeky. Dostupné z WWW: <[http://absolventi.gymcheb.cz/2007/pakroca/download/SP\\_VOS\\_Audiokodeky.doc](http://absolventi.gymcheb.cz/2007/pakroca/download/SP_VOS_Audiokodeky.doc)>.
12. Linkovnik.cz [online]. 2008 [cit. 2011-03-13]. MP3. Dostupné z WWW: <<http://linkovnik.wz.cz/mp3/>>.
13. LIST, Martin. Kompresie multimédií [online]. Plzeň : SPŠE Plzeň, 2006. 24 s. Ročníková práce. SPŠE Plzeň. Dostupné z WWW: <[http://www.jenseneboj.com/pracestudentu/data/Kompresie%20multim%EDi%ED.htm#\\_Toc131170347](http://www.jenseneboj.com/pracestudentu/data/Kompresie%20multim%EDi%ED.htm#_Toc131170347)>.
14. Litexmedia.com [online]. 2007 [cit. 2011-01-11]. Audio formats. Dostupné z WWW: <[http://www.litexmedia.com/article/audio\\_formats.html](http://www.litexmedia.com/article/audio_formats.html)>.
15. MALVAR, Henrique S. Signal Processing with Lapped Transforms. Norwood, MA, USA : Artech House, Inc. , 1992. 379 s. ISBN 0890064679.
16. MCGOODWIN, Michael. Mcgoodwin.net [online]. 2000 [cit. 2011-03-13]. Digital Audio CD and Other Selected Digital Technologies. Dostupné z WWW: <<http://www.mcgoodwin.net/digitalaudio/digitalaudio.html>>
17. Mndigital.org [online]. 2006 [cit. 2010-10-10]. Digital Audio Best Practices. Dostupné z WWW: <<http://www.mndigital.org/digitizing/standards/audio.pdf>>.
18. MYSLÍK, Jan. Srovnávací studie vestavěných operačních systémů Embedded OS - A Comparison Study. Praha, 2008. 59 s. Bakalářská práce. ČVUT v Praze. Dostupné z WWW: <[https://dip.felk.cvut.cz/browse/pdfcache/xsmitka\\_2009bach.pdf](https://dip.felk.cvut.cz/browse/pdfcache/xsmitka_2009bach.pdf)>.
19. Navajo.cz [online]. 2010 [cit. 2011-02-01]. Hlasitost. Dostupné z WWW: <<http://hlasitost.navajo.cz/hlasitost.png>>.

20. PanWiki [online]. 2010 [cit. 2010-11-12]. Akustické pojmy. Dostupné z WWW: <[http://panwiki.panska.cz/index.php/Akustické\\_pojmy](http://panwiki.panska.cz/index.php/Akustické_pojmy)>.
21. PanWiki [online]. 2010 [cit. 2010-11-12]. Zvuk a sluch. Dostupné z WWW: <[http://panwiki.panska.cz/index.php/Zvuk\\_a\\_sluch](http://panwiki.panska.cz/index.php/Zvuk_a_sluch)>.
22. Premiumbeat.com [online]. 2010 [cit. 2011-01-11]. Which Audio Format is best?. Dostupné z WWW: <<http://www.premiumbeat.com/blog/which-audio-format-is-best/>>.
24. PŘIBYL, Michal. Kompresní algoritmus pro bezdrátový mikrofon [online]. Praha : ČVUT v Praze, 2007. 64 s. Bakalářská práce. ČVUT v Praze. Dostupné z WWW: <[https://dip.felk.cvut.cz/browse/pdfcache/pribym2\\_2007bach.pdf](https://dip.felk.cvut.cz/browse/pdfcache/pribym2_2007bach.pdf)>.
25. PŘICHYSTAL, Jan . Mendelu.cz [online]. 23.2.2009 [cit. 2011-03-22]. Komprimace v praxi. Dostupné z WWW: <[https://akela.mendelu.cz/~jprich/predn/komprimace\\_v\\_praxi.pdf](https://akela.mendelu.cz/~jprich/predn/komprimace_v_praxi.pdf)>.
26. REICHL, Jaroslav; VŠETIČKA, Martin. Encyklopedie fyziky [online]. 2006 [cit. 2011-01-22]. Weber - Fechnerův psychofyzikální zákon. Dostupné z WWW: <<http://fyzika.jreichl.com/index.php?sekce=browse&page=210>>.
27. Slaboproud.sweb.cz [online]. 2001 [cit. 2011-03-13]. Záznam zvuku. Dostupné z WWW: <<http://slaboproud.sweb.cz/elt2/stranky1/elt042.htm>>.
28. SVOBODA, Radek. Experimenty s psychoakustickou kompresí zvuku [online]. Praha : ČVUT v Praze, 2010. 70 s. Bakalářská práce. ČVUT v Praze. Dostupné z WWW: <[https://dip.felk.cvut.cz/browse/pdfcache/svobora4\\_2010bach.pdf](https://dip.felk.cvut.cz/browse/pdfcache/svobora4_2010bach.pdf)>.
29. ŠVEC, Jiří. Kompresie zvuku? Jen podvod na uši!. E15.cz [online]. 10.6.2009, [cit. 2011-03-21]. Dostupný z WWW: <<http://avmania.e15.cz/kompresie-zvuku-jen-podvod-na-usi>>.
30. Technet.idnes.cz [online]. 2000 [cit. 2011-03-21]. Kdo nahradí MP3? - přehled kompresních audio formátů. Dostupné z WWW: <[http://technet.idnes.cz/tec\\_technika.asp?r=digital&c=A000326\\_0012928\\_digital](http://technet.idnes.cz/tec_technika.asp?r=digital&c=A000326_0012928_digital)>.

31. TIŠNOVSKÝ, Pavel. General MIDI a formát souborů SMF. Root.cz [online]. 5.3.2009, [cit. 2011-03-21]. Dostupný z WWW: <<http://www.root.cz/clanky/general-midi-a-format-souboru-smf/>>.
32. ULMANN, Vojtěch. Astronuklfyzika.cz [online]. 2007 [cit. 2011-03-01]. Minidisky. Dostupné z WWW: <<http://astronuklfyzika.cz/Minidisk.htm>>.
33. VEČERKA, Arnošt. Kompresi obrazu a zvuku. In Kompresi dat [online]. Olomouc : Univerzita Palackého, 2008 [cit. 2011-03-13]. Dostupné z WWW: <<http://phoenix.inf.upol.cz/esf/ucebni/kompresi.pdf>>.
34. WATKINSON, John. The art of digital audio. Oxford: Focal Press, 2001. ISBN 0-240-51587-0
35. Wav.name [online]. 2011 [cit. 2011-03-22]. WAV. Dostupné z WWW: <<http://wav.name/>>.
36. Xiph.Org [online]. 2010 [cit. 2010-03-11]. Vorbis I specification. Dostupné z WWW: <[http://www.xiph.org/vorbis/doc/Vorbis\\_I\\_spec.html#x1-70001.1.4](http://www.xiph.org/vorbis/doc/Vorbis_I_spec.html#x1-70001.1.4)>.
37. Zdnet.com [online]. 2005 [cit. 2010-06-21]. 61.44% of P2P traffic is video, 11.34% is audio. Dostupné z WWW: <<http://www.zdnet.com/blog/itfacts/6144-of-p2p-traffic-is-video-1134-is-audio/8641>>.

## 7 Přílohy

### 7.1 Porovnání všech ztrátových formátů



Graf 5 - Porovnání ztrátových formátů celkem

## 7.2 Loga jednotlivých formátů

Bezztrátové:



WAV



FLAC



AIFF



APE

Ztrátové:



MP3



WMA



OGG Vorbis



AAC



MPC