

Digitální zpracování audio signálu – ztrátová komprese (MP3)

(Materiál pro potřeby výuky předmětu KMT/UE – Pavel Kratochvíl, kratinek@kmt.zcu.cz)

Soubory v CD kvalitě mají datový tok cca 1,4 Mb/s. Datový tok je třeba zmenšit - ať už z důvodu snížení náročnosti při datových přenosech (rádio, televize, internet), nebo zmenšení souborů pro uložení na disk.

Existuje mnoho kompresních metod. Obecně lze říci, že většího stupně komprese se dosahuje ztrátovou kompresí. Na rozdíl od bezztrátové komprese jsou ale data degradována a nelze zpětně získat původní plnohodnotný signál.

Kompresní metody (ztrátové) fungují na principech vynechání nadbytečných (redundantních) a zbytečných (irelevantních) dat.

Nadbytečná data – např. stereofonní signál se shodnou informací v obou kanálech.

Zbytečná data – lze vynechat slabší zvuky, které jsou maskovány silnějšími; lze snížit počet kvantizačních stupňů v pásáží, kdy je kvantizační šum maskován vyšší úrovní užitečného signálu.

Maskování

Nechť zní tón o frekvenci f a prahovém tlaku p_0 . Zazní-li druhý tón o frekvenci f' s efektivním tlakem takovým, že sluch přestane vnímat první tón, říkáme, že tón o frekvenci f je maskován tónem o frekvenci f' .

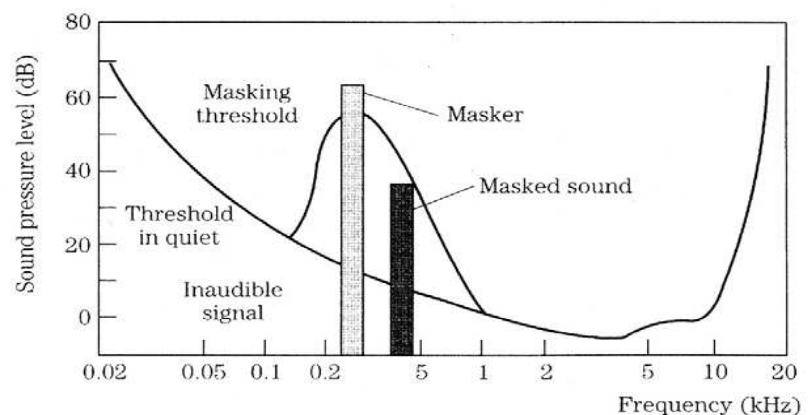
Poté zvyšujeme tlak p_0 až k hodnotě p , kdy ho ucho opět postřehne. Poměr p/p_0 nazýváme stupeň maskování.

Pro potřeby fyzikálního popisu mluvíme o frekvenčním a časovém maskování. Ve skutečnosti se nejedná o dva jevy - jsou to pouze frekvenční a časová specifika téhož jevu!

Frekvenční maskování:

U sinusových průběhů je největší hodnota maskování pro blízké frekvence a frekvence blízké jejich násobkům. Pokud jsou frekvence velmi blízké, maskování mírně klesá (vlivem rázů).

Tóny mohou být maskovány hlukem. Aby byla řeč srozumitelná, musí se její hladina tlaku zvýšit alespoň o 10 dB nad hladinu maskujícího hluku.

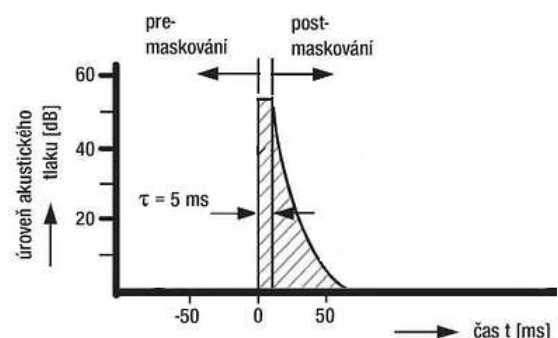
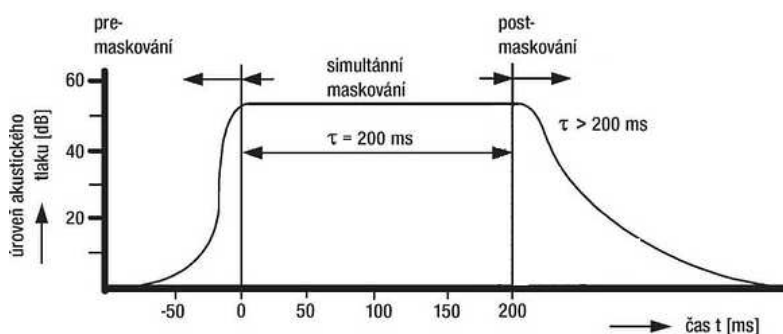


Časové maskování:

Po silném zvuku ucho nevnímá až 200 ms zvuky slabší – tzv. post-maskování.

Také jsou přemaskovány slabé tóny, po nichž zazní silný tón – tzv. pre-maskování.

Délka pre-maskování i post-maskování závisí na délce znění maskovacího zvuku.



Slabé zvuky, které jsou maskovány hlasitějšími, je možné z nahrávky vynechat – lidské ucho tuto redukci nezachytí. Kromě maskování slabšího zvuku silnějším zvukem využívají kompresní metody ještě:

Maskování šumu zvukovým signálem – Kvantování signálu s sebou přináší kvantizační šum (odchylky vlivem zaokrouhlování). Šum je stejně jako slabší zvuky maskován silnějším zvukem. Pokud je maskující zvuk hlasitý, může zamaskovat i větší šum. Proto je možné hlasitější zvuky kvantovat menším počtem bitů.

Kompresce MP3

Algoritmus standardizovaný jako *ISO-MPEG Audio Layer-3* byl vytvořen v roce 1987 Fraunhoferovým Institutem pro integrované obvody ve spolupráci s univerzitou v Erlangenu. Původně metoda pro zpracování zvukové stopy video souborů. Jedná se o ztrátový algoritmus s velmi příznivým kompresním poměrem. Pro kompresi hudby dosahuje uspokojivých výsledků, v oblasti mluveného slova je výsledek horší – často je utlumena první nebo poslední slabika, zkráceny pomlky mezi slovy.

Plná CD informace: $44\,100 \cdot 32 = 1,4$ Mb/s

Standardní MP3: 128 kb/s ... kompresní poměr 11:1 (soubor zabírá cca 11x menší místo na disku)

Je možné používat i větší nebo menší kompresní poměr např.:

256 kb/s (kvalitnější záznam 5:1)

8 kb/s (telefonní kvalita 96:1)

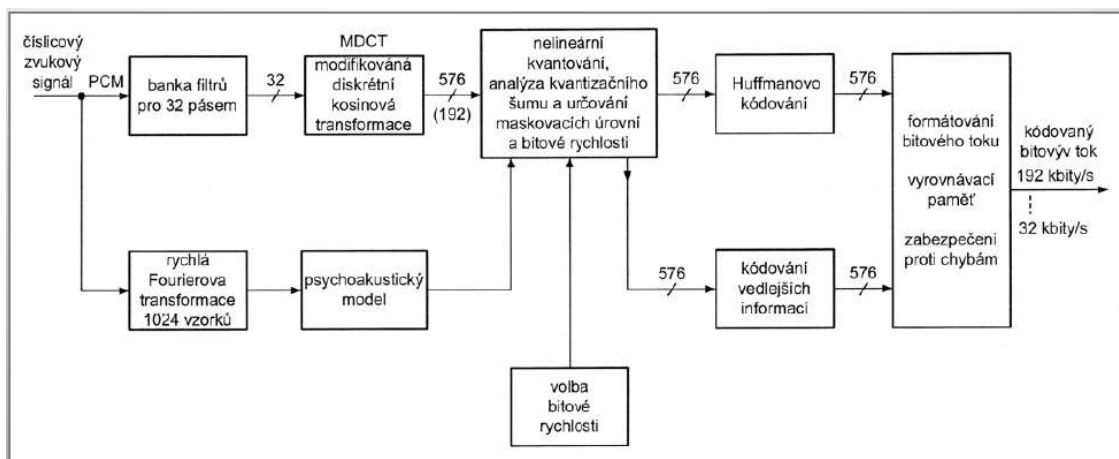
Postup kódování MP3

Nejprve je nahrávka rozdělena na základní skupiny sériových dat – tzv. datové rámce.

Soubor frekvenčních složek každého rámce se rozdělí pomocí banky filtrů do 32 frekvenčních pásem.

Následuje jemná frekvenční analýza každého pásma metodou MDCT (modifikovaná diskretní kosinová transformace) – příliš úzká frekvenční filtrace může způsobit chyby výstupního signálu, proto se počet spektrálních koeficientů může snížit z 18 až na 6.

V každém pásmu je vybrán vzorek s maximální amplitudou. Podle něj jsou porovnáním s psychoakustickým modelem určeny maskovací prahy v jednotlivých pásmech. Jsou vypuštěny maskované vzorky. Je přidělen počet kvantizačních bitů pro každé pásmo tak, aby byl kvantizační šum maskován výší užitečného signálu.



Další redukce datového toku je dosaženo kompresí stereofonního efektu:

Často má skladba totožný signál levého a pravého kanálu – stačí kódovat pouze jeden kanál (mód Dual Chanel).

V pasážích, kdy se levý a pravý kanál velmi liší, je kódován každý samostatně (mód Stereo).

Většinou je při stereo záznamu obsah levého L a pravého R kanálu hodně podobný. Proto se L a R kanály převedou na středový $M=L+R$ a postranní $S=L-R$ (mód Joint Stereo). Více bitů se při kódování přidělí středovému, který obsahuje více důležité informace, postranní udává rozdíly mezi pravým a levým. Při přehrávání jsou opět zpětně dopočítány obsahy levého a pravého kanálu. Obecně platí, že joint sterea je vhodné použít při nižších bitratech.

Skladba je rozdělena na krátké úseky (rámce) – každý rámeček je kódován metodami, které jsou pro něho nejvýhodnější (různé metody komprese stereo signálu, různé vzorkovací frekvence...

Každý rámeček má svoji hlavičku – identifikaci použitého systému kódování, díky tomu lze bezchybně přehrát libovolnou pasáž.

délka [bit]	poz. [bit]	popis
11	31-21	Synchronizace rámce (všechny bity nastaveny)
2	20-19	Verze MPEG: 00 - MPEG verze 2.5 01 - rezervováno 10 - MPEG verze 2 11 - MPEG verze 1
2	18-17	Layer: 00 - rezervováno 01 - Layer-3 10 - Layer-2 11 - Layer-1
1	16	0 - chráněno pomocí CRC, 16bitové CRC následuje hlavičku 1 - není chráněno
4	15-12	Bitrate, viz následující tabulku
2	11-10	Vzorkovací frekvence [kHz] bity MPEG1 MPEG2 MPEG2.5 00 44100 22050 11025 01 48000 24000 12000 10 32000 16000 8000 11 rezervováno
1	9	Doplňující bit (padding) 0 - rámeček není doplněn 1 - rámeček je doplněn jedním bitem navíc
1	8	Tajný bit (neznámý účel)
2	7-6	Mód kanálů 00 - stereo 01 - Joint stereo 10 - dual channel (stereo) 11 - single channel (mono)
1	5	MS stereo (jen pro Joint stereo) 0 - ne 1 - ano
1	4	Intensity stereo (jen pro Joint stereo) 0 - ne 1 - ano
1	3	0 - je chráněno autorským právem 1 - není chráněno autorským právem
1	2	0 - kopie originálního média 1 - originální médium
2	1-0	Zvýraznění: 00 - žádné 01 - 50/15 ms 10 - rezervováno 11 - CCIT J.17

https://dk.upce.cz/bitstream/handle/10195/37622/SedivyM_DekoderFormatu_MH_2010.pdf?sequence=7&isAllowed=y

http://www-kiv.zcu.cz/~herout/html_sbo/mp3/toc.html