

Od analogového signálu k digitálnímu

Zdroj: <http://www.multi-media.kvalitne.cz/>

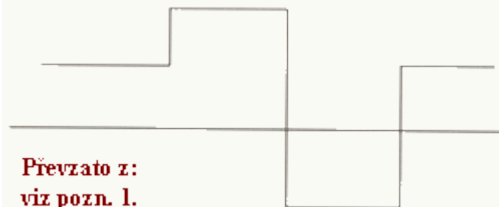
Analogový signál



Obrázek grafu analogového signálu

Obrázek analogového signálu vidíme nalevo. Může se jednat o jakýkoliv jiný graf, ale podstata analogového signálu je v tom, že teoreticky nikdy nebude obsahovat svislé či kolmé úseky (úmyslně píší teoreticky, protože záleží na zvětšení obrázku grafu a pokud by bylo nedostačující, tak by se na první pohled mohl jevit opak).

1.2.5 Digitální signál



Obrázek grafu digitálního signálu

Naopak digitální signál se vyznačuje tím, že se jeho graf skládá jen z vodorovných či svislých úseků. Obrázek grafu digitálního zvuku vidíme napravo. Zase opakují, že vzhled grafu závisí na konkrétním zvuku, pro který daný graf vidíme.

A jak to skloubíme dohromady?

Možná si právě říkáte, proč tohle všechno člověk musí vědět. Veškerá hudba, která je v počítači nebo v MP3 přehrávačích, DVD aj., je právě uložena digitálně. Hodně hudby vzniká analogově (kromě profesionální hudby), např. gramofony, magnetofony aj. A pokud si chceme přehrát svoji oblíbenou skladbu v MP3 přehrávači či na PC, musíme analogově zaznamenaný zvuk převést do digitálního. Také výstupem mnoha mixážních pultů je analogový signál.

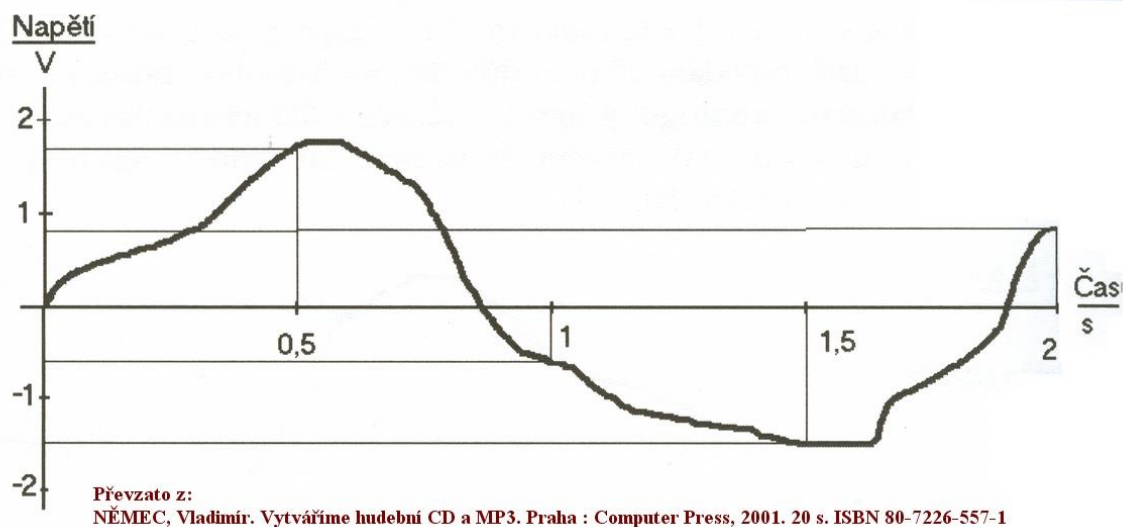
Mixážní pult je zařízení, které nám nejčastěji směšuje více stop do jedné, neboli vytváří jednu stereostopu (= "dává dohromady"). Např. kapela obsahuje bicí, kytaru, keyboard (klavír). Pokud by každý

nástroj měl jeden samostatný mikrofon, vzniknou nám minimálně tři vstupy. Mixážní pult všechny tyto vstupy spojí do jednoho a výsledkem je tedy jedna stereo stopa.

Mixážní pulty se také užívají pro různé úpravy zvuku. Jejich výhoda spočívá v tom, že dokáží upravovat každý vstupní signál (např. mikrofony, keyboardy atd.) zvlášť a přidávat jim různé efekty (mezi základní patří ekvalizace - úprava výšek a hloubek signálu a jiné).

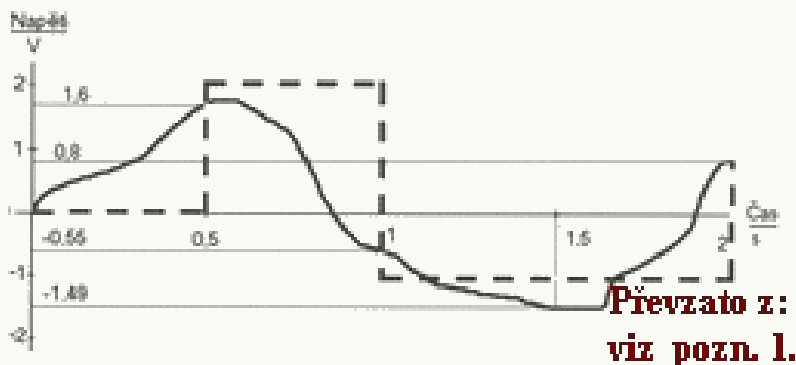
Digitalizace analogového signálu zvuku

Nyní bude řeč o tzv. "pulzně-kódové modulaci (= PCM)" neboli o bezztrátovém způsobu uložení analogových zvukových dat do digitálních. A nejdříve začneme u grafu analogového signálu, kam přiřadíme osu "x" a "y". Na ose x si zvolíme čas (např. po 0,5 sekundách) a na ose "y" napětí ve voltech (např. po 1V).

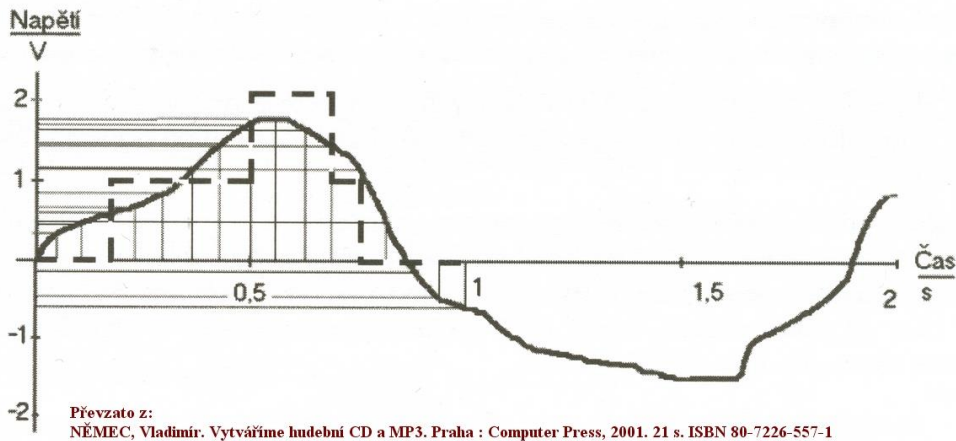


🕒 Obrázek analogového signálu zvuku s osou x a y.

Pokud chápeme předchozí odstavec, tak nyní utvoříme z analogového signálu digitální. A to tak, že se vžijeme do role analogově-digitálního převodníku (A/D), který tuto práci dělá. V čase 0 sekund naměří 0 V. Po 0,5 s znovu změří hodnotu analogového signálu a naměří hodnotu 1,6 V. Ale protože hodnoty na ose y nemůžou nabývat jiných než nadefinovaných hodnot, zaokrouhlí se na 2 V. Další měření se provede zas po 0,5 s (čili po jedné sekundě od začátku měření) a naměří -0,55 V a tuto hodnotu zas zaokrouhlí na -1 V. Stejně to udělá v čase 1,5 s a naměří hodnotu -1,49 V a zaokrouhlí ji na 1 V a v čase 2 s a naměří hodnotu 0,8 V a zaokrouhlí na 1 V. Výsledkem je digitální signál znázorněn přerušovanou čarou.



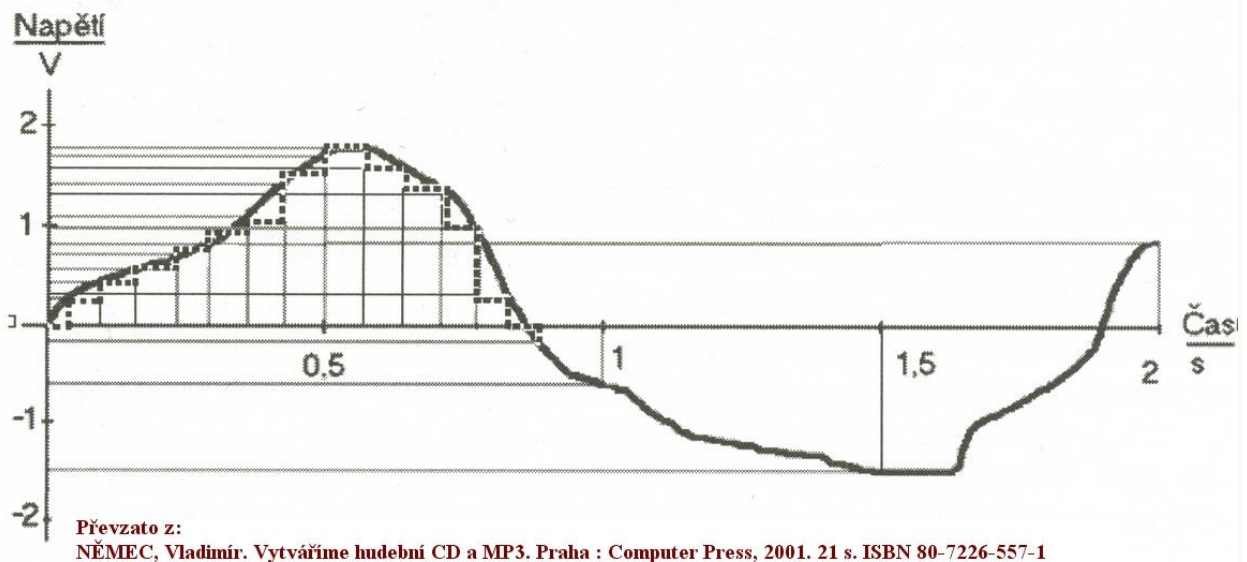
🕒 Obrázek grafu převezení digitálního zvuku z analogového s měřením po 0,5 s.



😊 Obrázek grafu převedení digitálního zvuku z analogového s měřením 32x za sekundu.

Ovšem nějak se nám ten digitální signál změnil, oproti tomu původnímu analogovému. Na obrázku vlevo měříme již 8x častěji (tj. 32x/s), ale přesto výsledek není příliš podobný originálnímu analogovému signálu. Důvod je ten, že máme málo nadefinovaných hodnot na svislé ose y (jen -2, -1, 0, 1, 2 V), a proto má několik měření stejnou hodnotu. Musíme mít tedy na této ose více hodnot, kde měříme s přesností na 0,1 V. Zde se již trochu digitální signál tomu originálnímu analogovému podobá.

Ale pořád bychom toho moc neslyšeli. A co s tím uděláme? To je velice snadné. Zase budeme měřit častěji, třeba 44 100x za sekundu a tím pádem budeme mít na svislé ose 65 000 hodnot. Zde se již bude digitální signál hodně podobat tomu původnímu analogovému. Ještě je třeba podotknout, že digitální signál je posléze ještě vyhlazen, aby nebyl tak schodovitý.



😊 Obrázek grafu převedení digitálního zvuku z analogového při měření po 0,1 V.

Vzorkovací frekvence

A proč právě měříme 44 100x? Tato hodnota je dohodnuta, ale vychází z toho, že lidské ucho je schopno maximálně slyšet 20 000 Hz (20 kHz) a tato hodnota se násobí dvěma (tzv. Shannonův teorém) s ještě nějakou rezervou. A proto hodnota 44 100, která se označuje jako vzorkovací frekvence, znamená počet měření na ose x za určitý čas a udává se v Hz (= Hertz = 1 kmit za sekundu).

Vzorkovací frekvence u lisovaných CD (běžná audio CD profesionálně vyrobena) je 44 000 Hz, přesněji řečeno 44 100Hz. Ovšem v počítači můžeme hudbu ukládat i na nižší hodnoty, např.: 22 050 či 11 025Hz. Zde platí: Čím nižší vzorkovací frekvence, tím horší kvalita nahrávky, ale zase tím menší místo nám zabere např. na pevném disku (na HDD). Je tedy na nás, co si zvolíme za kompromis. Obecně platí, že programy na vypalování CD si automaticky upraví vzorkovací frekvenci na CD kvalitu, čili na 44 100, i když by původní byla nižší. Ovšem i tak se nám kvalita nezvýší, protože počítač hudbu jen dopočítá podle obecně platných pravidel (algoritmů) a ne podle toho, co chceme slyšet ve výsledku.

A co znamená bitová hloubka, tedy bitrate (bit rate)?

Nyní jsme se bavili jen o hodnotách na vodorovné ose (osa x), ale nyní se pojdme podívat na hodnoty na ose y. Na ose y je počet bitů, které jsou zpracovány za určitou časovou jednotku (obvykle sekunda). Můžeme se setkat se zápisem kbps (= kilobit per second), kbit/s (= kilobitů za sekundu). Vzorkovací frekvence 44 100 Hz odpovídá 16 bitům a 16 bitů je přibližně 65 000 hodnot ($2^{16} = 65536$, umocňuje se dvojkou, protože jeden bit může nabýt dvou hodnot. A to 1 nebo 0). A právě počet bitů (počet max. hodnot na ose y, kterých může funkce nabýt - graf zvuku) zpracovávaných za určitý čas je bit rate.

Kompresce zvukových souborů

Kompresce obecně je proces, kdy se zmenšuje objem, a ani v počítačové terminologii tomu není jinak. Zde v tomto konkrétním případě zmenšujeme velikost zvukového souboru, respektive dat, aby nezabíral tolik místa na pevném disku či jiném médiu a dalo se s ním lépe manipulovat. Rozlišujeme tyto základní typy:

- Kompresce neztrátová
- Kompresce ztrátová

Kompresce neztrátová

Tento druh komprese znamená, že se neztrácejí data a lze i po dokončení komprimace znovu obnovit původní data beze ztráty. Hodně lidí si říkají, jak je možné, něco komprimovat beze ztráty. Zkusme si to přiblížit na následujícím příkladu:

Potřebujeme koupit 5 housek. A můžeme jít do obchodu a říci, že chci 1 + 1 + 1 + 1 + 1 housek a nebo říci rovnou, že chci 5 housek. V prvním případě by z nás prodavač moc velkou radost neměl. Ale kdybychom řekli rovnou, že chceme 5 housek, tak ušetříme čas sobě i druhým. To určitě každý cítíme. A to máme jen 5 housek, teď si představme, že bychom jich chtěli koupit stovky, to bychom nakupovali celý den jen housky. A přesně takhle to funguje při bezztrátové kompresi. Data, která potřebujeme uložit, jsou ty naše "housky". A pak záleží na konkrétním kodeku (*kodek pochází ze slov „kodér a dekodér“, respektive komprese a dekomprese, neboli program schopný komprimovat či dekomprimovat (= transformovat) data (datový proud)*), který uijeme a on rozhodne, jak bude probíhat komprese. Samozřejmě se pak zohledňuje kvalita a typ dat, co ukládáme. Kompresních algoritmů je celá řada a není cílem je tu popisovat, jde mi jen o určitý platný princip.

Kompresa ztrátová

Naopak komprese ztrátová je příklad, kdy se data ztrácejí nenávratně pryč a nelze je obnovit do původní podoby.

Samozřejmě, že pokud budeme chtít obrázek např. zvětšit, existuje spousta možností, jak "dodělat" ztracená data. Tento postup se nazývá převzorkování či interpolace (hlavně u grafických souborů). Je nutné však podotknout, že takto vzniklá data jsou jakým si "výmyslem" algoritmů, který na základě obecně platných pravidel dané hodnoty dopočítává. Tato data se ovšem neshodují s původními daty, jen se je snaží napodobit. Je řada ztrátových kodeků, které se liší kvalitou komprese a typu komprimovaných dat. Neboli každý kodek má své optimální použití, např. AMR je dobrý na zachycení lidského hlasu a WMA je dobrý na hudbu (podrobněji dále).

Kompresní poměr

je poměr velikosti souboru před kompresí a velikosti souboru po kompresi v bajtech. Udává účinnost kompresního algoritmu.

Formáty souborů

Nyní si představíme základní typy formátů souborů a rozdílů mezi nimi. Cílem tohoto přehledu je stručný přehled formátů běžné praxe na středoškolské úrovni a vysvětlení základních rozdílů mezi nimi.

- **WAVE (.wav)**

V praxi je formát WAVE převážně v nekomprimované formě (samozřejmě zde již hovořím o digitálně uloženém zvuku a neberu zde v úvahu ztrátu způsobenou převodem z analogového do digitálního signálu) v tzv. [pulzně kódové modulaci \(= PCM\)](#). Přípona wav je odvozena z "Waveform audio format", který vytvořily firmy IBM a Microsoft. Vzhledem k tomu, že zvuk takto uložený je bezztrátový, používá se nejčastěji při zpracování zvuku. Ovšem velkou nevýhodou je jeho náročnost na množství uložených dat, a proto se velikost výsledného souboru pohybuje v řádech desítek MB. Proto se raději volí komprimované soubory, např. formát MPEG (.mp3).

- **MPEG (.mp3)**

Formát mp3 je založený na kompresním algoritmu MPEG (Motion Picture Experts Group). Byl vyvinut hlavně proto, aby se minimalizovala velikost výsledného souboru. Kompresa je ztrátová, proto se nehodí pro profesionální práci s hudbou, ale při dostatečně velkém [bit rate](#) (alespoň tak 320 kbit/s) jen málo kdo uslyší rozdíly oproti wav formátu. Ovšem podstatně menší velikost souboru přehlédne jen málo kdo. Tento formát je dnes populární pro své malé nároky na velikost a s tím spojenou dobrou manipulací. Kompresa probíhá podle teorie "smažu, co člověk stejně neuslyší". S tímto způsobem uložení dat přišel německý vědec Karlheinz Brandenburg a jeho vědecký tým. Mezi jeho klady patří i přidání tzv. [ID3 tagu](#), který slouží k uložení různých informací přímo do souboru mp3 (např.: titul, interpret, album, rok vydání, číslo stopy, žánr a komentář) a tyto informace se pak následně zobrazují při přehrávání. Ovšem mezi jeho slabiny patří ukládání mluveného slova, protože může zkracovat pauzy mezi jednotlivými slovy či potlačit první nebo poslední slabiky slov. Pro účely uložení řeči lze použít kodek AMR.

- **AMR (.amr)**

AMR (= Adaptive Multi-Rate compression) je kodek pro ukládání zvuku především pro mluvené slovo. Jeho největší užití je v telekomunikacích (např. mobilní sítě). Proto dnes většina mobilních zařízení umí s tímto formátem pracovat, protože dosahuje dobrých výsledků. Vzorkovací frekvence je 8 kHz při 13 bitech a po každých 20 ms je možné tento datový tok (= [bit rate](#)) změnit, proto název Adaptive Multi-Rate. Ovšem jeho slabou stránkou je ukládání hudby. Pro tyto účely se používá již zmíněný MPEG kodek nebo Windows media audio (.wma).

- **WMA (.wma, .asf)**

Windows Media Audio je formát zvuku, který byl standardně zamýšlen pro platformu Windows, kde je také tento kodek implementován do hudebního přehrávače Windows Media Player. Standardně tento formát je ztrátový, ale jsou i zvláštní kodeky z této rodiny pro bezztrátovou kompresi. Nyní se ale postupně dostává i do jiných prostředích, a to jak softwarových (operační systém Linux), tak i hardwarových (různá zařízení - DVD přehrávače, aj.), možná i díky rozšíření platformy Windows. Ovšem nevýhodou je, že není volně šiřitelný a nemá otevřený zdrojový kód, čili uživatel je závislý na tom, jaké možnosti nastavení "dostane" od Microsoftu. A těmi právě wma příliš nedisponuje. Říká se, že vznikl proto, aby Microsoft nemusel platit integraci mp3 do Windows. Dnes je považován za běžný standard ve světě hudebních souborů a spíše "soupeří" s Applovým AAC formátem.

Informace z jiných zdrojů:

<http://fonoklub.wz.cz/dig.html>

Digitální záznam zvuku je metoda kdysi používaná jen profesionály a dnes se již bez ní neobejdeme, pokud chceme něco kvalitně nahrát třeba i doma. Zatímco u analogového záznamu je kvalita záznamu daná souhrnem nedokonalostí celého přenosového řetězce (např. šum a zkreslení magnetofonů a pásků použitých pro prvotní záznam a kopírování), u digitálního záznamu bez komprese dat je kvalita záznamu daná převážně nastavením parametrů záznamu. Při standardní vzorkovací rychlosti 44,1 KHz a 16 bitovém kódování je kvalita omezena zpravidla pouze vlastnostmi mikrofonu a nežádoucími hluky v místnosti, kde nahráváme.

Princip digitalizace zvuku

Signálové napětí je v pravidelných časových intervalech vzorkováno a převáděno na odpovídající číslo. Zatímco signálové napětí má libovolnou hodnotu, převodem na celé číslo dochází k zaokrouhlování. To se projeví jako kvantizační zkreslení, nebo též kvantizační hluk. Poměr maximálního signálového napětí a kvantizačního hluku je dán maximálním číslem, které je možno převodníkem vytvořit a zaznamenat. Pro kvalitní záznam zvuku se používá 16 bitové kódování, kde maximální číslo 65535 znamená, že převodník rozliší 65535 napěťových úrovní a to odpovídá rozsahu zaznamenaných úrovní 96 dB. Pro méně náročné účely, např. při telefonním spojení v pevné síti, se používá 8 bitové kódování. Rozsah úrovní je pouze 48 dB, ale množství přenášených dat je poloviční. Skutečný rozsah je cca o 3 dB menší, protože kvantizační hluk se v dekodéru záměrně překrývá šumem.

Maximální přenášený kmitočet je dán rychlostí vzorkování převodníku. Pro nejvyšší přenášený kmitočet se musí přenést minimálně dva vzorky na periodu signálu. Vzorkovací kmitočet 44,1 KHz použitý u CD dává frekvenční rozsah do 20 KHz, ale pro běžnou rozhlasovou kvalitu s frekvenčním rozsahem do 15 KHz se používá též vzorkování 32 KHz. Naopak pro dosažení ještě vyšší kvality záznamu se používá vzorkování 48 KHz. Některé moderní systémy podporují též 32 bitové kódování a vzorkovací kmitočet až do 192 KHz. Kvalitativně to ale není podstatný přínos a záznam pouze zabírá nesrovnatelně větší místo na záznamovém médiu. Pro porovnání, v telefonech pevné sítě se používá vzorkování pouze 8 KHz a 8 bitové kódování.

Kvalita převodníků - aneb svět chce být klamán

Pokud by převodníky fungovaly tak jak bychom očekávali, tedy při vzorkovací frekvenci 44,1 KHz by převodník udělal 44100 samostatných převodů např. s 16 bitovým rozlišením, byla by digitální technika opravdu kvalitní. Takové převodníky se běžně používají v měřicí technice, kde obvykle buď nejde o rychlost, nebo je nižší rozlišení. Audio převodníky obvykle pracují na principu sigma-delta, protože je to jednoduché. Komparátor má výstup připojený na integrační článek a podle toho, jestli napětí na jeho výstupu je menší nebo větší než vstupní, komparátor se v pravidelných intervalech překlápí na 0 nebo na jedničku. Pak už se jen spočítá poměr jedniček a nul a převede se na číslo. Problém je ta vynucená pravidelnost, dochází tam proti vstupnímu signálu k malým rozdílům (kvantizační šum).

Pokud chceme rozlišit na vzorkovací frekvenci 44,1 KHz 65000 úrovní, museli bychom v rámci jednoho převodu vytvořit 65000 časových intervalů a z nich počítat onen průměr. Vzorkovací kmitočet by vyšel 2,87 GHz a komparátor by při tom musel mít požadovanou přesnost, aby dokázal rozlišit 65000 úrovní (těch 16 bitů). Tohle je když už snadné nerealné, tak alespoň hodně drahé. A tak se spokojíme s podstatně (100x až 1000x) nižším kmitočtem a s tím, že se integruje přes více vzorků a zavádějí se různé zpětné vazby vyšších řádů. Na vyšších kmitočtech podstatně stoupá šum a zavádějí se různá opatření, aby převodník neprodukoval příliš rušivé hluky. A tak máme převodníky levné a kvalitnější a nejsou úplně mimo mísu ani názory, že kvalitní studiový analogový magnetofon není až tak úplně špatný. U kvalitních převodníků je sice určitá možnost zlepšení při vyšším vzorkovacím kmitočtu a vyšším počtu bitů, ale není to o tolik lepší, jak bychom čekali.

Kvalita zvuku a objem dat

Zatímco kvalitní analogový magnetofon zaznamená maximální frekvenci 15 až 20 KHz, u kvalitního digitálního záznamu s 16 bitovým kódováním a vzorkováním 44,1 KHz potřebujeme přenést $44100 \times 16 = 705600$ bitů za vteřinu při mono a 1411200 bitů za vteřinu při stereo záznamu. K tomu je zapotřebí šířka pásma alespoň 800 KHz. Na rozdíl od analogového záznamu zde stačí rozlišit pouze logickou jedničku a nulu a na to stačí poměr signálu k šumu jen 15 až 20 dB. Proto se pro digitální záznam používají magnetofony s rotační hlavou, které zaznamenávají vysokou rychlostí do velmi úzké stopy. Existovaly převodníky, které lze připojit k videorekordéru a zvukový signál se zaznamenával místo obrazu. Videorekordéry s HIFI stereo zvukem zaznamenávaly zvuk digitálně do šikmé stopy společně s obrazem. Zmenšením, zjednodušením a zdokonalením tohoto systému vznikl systém DAT (Digital Audio Tape) používaný pro digitální záznam zvuku na kazety.

Tyto systémy, stejně jako minidisk s magnetooptickým záznamem a s kompresí Atrac, jsou již minulost. V současné době se používá záznam do flash paměti a využívají se paměťové karty používané také ve foťácích. Je to malé, levné a praktické. Kvalitní přístroje umožňují i záznam bez komprese, diktafony používají kompresi MP3. Záznamník bez pohyblivých částí je nehlukný a proto nahrává kvalitně i přes vestavěný mikrofon. Data se jednoduše přepíší do počítače přes USB.

Na co si dát pozor

Při digitálním zpracování zvuku můžeme záznam snadno znehodnotit. Digitální záznam můžeme na rozdíl od analogového kopírovat neomezeně a bez ztráty kvality. Musí se však jednat o digitální kopírování a musí být zachován formát záznamu. Pokud např. záznam z CD se vzorkovací frekvencí 44,1 KHz převzorkujeme na "kvalitnější" frekvenci 48 KHz, záznam tím můžeme znehodnotit, protože rozdíl vzorkovacích frekvencí (v tomto případě 3,9 KHz) se v převzorkovaném zvuku může projevit jako nežádoucí interference. Výsledná kvalita snímku závisí na kvalitě přepočítacího algoritmu.

Dalším problémem je analogové kopírování. Běžná CD mechanika v počítači má analogový převodník a ve starších počítačích je se zvukovou kartou spojena analogově. Při kopírování CD na disk je takto získaný analogový signál zvukovou kartou opět digitalizován, čímž se snižuje jeho kvalita. Pro získání původní kvality je nutné použít grabovací program, který čte data z audio CD přímo a ukládá je na disk. Novější systémy už používají grabování automaticky, protože pod Windows XP hraje CD mechanika přes zvukovou kartu, i když jsem zapomněl zapojit propojku.

Kompresce dat

Poměrně velkou šířku pásma potřebnou pro digitální záznam je možné snížit vhodnou kompresí dat. V zásadě je možné použít bezztrátovou kompresi používanou pro počítačová data. Výhodou je možnost obnovení původního obsahu dat, tedy komprese nesnižuje kvalitu záznamu. Nevýhodou je malý dosažitelný poměr komprese, proto se tato metoda nepoužívá.

Radikální snížení objemu dat, běžně 1:10, se dosahuje ztrátovou kompresí. Ta vychází z principu, že je potřeba přenést jen významné složky zvuku, které nejsou jinými zvuky maskovány. Nejrozšířenější metoda komprese dat [MP3](#) vychází z komprese MPEG používané též u digitálního zpracování obrazu. Jiná, ale principem podobná komprese ATRAC se používá u minidisků. Proto dosahují o něco nižší kvalitu zvuku, než CD a magnetofony DAT.

Při kódování MP3 kvalita výsledného snímku silně závislá na použitém kodéru, zatímco všechny dekodéry pracují stejně a při bezchybné funkci nemají na kvalitu vliv. Tato jednotnost datového formátu nejvíce přispěla k jeho rozšíření. Zajímavý projekt je [kodér LAME](#). Zatímco většina kódovacích programů je šířena za peníze, LAME je dnes již kolektivní práce převážně různých univerzit. Bylo to umožněno zveřejněním zdrojového kódu a šířením programu podle licence GNU. Podle této licence program může každý používat bezplatně a má přístup k jeho zdrojovému kódu. Pokud na základě tohoto zdrojového kódu vytvoří novou verzi, musí zveřejnit též zdrojový kód nové verze. Stejným způsobem je též šířen a vyvíjen operační systém [LINUX](#).

Komprese MP3 ve spojení s technikou modemů využívajících maximálně rychlost telefonní linky dokáže téměř nemožné. Zvuk s frekvenčním rozsahem 15 KHz zakódovaný do MP3 s bitovou rychlostí 48000 bitů/sec lze při rychlosti připojení na internet 56000 Bd s rezervou přenést v reálném čase po telefonu. Při tom šířka pásma telefonu je pouze 3 KHz.

Originální způsob komprese dat je používán v mobilních telefonech [GSM](#). Jedná se v podstatě o napodobení lidského hlasu, který má vždy pouze jeden základní kmitočet a jeho násobky. Proto když si u mobilu necháme hrát rádio, na druhé straně z toho vylezou neuvěřitelné zvuky, které se těm původním ani trochu nepodobají.