

DIGITÁLNÍ ZVUK NA POČÍTAČÍCH

Technologický vývoj a výroba dostatečně rychlých, přesných a levných A/D a D/A převodníků umožnily rozvoj digitálního zpracování zvuku tak, jak jej známe dnes. V oblasti komerčního využití je to především kompaktní disk (CD) následovaný technologií DAT, posléze minidisk (MD) a konečně CD-ROM a digitální videodisk (DVD). Kromě zásadní změny v kvalitativních parametrech, kterou digitalizace zvukového signálu a digitalizace obecně přinesla, to znamená především zcela nový, principiálně odlišný přenos a záznam zvukového signálu a také jeho zpracování, které se díky své povaze přesunulo na platformu počítačů a příslušného softwaru.

Formy záznamu zvukového signálu

Připomeňme si jen ve stručnosti některé „nešvary“ analogového signálu a problémy, které na něj čekají během jeho zpracování a případného záznamu. Je to především všudypřítomný šum a brum (noise), nelineární zkreslení, omezený kmitočtový rozsah a jeho nelinearita v daném pásmu. Všechny tyto parametry jsou ovlivňovány zásadním způsobem při každé operaci: zesílení, kmitočtové a dynamické korekci, záznamu a jeho přepisu, editaci. Zjednodušeně řečeno, vliv na všechny tyto parametry má kvalita jednotlivých prvků celého analogového řetězce (zesilovače, korektory, zařízení pro dynamickou úpravu signálu – limity, kompresory, expandéry a záznamová zařízení). Skutečně kvalitní profesionální zařízení jsou proto velmi drahá, neboť v takto pojetém analogovém řešení nelze nic ošidit. Platí téměř přímá úměra kvalita – cena.

U digitálního zpracování je to trochu jinak. To, co zásadním a podstatným způsobem určuje výslednou kvalitu zvukového záznamu, je především vstupní převod analogové veličiny na digitální. Zde jsou kladeny hlavní požadavky na kvalitu, kterou takový převod musí splnit. Kromě patřičné linearit v daném dynamickém rozsahu příslušných veličin, je to pak hlavně rychlost převodu, který je třeba provádět v reálném čase. Při dnešních technologiích však už není problém zpracovávání akustických signálů mnohem náročnější je problematika videa, jeho střihu a efektů, kde jsou datové toky ještě zhruba o dva řády vyšší. Všechny další operace jako např. editace, kmitočtové a dynamické úpravy včetně přenosu záznamu a jeho přepisu, pak už nemají při zachování daného formátu digitálního signálu vliv na jeho kvalitu. Problémy šumu, nelineárního zkreslení, omezení a deformace dynamiky

zůstaly na analogové straně a tak se jimi při digitálním zpracování nemusíme zabývat. Proto se třeba první a desátá kopie v podstatě neliší, což u analogového zpracování (i profesionálního) nebylo myslitelné. To samozřejmě neznamená, že digitální zpracování je bez problémů, ty se však řeší v realizovatelné jiné, kvalitativně vyšší hladině, nehledě na nové možnosti, které v analogové sféře byly jen těžko nebo vůbec. V současné době je i cena kvalitativně srovnatelných digitálních zařízení oproti zařízením analogovým nižší.

Formáty digitálního zvukového signálu

Úvodem připomeňme, že u digitálního signálu se na rozdíl od analogového, který přenáší spojitou, kontinuálně se měnící hodnotu dané veličiny, přenáší v jednotce času určitý počet nespojitých (diskrétních) hodnot, které tento signál popisují. Základem digitálního zpracování je tedy „rozsekání“ spojitého signálu na jednotlivé vzorky a jejich následné kódování do určitých standardů tak, aby byly později snadno zpracovatelné. Toto kódování rovněž podstatnou měrou přispívá k zabezpečení a případné opravě chyb u takto získaných dat i k jejich odolnosti proti poruchám při přenosu a záznamu (různá rušení, drop-outy ap.).

Z hlediska digitálního signálu a jeho kvality jsou určující vzorkovací kmitočty a bitová šířka nebo hloubka jednotlivých vzorků. Nejrozšířenější a neznámější je formát užívaný u CD zvukových kompaktních disků, kterým je v zásadě definován často používaný termín CD kvalita (např. u zvukových karet počítačů a programových produktů). Jde o šestnáctibitové vzorkování 44,1 kHz, které zaručuje lineární kmitočtový rozsah do 20 kHz, dynamiku 96 dB a nelineární zkreslení řádu tisícín procenta. Tak jako vždy při zrodu světových standardů a norem nebyly rozhodující jen technické, případně konkurenční, ale často důležitější ekonomické a „obchodní“ důvody. Proto vznikly ještě formáty 48 kHz/16 bitů a úsporný 32kHz/14/12 bitů užívané pro satelitní rozhlas a long play režim DAT magnetofonů a minidisků. Při nástupu zvuku do oblasti počítačů (nikoli jeho profesionálního zpracování) pak přibyly další formáty, tj. nejprve osmibitové a pak i šestnáctibitové při vzorkování 11 kHz a 22 kHz. Se zdokonalováním příslušných čipů a zvukových karet pak již zmíněná plná CD kvalita, tj. 44,1 kHz/16 bitů stereo.

Pro běžnou praxi bylo nutné vytvořit ještě další normy, které popisují formáty pro přenos digitálních signálů a vzájemné propojování přístrojů. Mezi nejužívanější standardy patří např. AES/EBU a S/PDIF, které popisují formát signálu pro přenos. Pro definici

struktury datových souborů se nejčastěji používají formáty WAV, AIFF, SDII, QuickTime, 'snd', Rew.

V následujících ukázkách uslyšíte dva jednoduché testy, které umožní subjektivní porovnání kvality různého vzorkování a základní prověření zvukové cesty vašeho PC. Zda však uslyšíte vše (a v odpovídající kvalitě) záleží také na technických parametrech vašeho počítače, především zvukové karty a reproduktorů.

Jako první uslyšíte zkušební sinusový signál o kmitočtu 1 kHz s plnou úrovní v obou stereofonních kanálech ve fázi i protifázi a v jednotlivých kanálech separátně. [TEST1] Můžete tak jednoduše otestovat svůj počítač případně i s připojeným zesilovačem. Další test obsahuje krátké zvukové ukázky, které umožní bezprostřední porovnání kvality jednotlivých kvantizací a vzorkování na konkrétní zvukové kartě, a to ve stereo- i monoverzi. [TEST2]

Následují krátké stereo a mono hudební ukázky s rozdílným vzorkovacím kmitočtem a bitovou šířkou (v pořadí od nejvyšší kvality):

- standard, vzorkování 44,1 kHz, 16 bitů, stereo [44_16_S.WAV]
- vzorkování 22 kHz, 16 b, stereo [22_16_S.WAV]
- vzorkování 11 kHz, 16 b, stereo [11_16_S.WAV]
- vzorkování 22 kHz, 8 b, stereo [22_8_S.WAV]
- vzorkování 11 kHz, 8 b, stereo [11_8_S.WAV]
- komprimovaný signál 22 kHz, 16 bitů do 4 bitů, stereo [22_4_S.WAV]
- vzorkování 44,1 kHz, 16 b, mono [44_16_M.WAV]
- vzorkování 22 kHz, 16 b, mono [22_16_M.WAV]
- vzorkování 22 kHz, 8 b, mono [22_8_M.WAV]
- vzorkování 11 kHz, 8 b, mono. [11_8_M.WAV]

Jak jste si možná povšimli, vzorkování určuje především nejvyšší přenášený kmitočet zvukového signálu. Teoreticky platí, že nejvyšší přenášený kmitočet zvukového signálu je roven polovině vzorkovacího kmitočtu. Lidově řečeno, s nižším vzorkovacím kmitočtem se ve zvuku ztrácí „výšky“.

Bitová hloubka digitálního signálu určuje strukturu dynamického rozlišení a zkreslení zvukového signálu. Při osmibitové struktuře je to pouze 256 kvantizačních stupňů, do kterých

se musí rozložit celá amplituda signálu. Navíc, při lineární kvantizaci je tento rozklad při malých amplitudách zvukového signálu odpovídajícího tichým pasážím příliš hrubý, což se projevuje tzv. kvantizačním šumem, ve zvuku vnímaným jako určitá zrnitost nebo hrubost. To je přijímáno jako určité zkreslení, které původní signál neobsahoval. U 16bitového rozlišení je pak těchto hladin 65 536 a problémy kvantizačního šumu a zkreslení jsou řádově menší, v běžné praxi zanedbatelné. S kvantizací také souvisí dosažitelný odstup signál / šum a dynamika signálu, která u 16 bitů dosahuje hodnoty 96 decibelů, zatímco osmibitové rozlišení poskytuje dynamiku jen 48 dB. Tato hodnota by se mohla zdát dostačující, avšak subjektivní dojem je zhoršen vysokou úrovní kvantizačního šumu. Tuto situaci částečně řeší nelineární kvantizace, kdy jsou jednotlivé kvantizační stupně nestejně velké a úměrné amplitudě zpracovávaného zvukového signálu. Pro malé amplitudy je toto rozlišení jemnější a s rostoucí velikostí roste i jeho hrubost. Tato metoda se používá u 12- a 14bitových hloubek, u osmibitového rozlišení se neuzívá. Subjektivně poskytuje dobré výsledky, viz např. long play režim DAT magnetofonů, kde přináší prodloužení hrací doby na dvojnásobek.

Základní údaje o formátech, datových objemech a časové délce jednotlivých ukázek, které jste si poslechli, uvádí tabulka:

Název souboru (WAV)	Vzorkovací kmitočet [kHz]	Bitová šířka (hloubka) [bit]	Stereo - Mono	Velikost souboru [MB]	Hrací doba [sec]
44_16_S	44,1	16	S	5,7	34,0
22_16_S	22,05	16	S	2,6	31,7
11_16_S	11,025	16	S	1,3	31,8
22_8_S	22,05	8	S	1,3	31,8
11_8_S	11,025	8	S	0,704	31,5
22_4_S	22,05	4	S	0,717	32,2
44_16_M	44,1	16	M	3,2	38,1
22_16_M	22,05	16	M	1,5	35,9
22_8_M	22,05	8	M	0,8	35,8
11_8_M	11,025	8	M	0,416	36,1

Zde uvidíte, jaký vliv na objem dat mají, při obdobné délce ukázek, jednotlivé parametry – vzorkovací kmitočet, bitová šířka a stereo/mono. Poznamenejme, že 1 minuta stereozáznamu 44 KHz/16 bitů potřebuje prostor 10 MB. Pro poloviční vzorkovací kmitočet nebo poloviční bitovou šířku, případně jednokanálový zvuk (mono) je tento objem poloviční, příp. čtvrtinový, osminový, atd.

To bylo jen letmé ohlédnutí a dílčí přiblížení problematiky digitálního zvuku na platformě multimediálního využití. Jeho možnosti, zdá se, však teprve nastupují – hlasová pošta a komunikace s počítačem, použití Internetu i lokálních sítí a třeba i poslech digitálního satelitního vysílání.

Jindřich Michalík