

Počítačové zpracování řeči

Luděk Bártek

Fakulta Informatiky
Masarykova Univerzita Brno

podzim 2014

Obsah

1 Digitalizace akustického signálu

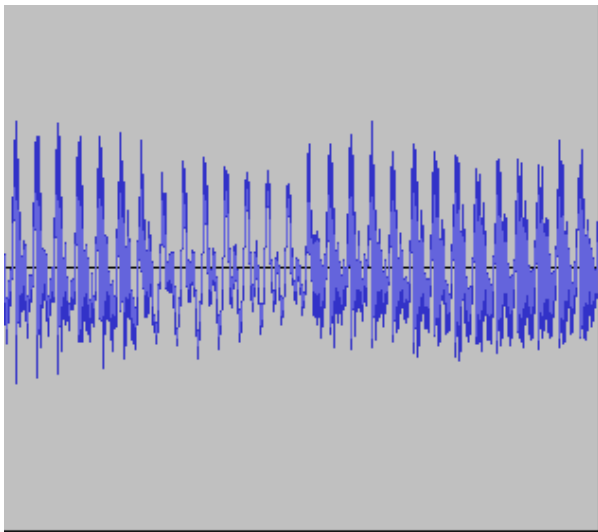
Digitalizace zvuku

- Cíl - převod spojitého signálu na posloupnost digitálních hodnot vhodných pro uchování v počítači.
- Postup digitalizace:
 - 1 Vzorkování - převod reálných vstupních hodnot na posloupnost diskrétních reálných čísel.
 - 2 Kvantizace - převod posloupnosti reálných čísel na posloupnost celých čísel.
 - 3 Kódování - způsob uložení a kódování posloupnosti celočíselných hodnot získaných v kroku 2.

Vzorkování

- Transformace spojitého časově závislého signálu $s(t)$ na časově diskrétní posloupnost $s_n(T) = 0, 1, 2, \dots$
 - T - perioda vzorkování.
 - Pokud nemá dojít ke ztrátě informace, musí být vzorkovací frekvence aspoň dvojnásobkem nejvyšší frekvence, která je signálu obsažena.
- Po čase T je sejmuta a dána na výstup (ke kvantizaci) hodnota ze vstupního snímače.
 - většinou okamžitá úroveň napětí nebo proudu na vstupu.
- Oblasti použití
 - digitální zpracování zvuku
 - audio CD
 - mp3 - navíc použita ztrátová komprese
 - miniDisc - navíc použita ztrátová komprese ATRAC
 - DAT
 - ...
 - digitální zpracování signálu obecně (digitalizace dat z různých analogových měřicích zařízení, digitální zpracování obrazu, ...)

Ukázka digitalizovaného signálu



Shannonův vzorkovací teorém

- Analogový signál $s(t)$ lze rekonstruovat z hodnot vzorků $s_n(T)$ následovně:

$$s(t) = \sum_{n=-\infty}^{\infty} s_n(T) \frac{\sin(\pi(\frac{t}{T} - n))}{\pi(\frac{t}{T} - n)}$$

právě tehdy když je vzorkovací frekvence alespoň dvojnásobkem nejvyšší frekvence obsažené ve vstupním signálu.

- Důsledky:
 - Vzorkovací frekvence by měla být alespoň dvojnásobkem nejvyšší frekvence vstupního signálu.
 - Je-li menší dochází ke zkreslení složek vyšších frekvencí.
 - Spor příznivců a odpůrců audio CD - je 44kHz dostačující vzorkovací frekvence pro hudbu?

Kvantizace

- Převod reálných navzorkovaných hodnot na celočíselné hodnoty.
- Počet celočíselných hodnot = počet úrovní kvantování
 - 256
 - 65 536
 - 16 777 216
- Kvantizační krok - reálný interval přiřazený kvantizované jednotce.
 - Na vstupu je signál s amplitudou 128 mA (-128 - 127 mA).
 - 8bitová kvantizace - 256 kvantizačních úrovní
 - kvantizační krok = $\frac{256[mA]}{256} = 1[mA]$.
- Běžně používané kvantizace – 8, 16, 24, 32 bitů.
- Realizováno pomocí A/D převodníků
 - součást zvukových karet
 - mobilních telefonů
 - ...

Běžně používané parametry digitalizace zvuku

- Vzorkovací frekvence:
 - 8 kHz - telefonní kvalita
 - 16 kHz - běžná řeč
 - 22 kHz - rozhlasová kvalita
 - 44 kHz - audio CD
 - 48 kHz - DVD
- Kvantizace:
 - 8 bitů
 - 16 bitů
 - 24 bitů
 - 32/64 bitů v pohyblivé řádové čárce
- Počet audio kanálů
 - 1
 - 2
 - 4
 - 6 (5.1, 5 směrových kanálů + basy)

Způsoby kódování signálu

- PCM - přímé ukládání hodnot získaných kvantizací.
 - Výhody – jednoduché na zpracování, nedochází k další ztrátě informací.
 - Nevýhody:
 - často malé rozdíly mezi hodnotami sousedních vzorků – značná redundance dat,
 - konstantní hodnota kvantizačního kroku (závisí na parametrech AD převodníku) – v případě malé amplitudy vstupního signálu – ztráta informace (signál nepřekročí kvantizační krok), v případě velké amplitudy – hodnota překročí rozsah – zkreslení signálu. Oba případy brání kvalitní rekonstrukci původního signálu.

Kódování průběhu vlny

Řešení nevýhod PCM

- Diferenční PCM
 - Uchovávání rozdílů sousedních vzorků místo uchovávání jejich hodnot.
 - Hodnota rozdílu bývá podstatně menší než hodnota vzorku - lze uchovat pomocí méně bitů.
- Adaptivní PCM
 - Kvantizační krok se určuje na základě amplitudy vstupního signálu.