

# METHODS FOR LOGARITMICAL COMPANDING OF SPEECH SIGNALS

Ondřej HÜTTL, Bachelor Degree Programme (3)  
Dept. of Radio Electronics, FEEC, BUT  
E-mail: xhuttl00@stud.feec.vutbr.cz

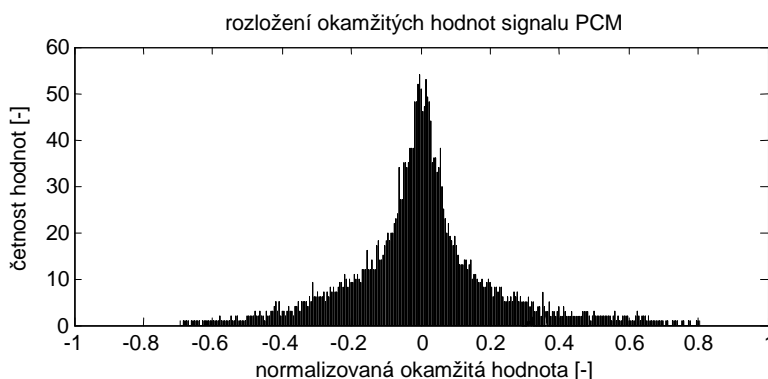
Supervised by: Dr. Milan Sigmund

## ABSTRACT

The presented paper deals with two most common methods of logarithmical coding used for compression of speech signals, namely the  $\mu$ -Law and the A-Law.

## 1 ÚVOD

Logaritmičká kompanze signálů je jednou z prvních metod, vyvinutých v USA, ke komprimaci především řečových signálů. Využívá logaritmičkého vnímání hlasitosti lidským sluchem, které způsobuje, že změna intenzity je více vnímána při nízkých úrovních signálu. Dalším důvodem je obvyklé rozložení okamžitých hodnot řečového signálu, které připomíná průběh gama kolem střední (nulové) hodnoty [2]. Příklad rozložení okamžitých hodnot úseku běžného řečového signálu je na Obr. 1:



**Obr. 1:** Rozložení okamžitých hodnot 2 s řečového signálu,  $f_{VZ}=44,1$  kHz,  $N=16$  bit, mono

## 2 PRINCIP KOMPRESI

Zvýšení informativní účinnosti kódování je dosaženo logaritmičkou kompanzí předcházející uniformní kvantizaci. Tato operace více zesílí malé hodnoty okamžitých

výchylek signálu, velké hodnoty výchylek signálu jsou zesíleny jen nepatrně. Při obnovování signálu musí být za dekodérem zařazen expandér, který obnoví původní rozsah signálu a vrátí vzorkům správný vzájemný poměr velikostí. Převodní amplitudová charakteristika komandéru se nazývá *kompresní charakteristika*, převodní amplitudová charakteristika expandéru se nazývá *expanzní charakteristika* a má inverzní průběh ke kompresní charakteristice.

### 3 NEJČASTĚJI POUŽÍVANÉ METODY

V mezinárodním telekomunikačním standardu ITU G.711 jsou definovány 2 nejčastěji používané kompresní charakteristiky:

- $\mu$ -Law (u-Law, mu-Law)
- A-Law.

**$\mu$ -Law** je nejčastěji používanou metodou nelineárního kódování v USA a Japonsku. Kompresní a expanzní charakteristiky [3] této metody jsou vyjádřeny vztahy:

$$y(n) = \text{sgn}(n) \cdot \frac{\log(1 + \mu \cdot |s(n)|)}{\log(1 + \mu)} \quad (1)$$

$$s(n) = \text{sgn}(n) \cdot \frac{1}{\mu} \cdot \left[ (1 + \mu)^{|y(n)|} - 1 \right], \quad (2)$$

kde  $y(n)$  je výstupní posloupnost komandovaných vzorků,  $s(n)$  je posloupnost vstupních vzorků z intervalu -1 až +1,  $\text{sgn}(n)$  je znaménková funkce nabývající hodnot -1 pro zápornou hodnotu  $s(n)$  a +1 pro kladnou hodnotu  $s(n)$ ,  $\mu$  je parametr určující míru komprese. Nejlepších výsledků [1] se dosahuje s hodnotou kompresního parametru  $\mu = 255$ , proto je tato hodnota v praxi nejpoužívanější.

**A-Law** je standardní metoda nelineárního kódování používaná v evropských telekomunikacích. Její princip je velmi podobný metodě  $\mu$ -Law. Kompresní charakteristika [3] je vyjádřena vztahy:

$$y(n) = \text{sgn}(n) \cdot \frac{A|s(n)|}{1 + \log(A)} \quad \text{pro } 0 \leq |s(n)| < \frac{1}{A} \quad (3)$$

$$y(n) = \text{sgn}(n) \cdot \frac{1 + \log(A \cdot |s(n)|)}{1 + \log A} \quad \text{pro } \frac{1}{A} \leq |s(n)| \leq 1, \quad (4)$$

kde  $A$  je parametr určující míru komprese. Nejčastěji je podle [3] používána hodnota kompresního parametru  $A = 87,7$  nebo  $A = 87,6$ .

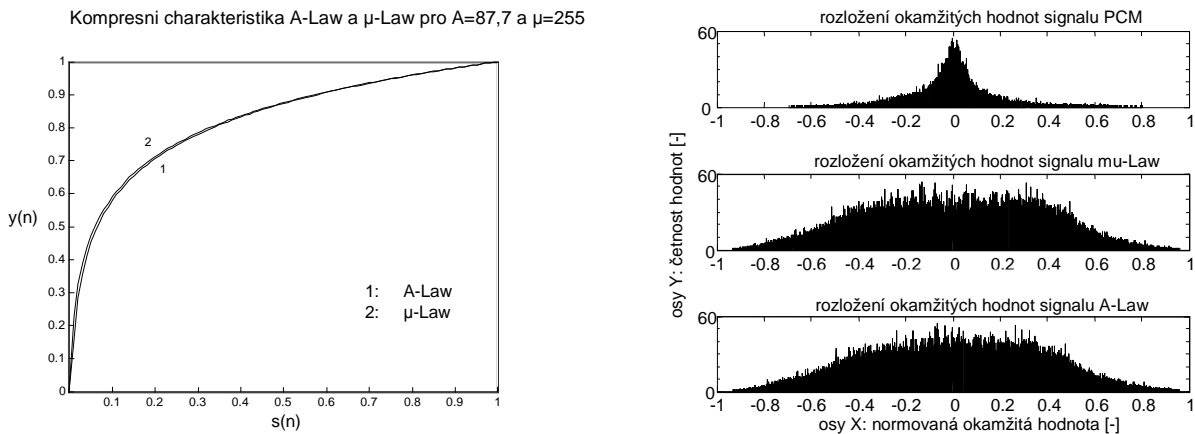
Expanzní charakteristika metody A-Law je dána vztahy:

$$s(n) = \text{sgn}(n) \cdot \frac{|y(n)| \cdot (1 + \log(A))}{A} \quad \text{pro } 0 \leq |y(n)| < \frac{1}{1 + \log(A)} \quad (5)$$

$$s(n) = \text{sgn}(n) \cdot \frac{\exp(|y(n)| \cdot (1 + \log(A)) - 1)}{A} \quad \text{pro } \frac{1}{1 + \log(A)} \leq |y(n)| < 1. \quad (6)$$

#### 4 POROVNÁNÍ A ZHODNOCENÍ PŘÍNOSU LOGARITMICKÉHO KÓDOVÁNÍ

Logaritmickou kompanzí signálu dojde k zesílení nízkých hodnot vzorků, které jsou tak vlastně uniformním kvantizérem zakódovány s větší přesností, než kdyby kompanze nebyla provedena. Pro kvalitní kvantizaci pak postačí méně bitů. Tvar kompresních charakteristik obou metod je pro nejčastěji používané hodnoty kompresních parametrů velmi podobný, jak vidíme na Obr. 2A.



**Obr. 2:** A) Tvar kompresních charakteristik B) Rozložení okamžitých hodnot signálu

Kompanzí také dosáhneme lepšího využití celého dynamického rozsahu signálu, což vidíme na Obr. 2B. Zde byl použit stejný signál, jako na Obr. 1, tedy 2s dlouhý řečový signál s  $f_{VZ}=44,1$  kHz a  $N=16$  bit. Zatímco rozložení okamžitých hodnot lineárně kódovaného signálu má výrazné maximum v okolí nulových hodnot, rozložení hodnot logaritmicky komandovaného signálu je mnohem rovnoměrnější, v ideálním případě téměř lineární.

Logaritmická kompanze řečového signálu je velmi efektivní. Rekonstrukcí 8 bitového logaritmicky kódovaného signálu získáme signál kvalitativně ekvivalentní 13 až 14 bitovému lineárně kódovanému PCM signálu [3]. Obě popsané metody jsou používány především v digitálních telekomunikacích, ale jsou k dispozici i v běžně vybavených počítačích typu PC.

#### LITERATURA

- [1] Levický, D. Multimediálne telekomunikácie, Košice, Elfa, 2002, ISBN 80-89066-58-5
- [2] Sigmund, M. Analýza řečových signálů, Brno, MJ Servis 2000, ISBN 80-214-1783-8
- [3] webové stránky firmy CISCO, Document ID: 8123, Waveform Coding Techniques [http://www.cisco.com/warp/public/788/signalling/waveform\\_coding.html](http://www.cisco.com/warp/public/788/signalling/waveform_coding.html) (březen 2006)