

MERANIE ZÁKLADNÉHO TÓNU REČI WAVELET TRANSFORMÁCIOU

J. Zimmermann, J. Volín

Fonetický ústav Filozofickej fakulty Univerzity Karlovej v Prahe

Abstract

The measurement and analysis of phonetic parameters from the acoustic speech signal finds its applications not only in linguistics itself but also in technical fields whose objective is solving the problem of machine-human communication. The paper presents a method of fundamental frequency measurement in speech using the wavelet transform. Graphical display of the results, calibration into semitones and relevance to speech are also discussed.

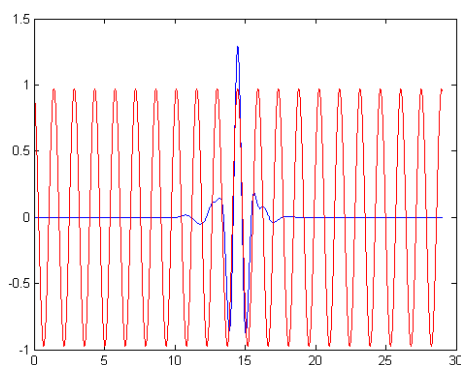
1. Princípy použitej metódy

Experimentálna fonetika, skúmajúca akustický rečový signál, využíva viacero metód v časovej aj frekvenčnej doméne, ktorými sa dajú merať a graficky zobrazovať jeho segmenty (hlásky) a supra-segmenty (časové, tónové a silové modulácie rečového prúdu). Tónová modulácia – zmena výšky hlasu – je prejavom zmien frekvencie kmitania hlasiviek, na oscilograme sa ukazuje ako kolísajúca perióda základného harmonického tónu v znelých segmentoch reči. Na meranie tejto periódy jestvuje množstvo metód s rôznou robustnosťou a schopnosťou spracúvať aj degradovaný signál. Prax experimentálnej fonetiky ukazuje, že by bolo osožné nájsť taký spôsob merania základného tónu, ktorý by bol odolný voči zašumeným signálom a zachytil by aj krátke rozkmity hlasiviek.

Jedným z riešení môže byť využitie Wavelet Toolboxu v prostredí jazyka Matlab, „nespojité“ využitie spojitej vlnkovej (wavelet) transformácie pri konštantnej hodnote mierky [2]. V princípe pôjde o zisťovanie zhody signálu s vlnkou krátkeho trvania, ktorej tvar naznačuje periodické zakmitanie základného tónu reči. Ak sa toto zisťovanie zhody vykoná na celom trvaní výpovede posúvaním vlnky po časovej osi, dá sa z extrahovaných koeficientov zhody vypočítať a skonštruovať priebeh melódie. Presnejšie povedané, pôjde o zisťovanie koeficientov zhody W medzi k vzorkami vlnky $\Psi(k)$ typu Coiflet, ktorej centrálna frekvencia bude nastavená mierkou s do pásma základného tónu hovoriaceho, a medzi vzorkami rečového signálu $x(n)$ podľa vzťahu:

$$W(s) = \sum_k x(n+k) \frac{1}{\sqrt{s}} \psi(k)$$

Centrálna frekvencia vlnky typu Coiflet 5. rádu je 0,68966 Hz. Na obr. 1 je zachytená funkcia vlnky a aproximácia jej centrálnej frekvencie sínusovou funkciou.



Obr. 1. Centrálna frekvencia vlnky Coiflet 5. rádu.

Aby bola centrálna frekvencia F_c vlnky položená do oblasti základného tónu F rečníka, musí sa jej mierka nastaviť tak, aby vyhovovala vzťahu:

$$s = \frac{F_c}{F \cdot T_0}$$

kde T_0 je vzorkovacia perióda signálu. Rozpätie výšky analyzovaného hlasu rečníka vo verbálnej komunikácii je 100 – 140 Hz, teda ide o rozpätie mierok vlnky 235 – 330. Z rozsahu vyplýva, že mierka môže byť nastavená do stredu tohto intervalu, na hodnotu 280.

Sekvencia koeficientov zhody pri posúvaní takto nastavenej vlnky vytvorí funkciu, časový priebeh tejto funkcie zachytáva nielen fázu výdrže, ale aj fázu nakmitávania a dokmitávania hlasiviek, čo je osobitným prínosom pre experimentálnu fonetiku pri skúmaní krátkych akustických udalostí.

Aby sa nameraný časový priebeh základného tónu mohol pokladať za relevantný pre fonetiku, musí sa kalibrovať jeho amplitúda aj jeho tempo s ohľadom na percepciu. Kalibrácia amplitúdy je nájdenie tej hladiny, ktorá sa nachádza na rozhraní medzi vnemom znelosti a neznelosti. Štatistickým vyhodnocovaním posluchového testu a ich porovnávaním s nameranými amplitúdami základného tónu vo fáze nakmitávania a dokmitávania sa určuje tá hodnota amplitúdy, ktorá rozdeľuje kmity na znelé a neznelé. Druhá kalibrácia, ktorou sa prispôbuje merací algoritmus vlastnostiam hlasu rečníka, je vylúčenie tých lokálnych maxím základného tónu, ktoré svojou základnou periódou nepatria do prirodzeného rozsahu výšky hlasu analyzovaného rečníka. Určujúcim parametrom je v tomto prípade dĺžka základnej periódy.

Fyzikálna jednotka výšky tónu Hertz nerespektuje nelineárne vlastnosti vnemu sluchového orgánu. Aby bol aj tento nameraný parameter smerodajný vo fonetickom výskume, musí sa konvertovať na inú stupnicu. Zdá sa, že najpriateľnejšie výsledky pri skúmaní základného tónu sa dosahujú aplikáciou nelineárnej hudobnej poltónovej stupnice [1]. Frekvencia základného tónu F v hertzoch, odvodená od trvania periódy, je prepočítaná na poltóny (ST) podľa vzťahu:

$$ST = \frac{12 \cdot \ln\left(\frac{F}{100}\right)}{\ln(2)}$$

Vzostup tónu o poltón je v muzikológii vzostup tónu o dvanástinu oktávy; oktáva je rozpätie dvoch tónov v pomere 1:2. Ak sa vo fonetike určuje vzostup základného tónu v poltónoch, musí byť arbitrárne stanovený začiatok stupnice v hertzoch, od ktorého sa poltóny počítajú. V uvedenom vzorci je tento začiatok stanovený na hodnotu 100 Hz.

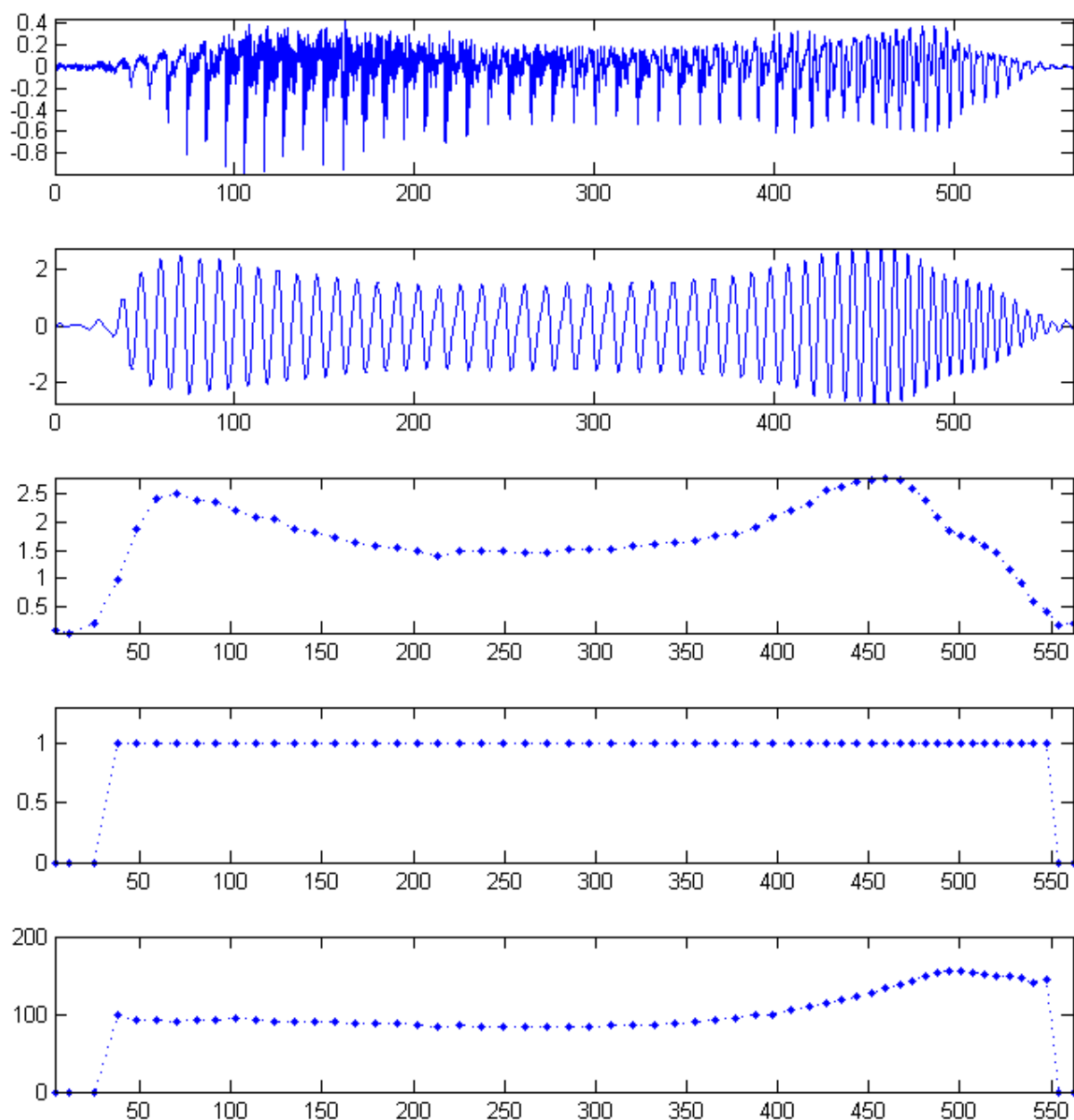
Podobne, zo známeho rozpätia kolísania základného tónu v hertzoch sa vypočíta počet poltónov v tomto rozpätí podľa vzťahu

$$\text{počet poltónov} = 39,863 \cdot \log\left(\frac{F_h}{F_d}\right)$$

kde F_h je horná hranica a F_d je dolná hranica frekvenčného rozpätia v Hz.

2. Výsledky experimentu

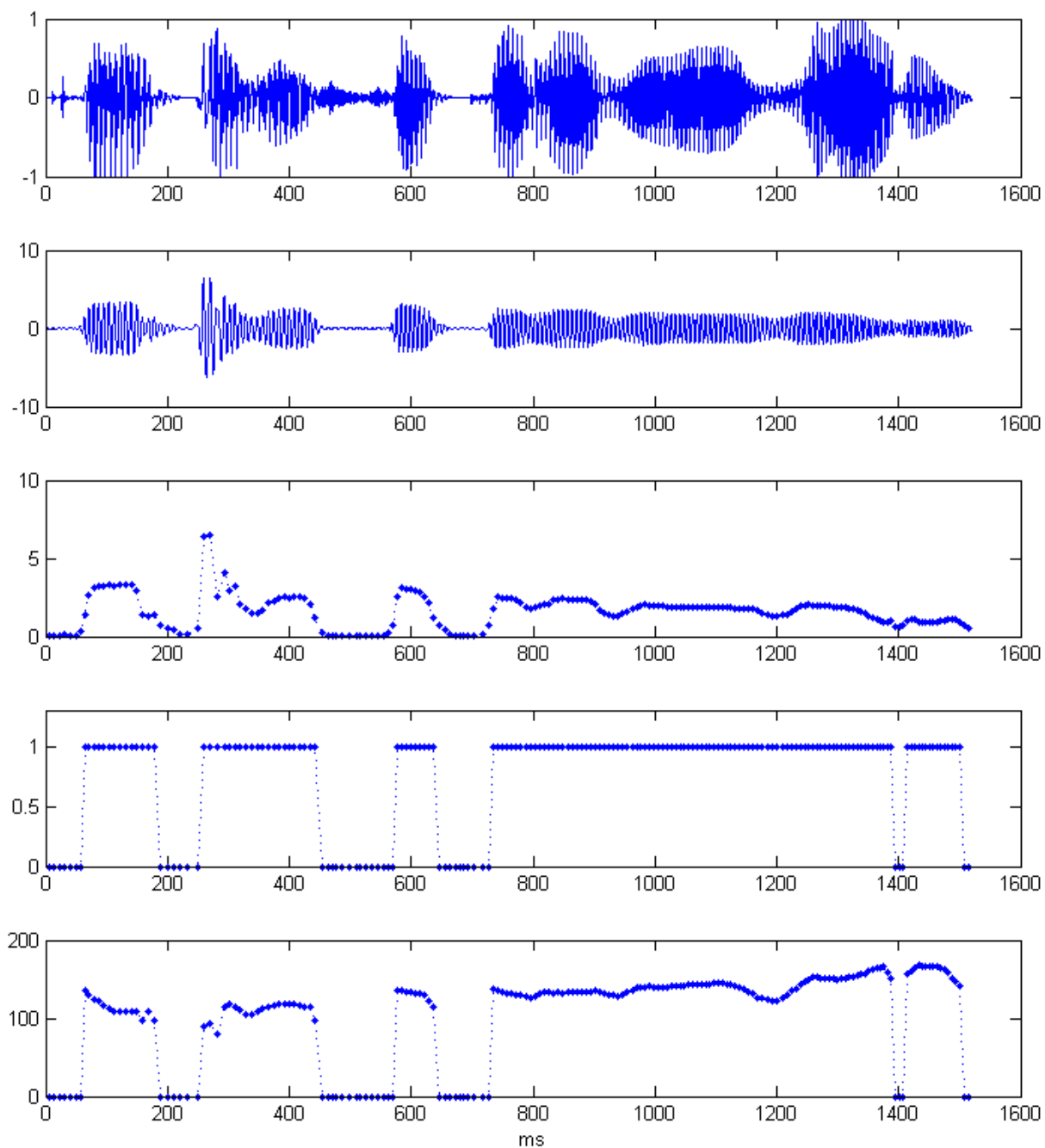
Na podrobnejšie vysvetlenie uvedeného postupu najprv uvidíme experiment analýzy krátkej výpovede – otázky *Hej?*, aby sme ukázali jednotlivé kroky algoritmu a ich účinky na vzorky signálu. Na obr. 2 sú znázornené jednotlivé štádiá. Oscilogram (prvá krivka zhora) ukazuje, že celá výpoveď sa skladá z troch znelých hlások, neznelý je len úsek nakmitávania a dokmitávania výpovede. Druhá krivka zhora je funkciou koeficientov zhody vlnky typu Coiflet so signálom pri mierke 280, jej pseudofrekvencia sa približne zhoduje s frekvenciou základného tónu rečníka. Krivka je teda kmitaním základného tónu, možno na nej pozorovať veľkosť amplitúdy a jednotlivé periódy. Tretia krivka zhora zachytáva nájdené lokálne maximá kladných amplitúd funkcie koeficientov zhody. Tie sú ďalej vyhodnocované tak, že sú podľa ich veľkosti rozdelené na znelé a neznelé (štvrtá krivka zhora) pomocou orezávacej hladiny, ktorá je na základe posluchového testu stanovená na hodnotu 15% z najvyššej hodnoty lokálnych maxím. V tomto experimente sú len tri neznelé lokálne maximá na začiatku výpovede a dve na konci. Najspodnejšia krivka – priebeh kolísania základného tónu v Hertzoch, je skonštruovaná tak, že sú najprv odmerané vzdialenosti medzi jednotlivými lokálnymi maximami a tieto vzdialenosti sú ďalej prepočítané na frekvenciu.



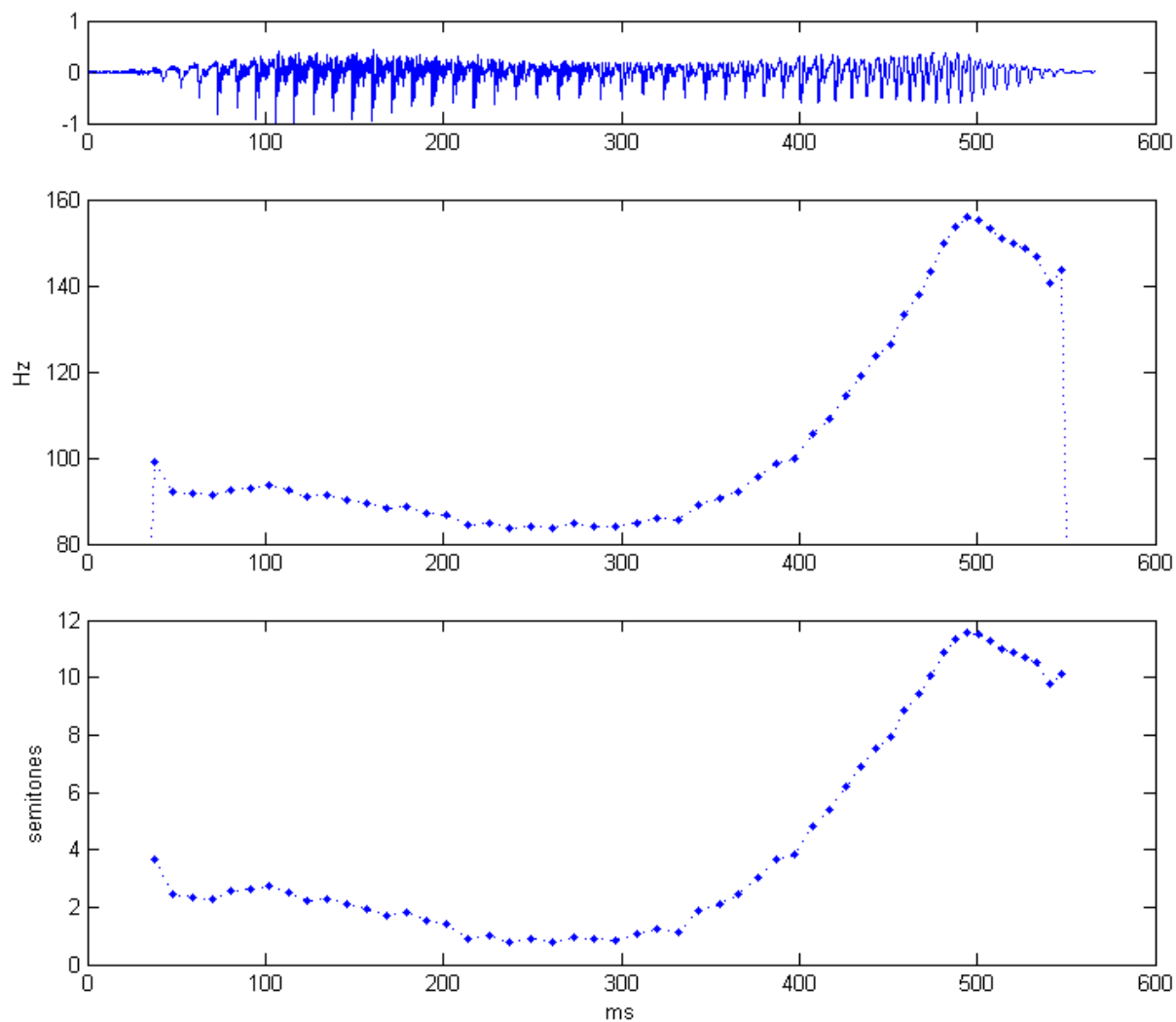
Obr. 2. Extrakcia základného tónu zo signálu slova *Hej?*. Krivky zhora (časové osi v ms): oscilogram, priebeh kmitov základného tónu, lokálne maximá kladných amplitúd, klasifikácia znelé (1) a neznelé (0), priebeh kolísania základného tónu v Hz.

Naúčinnnejším overením tejto metódy môže byť analýza výpovede, v ktorej sa nachádzajú šumové spoluhlásky, explozívny a vibranty, kde sa nachádzajú fázy nakmitávania a dokmitávania v zložitejších akustických udalostiach, akými môžu byť aj tranzienty a akustické kmity jazyka. Na obr. 3 uvádzame analýzu výpovede *Kúpili ste televízor?*. Jednotlivé grafy znázorňujú rovnaké štádiá algoritmu ako na obr. 2. Pozoruhodné na tejto analýze je to, že neznelé úseky reči sú presne zachytené vo výslednej (najspodnejšej) krivke poklesom kontúry výšky základného tónu na nulu.

Poslednou časťou experimentu je konverzia frekvenčnej stupnice v Hertzoch v grafe kolísania výšky základného tónu na poltónovú stupnicu, aby sa dala kontúra tejto krivky pokladať za priebeh melódie reči. Na ilustráciu tohto prevodu uvádzame na obr. 4 analýzu výrazu *Hej?*, kde je v druhej krivke zhora použitá frekvenčná stupnica a v najspodnejšej krivke poltónová stupnica, jej začiatok je stanovený na hodnotu 80 Hz.



Obr. 3. Analýza vety *Kúpili ste televízor?*. Krivka priebehu základného tónu (úplne dole) je nulová v úsekoch, kde sú explozívny, sykavka a nákmít vibranty.



Obr. 4. Konverzia frekvencnej stupnice v Hz na poltónovú stupnicu priebehu melódie výrazu *Hej?*.

Literatúra

- [1] J. VOLÍN. *Statistické metody ve fonetickém výzkumu*. Praha, EPOCHA 2007.
 [2] J. ZIMMERMANN. *Spektrografická a škálografická analýza akustického řečového signálu*. Prešov, Náuka 2002.

Jan Volín: jan.volín@ff.cuni.cz

Július Zimmermann: julius.zimmermann@ff.cuni.cz