

AKUSTICKÁ ANALÝZA INTENZITY A RYCHLOSTI ŘEČI U PARKINSONOVY NEMOCI

Jan Rusz¹, Roman Čmejla¹, Hana Bachurová², Jan Janda¹

¹Fakulta elektrotechnická, ČVUT, Praha

²Neurologická klinika 1. LF UK a VFN, Praha

Abstrakt

K příznakům Parkinsonovy nemoci patří snížená intenzita fonace a změna tempa a plynulosti v řeči. Tato práce se zabývá analýzou výše uvedených jevů v řeči pacientů na počátku onemocnění a přináší použití nových původních charakteristik pro jejich hodnocení. Analýzy byly provedeny v programovém prostředí MATLAB na skupině 16 pacientů s Parkinsonovou nemocí (PN) a na kontrolní skupině 14 zdravých mluvčích podobného věkového složení (KS).

Úvod

Parkinsonova nemoc je chronické, pomalu se rozvíjející onemocnění, jehož rozvoj lze při včasné detekci potlačit či omezit, a to dlouhodobě. Parkinsonova nemoc obvykle začíná ve středním věku, průměrný stáří pacientů při prvotní diagnóze onemocnění se pohybuje okolo 50 až 60 let. V České republice je postiženo řádově tisíc obyvatel trpících touto nemocí, u osob starších 60 let se jedná až o jedno procento této populace. Na riziko onemocnění nemá vliv společenská vrstva, vzdělání, strava, zaměstnání, kontakt se zvířaty, očkování, životní standard či příjem alkoholu. Riziko výskytu nemocí narůstá v dalších generacích. Zhoršení kvality řeči je obvyklým a jedním ze základních symptomů této nemoci. Cílem této práce je poukázat na nezanedbatelný význam analýz řeči, který může vést ke včasné identifikaci této nemoci, přičemž tato práce se věnuje snížené intenzitě fonace, změně tempa, plynulosti řeči a algoritmům navrženým pro jejich automatickou detekci.

1. Artikulační rychlost

Prvním analyzovaným parametrem je artikulační rychlost řeči [1], [2], [3]. Ta je počítána jako poměr délky pauz a celkové promluvy, kdy je skupina pacientů s Parkinsonovou nemocí a kontrolní skupina zdravých lidí požádána o přečtení stejného referenčního textu „Zahradníkův rok“ převzatého z knihy Karla Čapka v délce trvání 60 sekund (parametr AR1 [-]) a v druhém případě jsou požádáni o volné vlastní vyprávění ve stejné délce (parametr AR2 [-]). Oba parametry byly získány stejným postupem, kdy pauzy byly stanovovány v časové oblasti detektorem založeným na prahování energie signálu x (1.1) a jeho počtu průchodů nulou (1.2)

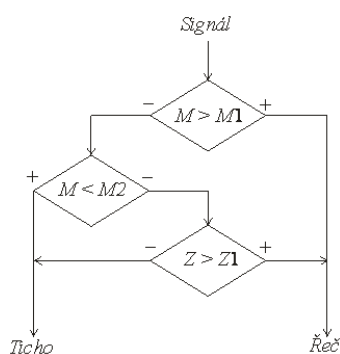
$$E = \sum_{n=1}^N x^2[n] \quad (1.1)$$

$$Z = \frac{1}{N} \sum_{n=1}^N \frac{|\operatorname{sgn}(x[n]) - \operatorname{sgn}(x[n-1])|}{2} \cdot f_s \quad (1.2)$$

kde E je energie, Z je počet průchodů nulou, n je aktuální vzorek, N počet vzorků signálu a f_s vzorkovací frekvence.

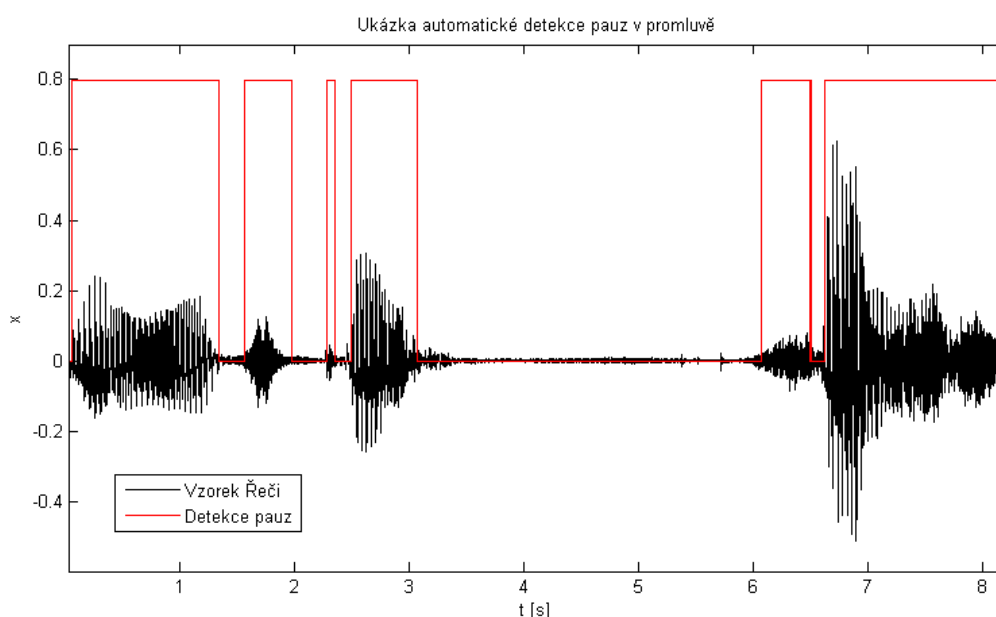
Algoritmus detektoru „Řeč – Pauza“ (obr. 1.1) v prvním kroku porovná aktuální vzorek signálu M se střední hodnotou signálu $M1$. Jestliže je výsledek pozitivní, tzn. aktuální vzorek je větší hodnoty než střední hodnota intenzity signálu, může algoritmus s velkou přesností klasifikovat signál jako řeč. V případě negativního výsledku porovná aktuální vzorek s rozdílem střední hodnoty signálu a jeho směrodatné odchylky $M2$. V případě, že je aktuální vzorek menší než tento rozdíl, klasifikuje signál jako pauzu. Jestliže se aktuální vzorek nachází v intervalu $\langle M2, M1 \rangle$, využije znalost o počtu průchodů signálu nulou, která spočítána jako klouzavý průměr s velikostí okna dle vzorkovací frekvence. Nakonec porovná aktuální počet vzorků Z se střední

hodnotou počtu vzorků získaných využitím klouzavého průměru $Z1$ a klasifikuje, zda se jedná o řeč, nebo pauzu.



Obr. 1.1: Vývojový diagram automatické detekce Řeč – Pauza

Obr. 1.2 znázorňuje výsledek automatické detekce Řeč – Pauza na vybraném úseku signálu.



Obr. 1.2: Výsledek automatické detekce Řeč -Pauza na vybraném úseku signálu

2. Počet pauz

Dalším parametrem, který přímo navazuje na předchozí studii, je počet pauz (parametr PP [-]), mající vztah k rychlosti a rytmu řeči. Rozbor výsledků ukázal, že pacienti s Parkinsonovou nemocí dělají ve své řeči méně, avšak delších pauz [2]. U tohoto parametru je důležité, aby byl získán ze stejné referenční promluvy, proto byla taktéž využita pasáž „Zahradníkův rok“ po eliminaci pauz kratších než 60 ms.

3. Diadochokinéze

Ve foniatrické praxi se obvykle pro posouzení motorických schopností artikulačního ústrojí měří, jak rychle je pacient schopen opakovat počet kombinací souhláska – samohláska (C - V). Aby takovéto měření postihlo motorickou zdatnost celého artikulačního aparátu, je obvykle voleno více různých souhlásek s různým místem artikulace. Často je pacient požádán, aby opakoval nejvyšší možnou rychlostí trojici slabik /pa/ - /ta/ - /ka/. Takováto kombinace souhlásek rovnoměrně zatěžuje hlasový trakt od artikulace oboustranné okluzivy /p/ přes předodásňové /t/ až po měkkopatrové /k/. Počet C - V kombinací vyřčených za jednu sekundu se vyjadřuje mezinárodně používanou veličinou diadochokinéze (diadochokinetic rate - parametr DDK [s]), která je dalším parametrem pro popis rychlosti řeči [3], [4], [5]. Přídavný

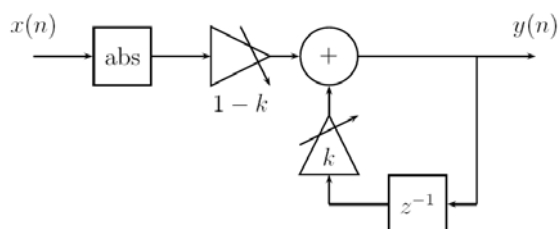
parametr vyjadřuje míru kolísání, tedy schopnost udržet konstantní rychlost kontextu souhláska-samohláska (parametr $\text{var}(DDK)$ [-]).

Algoritmus pro výpočet DDK detekuje jednotlivé slabiky z energetické obálky, která je počítána pomocí špičkového detektoru (obr. 3.1.) Ten je realizován jako integrátor s proměnnou integrační konstantou řízenou podmínkou. Pro detekci maxim signálu je zvolena integrační konstanta

$$k = 0,5 \quad \text{pokud} \quad x(n) \geq y(n - 1) \quad (3.1)$$

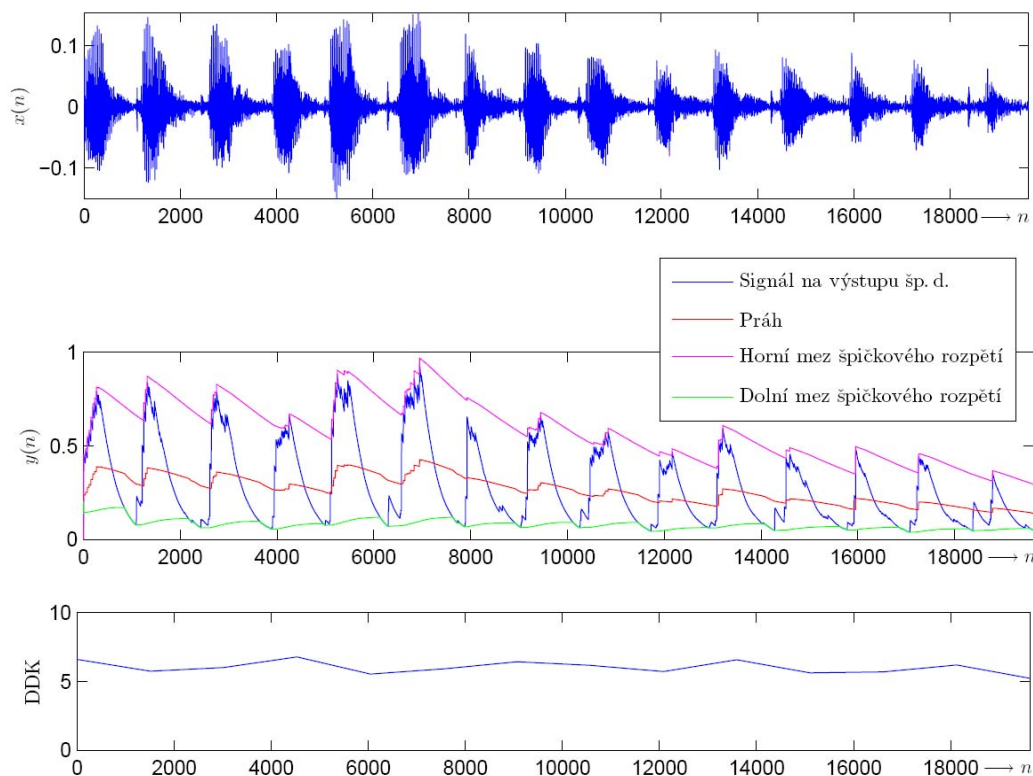
$$k = 0,9971 \quad \text{pokud} \quad x(n) < y(n - 1) \quad (3.2)$$

Je-li potřeba detekovat minima signálu, podmínky se znegují.



Obr. 3.1: Schéma špičkového detektoru

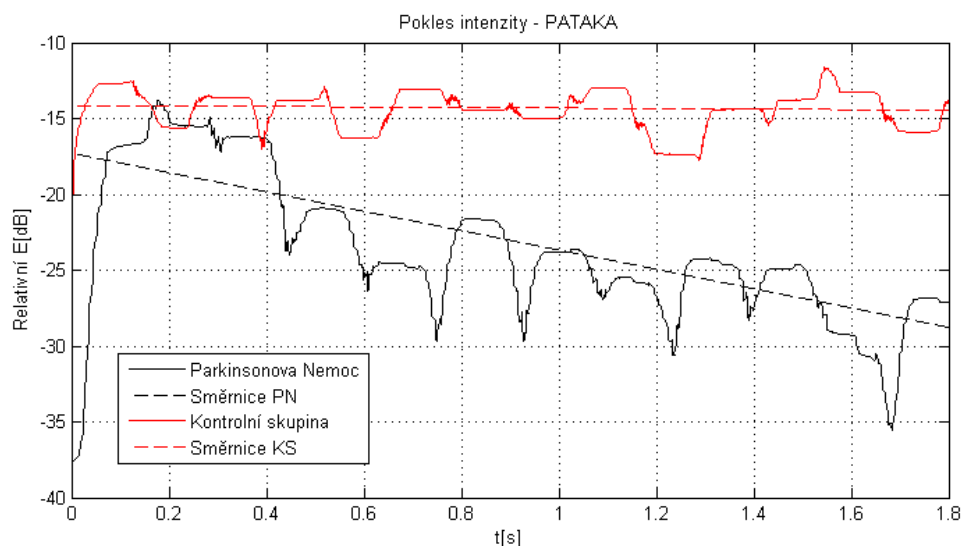
Detektor lze použít znovu k určení úrovně pro prahování. Detekcí maxim a minim špičkovým detektorem s velkou integrační konstantou určíme v signálu špičkové rozpětí. Relativně k němu pak nalezneme úroveň prahu (obr. 3.2b). Když známe hranice jednotlivých slabik, stanovíme u promluvy průběh DDK (obr. 3.2c).. Dále pak určíme parametry promluvy jako střední hodnotu DDK a $\text{var}(DDK)$ jako její rozptyl. Hodnoty těchto parametrů pro jednotlivé pacienty porovnáme.



Obr. 3.2: a) Promluva, b) Výstup špičkového detektoru a určení prahu, c) DDK

4. Časový průběh intenzity

Měření DDK a poklesu intenzity [6] vedlo k novému poznatku, a to, že u skupiny lidí s příznaky Parkinsonovy nemoci právě při rychlém opakování trojice slabik /pa/ - /ta/ - /ka/ klesá intenzita promluvy mnohem strměji než u zdravých mluvčích (parametr Int [dB/s]). Parametr je počítán jako směrnice (obr. 4.1) přes celou promluvu, která nám určuje pokles decibelů za sekundu. Je vypočítán jako robustní lineární regrese s využitím Matlabovského příkazu *robustfit* ze Statistics toolboxu.



Obr. 4.1: Typický záznam časového průběhu ilustrující intenzity u pacientů s PN a u kontrolní skupiny

5. Experimentální výsledky

Tab. 1: SHRNUTÍ VÝSLEDKŮ (STŘEDNÍ HODNOTY A SMĚRODATNÉ ODCHYLKY PARAMETRŮ)

		AR1	AR2	PP	DDK	var(DDK)	Int
PN (Parkinsonova nemoc)	<i>mean</i>	31,45	33,46	129	0,16	0,086	3,16
	<i>(std)</i>	(1,93)	(2,26)	(22)	(0,02)	(0,057)	(1,92)
KS (Kontrolní skupina)	<i>mean</i>	29,02	29,51	156	0,14	0,072	1,18
	<i>(std)</i>	(2,42)	(3,47)	(12)	(0,02)	(0,054)	(0,67)

V tab. 1 vidíme nižší rychlost (delší čas) ve skupině PN jak při čtení textu, tak i při běžné promluvě. Výše popsanou úpravou algoritmu je také prokázáno, že skupina PN dělá ve své řeči méně pauz, které jsou však delší. U promluv /pa/ - /ta/ - /ka/ má skupina PN rovněž nižší schopnost udržet konstantní rychlost kontextu C - V a dochází u ní kvýraznému poklesu intenzity.

6. Závěr

Výhodou prezentovaných charakteristik, na rozdíl od řady dalších založených na průběhu fundamentální frekvence F_0 je, že nejsou závislé na pohlaví. Výsledky akustické analýzy pacientů s Parkinsonovou nemocí potvrzují, že změny v řeči se objevují již v rané fázi onemocnění a nemusí být rozpoznány při běžném verbálním projevu.

Poděkování

Tato práce je podporována z výzkumného záměru „Transdisciplinární výzkum v oblasti biomedicínského inženýrství“ (č. MSM6840770012), a grantů „Analýza a modelování biologických a řečových signálů“, GAČR č. 102/08/H008, a "Rozpoznávání mluvené řeči v reálných podmínkách" ("Speech Recognition under Real-World Conditions"), GACR 102/08/0707.

Reference

- [1] Goberman M. *Correlation between acoustic speech characteristics and non-speech motor tasks in Parkinson's disease*. Med. Sci. Monit, 2005; 11(3): CR109-116
 - [2] Skodda S, Schlegel U. *Speech rate and rhythm in Parkinson's disease*. Movement Disorders Vol. 23, No. 7, 2008, pp. 985–992
 - [3] Rosen KM, Kent RD, Delaney AL. *Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers*. Journal of Speech, Language, and Hearing Research, 2006, Vol. 49, 395–4.
 - [4] Tjaden K, Watling E. *Characteristics of Diadochokinesis in Multiple Sclerosis and Parkinson's disease*. Folia Phoniatica et Logopaedica, 2003; 55:241-259.
 - [5] D'Alatri L. et al. *Effects of Bilateral Subthalamic Nucleus Stimulation and Medication on Parkinsonian Speech Impairment*. The voice foundation, Journal of Voice, 2006.
 - [6] Rosen KM, Kent RD, Duffy JR. *Task-based profile of vocal intensity decline in Parkinson's disease*. Folia Phoniatica et Logopaedica, 2005; 57,1.
-

Autor1

Jan Rusz

ruszj1@fel.cvut.cz

Autor2

Roman Čmejla

cmejla@fel.cvut.cz