



III. LETNÍ DOKTORANDSKÉ DNY 2013

SBORNÍK PŘÍSPĚVKŮ

Editori sborníku

Doc. Ing. Roman Čmejla, CSc.

Ing. Jan Rusz, Ph.D.

Ing. Daniel Špulák



květen 2013

Název: III. LETNÍ DOKTORANDSKÉ DNY 2013

třetí ročník prezentací doktorandů Katedry teorie obvodů

23. – 24. května 2013

ČVUT FEL, blok C3, posluchárna T2:C3-54

Editoři sborníku: Doc. Ing. Roman Čmejla, CSc.; Ing. Jan Ruzs, Ph.D.; Ing. Daniel Špulák

Vydalo: České vysoké učení technické v Praze

Zpracovala: Fakulta elektrotechnická

Kontaktní adresa: Technická 2, 166 27 Praha 6 – Dejvice

Telefon: 224 352 236

Web: <http://sami.fel.cvut.cz/LDD13/>, <http://obvody.feld.cvut.cz/>

Počet stran: 130

Datum vydání: květen 2013

Vydání: 1.

ISBN: 978-80-01-05251-8

Poděkování:

Děkujeme Ing. Adamu Stráníkovi za technickou pomoc při přípravě sborníku.

PROGRAM LDD'13 – seznam příspěvků

čtvrtek, 23. května 2013, 10:00 – 18:05, přednášky

10:00 Zahájení LDD 2013: Prof. Ing. Václav Havlíček, CSc. – rektor ČVUT

téma: **TEORIE A APLIKACE SIGNÁLŮ A SYSTÉMŮ**

10:05 Ing. Jiří Náhlík
Realizace banky filtrů technikou spínaných kapacitorů.....6

téma: **BIOLOGICKÉ SIGNÁLY**

10:30 Ing. Antonín Hlaváček
Studium neurosvalové aktivity a souvisejících metod zpracování signálů

10:55 Ing. Matouš Pokorný
Analýza chování biomedicínského modelu Huntingtonovy choroby.....13

11:20 Ing. Miroslav Ložek
Optimalizace parametrů modelu KVS.....17

11:45 Ing. Jan Dvořák
Měření hemodynamických parametrů – metody, realizace, problémy.....22

12:10 *Přestávka na oběd*

13:00 Ing. Michaela Nerudová
Měření ultraslabé emise fotonů z biologických vzorků.....26

13:25 Ing. Lukáš Bauer
Měření sluchových korových odezev u pacientů se sluchovou náhradou32

13:50 Ing. Daniel Špulák
Využití povrchových elektromyografických signálů při analýzách svalové koordinace.....36

14:15 Ing. Jan Sedlák
Využití videozáznamu při zpracování povrchového EMG.....40

14:40 *Přestávka*

téma: **ZPRACOVÁNÍ EEG SIGNÁLŮ**

14:55 Ing. Jan Šebek
Možná řešení problému přeučení algoritmu FastICA při zpracování EEG.....46

15:20 Ing. Vladimír Černý
Identifikace subjektu v reálném čase pro BCI.....53

15:45 Ing. Martin Dobiáš
Validace výběru parametrů pro filtrování podprostorů při zpracování pohybového EEG....59

16:10	Ing. Milan Kostílek Klasifikace jemných pohybů z EEG signálu.....	65
16:35	<i>Přestávka</i>	
16:50	Ing. Radek Janča Automatická detekce a prostorové klastrování interiktálních výbojů v invazivním EEG.....	72
17:15	Ing. Tomáš Havel Detektor epileptiformních vysokofrekvenčních oscilací v intrakraniálních elektroencefalografických záznamech.....	74
17:40	Ing. Jiří Balach Využití Hilbert-Huangovy transformace v analýze epileptického EEG – pilotní studie.....	80

pátek, 24. května 2013, odpoledne 13:00 – 17:15, přednášky

téma: **ZPRACOVÁNÍ ŘEČI**

13:00	Ing. Michal Borský Akustické modelovanie trifónov na bázi HMM.....	86
13:25	Ing. Petr Mizera Zlepšení přesnosti fonetické segmentace na bázi HMM s akustickými modely trifónů.....	92
13:50	Ing. Jan Bartošek Prozodie a modelování přízvukových taktů.....	98
14:15	<i>Přestávka</i>	

téma: **ANALÝZA A HODNOCENÍ PATOLOGICKÉ ŘEČI**

14:30	Ing. Josef Vavřina Význam difúzní magnetické rezonance u dětí s vývojovou dysfázií.....	102
14:55	Ing. Martina Nejepsová Klasifikace promluv pacientů s vývojovou dysfázií.....	108
15:20	Ing. Adam Stráník Hodnocení promluv pacientů s dysfonií – subjektivní testy.....	112
15:45	<i>Přestávka</i>	
16:00	Ing. Tomáš Lustyk Hodnocení koktavosti pomocí automatických algoritmů ve čtených promluvách.....	118
16:25	Ing. Michal Novotný Charakteristiky promluv pacientů s Parkinsonovou nemocí extrahované z řečové diadochokinetické úlohy.....	122
16:50	Ing. Tereza Tykalová Objektivní metody hodnocení důrazu u Parkinsonovy nemoci.....	128

Realizace banky filtrů technikou spínaných kapacitorů

Jiří NÁHLÍK¹, Jiří HOSPODKA¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická,

Katedra teorie obvodů

nahlijir@fel.cvut.cz, hospodka@fel.cvut.cz

Abstrakt: Technika spínaných kapacitorů se často využívá při konstrukci filtrů integrovaných na čipu. Obvykle však tyto filtry pracují s řádově větším hodinovým kmitočtem, než je nejvyšší kmitočet zpracovávaného signálu. Tento příspěvek popisuje realizaci dvoukanálové banky filtrů se spínanými kapacitami, která zpracovává signál až do poloviny svého spínacího kmitočtu 16 kHz. Banka filtrů byla realizována na plošném spoji a výsledné naměřené frekvenční charakteristiky byly porovnány se simulovanými.

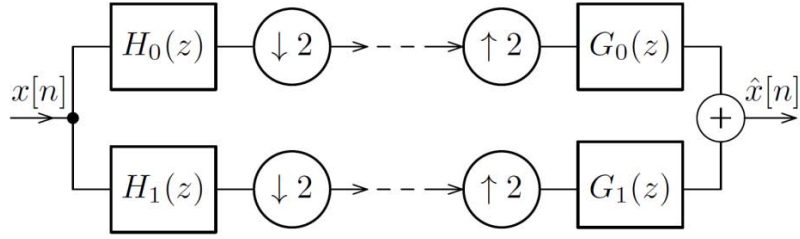
1. Úvod

Analogové a digitální banky filtrů se používají pro zpracování signálu již řadu let. Metody syntézy jak digitálních, tak hybridních forem byly popsány již dříve např. v [2,3]. U diskrétně pracujících systémů ale vyvstává problém aliasingu. Nepřesnosti analogových filtrů mohou být zanedbány, pokud je problém aliasingu řešen dostatečným útlumem v nepropustném pásmu jednotlivých filtrů. Při konstrukci banky filtrů pomocí techniky spínaných kapacitorů je nutné problém aliasingu řešit snížením zesílení (gain cancelation) v nepropustném pásmu.

Pro návrh a realizaci byla vybrána dvoukanálová banka filtrů. Ta pomocí filtru typu dolní propust s přenosovou funkcí $H_0(z)$ a filtru typu horní propust s přenosovou funkcí $H_1(z)$ rozkládá vzorkovaný vstupní signál ($f_c = 16$ kHz) na dva subpásmové signály s polovičním vzorkovacím kmitočtem. Tyto signály jsou pak zpětně složeny na signál s původním vzorkovacím kmitočtem 16 kHz. Principiální blokové schéma je uvedeno na obr. 1.

2. Návrh banky filtrů

Přenosové funkce jednotlivých filtrů musí být navrženy tak, aby byl výstupní signál co nejvíce podobný vstupnímu signálu. Zároveň požadujeme, aby zvlnění v nepropustných pásmech jednotlivých filtrů bylo minimální



Obr. 1: Principiální blokové schéma dvoukanalové banky filtrů

a filtry byly symetrické. Z těchto podmínek je možné sestavit vztahy mezi přenosovými funkcemi jednotlivých filtrů:

$$H_1(z) = H_0(-z), \quad (1)$$

$$G_0(z) = 2H_1(-z) = 2H_0(z), \quad (2)$$

$$G_1(z) = -2H_0(-z) = -2H_1(z). \quad (3)$$

Dle (1)-(3) je tedy potřeba navrhnout pouze jednu přenosovou funkci. Pro realizaci byl zvolen eliptický IIR filtr třetího řádu (nejmenší řád filtru splňující všechny podmínky) s přenosovou funkcí:

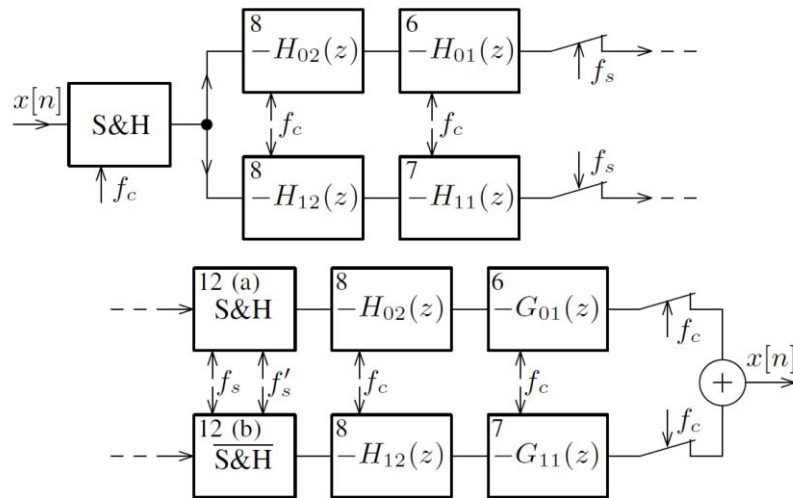
$$H_0(z) = \frac{1}{2} \left(\frac{\alpha + z^{-2}}{1 + \alpha z^{-2}} + z^{-1} \right), \quad (4)$$

kde koeficient α určuje polohu nul a pólů.

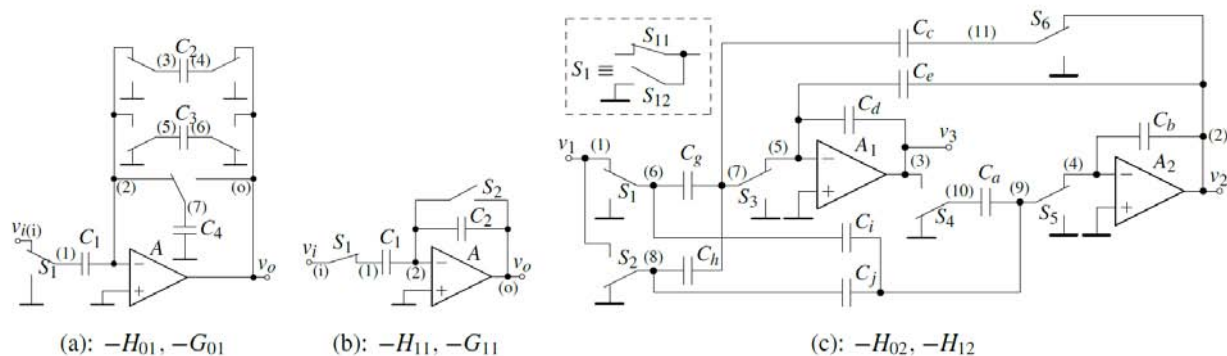
Modulová charakteristika filtru s přenosovou funkcí (4) je konstantní pro všechna $\alpha \in (0,1)$. Fázová charakteristika filtru je naopak nelineární a závislá na parametru α . Koeficient α má také vliv na velikost zvlnění v nepropustném pásmu. Pro realizaci byl pak zvolen $\alpha = 0,5$. Určení vhodného koeficientu α je detailně popsáno v [1].

3. Realizace banky filtrů pomocí obvodu se spínanými kapacitami

Přenosové funkce filtrů dolní a horní propusti (H_0 , H_1 , G_0 , G_1) lze realizovat kaskádní syntézou pomocí bikvadratické sekce následované filtrem prvního řádu. Tomuto případu odpovídá blokové schéma uvedené na obr. 2.



Obr. 2: Blokové schéma realizované banky filtrů



Obr. 3: Zapojení SC integrátoru (a), derivátoru (b), bikvadratické sekce (c)

Přenosové funkce bikvadratické sekce H_{02} a H_{12} jsou realizovány obvodem dle obr. 3c [4]. Přenosovou funkci prvního řádu (integrátoru) H_{01} a G_{01} lze realizovat obvodovým zapojením dle obr. 3a. Přenosovou funkci derivátoru H_{11} a G_{11} pak realizujeme zapojením dle obr. 3b [4].

Zapojení dále obsahuje S&H obvody, které jsou nutné pro správnou funkci bikvadratických sekcí a součtový zesilovač. Normalizované hodnoty všech kapacit jsou uvedeny v [1].

Obvodová zapojení většiny funkčních bloků neobsahují stejnosměrnou zpětnou vazbu a tedy nelze zajistit, aby například vlivem stejnosměrné složky na výstupu předchozího bloku nebyl operační zesilovač následujícího bloku v saturaci. Proto by měla být zavedena dodatečná stejnosměrná zpětná vazba.

3.1. Výsledky simulace

Celá struktura banky filtrů byla simulována v simulátoru WinSpice a simulátoru ELDO firmy Mentor Graphic development enviroment. Kapacitory byly modelovány pomocí ideálního kapacitoru s paralelním rezistorem. Spínače pak byly modelovány pomocí jejich odporu v sepnutém a rozepnutém stavu a vliv nábojové injekce byl zahrnut přemostěním spínače kapacitorem.

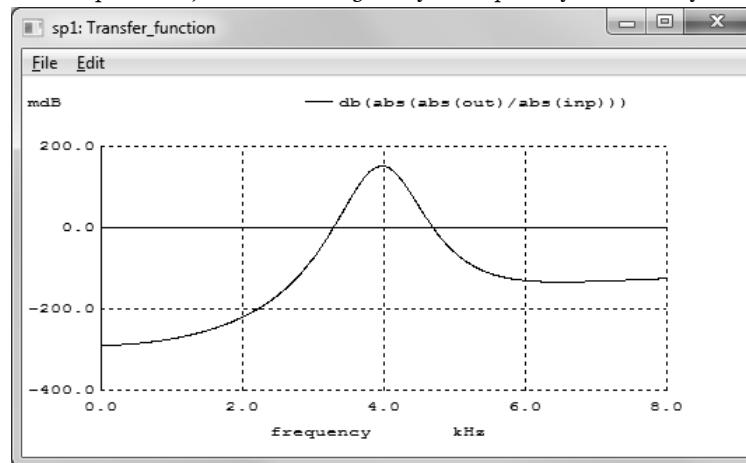
Obvod byl simulován v časové oblasti a jeho frekvenční charakteristika pak byla vypočítána z jeho odezvy diskrétní Fourierovou transformací [5].

Obvod byl simulován pro různé operační zesilovače. Frekvenční charakteristiky se pak lišily především v okolí kmitočtu 4 kHz. Frekvenční charakteristika pro model operačního zesilovače AD8033 [6,7] je uvedena na obr. 4.

3.2. Výsledky měření

Navržená banka filtrů byla realizována na plošném spoji. Jako operační zesilovač byl při konstrukci použit zesilovač AD8034 [6], který má oproti AD8033 dva zesilovače v jednom pouzdře. Pro realizaci spínačů byl použit integrovaný obvod DG456 [8].

Při měření je potřeba oproti simulaci pozměnit časovací signály tak, aby vlivem pomalejšího spínání a rozpínání použitých spínačů nedocházelo k současnému sepnutí spínačů, které mají být sepnuty v různých fázích.



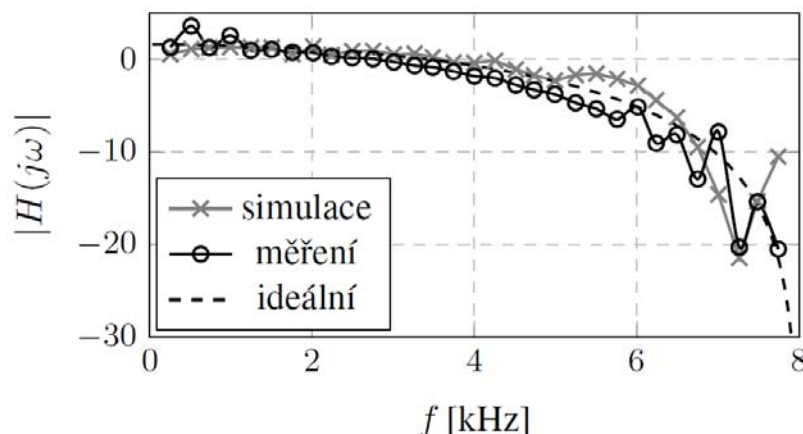
Obr. 4: Simulovaná modulová frekvenční charakteristika banky filtrů pro operační zesilovač AD8033

Nastavení dead-time mezi jednotlivými signály na 50 ns ale ovlivní funkci SC integrátoru. Jeho konstrukce (viz. obr. 3a) umožňuje při nastavení dead-timů stav, kdy operační zesilovač nemá zavedenou žádnou zpětnou vazbu a přechází do saturace. Při následném připojení zpětnovazebních kapacitorů to pak způsobí chybu.

Nastavení časovacích signálů má naopak malý vliv na funkci SC derivátoru. Změřená frekvenční charakteristika se jen velmi málo liší od charakteristiky ideální.

Neideální chování jednotlivých funkčních bloků banky filtrů může nepříznivě ovlivnit následující blok. Proto byly změřeny frekvenční charakteristiky jednotlivých funkčních bloků nejprve samostatně.

Vliv nastavení dead-timů na frekvenční charakteristiku SC integrátoru (viz. obr. 3a) je ukázán na obr. 5. Pro srovnání je uvedena simulovaná charakteristika s nastavením odpovídajícím měřenému průběhu a charakteristika pro idealizovaný obvod a ideální časovací signály.



Obr. 5: Měřená, simulovaná a ideální frekvenční charakteristika SC integrátoru

Problém s nezavedenou zpětnou vazbou SC integrátoru lze řešit několika způsoby:

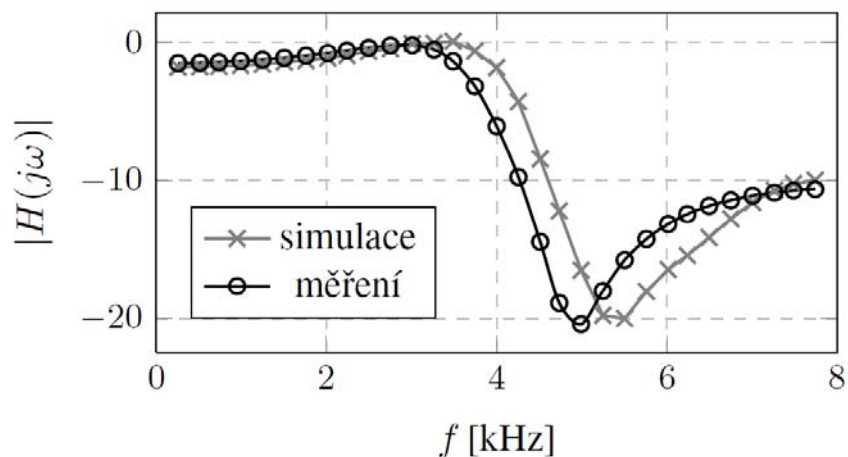
- Zapojením rezistoru s velkou hodnotou elektrického odporu mezi výstup a invertující svorku zesilovače. Toto řešení je ale nepřijatelné v případě realizace na čipu.
- Nastavení přesně navazujících časovacích signálů bez dead-timů. Frekvenční charakteristika v tomto případě poklesne o několik decibelů.
- V poměru zvětšit všechny kapacity obvodu. Chyba způsobená saturací operačního zesilovače se tak zmenší. Pro integraci na čipu je tato možnost také nepřijatelná, neboť se výrazně zvětší jeho plocha.
- Použít vhodnější topologii SC integrátoru.

Při realizaci na plošném spoji bylo nejlepších výsledků dosaženo zvětšením kapacit SC integrátoru stonásobně.

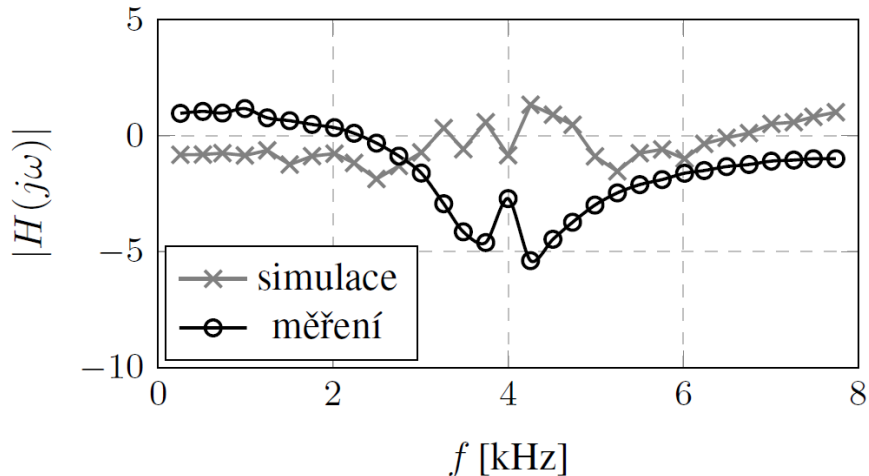
Posledním samostatně měřeným blokem je SC bikvadratická sekce, jejíž frekvenční charakteristika pro filtr typu dolní propust je uvedena na obr. 6. Změřená frekvenční charakteristika je oproti simulované mírně posunuta směrem k nižším frekvencím. Posun frekvenční charakteristiky bikvadratické sekce má vliv na frekvenční charakteristiku celé banky filtrů uvedenou na obr. 7.

Měřená frekvenční charakteristika celé realizované banky filtrů se od simulované liší především v okolí kmitočtu 4 kHz, kde rozdíl činí až 6 dB. Tento rozdíl je především způsoben posunutou frekvenční charakteristikou bikvadratických sekcí. Špička měřené frekvenční charakteristiky na kmitočtu 4 kHz je způsobena nastavením časovacích signálů vzorkovacích obvodů syntetizující části banky filtrů.

Měřená frekvenční charakteristika je také částečně ovlivněna 10 ns jitterem použitých generátorů realizujících časovací signály.



Obr. 6: Měřená a simulovaná frekvenční charakteristika SC bikvadratické sekce



Obr. 7: Měřená a simulovaná frekvenční charakteristika realizované banky filtrů

4. Závěr

Příspěvek se zabývá realizací a vlastnostmi dvoukanálové banky filtrů realizované technikou spínaných kapacitorů. Výsledné frekvenční charakteristiky jak jednotlivých funkčních bloků, tak celé banky filtrů jsou porovnány se simulacemi. Simulovaná a změřená frekvenční charakteristika banky filtrů se liší až o 6 dB. To je oproti simulaci způsobeno především posunutými frekvenčními charakteristikami bikvadratických sekcí. Pro dosažení konstantní frekvenční charakteristiky je tedy třeba obvod zoptimalizovat.

Poděkování

Tato práce byla podpořena grantem Studentské grantové soutěže ČVUT číslo SGS12/143/OHK3/2T/13.

Reference

- [1] Náhlík, J.; Hospodka, J.; Sovka, P.; Pšenička, B. Implementation of a two-channel maximally decimated filter bank using switched capacitor circuits. *Radioengineering* 2013, vol. 22, no. 1, p. 167–173, ISSN 1210-2512.
- [2] Phong S.; Kim C. W.; Vaidyanathan P. P.; Ansari I. RA New Class of Two-Channel Biorthogonal Filter Banks and Wavelet Bases. *IEEE Transactions on Signal Processing*, Vol. 43, No. 3, March 1995, pp. 649–665.
- [3] Lowenberg P., Johanson H., Wanhammer; Two-Channel Hybrid Analog/Digital Filter Banks with Alias-Free Subbands.. *Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems.*, 2000, Vol. 3, pp. 1162–1165.

- [4] Ananda Mohan P. V., Ramachandran V., Swamy M. N. S. ; SWITCHED CAPACITOR FILTERS Theory, Analysis and Design. *Prentice Hall 1995* ISBN 0-13-879818-4.
- [5] Bicak J., Hospodka J. ; Frequency Response of Switched Circuits in SPICE. *ECCTD '03* Cracow 2003, Vol. I, pp. 333-336, Poland, ISBN 83-88309-95-1.
- [6] Analog devices, Inc., *AD8033 Data Sheet*, Rev C [online] April 2008 [cit. 2012-10-06], URL: <http://www.analog.com/>
- [7] Analog devices, Inc., *AD8033 Spice Macro-model*. Rev E [online] August 2012 [cit. 2012-10-06], URL: <http://www.analog.com/>
- [8] Vishay Intertechnology, Inc., *DG456 Data Sheet*, Rev F [online] October 2012 [cit. 2013-01-14], URL: <http://www.vishay.com/>

Analýza chování biomedicínského modelu Huntingtonovy choroby

Matouš POKORNÝ¹, školitel specialista: Jan HAVLÍK¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

matous.pokorny@fel.cvut.cz

Abstrakt: Huntingtonova choroba (Huntington's disease, HD) je dědičné neuropsychiatrické onemocnění s fatálním průběhem. Hlavními příznaky nemoci jsou poruchy hybnosti (poruchy volní motoriky a mimovolní pohyby), úbytek intelektových schopností progredující do demence a poruchy chování (úzkost, deprese, agresivita). HD vzniká mutací genu IT15 na krátkém raménku 4. chromozomu. Tato mutace mění strukturální vlastnosti proteinu huntingtinu, což vede ke změně jeho funkce s toxickými důsledky. Nemoc se začíná projevovat typicky mezi 35. – 50. rokem života, vzácněji pak dříve (do 20 let) či později (po 60. roce). Průměrná doba přežití je zhruba 15 – 20 let. Výskyt HD se v Evropě a Americe v současnosti odhaduje na zhruba 5 – 10 případů na 100 tisíc obyvatel.

V rámci výzkumu Ústavu živočišné fyziologie a genetiky AV ČR vzniká biomedicínský model HD (projekt PIGMOD – Pig Models of Diseases). Jedná se o geneticky modifikovaná miniprasata, která lze využít pro výzkum diagnostiky nemoci a testování všech nových léčebných postupů. Miniprasata lze umístit do kotců s kamerovým systémem a monitorovat je během experimentu. Vyhodnocení kamerových záznamů by mělo dokázat, zda došlo ke změně chování (reakce na přísun potravy, spánkový režim) miniprasat během vývoje nemoci. To je důležité pro další pochopení vývojových fází HD.

Základní představu o chování modelu během experimentu může dát sledování pohybu prasete po kotci. Digitální kamera pořizuje záznam z experimentu, který je možno zpětně analyzovat a vyhodnotit tak pohyb (chování) modelu během experimentu. Kamera je umístěna shora, zorné pole kamery tedy obsáhne celý prostor, kde se zvíře pohybuje. Pro následnou analýzu záznamu (signálu) se uvažuje použití algoritmu Gaussian Mixture Model (GMM) v prostředí Matlab. Tento algoritmus umožňuje klasifikaci obrazu na popředí a pozadí, je vhodný pro zpracování signálu ze staticky umístěné kamery s poměrně stálým pozadím a neznámým popředím. Jako pozadí je v tomto případě chápán kotec, takže pozadí je tvořeno jednolitou barevnou plochou (ohrádka) nebo barevně nevýraznou texturou (podestýlka). Neznámým popředím se rozumí miniprase, které může mít různou barvu srsti, velikost, atd. Algoritmus je imunní vůči pomalým nebo periodickým změnám v obraze, např. měnící se stín. Klasifikace částí obrazu je závislá na správném nastavení parametrů. Po klasifikaci obrazu na popředí a pozadí se dále zaznamenává změna pozice popředí v obraze. V první fázi se uvažuje statistická analýza získaných trajektorií pohybu miniprasete po kotci, tj. jakou dobu je zvíře v klidu / pohybu, v jaké části kotce se nejčastěji vyskytuje. Dále je možná sofistikovanější klasifikace trajektorií na specifické případy, případně použití optických značek na sledovaném minipraseti.



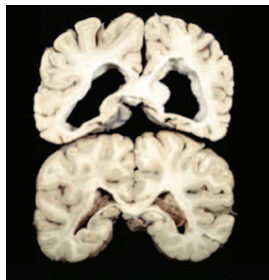
Analýza chování biomedicínského modelu Huntingtonovy choroby

Matouš Pokorný
Biomedical Electronics Group
Katedra teorie obvodů

2

Huntingtonova choroba

- Dědičné neuropsychiatrické onemocnění
- Mutace genu IT15 mění strukturu proteinu huntingtinu
- Toxické následky
- Degenerace bazálních ganglií



The human brain, showing the impact of HD on brain structure in the basal ganglia region of a person with HD (top) and a normal brain (bottom).
<http://hobdtek.msu.edu>

3

Úvod

- Huntingtonova choroba
- Biomedicínský model Huntingtonovy choroby
- Analýza chování modelu
- Sledování pohybu
- Další vývoj práce

Huntingtonova choroba

- Hlavní příznaky:
 - Poruchy hybnosti – poruchy volní motoriky a mimovolní pohyby
 - Psychické poruchy – úbytek intelektových schopností progredující do demence a poruch chování (úzkost, deprese, agresivita)

4

Huntingtonova choroba

- Projevy nemoci:
 - Mezi 35. – 50. rokem života
 - Vzácněji dříve (do 20 let) či později (po 60. roce)
- Výskyt 5 – 10 případů na 100 000 obyvatel (Evropa a Amerika)

5

Biomedicínský model Huntingtonovy nemoci

- Miniprasata, projekt ÚŽFG AV ČR, PIGMOD – Pig Models of Diseases
- Choroba se přenáší po generacích (několikátá generace, desítky kusů)
- Výhodnější než jiná zvířata – ovce, krávy
- Výzkum diagnostiky HD a všech druhů léčby



zdroj: <http://www.iapg.cas.cz/uzfg>

6

Analýza chování modelu

- Snaha detekovat příznaky HD u miniprasat
- Sleduje se, zda dojde ke změně chování během vývoje nemoci, reakce na farmaka ...
- Celková aktivita
- Spánkový režim
- Reakce na přísun potravy ...
- Lepší pochopení vývoje nemoci

7

Analýza chování modelu

- Videozáznam experimentů ze statické kamery
 - ~ 7 měsíců
 - ~ 3 dny v týdnu
 - ~ 1 hodina záznamu
- Detekce změny chování na základě sledování pohybu (motion tracking) miniprasete v kotci
- Offline analýza

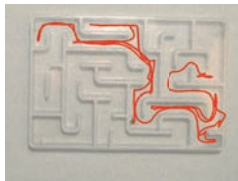


zdroj: ÚŽFG AV ČR

8

Sledování pohybu

- Sledování pohybu založené na Gaussian Mixture Model (GMM)
 - Detekce popředí
 - Sledování objektu
- Aplikace:
 - Kontrola aut na silnici, sledování zvířat v bludišti ...
- Nepotřebuje trénovací data, výpočetně náročný (po pixelech), statická kamera



zdroj: <http://www.cs.berkeley.edu/~flw/tracker/>

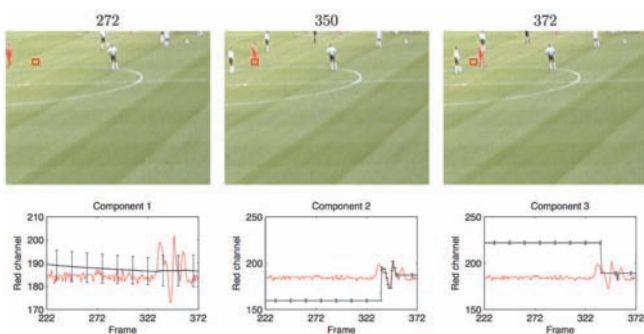
9

Sledování pohybu – detekce popředí

- Klasifikace popředí / pozadí
- Model pozadí (pro každý pixel, aktualizace každý frame)
 - Směs 3 – 5 Gaussovských rozdělání (1 směs na kanál)
- Pozadí je jednodílná plocha jedné barvy a nebo textura
- Libovolné popředí, nemusí být definováno
- Imunní vůči pomalým změnám osvětlení, lokálním pravidelným pohybům (listí ve větru)

10

Sledování pohybu – detekce popředí

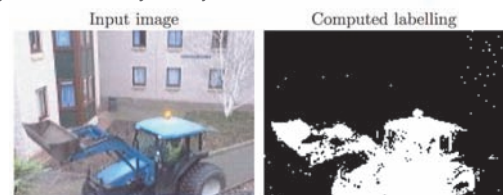


zdroj: <http://www.cs.berkeley.edu/~flw/tracker/tracker-synopsis.pdf>

11

Sledování pohybu – sledování objektu

- Spojení fragmentů popředí
- Vytvoření trajektorie pomocí Kalmanova filtru
- Je možné sledovat více objektů současně, jsou ale problémy s překryvem sledovaných objektů



zdroj: <http://www.cs.berkeley.edu/~flw/tracker/tracker-synopsis.pdf>

12

Další vývoj práce

- Statistická analýza
 - Zvíře v klidu / v pohybu
 - Kde tráví nejvíce času
- Sofistikovanější analýza
 - Reakce na přísun potravy
 - Detekce jednotlivých fází spánku (EEG, EMG)

Analýza chování biomedicínského modelu Huntingtonovy choroby

Matouš Pokorný
Biomedical Electronics Group

pokormat@fel.cvut.cz
bmeg.fel.cvut.cz

Prof. Ing. Pavel Sovka, CSc.
Ing. Jan Havlík, Ph.D.

školitel
školitel specialista

ČVUT – SGS12/143/OHK3/2T/13

granty

Ústav živočišné fyziologie a genetiky AV ČR
Neurologická klinika 1. LF UK a VFN v Praze

spolupráce

Optimalizace parametrů modelu KVS

Miroslav LOŽEK, Jan HAVLÍK¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

lozekmir@fel.cvut.cz

Abstrakt: Tento příspěvek se zabývá modelováním biologických systémů se zaměřením na kardiovaskulární systém (KVS). Projekt je rozdělen do několika etap. Nejprve byl zkonstruován obecný mechanický model základních částí KVS, jehož účelem bylo ověření možností simulace především pulzatilního toku kapaliny. Dále byla provedena optimalizace základních parametrů tak, aby bylo dosaženo alespoň parciálních shod s fyziologickým systémem. Dalším krokem bude navázání mechanického modelu KVS na model softwarový, který byl vyvinut ve výzkumném institutu kardiovaskulárních chorob v Maastrichtu. Poslední etapou (ve spolupráci s Dětským kardiocentrem ve FN v Motole) bude modifikace soustavy modelů na specifickou vrozenou komplexní vadu – Fallotovu tetralogii. Výsledkem tohoto snažení by mohl být nástroj, který dopomůže k lepší predikci vývoje choroby a volbě optimální terapie či chirurgické korekce.

Kardiovaskulární systém je soustava mnoha složitých orgánů a řídicích center, která umožňují distribuci životodárných látek do celého organismu. Hlavním elementem tohoto systému je srdce, které slouží jako pohonný zdroj cirkulace. Nedílnou součástí je také cévní síť, která svými parametry zásadně ovlivňuje tok krve. Mechanický model byl vyvinut s ohledem na tuto fyziologickou stavbu. Pohonným prvkem je zde mechanické čerpadlo, které je řízeno mikroprocesorovou jednotkou tak, aby bylo schopno práce v pulzatilním režimu. Dále jsou použity elektromagnetické ventily, které simulují funkci srdeční chlopně. Cévní systém je sestaven z umělých hadic a kapilárních filtrů s respektováním jejich hydraulických (hemodynamických) vlastností.

Mechanický model KVS je vybaven vlastním měřicím systémem, který umožňuje měření hemodynamických parametrů různými způsoby. Tlakové vlastnosti je možné měřit v různých částech vaskulárního systému současně až třemi implementovanými tlakovými senzory za pomoci klinických kateizačních setů. Stanovení průtoku (Cardiac Output) je zabezpečeno jednak přesným interním průtokovým senzorem a jednak konvenčními klinickými technikami – termodilučními metodami (přímá i transpulmonální – přes kapilární řečiště), barvivovou diluční metodou a dopplerovskou sonografií.

Za účelem optimalizace modelu bylo provedeno komplexní měření hemodynamických parametrů (tlak, průtok, cévní rezistence, vaskulární compliance a PWV) pro různá nastavení modelu. Tato studie potvrdila parciální správnost návrhu modelu, ale také odhalila výrazné nedostatky a dopomohla k jejich napravení (např. velká hydraulická rezistence elektromagnetického ventilu, což má v kardiologické praxi ekvivalent silné stenózy aortální chlopně).

Projekt mechanického modelu KVS je financován z grantu SGS12/143/OHK3/2T/13 a jeho využití ve výuce je podpořeno granty FRVŠ pro rok 2012 resp. 2013, jehož cílem je začlenění tohoto nástroje do praktické výuky předmětu Lékařská technika. Studentům je umožněno vyzkoušet si klinické měření hemodynamických veličin invazivní metodou za použití katetrizačního instrumentária.



ČVUT v Praze
Fakulta elektrotechnická
 Katedra teorie obvodů



Fakultní nemocnice v Motole
 Dětské kardiocentrum

OPTIMALIZACE PARAMETRŮ MECHANICKÉHO MODELU KVS

autor: Miroslav LOŽEK



Fakultní nemocnice v Motole
 Dětské kardiocentrum

ČVUT v Praze
 Fakulta elektrotechnická
 Katedra teorie obvodů



obsah prezentace

- úvodní slovo
- anatomie a fyziologie KVS
- hemodynamické parametry KVS
- mechanický model KVS
- měření a optimalizace
- pokračování projektu
- podpora a reference



Fakultní nemocnice v Motole
 Dětské kardiocentrum

ČVUT v Praze
 Fakulta elektrotechnická
 Katedra teorie obvodů



úvodní slovo

- disertační téma
 - modelování biologických systémů
- specifikace disertačního tématu
 - model kardiovaskulárního systému
- řešitelský tým
 - školitel: prof. Ing. Pavel Sovka, CSc.
 - školitel specialista: Ing. Jan Havlík, Ph.D.
 - Biomedical Electronics Group
 - Dětské kardiocentrum, Fakultní nemocnici v Motole
 - CARIM School for Cardiovascular Diseases, Maastricht University



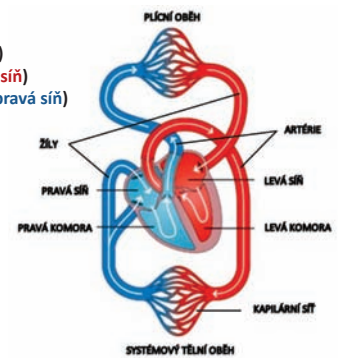
Fakultní nemocnice v Motole
 Dětské kardiocentrum

ČVUT v Praze
 Fakulta elektrotechnická
 Katedra teorie obvodů



anatomie kardiovaskulárního systému

- kardiální systém
 - čtyř-komorové srdce (2 síně, 2 komory)
 - plicní krevní oběh (pravá komora, levá síň)
 - systémový krevní oběh (levá komora, pravá síň)
- vaskulární systém
 - vysokotlaký systém – pulzatilní tok
 - Aorta (\varnothing cca 3 cm)
 - Arterie (\varnothing cca > 10 mm)
 - Arterioly (\varnothing cca > 100 μ m)
 - Kapiláry (\varnothing cca 10 μ m)
 - nízkotlaký systém – kontinuální tok
 - Venuoly (\varnothing cca > 20 μ m)
 - Vény (\varnothing cca > 0,5 mm)
 - Věna Cava (\varnothing cca 3 cm)



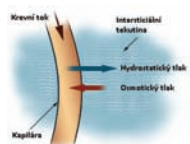
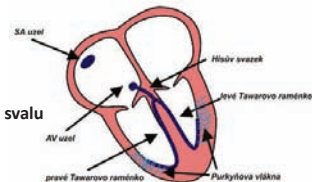
Fakultní nemocnice v Motole
 Dětské kardiocentrum

ČVUT v Praze
 Fakulta elektrotechnická
 Katedra teorie obvodů



fyziologie kardiovaskulárního systému

- převodní systém srdeční
 - specifické tzv. pacemakerové buňky
 - sekvence elektrických impulzů -> kontrakce svalu
- přenos krevních plynů
 - krevní hemoglobin – schopnost vázat a přenášet O_2 a CO_2
 - k přenosu dochází v kapilárním řečišti (parciální tlaky)
- regulace hemodynamických parametrů
 - vasodilatační a vazokonstriční mechanismy
 - regulace zajištěna hormonálním a nervovým systémem



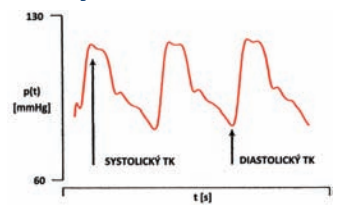
Fakultní nemocnice v Motole
 Dětské kardiocentrum

ČVUT v Praze
 Fakulta elektrotechnická
 Katedra teorie obvodů



hemodynamické parametry KVS

- krevní tlak
 - tlaková křivka – pulzatilní charakter
 - lokální extrémy systolický vs. diastolický TK
 - srdeční výdej (Cardiac Output)
 - objem vypuzené krve srdcem do krevního řečiště
 - vztaženo k srdečnímu cyklu = Stroke Volume [ml]
 - vztaženo k jednotce času = Minute Cardiac Output [l/min.]
- $Minute\ Cardiac\ Output = Heart\ Rate \cdot Stroke\ Volume$





hemodynamické parametry KVS

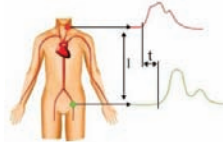
- cévní rezistence
 - síla působící proti toku krve v cévách
- cévní compliance
 - popisuje pružnost (elasticitu) cévy
 - změna objemu vlivem působícího tlaku
- cévní inertiance
 - reprezentuje setrvačnost
 - uplatňuje se při změnách průtoku
 - je ovlivněn geometrií a délkou cév
- rychlost šíření pulzní vlny (PWV)
 - vzruch vyvolaný rychlou změnou objemu
 - ovlivněno především compliance cévy
 - ukazatel aterosklerotických změn cévy

$$R = \frac{\Delta p}{Q} \left[\frac{\text{mmHg} \cdot \text{min.}}{l} \right]$$

$$C = \frac{\Delta V}{\Delta p} \left[\frac{\text{m}^3}{\text{N}} \right]$$

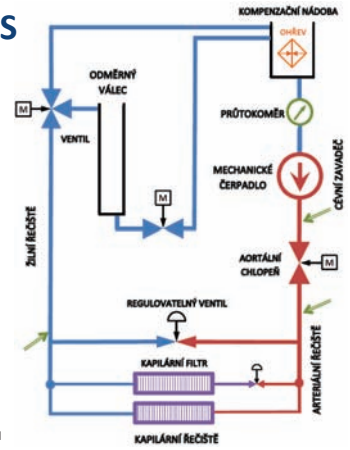
$$L = \frac{\rho \cdot l}{S} \left[\frac{\text{N} \cdot \text{s}^2}{\text{m}^5} \right] \quad \Delta p = L \cdot \frac{dQ}{dt}$$

$$PWV = \frac{l}{t}$$

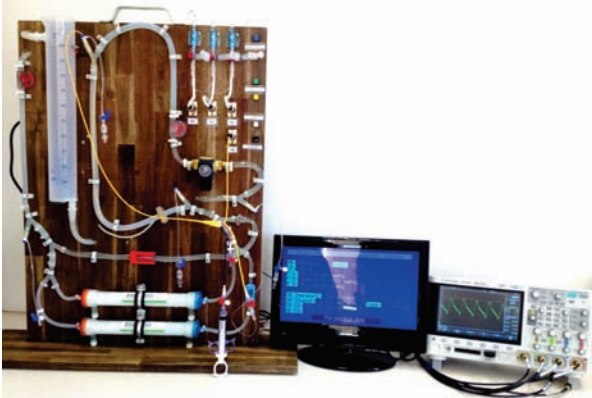


mechanický model KVS

- soustava hadic, ventilů a čerpadla
- arteriální řečiště – elastická hadice
- venózní řečiště – rigidní hadice
- kapilární řečiště - rezistence
- chlopeč – elektromagnetický ventil
- průtokoměr – přesné měření CO
- odměrný válec – měření CO
- regulovaný ohřev kapaliny (37 °C)
- klinické cévní zavaděče – vstup katetrů

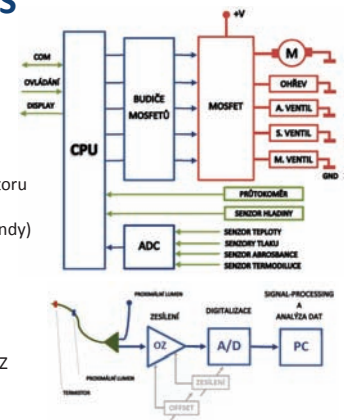


mechanický model KVS



mechanický model KVS

- mikroprocesorové řízení
- pulzní buzení čerpadla - PWM
- ovládání pomocí tlačítek (HR, EF, mód)
- zobrazení nastavení a měření na monitoru
- pulzní průtokoměr (virtulka, Hallovy sondy)
- AD převod všech měřených veličin
- měření tlaků, teplot a absorbance
- zpracování měřených veličin pomocí OZ
- zdroj napětí – 12 V / 300 W

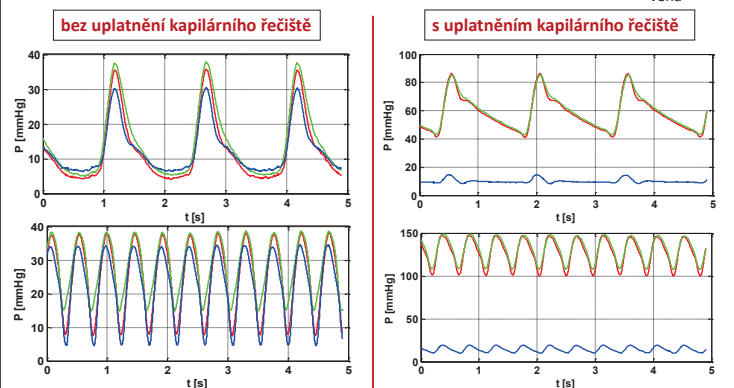


měření hemodynamických parametrů

- tlak
 - měření pomocí klinických převodníků
 - možnost měřit v různých částech řečiště
 - cévní zavaděč, katetizační set
- průtok (Cardiac Output)
 - průtokoměr (referenční)
 - termodiluční metoda (přímá i transpulmonální)
 - barviová diluční metoda
 - dopplerovská sonografie
- cévní compliance
 - měření Δp při ΔV
- rychlost šíření pulzní vlny
 - analýza křivky tlaku
 - měřeno ve dvou místech arteriálního řečiště

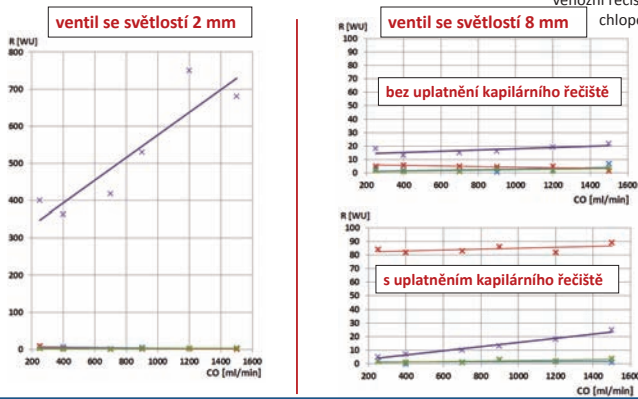


měření tlakových křivek

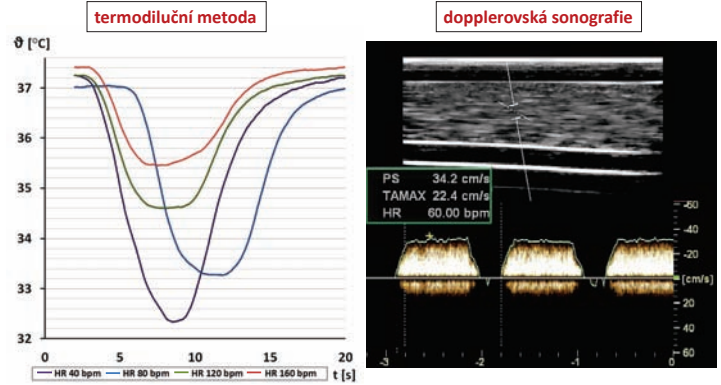




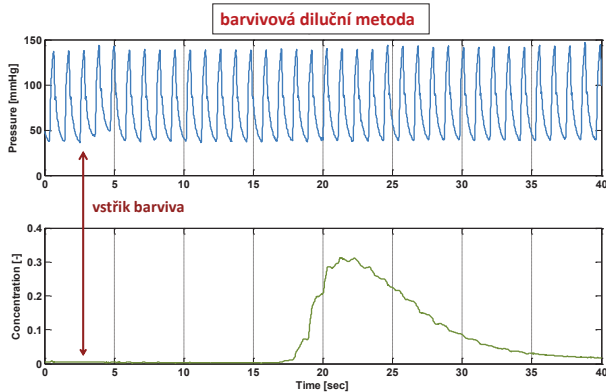
měření rezistence



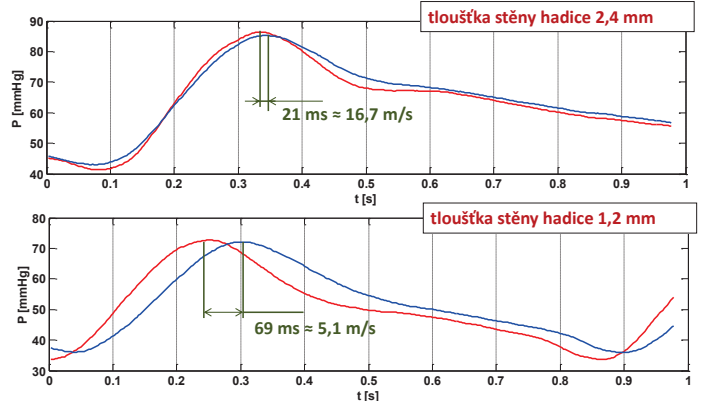
měření průtoku



měření průtoku



měření PWV



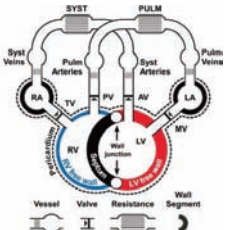
optimalizace parametrů

- volba čerpadla
 - vrtulkové x membránové – nedostatečný tlak x bez zpětného chodu
 - volba: simulace křivky komorové tlaku x tvrdý zdroj tlaku
- elektromagnetický ventil – aortální chlopeč
 - světlost ventilu – odstranění hydraulické rezistence ventilu
- typ hadice arteriálního řečiště
 - optimalizace PWV (5 m/s = fyziologická hodnota)
- úprava časování otvírání chlopečů
 - optimalizace arteriální tlakové křivky
- stanovení poměru parametrů vůči fyziologickému systému
 - tlakové poměry 1:1 (cca 100 mmHg)
 - poměr průtoku 1:10 (cca 550 ml/min : 5500 ml/min)
 - poměr systémové rezistence 10:1 (cca 150 WU : 10 – 20 WU)



pokračování projektu

- modifikace mechanického modelu
 - spolupráce s univerzitou v Maastrichtu (CARIM)
 - softwarový model KVS
 - pokus alespoň o parciální ekvivalenci
- modifikace komplexu modelů na TOF
 - modifikace obecného modelu na vrozenou vadu
 - Fallotova tetralogie
 - Defekt komorového septa
 - Dextropozice aorty
 - Stenóza plicnice
 - Hypertrofie RV
 - klinická data – Dětské kardiocentrum (FN v Motole)





podpora a reference

- grant FRVŠ 2012 resp. 2013
 - vývoj modelu pro praktickou výuku
 - Mechanický model KVS, Telemetricky řízený adaptivní model KVS
- grant SGS12/143/OHK3/2T/13
- reference
 - M. Ložek, M. Horáková, J. Havlík, "Mechanical model of the cardiovascular system: determination of cardiac output by dyE dilution," in Lékař a technika. 2012, vol. 42, no. 2, p. 77-80. ISSN
 - M. Ložek, "Model of the Cardiovascular System: Pump Control," in POSTER 2012, Czech Technical University in Prague, 2012, vol. 1, p. 1-4. ISBN 978-80-01-05043-9.
 - J. J. Batzel, F. Kappel, D. Schneditz, H. T. Tran, "Cardiovascular and Respiratory Systems: Modeling, Analysis, and Control," 2006. ISBN 978-0898716177.
 - N. Westerhof, N. Stergiopoulos, M. I. M Noble, "Snapshots of Hemodynamics," Basic Science for the Cardiologist, 2005.
 - C. Cobelli, E. Carsin, "Introduction to modeling in physiology and medicine," Academic Press.2008. ISBN:978-0-12-160240-6.
 - J. Lumens, T. Delhaas "Cardiovascular Modeling in Pulmonary Arterial Hypertension: Focus on Mechanisms and Treatment of Right Heart Failure Using the CircAdapt Model," The American Journal of Cardiology, Volume 110, Issue 6, Supplement, 15 September 2012.

Měření hemodynamických parametrů, metody, realizace, problémy

Jan DVORÁK¹ (školitel: Jan HAVLÍK¹)

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

dvoraj45@fel.cvut.cz

Abstrakt: Jednou z nejčastějších příčin úmrtí jsou v současnosti choroby kardiovaskulárního systému. Zvláště ateroskleróza představuje jeden z nejzávažnějších problémů, neboť je v raných stádiích obtížně diagnostikovatelná, přitom ale způsobuje postupné nevratné změny kardiovaskulárního systému. Při ateroskleróze dochází k ukládání tukových částic do stěn cév, ke snížení jejich pružnosti, k zužování jejich průměru a k omezení průtoku krve. V současnosti existuje několik klinických metod na záchyt pacientů s aterosklerózou a na zhodnocení míry poškození tepen tímto procesem. Žádná z doposud využívaných metod však není široce uplatnitelná. Je snahou, aby použitá metoda byla neinvazivní a co nejméně zatěžující pro pacienta, spolehlivá pro široký okruh pacientů a cenově dostupná pro masivní nasazení ve zdravotnickém systému.

Ukazuje se, že při měření krevního tlaku oscilometrickou metodou je možné při vhodném uspořádání měření získat některé sekundární parametry - hemodynamické parametry kardiovaskulárního řečiště. Všechny tyto parametry lze elegantně získat pomocí kombinace standardních lékařských vyšetření (elektrokardiografie, fotoplethysmografie, záznam oscilometrických pulzací při měření tlaku krve). Proto byla vytvořena databáze signálů s využitím komplexního měřicího zařízení, umožňujícího synchronně zaznamenávat EKG, PPG a oscilometrické pulsace během postupného napouštění a vypouštění manžety. Aktuálně signálová databáze obsahuje záznamy více než 70 pacientů ve věkovém rozmezí 19 až 94 let. Měřené signály v databázi byly doplněny údaji z anonymizovaných anamnestických dotazníků, které obsahují základní údaje o měřených osobách relevantní s ohledem na řešenou problematiku (pohlaví, věk, tělesné parametry, některé anamnestické údaje, informace o kouření a užívání drog a léků apod.).

Na základě provedených měření byly prokazatelně zjištěny signifikantní rozdíly středního arteriálního tlaku změřeného oscilometrickou metodou při nafukování a vyfukování manžety. Tento tlakový rozdíl se při první analýze dat zdá být u starších osob větší ve srovnání s mladšími osobami. Parametr delta MAP tedy podle prvních výzkumů koreluje s mírou poškození kardiovaskulárního řečiště. Dalším zajímavým markerem pro určení míry aterosklerotického procesu se jeví měření časové vzdálenosti mezi QRS komplexem (vlnou R) na EKG signálu a dobou odpovídající počátku růstu PPG signálu. Tento "fotoplethysmografický" čas šíření pulzní vlny (označován PTT) je tím kratší, čím je míra postižení tepen aterosklerózou větší.



Měření hemodynamických parametrů

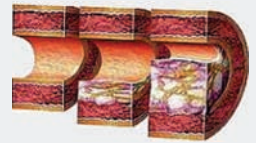
metody, realizace, problémy

Student: Jan Dvořák
Školitel spec.: Jan Havlík
Školitel: Pavel Sovka

LDD 2013

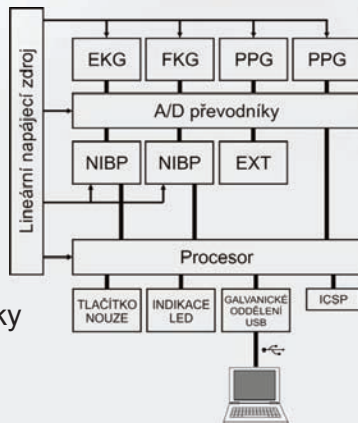
Hemodynamické parametry

- Hemodynamika studuje vliv fyzikálních vlastností krve a cév na tlak krve a její proudění cévami
- Zaměřeno na neinvazivní screening aterosklerózy
 - ukládání tukových látek způsobí zúžení stěny tepny, snížení její pružnosti a může omezit průtok krve
 - vede k řadě onemocnění (infarkt myokardu, mozková mrtvice)
- Většinou nepřímo měřené, odvozené parametry
 - Krevní tlak (SBP, DBP, MAP)
 - Index tuhosti cévní stěny (ASI)
 - Rychlost šíření pulzní vlny (PWV)
 - Tlakový index kotník-paže (ABI)
 - Index zesílení pulzní vlny (AI)
 - Index arteriální poddajnosti (CAVI)



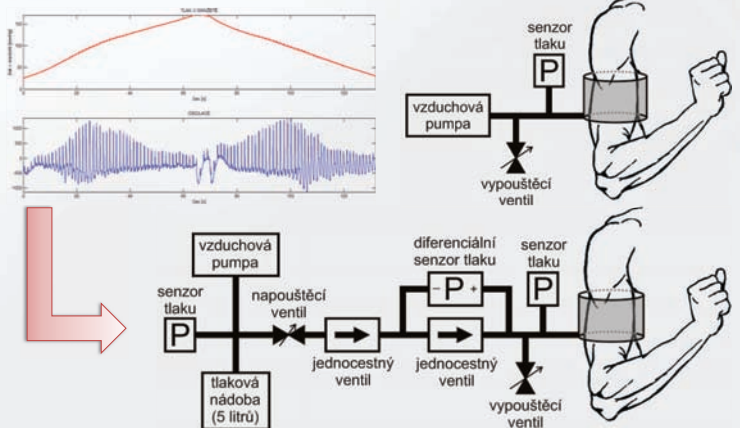
Čím měříme?

- 2 kanály EKG
- 2 kanály PPG
- FKG
- Klasický NiBP modul
- Speciální NiBP modul
- Externí vstupy



- 12-bitové A/D převodníky
- Galvanické oddělení

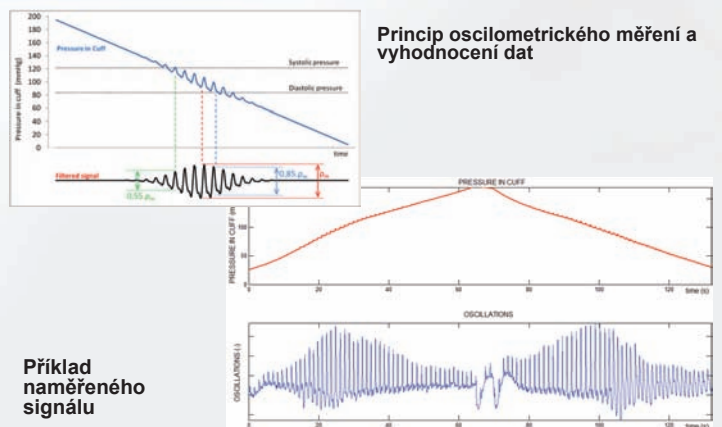
NiBP moduly



Metody

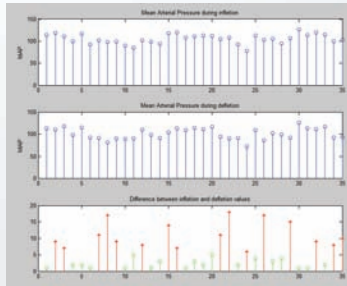
- Současné metody pro screening aterosklerózy nejsou široce uplatnitelné
- Snaha vyvinout metodu **neinvazivní, co nejméně zatěžující pro pacienta**, spolehlivou pro **široký okruh pacientů** a cenově dostupnou pro masivní nasazení ve zdravotnickém systému
- V současné době dva nekonvenční ukazatele
 - Rozdíl MAP při nafukování a vyfukování manžety
 - PPTT (photo-plethysmograph pulse transit time)

Rozdíl MAP

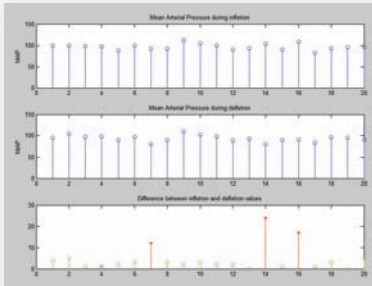


Rozdíl MAP

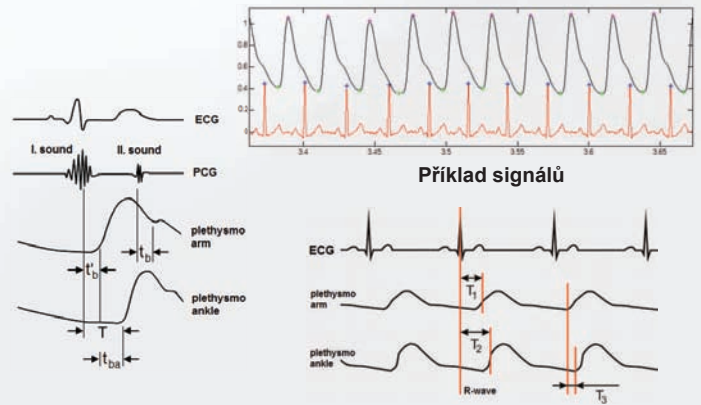
seniři



studenti

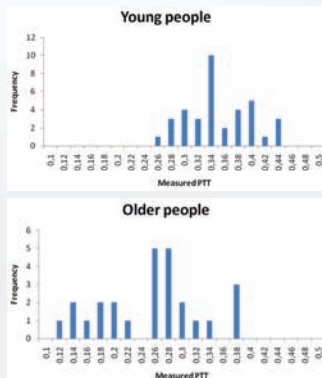


Rychlost šíření vlny (PTT)

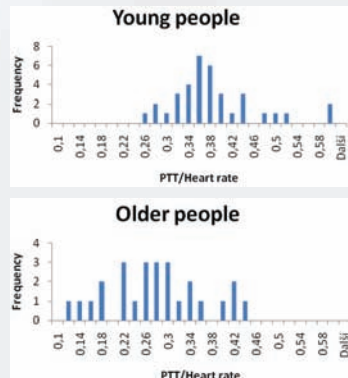


Rychlost šíření vlny (PTT)

Absolutní čas PTT



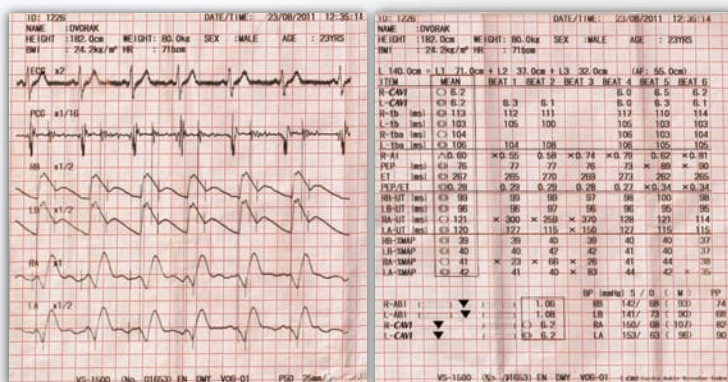
Přepočítané PTT na normální tep



Validace dat

- Část databáze je měřena naším zařízením a zároveň přístrojem VaSera VS-1500
- VaSera je v dnešní době „standardem“ v oblasti neinvazivního měření hemodynamických parametrů
- Používá 4 manžety, PCG a EKG
- VaSera poskytuje nejen výsledek měření ale také naměřené některé parametry (nejen odvozené)
- Některé naše algoritmy lze ověřovat

Validace

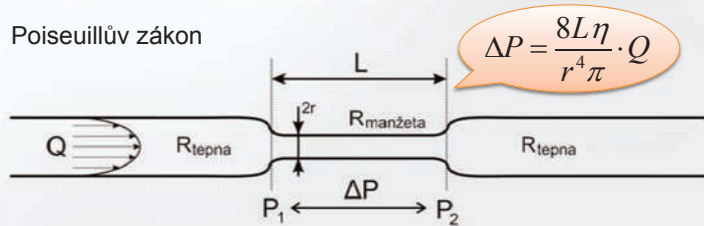


Další perspektivy

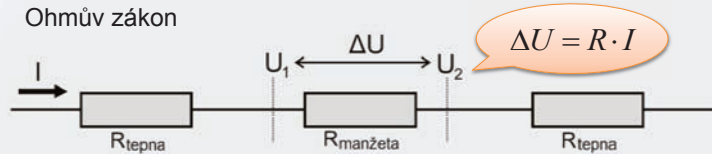
- Pokračování v tvorbě signálové databáze
- Validace algoritmů přístrojem VaSera
- Ověření vlivu velikosti manžety na přesnost měření
- Výzkum korelaci stavu KVS a rozdílu MAP v souvislosti s anamnestickými informacemi pacientů
- Zjištění množství nové užitečné informace při použití dvoumanžetového systému ruka-noha
- Výzkum závislosti tvaru diastolické fáze PPG signálu na věku a na tuhosti tepen (a tedy na míře aterosklerózy)
- Průzkum vypovídací hodnoty fonokardiogramu oproti EKG signálu při měření hemodynamických parametrů
- Opakovatelnost měření, odstranění artefaktů ze signálu

Ověření vlivu manžety na výsledky měření

Poiseuillův zákon



Ohmův zákon



Děkuji za pozornost

Měření ultraslabé emise fotonů z biologických vzorků

Michaela Nerudová^{1,2}, Kateřina Červinková^{2,3}

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra
teorie obvodů

² Ústav fotoniky a elektroniky, Akademie věd ČR, v.v.i.

³ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra
elektromagnetického pole

nerudmic@fel.cvut.cz, cervikat@fel.cvut.cz

Abstrakt: Tento článek se zabývá metodikou měření ultraslabé emise fotonů v optické oblasti. Procesy generující ultraslabou fotonovou emisi v biologických vzorcích souvisejí zejména s reakcemi reaktivních forem kyslíku, které generují elektronově excitované stavy. Tyto procesy jsou doposud objasněny jenom částečně. Hlavní motivací výzkumu ultraslabé emise z biosystémů je a) objasnění její role v biologii a b) využití jejích parametrů pro diagnostické účely v budoucí medicíně a biologii. V tomto článku je detailně popsán měřicí systém ultra slabé emise fotonů a návrhy pro jeho vylepšení. Účelem zde popsaných experimentů bylo prokázat schopnost změřit ultraslabou emisi fotonů z různých biologických vzorků. K měření byly použity buňky HMVECad, kvasinky (*Saccharomyces cerevisiae*) a semínka fazole munga. Naměřené signály jsou relevantní, jelikož odstup signálu z biologického vzorku ku signálu z pozadí(SNR) je příznivý, u semínek munga až 10.4 dB. V závěru jsou popsány možné optimalizace jednotlivých měření.

1. Úvod

Ve všech typech živých buněk probíhají metabolické procesy generující reaktivní kyslíkové formy (ROS - Reactive Oxygen Species) a volné radikály. Tyto procesy nacházíme například v mitochondriích [1], během respiračního vzplanutí [2], patofyziologických stavech [3] a dalších. Některé reakce volných radikálů a reaktivních kyslíkových forem vedou k nestabilním chemickým meziproductům (dioxetány, tetraoxidy lipidů a proteinů), z nichž po přeuspořádání jejich vazeb mohou vznikat elektronově excitované produkty [1]. Část těchto elektronově excitovaných molekul projde do nižšího stavu

vyzářením fotonů o vlnových délkách v rozmezí od 200 nm do 800 nm [1]. Na základě naměřených dat víme, že intenzita ultraslabé fotonové emise je kolem 10^2 fotonů/s*cm² vzorku [4]. Jelikož se volné radikály a ROSy účastní nejenom patologických procesů, ale i přirozené biologické signalizace [3], lze očekávat, že generovaná fotonová emise bude nést informaci o těchto procesech.

2. Měřicí aparatura

Měřicí aparatura (Obr. 1) se skládá z černé světlotěsné komory, fotonásobiče a jeho zdroje napětí, chladicí jednotky pro fotonásobič a jejího zdroje napětí, počítačí jednotky, teplotní regulační jednotky a počítače. Komora je vyrobena z 10 mm duralových desek a její vnitřek je pokryt izolačními deskami z extrudovaného polystyrenu. Desky jsou 5 mm tlusté a jejich součinitel prostupu tepla je 4,651 W/m²K. Použitým desek bylo dosaženo většího teplotního rozsahu uvnitř komory.

Použitý **fotonásobič** R4220P (Hamamatsu Photonics K. K.), má měřicí rozsah vlnových délek od 185 do 710 nm, přičemž nejvyšší citlovosti dosahuje při 410 nm. Tento fotonásobič má velmi citlovou katodu i anodu a nízké temné proudy, díky čemuž je vhodný pro měření ultraslabé fotonové emise. Jedná se o side-on typ detektoru, který má neprůhlednou fotokatodu a kruhovou klec dynod. Vstupní okénko detektoru je vyrobeno z křemenného skla, které propouští širší spektrum záření oproti běžnému příměsovému sklu, jež nepropouští záření pod 400 nm. K **napájení fotonásobiče** bylo použito vysoké napětí o hodnotě -1150V z napájecího zdroje PS350 (Stanford Research Systems).

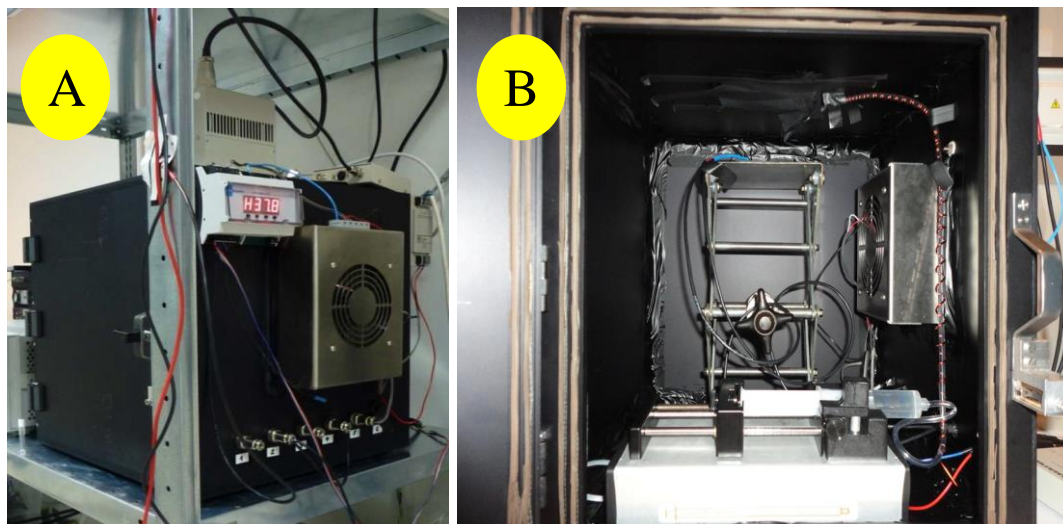
Chladicí jednotka fotonásobiče C9144-02 (Hamamatsu Photonics K. K.) funguje na principu Peltierova článku. Je složena z kontrolní jednotky a pouzdra, ve kterém je umístěna trubice fotonásobiče. Chladicí jednotka má teplotní rozsah od -30° C do -5° C, avšak výsledná teplota fotonásobiče je závislá na rozdílu teplot mezi okolním prostředím a chladicí jednotkou, přičemž tento rozdíl může být maximálně 50° C.

Teplota uvnitř boxu je regulována pomocí **termoregulace** (UWETRONIC), která se skládá z **ovládací jednotky** UETR-MOST-16A, **napájení** OMRON S8VS-12024 a **Peltierova článku** s dvěma větráky A2A-100W. Díky izolační vrstvě a termoregulaci je možné uvnitř komory dosáhnout teplot v rozmezí od 19,2° C do 38,6° C. Tento teplotní rozsah je pro měření tkáňových kultur (37° C) a kvasinek (28° C) dostačující.

Počítačí jednotka C9744 (Hamamatsu Photonics K. K.) umožňuje převod detekovaných fotonů na digitální signál. Počítačí jednotka se sestává ze zesilovače a diskriminačního okruhu. Signál z fotonásobiče je nejdříve zesílen a poté přiveden do komparátoru, kde je porovnán s prahovým napětím, v našem případě nastaveném na 450 mV. Signály o nízké intenzitě jsou zamítnuty a tím se snižuje příspěvek šumu a citlivost ve výsledném signálu.

Nezamítnuté impulsy jsou poté upraveny v tvarovacím obvodu tak, aby měly jednotlivé peaky konstatní šířku.

Počítač slouží k ovládání termoregulace komory, chladicí jednotky fotonásobiče a vlastního programu měření. Pro měření existuje možnost nastavení délky měření a časového úseku pro příjem fotonů.



Obr. 1: Měřicí systém a) pohled z venku, b) pohled dovnitř

3. Měření biologického materiálu

Měření biologických vzorků, je ovlivněno mnoha faktory. Před zahájením sběru dat je nutné charakterizovat laboratorní pomůcky (Petriho misky, skleněné kádinky a jiné.), ve kterých probíhá měření a média, jež zajišťují životnost buněk a rostlin. Z těchto důvodů byly proměřeny jak biologické vzorky, tak i použitá média a laboratorní pomůcky.

3.1. Měření kvasinek

K měření ultraslabé fotonové emise byl použit speciální typ β tubulinového mutantu tub2-401 rodu CUY67 genotypu Meta tub2-401 ura3-52 ade2-101, který nemůže polymerovat svoje mikrotubuly pod restriktivní teplotou 14° C. Jestliže teplota suspenze klesne pod tuto hranici, kvasinky ztratí schopnost vytvářet mitotická vřeténka, jejich buněčný cyklus se tedy zastaví na konci G2 fáze. Po zvýšení teploty nad permisivní hodnotu 25° C dojde k obnově buněčného vývoje a buňky v suspenzi přecházejí sjednoceně do M-fáze. V tomto případě je díky synchronizaci vývojových cyklů předpokládán nárůst intenzity záření. Kvasinky byly měřeny ve skleněných kádinkách (10 ml) při teplotě 28° C. Pro jejich měření bylo použito YPD médium (Yeast Peptone Dextrose), které obsahuje 1% kvasinkového extraktu (vitamíny a živiny), 2% peptonu (zdroj bílkovin), 1% dextrózy (D glukóza) a destilovanou vodu.

3.2. Měření tkáňové kultury

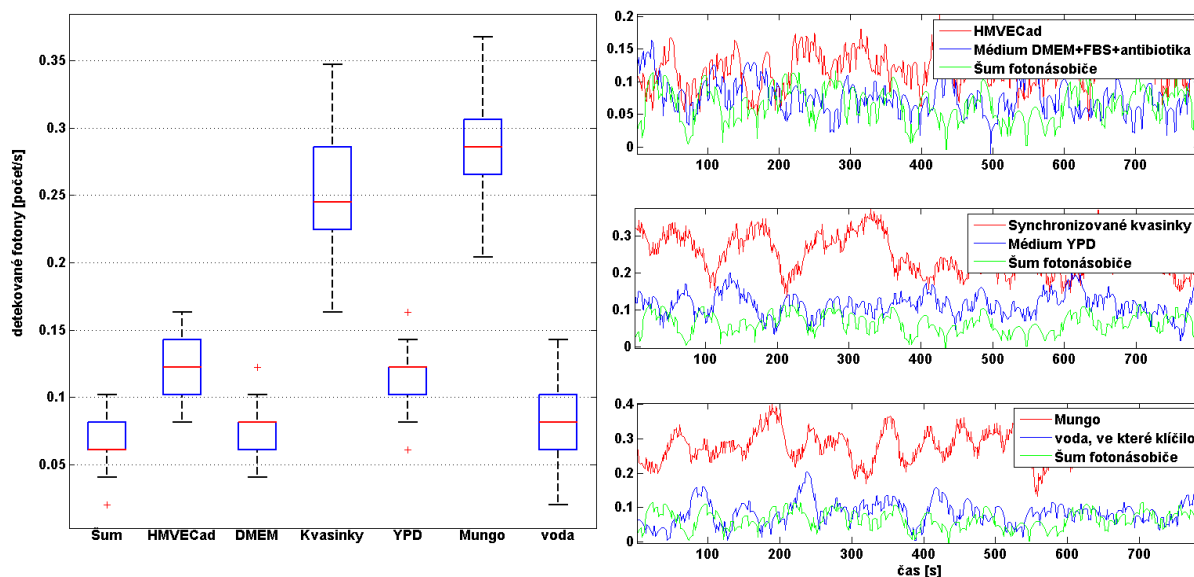
Měřicí aparaturou byla měřena fotonová emise z tkáňové kultury HMVECad (Human Microvascular Endothelial Cells, adult dermis), což jsou lidské mikrovaskulární endoteliální buňky. Buňky byly kultivovány v inkubátoru, jehož atmosféra obsahuje 5% CO₂. Kultivace a měření probíhalo ve skleněných Petriho miskách při teplotě 37° C. Pro měření bylo použito médium, které obsahuje 89% DMEM (Dulbecco's Modified Eagle Medium), 10% FBS (fetal bovine serum) a 1% roztoku L-glutamin-penicillin-streptomycin.

3.3. Měření klíčících semínek

Pro měření klíčících semínek byly použity klíčky fazole mungo. Semena byly 3 dny klíčeny ve tmě ve skleněných Petriho miskách s purifikovanou vodou.

4. Výsledky

U každého experimentu jsou patrné vyšší intenzity záření ze suspenzí s biologickým materiálem než ze samotného pozadí. Na obr. 2 vlevo jsou vidět statistické charakteristiky jednotlivých signálů. Na obr. 2 vpravo jsou vidět příklady naměřených signálů z biologických vzorků s médiem, samotného média a šumu měřicí soustavy.



Obr. 2: Statistická charakteristika signálů

Detekované fotony tvoří pouze malou část všech generovaných fotonů. Na počet detekovaných fotonů mají vliv fyzikální vlastnosti fotonásobiče, jako je jeho kvantová účinnost, tvarovost a jeho spektrální citlivost. Dalšími parametry ovlivňující počet detekovaných fotonů jsou účinnost a směrovost dopadu fotonů na fotokatodu, odrazivost na rozhraní dvou prostředí a útlum v jednotlivých prostředích. Obr. 2 vlevo vznikl odstraněním vysokonapěťového rušení a použitím filtru klouzavých průměrů se šířkou okna

50. Obr. 2 vpravo vznikl vyfiltrováním vysokonapětového rušení a použitím Savitzky-Golayeho filtru stupně 2 se šířkou okna 50.

Z obr. 2 je patrný rozdíl intenzit záření mezi biologickým vzorkem s médiem a samotného média. Tento rozdíl intenzit je charakterizován poměrem signálu ku pozadí (SNR), tedy mezi signálem z buněk a média ku samotnému médiu (Tab. 1). Toto pozadí obsahuje šum z měřicí soustavy a signál z média. Nejlepších výsledků bylo dosaženo u munga 10.4 dB. U synchronizovaných kvasinek výsledný poměr dosahuje 6.8 dB. Nejnižší intenzity dosáhly tkáňové buňky typu HMVECad, 3.9 dB.

TYP BUNEK/POZADÍ	SNR [dB]
HMVECad/Medium DMEM+FBS+antibiotika	3.8915
Synchronizované kvasinky/Medium YPD	6.8330
Mungo/Voda, ve které klíčilo	10.3878

Tab. 1: Poměr výkonu signálu k měřenému pozadí [dB]

5. Diskuze

Tento základní výzkum, měl prokázat možnost měřit ultraslabou emisi fotonů z biologických vzorků použitou měřicí aparaturou. K tomuto účelu byly zvoleny tři různé typy biologických vzorků, které budou biologickým základem budoucích experimentů. Všechny tři experimenty potvrdily teorii o generování fotonů živými organizmy. Velmi důležitou součástí výzkumu bylo prozkoumat měřené prostředí a co nejlépe tato měření optimalizovat.

Tkáňové buňky HMVECad se nejdříve kultivovaly a měřily na sterilních polystyrenových Petriho miskách. Pro tato měření se ve výsledných signálech objevovaly artefakty. Byla provedena série pokusů, která vedla ke zjištění, že tyto artefakty pocházejí z použitých kultivačních misek. Opodstatnění tohoto jevu bylo nalezeno ve článku [5], ze kterého vyplývá, že polystyren vykazuje vlastnosti scintilátoru - je schopen absorbovat energii vysokoenergetických částic a konvertovat ji na fotony. Z toho vyplývá, že polystyrenové misky nejsou vhodné pro měření fotonických signálů.

Při měření kvasinek se nejdříve používala jako médium sacharóza. Sacharóza je disacharid, který kvasinky neumí metabolizovat, výsledné signály synchronních kvasinek byly tedy velmi nízké. Z toho důvodu bylo pro další měření používáno výhradně YPD médium, které obsahuje všechny potřebné látky pro růst buněk, čímž je zachována metabolická funkce buněk a ty se mohou dělit v průběhu měření.

U měření klíčků munga byla zjištěna závislost intenzity naměřeného signálu na natočení klíčících výhonků. Největší intenzity bylo dosaženo při natočení všech klíčků směrem k fotonásobiči.

Kultivace a příprava všech vzorků probíhala ve tmě v zatemněné místnosti. Dodržením toho postupu byla snížena pravděpodobnost dosvitu u jednotlivých biologických vzorků.

6. Závěr

V současné době se připravuje změna konfigurace systému a základních experimentů, které by měly sloužit k ověření metabolických procesů biologických vzorků. U konfigurace systému se změní poloha fotonásobiče, který bude ve spodní stěně měřícího boxu, čímž se dosáhne nejmenší možné vzdálenosti vzorku od fotonásobiče a zároveň se eliminuje vliv sedimentace buněk na dno laboratorního skla.

V prvních experimentech se plánují dlouhá měření ultraslabé emise fotonů z biologických vzorků, kde by se měly projevit jednotlivé fáze růstu buněk. Největší intenzita by měla být získána na začátku experimentu, jelikož buněk bude málo a budou mít plné médium. S narůstajícím časem bude buněk přibývat, a pokud se nebude dodávat nové médium, nebudou mít buňky co metabolizovat, tudíž se nárůst jejich počtu bude snižovat a tím poklesne i intenzita jejich emise.

Druhá část výzkumu se bude zabývat oxidativním a hyperosmotickým stresem buněk. Jako oxidativní stresový faktor bude použit peroxid vodíku, vyvolávající nárůst reaktivních kyslíkových forem (ROS), které mohou vést až k usmrcení buněk. Odpověď buněk na oxidativní stres je spojena se zvýšenou produkcí fotonů. U hyperosmotického stresu dochází ke snižování obsahu vody uvnitř buněk, čímž dochází k jejich poškození, které může vést až k zániku buňky.

Poděkování

Na tomto místě bych chtěla poděkovat svým školitelům Ing. Michalovi Cifrovi, Ph.D., a prof. Ing. Pavlovi Sovkovi, CSc. za vedení a pomoc při řešení dílčích projektů, které budou naplní mé disertační práce. Tato práce vznikla za podpory projektů GAČR 13-29294S a SGS13/138/OHK3/2T/13.

Reference

- [1] PRASAD, A. and POSPÍŠIL, P., Photon source within the cell, In *Fields of the Cell*, Eds. D. Fels and M Cifra, Research Signpost, 2013
- [2] ROBINSON, John M. Reactive oxygen species in phagocytic leukocytes. *Histochemistry and cell biology*, 2008, 130.2: 281-297.
- [3] VALKO, Marian, et al. Free radicals and antioxidants in normal physiological functions and human disease. *International Journal of Biochemistry and Cell Biology*, 2007, 39.1: 44-84.
- [4] DEVARAJ, Balasigamani, et al. Biophotons: ultraweak light emission from living systems. *Current Opinion in Solid State and Materials Science*, 1997, 2.2: 188-193.
- [5] GEORGE, G. A.; HODGEMAN, D. K. C. Quantitative phosphorescence spectroscopy of polystyrene during photo-degradation and the significance of in-chain peroxides. *European Polymer Journal*, 1977, 13.1: 63-71.

Měření sluchových korových odezev u pacientů se sluchovou náhradou

Lukáš Bauer¹, Roman Čmejla¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

bauerlu3@fel.cvut.cz

Abstrakt: Měření ztráty sluchu v raném období vývoje dítěte je důležité pro jeho další rozvoj. Podle výsledků je dítě zvoleno za kandidáta sluchové náhrady, popřípadě adepta kochleárního implantátu. Při včasném odhalení sluchové poruchy a vybrání vhodné sluchové náhrady dochází k řádnému rozvoji sluchových center v mozku, tedy normálnímu začlenění dítěte do společnosti. Většina dětských pacientů však není schopna dát hodnotnou odezvu na sluchový podnět, z toho důvodu jsou k měření sluchu používané sluchové evokované potenciály, které nám informaci o sluchových centrech zprostředkují i bez odpovědi pacienta. Existující metody však využívají měření pouze z generovaných signálů jako je pípnutí nebo kliknutí. Metoda měření sluchových evokovaných potenciálů z kortexu (CAEP) se snaží snímat potenciály z nejvyšších korových center. V tomto případě se může jednat o detekci komplexnějšího signálu jako je například řeč. Jedním z problémů této metody u pacientů s kochleárním implantátem je však artefakt vznikající od implantátu, který zaruší mozkovou odezvu. Odstranění artefaktu je pak zásadním krokem pro vytvoření objektivní automatické metody hodnocení CAEP

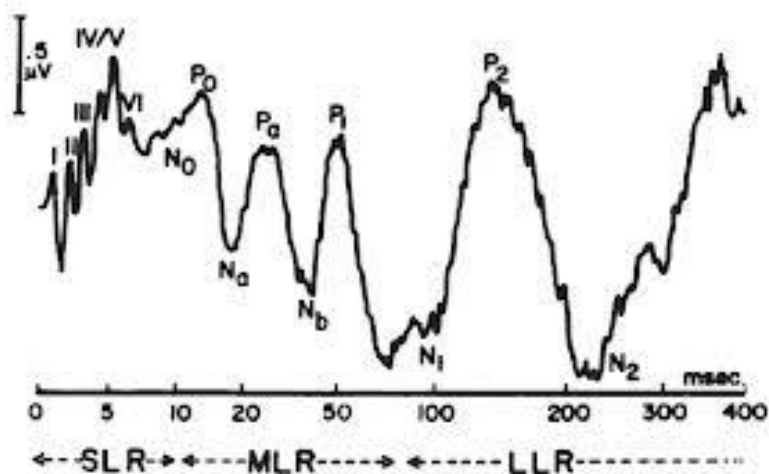
1. Úvod

U dětí narozených se ztrátou sluchu nedochází k rozvoji sluchových center. Pokud nedojde k včasnému umístění sluchové náhrady, popřípadě kochleární implantace, dojde k uzavření sluchových center a ta již nebudou nikdy využita k rozpoznání sluchu [1].

Na dětských pacientech se standardní měření ověřující ztrátu sluchu nedají provádět, protože dětští pacienti nejsou schopni spolupracovat. Z toho důvodu se využívá měření sluchových evokovaných potenciálů. Sluchové evokované potenciály se dělí podle doby odezvy, viz Obr. 1 na časné (I - VI), střednědobé (No - Nb) a pozdní (P1 - N2). V dnešní době se ve světě nejčastěji používá metoda měření Auditory Brainstem Response (ABR). Tato metoda využívá měření sluchových evokovaných potenciálů (EP) z mozkového

kmene, které jsou na začátku sluchové dráhy, tedy jedná se o časné EP (do 10 ms). Pro generování časných EP se používají tónové kliknutí. Při metodě CAEP se měří pozdní evokované potenciály (od 50 do 500 ms) z nejvyšších korových center. Díky těmto měřením jsme schopni rozeznávat odpověď na komplexní signál, mezi který patří například řeč. Předností techniky CAEP oproti ABR je amplituda měřených signálů, která je u CAEP vyšší. Další předností metody CAEP je možnost rozpoznání reakce mozku na slova. Při této metodě dochází také ale k problémům s detekcí promluv. Sluchová centra a s nimi i odezva na daný podnět se vyvíjí až do dospělosti [2]. Maximum P1 se po narození vyskytuje okolo 250 ms a v dospělosti se přesouvá až k 50 ms po zaznamenání stimulu. Maximum N1 se vyskytuje po narození okolo 500 ms, v dospělosti kolem 150 ms. Ostatní maxima se začínají vyskytovat až kolem 7 roku života.

Tato metoda naráží na další problém, kterým je zaznamenávání evokovaných potenciálů u pacientů s kochleárním implantátem. Ten způsobuje vznik artefaktu v zaznamenaném signálu. Signál je také samozřejmě zarušen svalovými artefakty a artefakty pohybu očí.



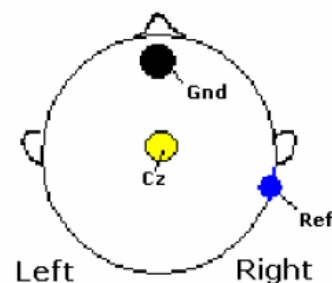
Obrázek 1: Sluchové evokované potenciály

2. Metodika měření

Signály jsou získávány pomocí přístroje HearLab od firmy Frye Electronics Inc. Přístroj využívá ke sběru dat 3 elektrod umístěných na pozicích Cz, Fpz, a mastoid za uchem (viz obr. 2).

Záznamy jsou měřeny pro písmena /m/, /t/ a /g/, kde délka trvání jednotlivých písmen je 30, 21 a 30 ms a interval mezi dvěma stimulacemi je zvolen jako 1125 ms. Písmena jsou vybrána tak, aby se vyskytoval v celém frekvenčním spektru

Databáze obsahuje jedinečné záznamy 16 dospělých pacientů, kde pro každého pacienta byly zaznamenány 2 nahrávky pro písmeno



Obrázek 2: Rozmístění elektrod

/m/, 2 nahrávky pro písmeno /g/ a 2 nahrávky pro písmeno /t/. nahrávky jsou vzorkovány 16 kHz a obsahují 200 ms před stimulací a 600 ms po stimulaci. Signál je dále filtrován dolní propustí na 30 Hz a je analyzován pomocí Hotellingova T^2 testu.

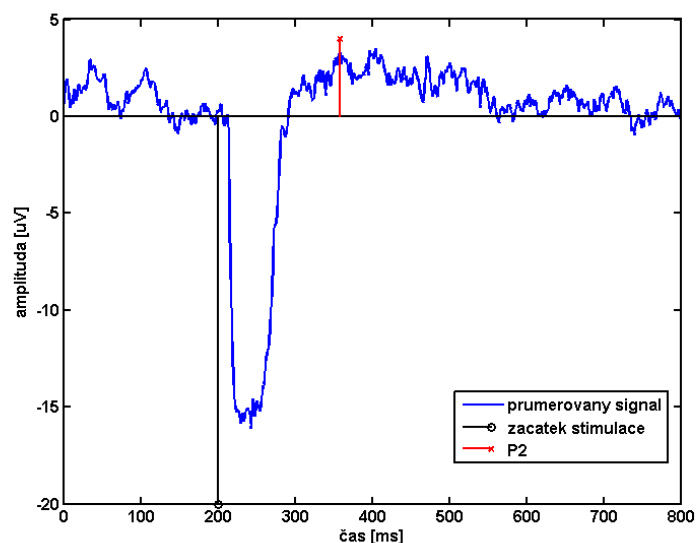
Kochleární implantát je přístroj sloužící k navrácení sluchu u pacientů, u kterých je porucha převodního systému mezi vnějším a vnitřním uchem, ale sluchový nerv a zbytek cesty je v pořádku. Do kochley je zaveden svazek elektrod, který dostává příkazy z vnějšího řečového procesoru. Při stimulaci sluchového nervu pak dochází k šíření proudu tkání, což působí rušení při měření CAEP.

3. Potlačení artefaktu

Největším problémem metody CAEP s aplikací na pacientech s kochleárním implantátem je výskyt artefaktu vznikající kochleárním implantátem viz Obr 3 a 4 sloužící pro porovnání záznamu s artefaktem a bez artefaktu. Ve světě dochází k dvěma přístupům potlačování artefaktu. První přístup využívá EEG záznam pacienta typu 10-20, ve kterém je využito 21 elektrod. Tyto systémy pak využívají robustní metody odstraňování artefaktu, jako je implementace ICA a PCA, popřípadě hledání isopotenciálních skalpových čar a umístění referenční elektrody na danou isopotenciálu [3]. Tyto přístupy mají velkou výhodu v možnosti odstraňování svalových artefaktů a pohybů očí, jejich velkou nevýhodou je ale velký počet elektrod, který je nutný umístit na pacienta. U implementace metody ICA pak není jednoznačné určení jednotlivých IC komponent a metodu je nutné kontrolovat. V druhém případě dochází k takové proměnlivosti isopotenciál, že během jednoho měření nelze nalézt správné umístění referenční elektrody, pomocí něhož by byl artefakt plně potlačen.

Dalším přístupem je detekce artefaktu za použití pouze 3 elektrod. V tomto případě se některé kliniky pokoušely měnit délky intervalu mezi stimulací dvou epoch, aby došlo k nárůstu odezvy a možnému odečtení daného artefaktu. Popřípadě

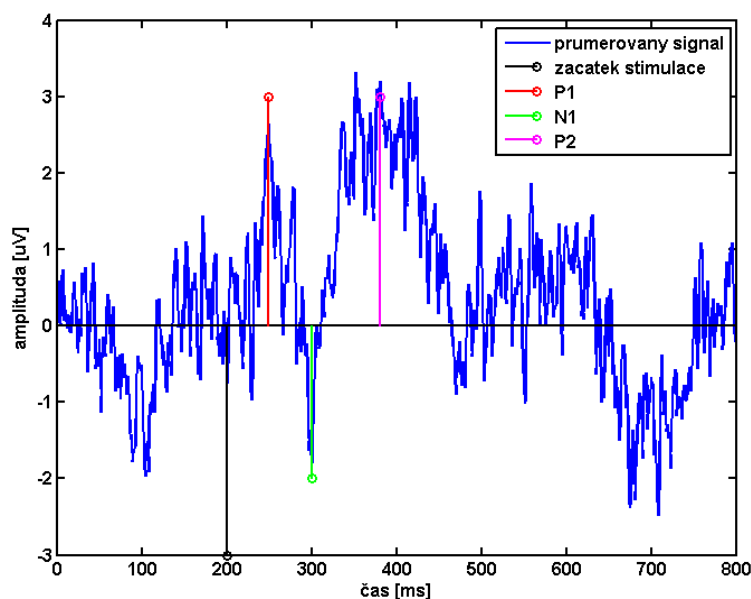
odečítání dvou záznamů po sobě [4]. Tyto metody jsou poměrně jednoduché a účinné, avšak dochází při nich k velkému nárůstu šumové složky.



Obrázek 3: Zobrazení výskytu artefaktu v záznamu CAEP

4. Závěr

Potlačení EEG artefaktu způsobeným kochleárním implantátem není jednoduchou úlohou. V dnešní době navržené metody nejsou stoprocentní a plně funkční, mají řadu chyb. Proto je potřeba navrhnout metodu, která by dokázala potlačit artefakt vznikající činností kochleárního implantátu. Díky takto funkčnímu algoritmu by pak mohli fungovat objektivní systémy detekující CAEP u dětských pacientů. Tento parametr by se potom mohl zařadit mezi jeden z faktorů předurčujících kochleární implantaci.



Obrázek 4: Zobrazení záznamu CAEP

Reference

- [1] Sharma, A. ; Dorman, M.F. Central Auditory Development in Children with Cochlear Implants: Clinical Implications. *Advances in otho-rhino-laryngology*, 64: 66-88, 2006.
- [2] Martin, B. ; Tremblay, K; Stapels, D. Principles and applications of cortical auditory evoked potentials. In R. Burkard, J. Eggermont, M. Don (Eds.), *Auditory evoked potentials: Basic principles and clinical application*, Philadelphia: Lippincott Williams and Wilkins, 482-507, 2007.
- [3] Gilley, P. M. ; Sharma, A. ; Dorman, M. ; Charles, C. F. ; Arunachalam, S. P. ; Martin, K. Minimization of cochlear implant stimulus artifact in cortical auditory evoked potentials, *Clinical Neurophysiology*, 1772-1782, 2006.
- [4] Friesen, L. M. ; Picton, W. P. A method for removing cochlear implant artefakt, *Hearing research*, 259: 95-106, 2010.

Využití povrchových elektromyografických signálů při analýzách svalové koordinace

Daniel ŠPULÁK¹, Roman ČMEJLA¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

spuladan@fel.cvut.cz

Abstrakt: V rámci výzkumu se zabýváme analýzou povrchových elektromyografických signálů. Přitom spolupracujeme s Fakultou tělesné výchovy a sportu (FTVS) Univerzity Karlovy v Praze, Katedrou sportů v přírodě.

Předmětem výzkumu kolegů z FTVS jsou rozborů pohybu během různých sportovních aktivit. Jedná se například o karate (úder naprázdno, do pytle, s expandérem či bez něj), lyžování, jízdu na rotopedu (s různými druhy treter), běh, chůze bez holí či s holemi (v různě skloněném terénu, s botami a bez bot), lezení na lezeckém тренаžeru, veslování, florbalové odpaly či výpady v šermu.

Ve všech případech se nahrává vícekanálový elektromyografický signál (EMG, signál elektrické aktivity svalů), a to výhradně pomocí povrchových (nalepovacích) elektrod. Ty jsou umístěny na pokožce nad sledovanými svaly a propojeny s přenosným záznamovým zařízením, které podle potřeby nahrává až 16 kanálů EMG.

Cílem je zejména sestavení grafů, z nichž je patrné pořadí svalových aktivit a deaktivací. V nich se pak hledají shody a odlišnosti u různých subjektů nebo různých pohybů.

Naše aktivity se zaměřují na tři problémy: segmentaci signálu z periodického pohybu, detekci nástupu a konce svalové aktivity a způsob prezentace výsledků s přihlédnutím k odlišnostem v provedení jednotlivých opakování pohybu.

K detekci nástupu a konce svalové aktivity používáme obálku EMG (tedy usměrněné EMG filtrované dolní propustí). Dobrých výsledků lze dosáhnout detekováním minima a maxima každé periody, přičemž za počátek aktivity se považuje takový bod na obálce EMG, který společně s minimem a maximem utvoří trojúhelník o maximálním obsahu. Předmětem výzkumu je srovnávání s jinak definovanými body trojúhelníka, případně zcela jinak řešené metody detekce nástupu. Sem patří zejména prahování stanovenou procentuální hodnotou maxima obálky. Při zpracování periodických pohybů se přikláníme k detekci uvedených bodů v každé periodě zvlášť, přestože obvyklým postupem je průměrování obálek z jednotlivých period a detekce prováděná na průměrovaném průběhu.

Během uplynulého roku jsme ke stále rozvíjenému algoritmu ve výpočetním prostředí MATLAB vytvořili i grafické uživatelské rozhraní. V grafickém rozhraní je možné zobrazit načítané signály, vybrat kanály obsahující EMG a kanály se signálem pro segmentaci a nastavit různé parametry pro segmentaci, výpočet obálek, detekci charakteristických bodů apod. Uživatelské rozhraní umožňuje uložit tyto parametry jako předvolby pro různé sportovní aktivity.

Pro lepší využití námi dodávaných výsledků (grafů a tabulek) jsme sepsali sedmnáctistránkový dokument s názvem „Jednoduché pokyny k interpretaci výsledků zpracování EMG signálů,“ který srozumitelnou formou vysvětluje, jakým způsobem lze z našich výstupů vyčíst potřebné informace.

Využití povrchových elektromyografických signálů při analýzách svalové koordinace

Ing. Daniel Špulák
vedoucí: doc. Ing. Roman Čmejla, CSc.
K13131
5/2013



Využití elektromyografie

- v medicíně
 - diagnostika neuromuskulárních chorob a poruch motorického systému
 - řízení umělých končetin
- ve sportu
 - analýza svalové únavy
 - tréninkové programy založené na detekci prahu únavy
 - detailní analýza pohybů
- v komerční sféře
 - rozhraní člověk/stroj u počítačových her

Elektromyografie

- snímání a vyhodnocení elektrické aktivity svalů
- elektromyografický/myoelektrický (EMG) signál snímáný pomocí povrchových nebo jehlových elektrod
- analýza v časové a frekvenční oblasti

Experimenty – sportovní aktivity

- karate
 - úderý naprázdno, do pytle, s expandérem či bez něj
- lyžování
- jízda na rotopedu
 - různé druhy treter
- chůze
 - bez holí či s holemi (Nordic walking)
 - různý sklon terénu
- běh
 - s botami či bez bot
- lezení na trenažéru
- florbalové odpaly
- výpady v šermu

Experimenty

- spolupráce s katedrou sportů v přírodě (Univerzita Karlova v Praze, fakulta tělesné výchovy a sportu)
- výhradně povrchové EMG
- studium činnosti jednotlivých svalů během pohybu při sportu

Experimenty – signály

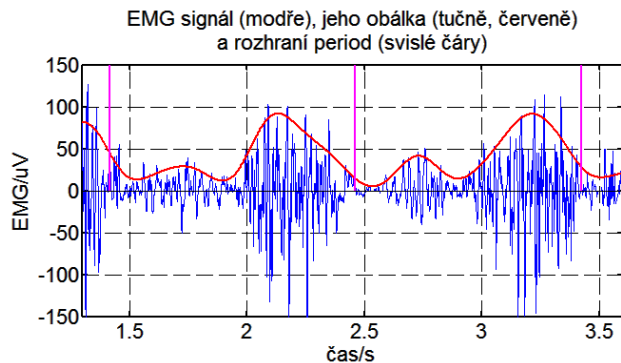
- přenosný nahrávací aparát ME 6000
- až šestnáctikanálové EMG (sledování celých svalových skupin)
- možnost připojení akcelerometrického či polohového snímače



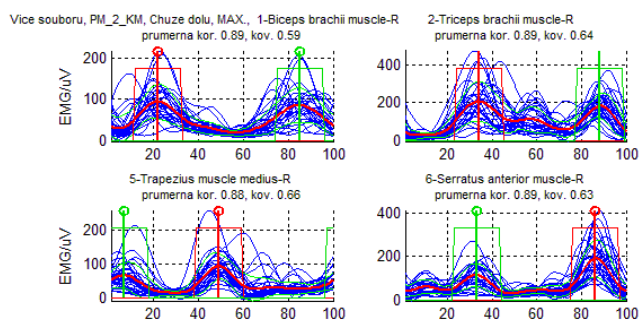
Experimenty – cíle

- sestavení grafů znázorňujících pořadí svalových aktivací a deaktivací během pohybu
- porovnání pořadí a časových rozestupů svalových aktivací a deaktivací
 - u různých osob při provádění téže aktivity
 - u stejné osoby při různých aktivitách
- zhodnocení podobnosti průběhu svalové aktivity v jednotlivých pohybových cyklech

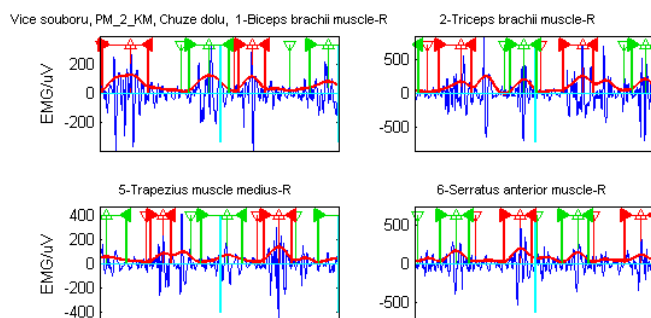
Zpracování EMG – obálka



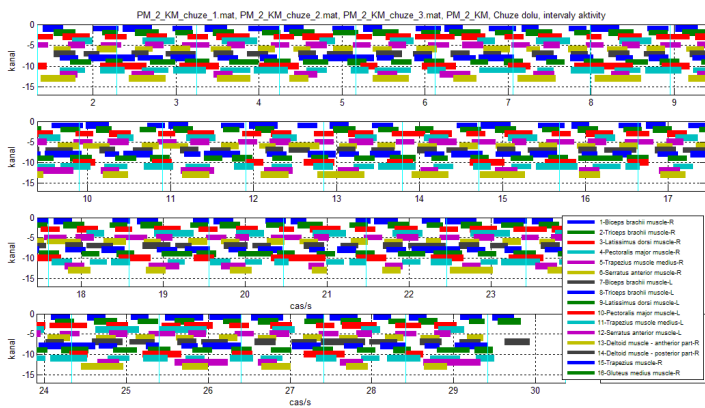
Zpracování EMG – průběhy z jednotlivých cyklů



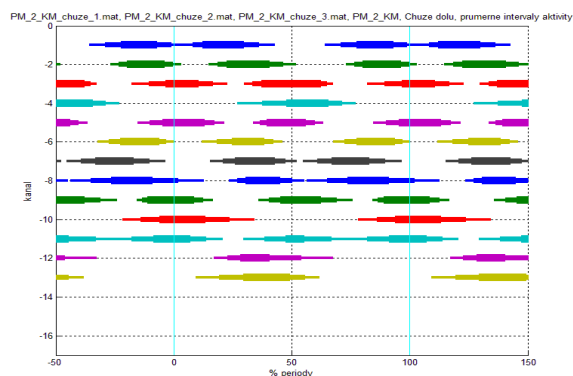
Zpracování EMG – detekované body



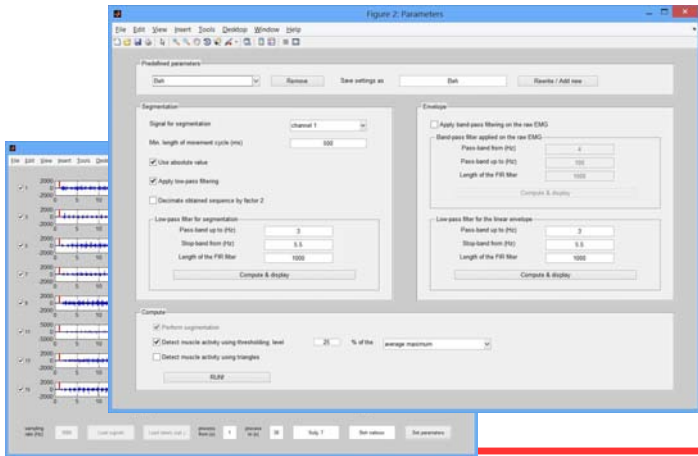
Zpracování EMG – intervaly svalové aktivity



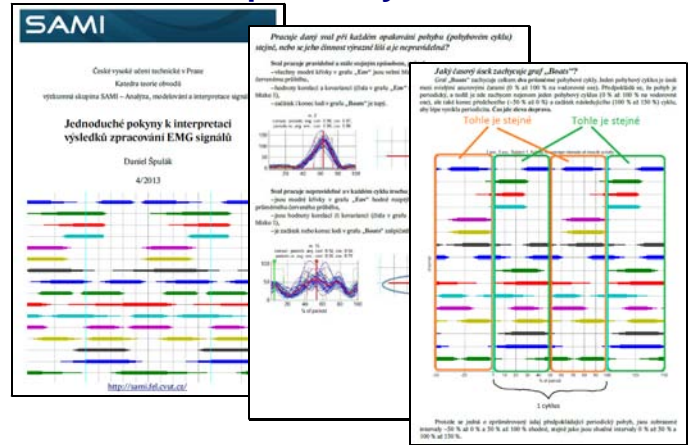
Zpracování EMG – průměrné intervaly aktivity (lodě)



Grafické uživatelské rozhraní



Návod k interpretaci výsledků



14

Shrnutí

- rozšíření množiny zpracovávaných pohybů
- vylepšení algoritmů detekce svalové aktivity
- vytvoření grafického uživatelského rozhraní
- sepsání pokynů k interpretaci výsledků

Děkuji za pozornost!

15

Využití videozáznamu při zpracování povrchového EMG

Jan SEDLÁK¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

sedlaj15@fel.cvut.cz

Abstrakt: Tento článek popisuje zpracování a využití videozáznamu sloužícího jako doplňková informace při signálovém zpracování povrchového elektromyografického (EMG) záznamu. Analyzované videozáznamy zachycují osobu během výkonu vybrané sportovní aktivity. Videozáznamy jsou synchronizovány s EMG záznamy svalové aktivity vybrané skupiny svalů. Tato práce se zaměřuje na segmentaci EMG záznamu na základě informace o pohybových cyklech zjištěné ve videozáznamu. Segmentace EMG signálu je žádoucí při analýze vzorů svalových aktivací. Výsledky dosažené pomocí navržené metody ukazují možnost využití videozáznamu při zpracování EMG signálů. Výsledky dosažené na základě postupu popsaného v tomto článku doplňují segmentační metody založené na zpracování obálky EMG záznamu. Výhodou segmentace na základě videozáznamu je nezávislost výsledků segmentace na rušení obsažené v EMG signálu. Další předností videozáznamu je snadná identifikace pohybových cyklů nevhodných pro vyhodnocení vzorů svalových aktivací. Nevýhodou využití videozáznamu je vyšší výpočetní náročnost než při segmentaci založené pouze na EMG signálu.

1. Úvod

Tato práce je součástí výzkumu zabývajícího se analýzou svalové koordinace během sportovních aktivit. Výsledky výzkumu jsou přínosné například v oblastech zkvalitňování tréninkových metod a rehabilitace. Nejčastějším úkolem analýzy EMG signálu je detekce svalových aktivací. Při analýze sportovní aktivity, která je složena z opakujících se vzorů pohybu, je možné navíc určení průměrné svalové aktivace během pohybového cyklu. Určení průměrné svalové aktivace vyžaduje segmentaci EMG signálů. Nejčastěji používaný postup segmentace je založen na vyhodnocení EMG záznamu svalu s vhodným vzorem aktivací. Technika segmentace založená na zpracování EMG signálu je blíže popsána v [1]. Cílem této práce je

otestovat použitelnost segmentace EMG záznamů na základě videozáznamu pořízeného během měření.

Problematika zpracování obrazových dat je velice rozsáhlá oblast vědy. Tento článek není primárně zaměřen na problematiku zpracování obrazových dat. Segmentace signálů EMG vyžaduje extrahování informace o pohybu aktéra zachyceného ve videozáznamu pořízeného jednou videokamerou. Informace o pohybu aktéra ve videozáznamu je reprezentována značkou. Jako značka slouží nejčastěji vhodně umístěný pásek antireflexní látky na těle aktéra. K analýze videozáznamu za účelem nalezení trajektorie značek byl otestován komerční software [2]. Tento software ale nesplňoval pro tyto účely potřebné požadavky, proto byl k analýze videozáznamů navržen jednoduchý algoritmus v programovém prostředí MATLAB.

2. Signálové zpracování dat

2.1. Analyzované data

Tato kapitola je věnována popisu záznamového zařízení a databáze záznamů použité pro testování navrženého algoritmu. Databáze je tvořena 10 záznamy pořízenými při běhu na lyžích a chůzi. Každý záznam databáze obsahuje EMG signály zkoumané skupiny svalů a videozáznam pořízený během měření. Délka záznamů je cca 30 sekund. Záznamy jsou pořízeny během sportovních aktivit v přirozených podmínkách.

Videozáznam byl pořízen kamerou SONY HDR-SR12 s rozlišením 576×720 px a snímkovací frekvencí 25 snímků za sekundu. Aktér ve videozáznamu má na těle umístěné barevně výrazné značky, které jsou následně využity při analýze videozáznamu. Stabilizaci obrazu aktéra ve středu videozáznamu zajišťuje obsluha kamery. V přirozených podmínkách není možné vždy zcela splnit požadavek na stabilitu a spojitě sledování obrazu aktéra. Nesplnění tohoto požadavku komplikuje analýzu videozáznamu.

Záznamy povrchového EMG byly zaznamenány pomocí zařízení KaZe05 se vzorkovací frekvencí 200 Hz a rozlišením 8 bitů. Zařízení zaznamenává přímo obálku EMG signálů vytvořenou usměrněním a filtrováním signálů dolní propustí s mezním kmitočtem 70Hz. Synchronizace s videozáznamem byla zajištěna pomocí filmové klapky spouštějící záznam EMG signálů. Měřená skupina svalů je: m. gluteus maximus dx., m. gluteus medius dx., m. vastus medialis dx., m. adductor longus dx., m. gastrocnemius caput medialis dx., m. tibialis anterior dx., m. gastrocnemius caput medialis sin. Správné umístění elektrod bylo zajištěno specialisty v oblasti fyziologie.

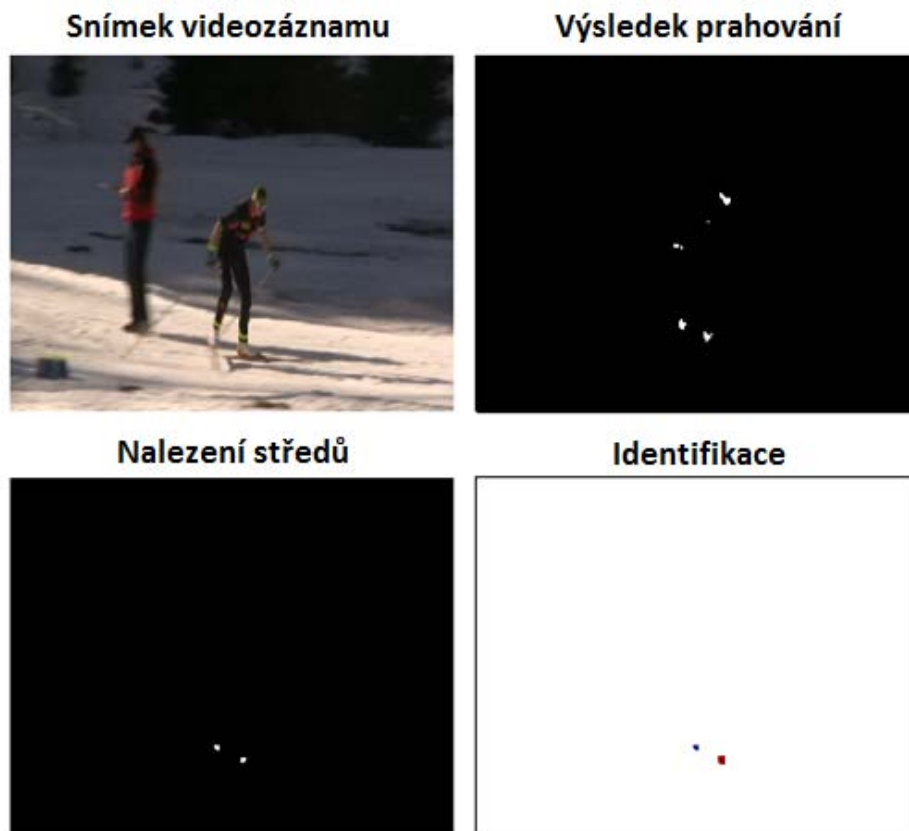
2.2. Zpracování videozáznamu

Ruční zpracování videozáznamu je vzhledem k objemu dat příliš zdoluhavé a neefektivní. Algoritmus popsáný v této kapitole byl navržen pro účely analýzy videozáznamu pořízeného během měření EMG záznamů. Cílem je zjistit trajektorie značek ve videozáznamu. K segmentaci EMG záznamu

popsaného v kapitole 2.3 je dostatečné sledování trajektorií dvou vhodně umístěných značek. Navržený algoritmus není plně automatický a v případě nejednoznačných výsledků je vyžadován zásah obsluhy.

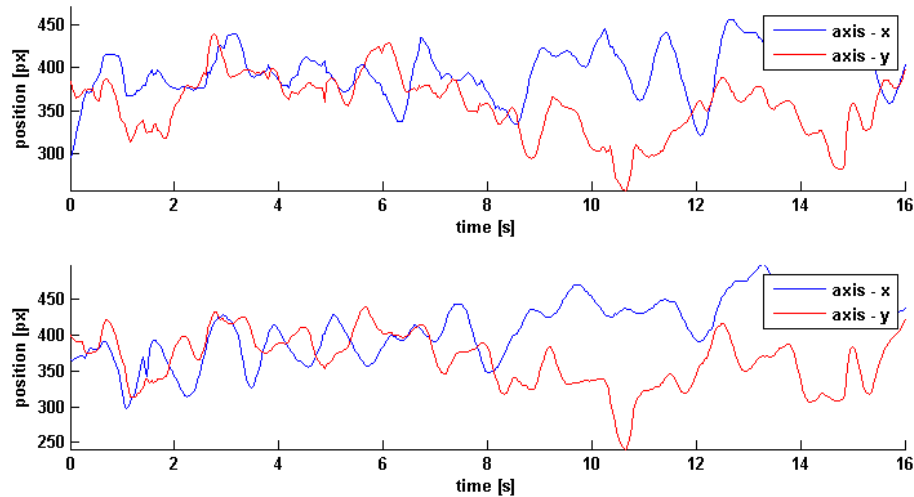
První částí zpracování je inicializace sledovaných značek v prvním snímku videozáznamu. Vyznačené obrazové body jsou vyhodnoceny a je vypočítán návrh parametrů filtru. Filtr je založen na principu prahování chromatické a jasové informace snímku. Obrazová data jsou při zpracování vyjádřena pomocí barevného modelu HSV popsaného v [3]. Výhodou reprezentace obrazových dat pomocí barevného modelu HSV je přirozenější kódování chromatické a jasové informace umožňující lepší nastavení prahů jednotlivých složek barevného modelu.

Navržený algoritmus postupuje v několika krocích. Ukázka průběhu zpracování snímku videozáznamu je uvedena na Obr. 1. Prvním krokem je detekce pixelů splňujících kritéria obrazového filtru. Výsledkem filtrace je snímek obsahující hodnotu 0 pro pozadí a hodnotu 1 pro pixely sledovaných značek. Dalším krokem je výřez kruhové oblasti kolem středu pozice značky nalezené v předešlém snímku. Výřez pouze omezené oblasti slouží k eliminaci nežádoucích značek a případného rušení ve snímku. K dalšímu zpracování je potřeba nalezení pozic středů sledovaných značek. K hledání středů shluků je využit algoritmus K-means. Správný výsledek tohoto postupu je podmíněn spojitým pohybem značek ve videozáznamu diskutovaným v kapitole 2.1. Pokud není automatickým zpracováním dosaženo jednoznačného výsledku, je vyžádán



Obr. 1: Ukázka průběhu algoritmu k detekci trajektorií značek.

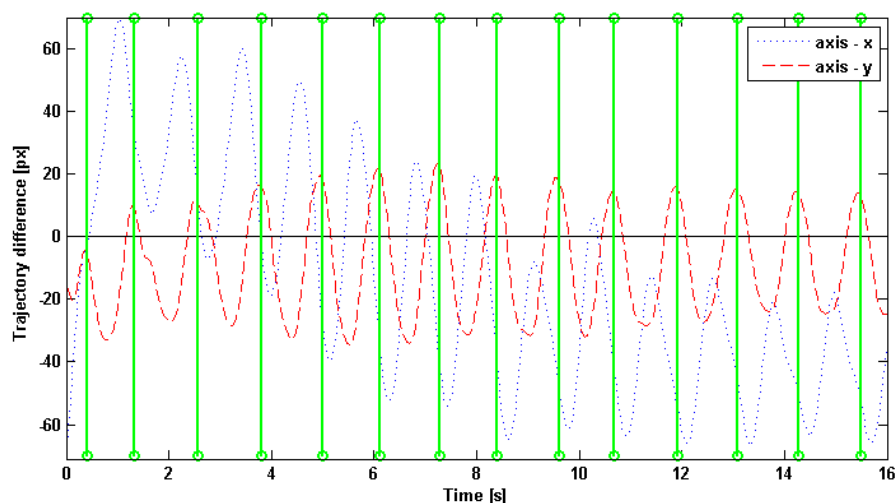
zásah obsluhy. Posledním krokem algoritmu k analýze videozáznamu je identifikace značek v po sobě jdoucích snímcích potřebné k určení trajektorii sledovaných značek. Identifikace značek je založena na hledání minimální vzdálenosti přiřazení značek v po sobě jdoucích snímcích. Při výskytu nejednoznačných situací jako je překrývání značek, je k vyznačení identity vyzvána obsluha. Ukázky trajektorii zjištěných ve videozáznamu jsou uvedeny na Obr. 2. Tento algoritmus je založen na základě technik signálového zpracování popsaných v [4] a [5].



Obr. 2: Ukázky detekovaných trajektorii značek.

2.3. Segmentace EMG záznamu

Trajektorie nalezené zpracováním videozáznamu, který je popsán v kapitole 2.2, jsou silně ovlivněny rušivými pohyby obsluhy videokamery. K eliminaci tohoto rušení je vypočten rozdíl trajektorii pohybu značek. Uvedená ukázka zpracování obsahuje trajektorie značek umístěných na levé a

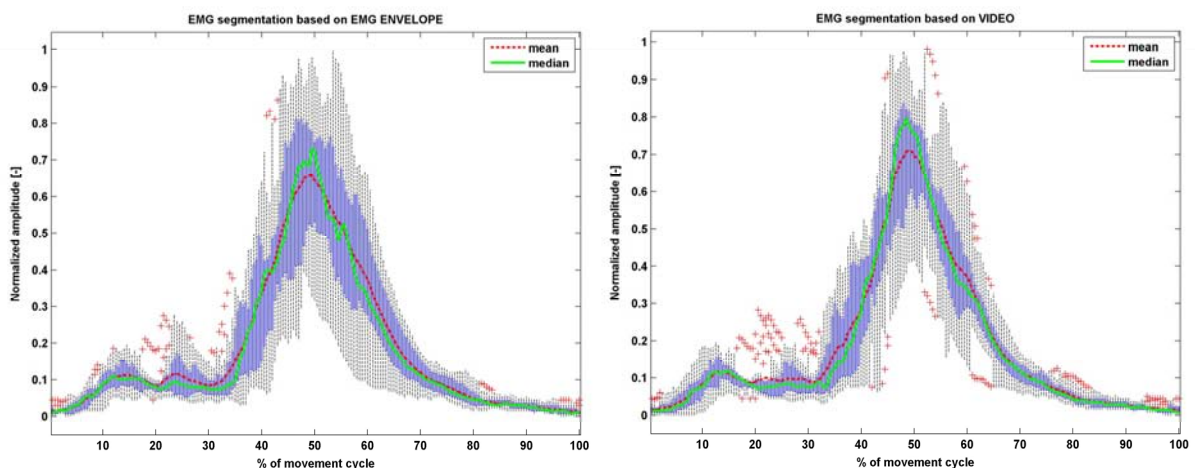


Obr. 3: Ukázka rozdílu trajektorií ve směru osy x a y, svislé čáry označují detekované hranice pohybových cyklů.

pravé noze během jízdy na běžkách. Signál vzniklý rozdílem trajektorii je lineárně interpolován tak, aby byl vyrovnán rozdíl vzorkovacích frekvencí videozáznamu a EMG signálu. Signál rozdílů trajektorií je dále filtrován klouzavým průměrem s délkou okna 200ms a poté je provedena detekce lokálních maxim k nalezení hranic mezi pohybovými cykly. Na základě dosažených výsledků bylo zvoleno k určení pohybových cyklů signálu rozdílu trajektorií ve směru osy y. Důvodem k nepoužití informace ve směru osy x je její ovlivnění proměnným pozorovacím úhlem při pořizování videozáznamu. Ukázka výsledků detekce je znázorněna na Obr. 3. Normované signály EMG jsou segmentovány na jednotlivé pohybové cykly na základě informace získané z videozáznamu.

3. Výsledky

Ukázka výsledku segmentace je uvedena na Obr. 4. Obrázek 4 obsahuje srovnání výsledků segmentace EMG pomocí dvou metod. První metoda je popsána v tomto článku, druhá metoda využívá k segmentaci informace získané přímo z EMG signálu, tato metoda je blíže popsána v [1]. Jednotlivé segmenty pohybových cyklů EMG signálu jsou v obou případech interpolovány na zvolenou jednotnou délku, aby bylo možné srovnání různě “rychlých” pohybových cyklů. Tento postup vede k určení průměrné obálky EMG během pohybového cyklu zkoumané sportovní aktivity. K prezentování výsledků je využito jednoduché statistické zpracování umožňující určení odchýlených hodnot na základě distribučního rozdělení dat. Výsledky segmentace EMG signálů ostatních svalů dosahovaly obdobných výsledků jako uvedená ukázka srovnání výsledků. Kvalitu segmentace lze posuzovat na základě rozptylu distribuce dat získané segmentací.



Obr. 4: Srovnání segmentačních technik vlevo na základě EMG a vpravo na základě videozáznamu. Průměrná obálka EMG je vyznačena červeně, medián zeleně. Svislé modré čáry vyznačují horní a dolní kvartil distribuce dat, svislé přerušované čáry označují hranice vychýlených hodnot, křížky označují vychýlené hodnoty.

4. Závěr

Na základě srovnání výsledků uvedených v Obr. 4 dosažených pomocí výše popsaných metod je možné konstatovat, že výsledky segmentace dosažené pomocí navržené metody jsou kvalitativně srovnatelné. Tato práce dokazuje použitelnost videozáznamu k získání doplňující informace vhodné ke zpracování EMG signálu, protože videozáznam je standardně pořizován při každém měření EMG signálu kvůli dokumentaci EMG záznamů. Využití videozáznamu odstraňuje některé nevýhody přítomny při segmentaci na základě EMG signálu. Nevýhodou segmentace založené na signálovém zpracování videozáznamu je vyšší výpočetní náročnost. Další předností využití videozáznamu při analýze svalové koordinace je snadná identifikace rušivých vlivů obsažených v záznamech. Využití videozáznamu je dále výhodné při prezentování výsledků zjištěné svalové koordinace u analyzovaných sportovních aktivit.

Dalším záměrem je ověření navržené segmentační techniky založené na videozáznamu i u jiných sportovních aktivit. Cílem další práce bude také srovnání navržené metody segmentace s jinými technikami založenými např. na zpracování signálu z akcelerometru umístěného na sportovci během výkonu sportovní aktivity. Po ověření vhodnosti zpracování videozáznamu k segmentaci EMG signálů bude zvažováno vylepšení algoritmu navrženého k detekci trajektorie značek.

Reference

- [1] ŠPULÁK, D., MIKULÍKOVÁ, P. Segmentation of Electromyographic Signal During Periodic Movements. In: *POSTER 2012 - 16th International Student Conference on Electrical Engineering*. Prague: Czech Technical University in Prague, 2012, p. 1-5. ISBN 978-80-01-05043-9.
- [2] DARTFISH SOFTWARE LTD. Dartfish 6.0 [software]. [cit. 2013-03-10]. In: <http://www.dartfish.com/>.
- [3] HSL and HSV. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-03-10]. In: http://en.wikipedia.org/wiki/HSL_and_HSV.
- [4] Šonka M., Hlaváč V., Boyle R.: *Image Processing, Analysis and Machine vision*, 3rd edition, Thomson Learning, Toronto, Canada, 2007.
- [5] Potůček, I.: Tracking movement objects in sequence pictures, In: *ElectronicsLetters.com*, Vol. 2003, No. 2, Brno, CZ, p. 1-15, ISSN 1213-161X.

Možná řešení problému přeučení algoritmu FastICA při zpracování EEG

Jan Šebek¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

sebekja4@fel.cvut.cz

Abstrakt: Cílem tohoto článku je upozornit na problém přeučení algoritmu FastICA při jeho použití na zpracování elektroencefalografického signálu (EEG) s relativně vysokým počtem kanálů vůči nedostatečnému počtu nezávislých vzorků ve zpracovávaném časovém úseku, shrnout dříve publikované metody řešení tohoto problému a zároveň popsat nově navržený přístup k řešení problému přeučení.

Byla navržena metoda, která efektivně potlačuje efekt přeučení i při zdánlivě nepříznivých poměrech počtu kanálů EEG signálu k počtu vzájemně nezávislých vzorků signálu ve zpracovávaném časovém segmentu. Jedná se o okénkovou metodu upravující činnost algoritmu FastICA. V závěru článku je tato metoda porovnána s metodami, které byly již dříve na toto téma publikovány, a zároveň je zde ukázáno, že poskytuje lepší výsledky než ostatní metody.

1. Úvod

V odvětví signálového zpracování dnes techniky slepé separace představují významný nástroj pro zpracování signálů získaných pomocí více senzorů. Jednou z oblastí, kde se použití slepé separace stalo běžné, je oblast zpracování elektroencefalografických (EEG) signálů, kde tyto techniky nacházejí uplatnění při odhadu zdrojových signálů ze vstupních dat. Tento odhad potom může posloužit při extrakci příznaků, či odstraňování rušivých složek (artefaktů) z užitečného EEG signálu. Mezi techniky slepé separace patří algoritmy analýzy nezávislých komponent (ICA), jejichž představitelem je známý a velmi často využívaný algoritmus FastICA.

Tento článek se zabývá problémem přeučení vznikajícím při aplikaci algoritmu FastICA v rámci algoritmu pro odstraňování svalových artefaktů (EMG signálů) z naměřených EEG signálů [7]. Algoritmus FastICA v rámci tohoto algoritmu plní nezastupitelnou funkci, protože umožňuje oddělit užitečný signál od artefaktů v prostorové oblasti.

Již od roku 1999 je známo, že pokud aplikujeme jakýkoliv ze známých algoritmů slepé separace, jako je např. algoritmus FastICA, na vícedimenzionální signál EEG, jehož časové trvání neposkytuje dostatečný počet vzorků pro zpracování v porovnání s dimenzí vstupního signálu, můžeme pozorovat jev přeučení algoritmu slepé separace [3], [4], [5]. Ve zmíněných publikacích je jev přeučení algoritmu ICA popisován jako generování rušivých složek, jimiž jsou zákmity (spikes) a lokální fluktuace střední hodnoty signálu (bumps) ve tvaru plochých kopců. Konkrétně byly tyto dva typy projevu přeučení označeny jako první typ přeučení v případě zákmitů a druhý typ přeučení v případě lokálních fluktuací střední hodnoty [3], [4], [5].

Tento jev přeučení algoritmu analýzy nezávislých komponent znemožňuje úspěšnou separaci zdrojových signálů, což ve výsledku znamená, že každá získaná nezávislá komponenta obsahuje jak užitečný EEG signál, tak rušivé složky svalových artefaktů.

Bylo již publikováno několik přístupů k řešení problému přeučení algoritmu analýzy nezávislých komponent.

V první řadě se důrazně doporučuje, aby byla v případě využití algoritmu FastICA použita robustní objektivní funkce, jako jsou různé aproximace negentropie, v roli míry odlišnosti konkrétního rozdělení od gausovského rozdělení signálu [1], [3]. Bylo ukázáno, že více robustní objektivní funkce založené na aproximacích negentropie jsou méně náchylné k přeučení než objektivní funkce založené na špičatosti [1], [3], [4].

Kromě použití vhodné objektivní funkce v případě využití algoritmu FastICA se nejběžněji můžeme setkat se zvyšováním počtu nezávislých vstupních vzorků každého kanálu zpracovávaného signálu [3], [4].

Další navržený přístup je dle [3], [4] vhodný k potlačení druhého typu přeučení s lokálními fluktuacemi střední hodnoty nezávislých komponent. Při tomto přístupu je signál rozdělen ve frekvenční doméně na dvě části pomocí filtrace horní propustí s mezní frekvencí 1 Hz nebo odečtením autoregresního modelu nízkofrekvenčních složek signálu. Algoritmus slepé separace je pak aplikován na takto modifikované kanály vstupního signálu.

Následující metoda je založena na snížení dimenze zpracovávaného signálu [1], [3], [4]. Snížení dimenze může být dle [1] dosaženo ve dvou krocích. Za prvé je nutné signál prostorově vybělit, což lze provést pomocí analýzy hlavních komponent (PCA). Za druhé jsou vyřazeny z dalšího zpracování ty hlavní komponenty, jejichž vlastní čísla, odpovídající jejich výkonu, jsou příliš malá v porovnání s největším získaným vlastním číslem.

Poslední z publikovaných řešení přeučení algoritmu ICA je technika pro potlačení artefaktů elektrookulogramu (EOG) [5]. Tato technika je také založena na snížení dimenze zpracovávaného signálu. Jedná se o kombinaci algoritmu FastICA a regresní metody. Aby nedošlo k přeučení algoritmu FastICA, jsou odhadovány pouze tři nezávislé komponenty z celkového množství dimenzí vstupního signálu. Snížení dimenze zpracovávaného signálu je zde tedy provedeno tím, že je odhadováno méně nezávislých komponent, než je dimenze vstupního signálu. Tři odhadnuté nezávislé komponenty představují odhad EOG artefaktů a slouží jako referenční kanály pro regresní metodu, která odhaduje, jakou měrou jsou tyto artefakty zastoupeny ve všech kanálech signálu. Jakmile je tato míra přítomnosti artefaktů ve všech kanálech odhadnuta pomocí váhovacích koeficientů, jsou odhadnuté artefakty těmito koeficienty násobeny a odečítány od jednotlivých příslušných kanálů EEG signálu.

I přesto, že se v publikacích o jednotlivých metodách tvrdí, že jev přeučení potlačují, každá z těchto metod má ve skutečnosti slabinu, a to ať už v nereálných požadavcích na signál, či ve velmi specializovaném použití.

První z výše uvedených přístupů, který doporučuje zvýšení počtu nezávislých vzorků signálu, je často nerealizovatelný z důvodu nemožnosti zvětšení časového intervalu, ve kterém jsou vzorky signálu shromažďovány. Platí, že je důležité zpracovávat signál v časovém úseku, kde je stacionární. Čím delší je tento časový interval, tím větší je pravděpodobnost, že je naměřený signál nestacionární, tj. že zdrojové signály se objevují a opět mizí v rámci vyšetřovaného časového úseku.

Druhé řešení problému přeučení, které je založeno na filtraci horní propustí a odečítání autoregresního modelu nízkofrekvenčních složek signálu, nezaručuje potlačení jevu přeučení. Během simulací, které byly provedeny autorem článku, k přeučení i přes filtraci došlo. Navíc v příslušné publikaci není uvedeno žádné teoretické odůvodnění ukazující na to, že to by tento přístup měl fungovat.

Třetí navržené řešení, které spočívá ve snížení dimenze zpracovávaného signálu pomocí procedury PCA, může být problematické. V [1] je uvedeno, že tato procedura může snížit

šum, který je většinou spojen s nižšími získanými hodnotami vlastních čísel, nicméně v případě zpracování EEG signálu získáme užitečné zdrojové signály, které mají v porovnání s odhadnutými EMG artefakty nižší výkon a jejich odstranění z dalšího zpracování by vedlo na ztrátu užitečné informace.

Poslední varianta řešení problému přeučení, kde se využívá kombinace algoritmu FastICA a regresní metody je sice slibná při odstraňování EOG artefaktů, nicméně její slabina spočívá ve velmi specializovaném použití. Pokud by byla tato metoda aplikována při odstraňování EMG artefaktů, pak by artefakty nemusely být vůbec potlačeny, protože počet zdrojových signálů odpovídajících EMG artefaktům může být daleko vyšší než tři odhadované nezávislé komponenty v této metodě.

V tomto článku je popsána další alternativa k řešení problému přeučení. Navržená metoda je založena na snížení dimenze zpracovávaného signálu způsobem, který je vhodný pro odhad zdrojových EMG signálů, u kterých neznáme jejich počet. První část článku se věnuje popisu algoritmu, který je založen na okénkové modifikaci algoritmu FastICA. V další části jsou ukázány výsledky potlačení EMG artefaktů a je provedeno jejich porovnání s ostatními publikovanými metodami.

2. Přístupy k řešení problému přeučení

Konkrétní podoba algoritmu na odstraňování svalových artefaktů bez zavedených opatření proti přeučení algoritmu FastICA sice nemá na podobu metod řešících přeučení vliv, nicméně poskytuje referenci pro porovnání kvality EEG signálu na výstupu tohoto algoritmu s kvalitou EEG signálu na výstupu algoritmu využívajícího metody pro potlačení jevu přeučení algoritmu FastICA. Tvoří tedy výchozí poznatky, na které je při vyhodnocování účinnosti navržených metod brán zřetel.

Při řešení problému přeučení tedy vycházíme z následujícího algoritmu [7], [8].

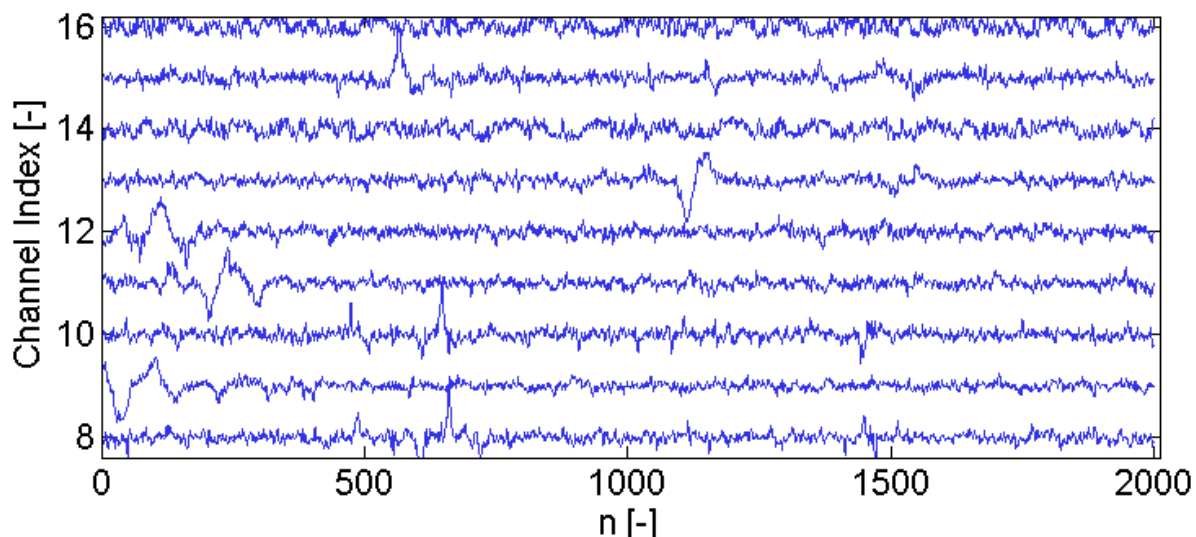
- Pomocí elektrodového systému s určitým počtem elektrod bylo naměřeno odpovídající množství EEG signálů. Množina těchto pro zde popisovaný algoritmus vstupních EEG signálů je v rámci tohoto článku vnímána jako jeden vícedimenzionální EEG signál s počtem dimenzí odpovídajících počtu vstupních EEG 1D signálů, nebo také počtu měřících elektrod u použitého měřícího systému. Tento signál tvoří matici je v rámci tohoto článku označen symbolem X , kde řádky představují jednotlivé EEG 1D signály z elektrod a sloupce jednotlivé časové okamžiky, ve kterých byly signály naměřeny.
- Vstupní signál EEG je podél časové osy segmentován takovým způsobem, aby daný časový úsek v případě výskytu EMG artefaktu tento artefakt obsahoval.
- V jednotlivých časových úsecích je dále řešen problém lineární slepé separace, který spočívá v následujícím. Předpokládá se, že signály (řádky) z matice X vznikly lineární kombinací zdrojových signálů, kterými jsou nezarušené 1D EEG signály a také 1D EMG signály z oblasti obličeje a krku, uložených v matici označené např. symbolem S , což lze zapsat pomocí vztahu

$$X = A \cdot S. \quad (1)$$

Pomocí algoritmu FastICA pak tedy odhadujeme matici A a také zdrojové signály S (nezávislé komponenty označované jako IC).

- Získané nezávislé komponenty v matici S_{est} jsou dále klasifikovány a ke zpětné rekonstrukci pro získání signálu X_c bez svalových artefaktů jsou propouštěny pouze komponenty v matici S_{eeg} klasifikované jako EEG signály.

Data, která byla v rámci metod uvedených v tomto článku zpracovávána, byla naměřena pomocí elektrodového systému, který obsahoval 111 elektrod, tzn. algoritmus FastICA odhadoval matici lineárních kombinací A o rozměrech 111×111 .



Obr. 1: Efekt jevu přeučení pozorovatelný v několika nezávislých komponentách

Při zpracování časového segmentu signálu obsahujícího 2000 vzorků pomocí výše zmíněného algoritmu byl jev přeučení pozorovatelný v odhadnutých nezávislých komponentách jako tzv. druhý typ přeučení [3], [4], [5], viz **Obr. 1**, kde jsou dobře vidět lokální fluktuace střední hodnoty v jednotlivých odhadnutých nezávislých komponentách.

2.1. Okénková úprava algoritmu FastICA

Pro tuto metodu zavádíme předpoklad, že vzájemná korelace dvou signálů měřených pomocí dvou různých elektrod na povrchu skalpu se se zvětšující se vzdáleností elektrod zmenšuje [7], [8].

Na základě tohoto předpokladu dosáhneme snížení počtu dimenzí zpracovávaného signálu pomocí provádění algoritmu FastICA na omezené množině kanálů (řádků matice \mathbf{X}) signálu EEG získaných z prostorově vzájemně blízkých elektrod.

Potřebujeme tedy vědět, jakým způsobem jsou elektrody na skalpu rozmístěny. Tuto informaci ovšem nezískáme z pořadí řádků v matici EEG signálu \mathbf{X} , protože tyto indexy řádků neodpovídají vzájemné prostorové vzdálenosti jednotlivých elektrod, kterými byly kanály EEG signálu naměřeny, ale pouze jejich označení. Pro každý použitý měřicí systém elektrod je tedy nutno definovat pozice jednotlivých elektrod v prostoru takovým způsobem, abychom měli u každé elektrody přehled o jejich sousedních elektrodách.

Na **Obr. 2** je znázorněna mapa elektrod, kterou navržený algoritmus v případě 111kanálového záznamu EEG signálu v podobě matice označené symbolem \mathbf{P} využívá. Tato mapa slouží pro označení užitečných kanálů EEG signálu nenulovým indexem, číslo elektrody v matici \mathbf{P} odpovídá pořadí řádku v matici \mathbf{X} signálu EEG.

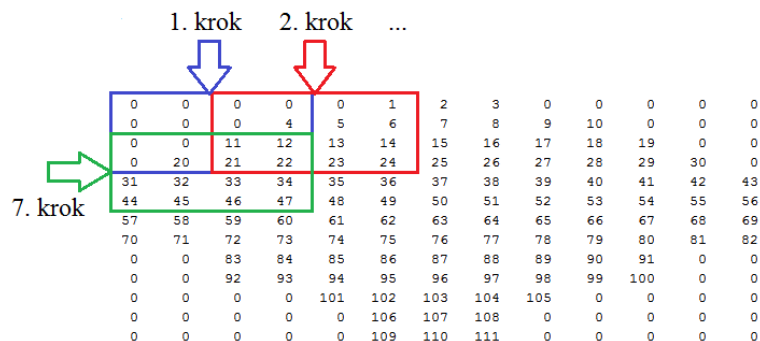
0	0	0	0	0	1	2	3	0	0	0	0	0
0	0	0	4	5	6	7	8	9	10	0	0	0
0	0	11	12	13	14	15	16	17	18	19	0	0
0	20	21	22	23	24	25	26	27	28	29	30	0
31	32	33	34	35	36	37	38	39	40	41	42	43
44	45	46	47	48	49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79	80	81	82
0	0	83	84	85	86	87	88	89	90	91	0	0
0	0	92	93	94	95	96	97	98	99	100	0	0
0	0	0	0	101	102	103	104	105	0	0	0	0
0	0	0	0	0	106	107	108	0	0	0	0	0
0	0	0	0	0	109	110	111	0	0	0	0	0

Obr. 2: Pozice elektrod zapsané do matice v Matlabu s doplněnými nulami

Aby bylo možno vyjádřit mapu elektrod, která má většinou hexagonální tvar, maticí, byly okrajové části mapy doplněny nulami tak, aby vznikla matice P znázorněná na **Obr. 2**. Horní řádek matice P odpovídá číslům elektrod umístěných na čele, spodní řádek pak odpovídá elektrodám umístěným na týlu hlavy.

Nyní, když známe prostorové uspořádání jednotlivých elektrod, definujeme vzájemně blízké elektrody pomocí výběrového čtvercového okna (viz **Obr. 3**). Velikost okna může být uživatelem změněna na jakoukoliv smysluplnou hodnotu (tedy celé číslo, menší než mapa). Okno z **Obr. 3** může maximálně označit 16 elektrod zároveň, přičemž se přes mapu elektrod pohybuje s překryvem 8 kanálů mezi sousedními oblastmi. Tento pohyb čtvercového okna po matici P je zobrazen na **Obr. 3**.

V každém kroku, kdy se okno přemístí do nové pozice (viz **Obr. 3**), je vyhodnoceno, jestli obsahuje dostatečný počet kanálů EEG signálu s nenulovým indexem, na kterých by mohla být provedena slepá separace. V současnosti je tento počet stanoven na 3 kanály EEG signálu, a to podle nejvyšší dosažené úspěšnosti separace komponent uvedené v kapitole Vyhodnocení algoritmu.



Obr. 3: Pohyb okna pro výběr kanálů k separaci po matici P

Jednotlivé vybrané kanály EEG signálu z dané oblasti jsou zpracovány algoritmem FastICA, který nyní odhaduje matici lineárních kombinací A_r s mnohem menšími rozměry (maximálně 16×16) oproti matici A (111×111). Během separace kanálů se pro jednotlivá čísla elektrod ukládají počty použití příslušných kanálů. Odhadnuté nezávislé komponenty S_{estr} jsou klasifikovány a komponenty odpovídající signálu EEG S_{egr} jsou zpětně rekonstruovány pomocí matice A_r jako signál X_{cr} .

Po průchodu čtvercového okna všemi oblastmi matice P tak kvůli vzájemnému překryvu oblastí, viz **obr. 3**, získáme pro i -tý kanál x_{ci} zrekonstruovaného signálu X_c sadu EEG kanálů s potlačeným artefaktem. Aritmetickým průměrem přes tyto realizace jednoho časového segmentu získáme zrekonstruovaný kanál x_{ci} signálu X_c .

3. Výsledky

Úspěšnost potlačení jevu přeučení algoritmu FastICA byla vyhodnocena pomocí následující metody. Z vlastních naměřených dat EEG signálu byla vybrána množina nezarušených úseků EEG signálu. K těmto úsekům byly přičteny stejně dlouhé úseky z množiny svalových artefaktů, a tím vznikla množina úseků EEG signálů se svalovými artefakty. Na každý úsek signálu z této množiny byl aplikován výše popsáný algoritmus pro odstraňování svalových artefaktů. Za prvé byl algoritmus použit bez zavedení metody potlačující jev přeučení, za druhé byla využita okénková metoda pro potlačení jevu přeučení. Pro porovnání výsledků byla také implementována metoda redukce dimenze s využitím metody PCA [1]. Mezi každou odpovídající dvojicí originálního kanálu EEG signálu a kanálu

EEG signálu s odstraněným svalovým artefaktem pomocí algoritmu pro odstranění svalových artefaktů s danou modifikací algoritmu FastICA je následně vyhodnocena podobnost pomocí korelačního koeficientu definovaného dle vztahu (2).

$$r_{xy} = \frac{R_{xy}}{\sqrt{R_x \cdot R_y}} \quad (2)$$

kde R_{xy} je vzájemný výkon originálního EEG signálu a EEG signálu zpracovaného pomocí algoritmu pro odstranění svalových artefaktů, R_x je výkon originálního EEG signálu a R_y je výkon EEG signálu zpracovaného pomocí algoritmu pro odstranění svalových artefaktů.

V **Tab. 1** jsou shrnuty dosažené průměrné korelační koeficienty a jejich rozptyl získané z množin signálů zpracovaných algoritmem pro odstranění svalových artefaktů s příslušnou modifikací algoritmu FastICA.

Tab. 1: Dosažené velikosti korelačních koeficientů u jednotlivých metod

Použitá úprava alg. FastICA	Střední hodnota r_{xy}	Rozptyl r_{xy}
Bez úpravy	0.6166	0.0027
PCA: Redukce dimenze: 75 kanálů	0.6777	0.0015
PCA: Redukce dimenze: 50 kanálů	0.7106	0.0013
PCA: Redukce dimenze: 25 kanálů	0.7482	0.0032
Okénková metoda, okno 4x4	0.8969	4.1808e-04

4. Diskuze

Z výsledků v **Tab. 1** je patrné, že okénková úprava algoritmu FastICA jev přeučení efektivně potlačuje a algoritmus pro odstranění svalových artefaktů tak může plnit svou funkci. V porovnání s ostatními metodami tato metoda dosahuje nejvyšších korelačních koeficientů, které se navíc od sebe v jednotlivých případech zpracování různých signálů příliš neliší.

U referenční metody snížení dimenze zpracovávaného signálu pomocí PCA [1] se sice korelační koeficient se snižujícím se počtem ponechaných dimenzí zvyšuje, nicméně ani při malém počtu ponechaných kanálů EEG signálu zdaleka nedosahuje úrovně koeficientu r_{xy} při využití okénkové metody. Při snižování počtu kanálů pro zpracování navíc roste i rozptyl dosažených korelačních koeficientů a při příliš nízkém počtu ponechaných kanálů se objevuje problém se samotným potlačením artefaktů.

5. Závěr

V článku byla představena navržená metoda pro potlačení jevu přeučení u algoritmu slepé separace FastICA a porovnána s ostatními metodami, které byly již dříve publikovány.

Článek nejdříve popisuje již dříve publikované metody řešící problém přeučení a zdůrazňuje jejich slabiny oproti navržené okénkové metodě úpravy algoritmu FastICA.

Hlavní část článku stručně popisuje algoritmus pro odstraňování svalových artefaktů, kde je použití algoritmu FastICA zásadní a jev přeučení činnost tohoto algoritmu znemožní. Dále je popsán princip okénkové metody úpravy algoritmu FastICA.

V následující kapitole jsou uvedeny dosažené výsledky rekonstrukce nezarušeného EEG signálu s potlačením svalových artefaktů s ohledem na kvalitu zrekonstruovaného signálu. Jsou zde porovnány případy, kdy nebyla použita žádná metoda pro zamezení jevu přeučení u algoritmu FastICA, kdy byla použita okénková metoda úpravy algoritmu FastICA a na závěr je zde pro porovnání uvedena metoda snížení dimenzionality vstupního signálu pomocí PCA procedury [1]. Dosažené výsledky ukazují, že navržená okénková metoda poskytuje vyšší kvalitu zpracovaného signálu v porovnání s ostatními metodami.

Reference

- [1] Hyvarinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*. Wiley, 2001, ISBN 0-471-22131-7.
- [2] Hyvarinen, A.; Oja, E. Independent Component Analysis: algorithms and applications. *Neural Networks*, 2000, vol. 13, no. 4-5, pp. 411-430.
- [3] Sarela, J.; Vigário, R. *Overlearning in Marginal Distribution-Based ICA: Analysis and Solutions*. *Journal of Machine Learning Research*, 2003, vol. 4, pp. 1447-1469.
- [4] Sarela, J.; Vigário, R. The Problem of Overlearning in High-Order ICA Approaches: Analysis and Solutions. *IWANN*, 2001, LNCS 2085, pp. 818 -825.
- [5] Hyvarinen, A.; Sarela, J.; Vigário, R. Spikes and Bumps: Artefacts generated by Independent Component Analysis with insufficient sample size. *Proc. Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99)*, pp. 425-429, Aussois, France, 1999.
- [6] Ng, S.C.; Raveendran, P. Removal of EOG Artifacts Using ICA Regression Method. *In 4th Kuala Lumpur International Conference on Biomedical Engineering (IFMBE)*, 2008, vol. 21/I, ISBN-13 978-3-540-69139-6.
- [7] Mammone, N.; Morabito, F.C.; Enhanced automatic artifact detection based on independent component analysis and Renyi's entropy. *Neural Networks*, 21 (2008), 1029-1040.
- [8] Delorme, A.; Sejnowski, T.; Makeig, S.; Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, 34 (2007), 1443-1449.

Identifikace subjektu v reálném čase pro BCI

Vladimír ČERNÝ

České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

cernyvl3@fel.cvut.cz

Abstrakt: Funkce rozhraní mozek-stroj (The Brain Computer Interface - BCI) velmi závisí na subjektu, proto je vhodné mít informaci o tom, jaký subjekt právě systém používá. Tato práce popisuje algoritmus pro identifikaci subjektu přímo z EEG signálu, který je už použitý pro vlastní BCI, tudíž algoritmus nepotřebuje další vstupní kanál. Algoritmus využívá frekvence v okolí μ rytmu jako klasifikační příznaky. Frekvence jsou odhadovány z FZ-AR modelu natrénovaného pomocí NLMS. Jako klasifikátor je využita Mahalanobisova vzdálenost. Algoritmus byl implementován tak, aby pracoval v reálném čase. Funkčnost algoritmu byla ověřena pomocí křížové validace na datech nahraných při podobných podmínkách jaké budou při plánovaném využití. Průměrné skóre zjištěné pomocí křížové validace bylo 87%, což je pro zamýšlené využití postačující. Algoritmus bude zabudován do již existujícího BCI systému.

1. Úvod

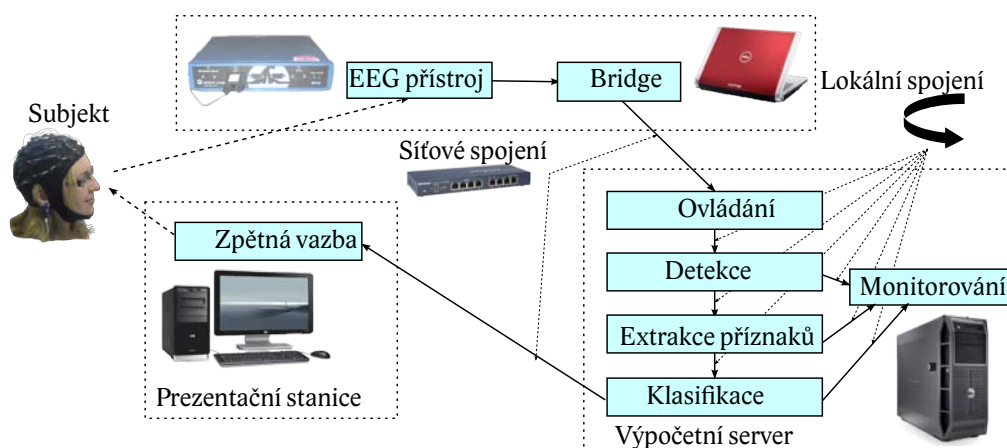
Funkce rozhraní mozek-stroj (The Brain Computer Interface - BCI) velmi závisí na subjektu, proto je vhodné upravovat chování BCI podle konkrétního subjektu. Individuálně mohou být vybírány předtrénované koeficienty klasifikátorů nebo dokonce celé klasifikační metody. I způsob jakým je prezentována zpětná vazba, je vnímán subjektivně a může být vybírán podle konkrétního subjektu.

BCI systém potřebuje mít informaci o tom, kdo ho právě používá. Informaci může snadno poskytnout operátor, tím že ji manuálně zadá. Tím se ale zavádí další krok a tedy i prostor pro chybu. Přitom sám systém může identifikovat subjekt. A to může být dokonce provedeno přímo z EEG [1]. Nahrávání EEG je už potřeba pro vlastní BCI, takže se nemusí zavádět další vstupní kanál.

1.1. Zpracování v reálném čase

Naše skupina se zabývá výzkumem v oblasti rozpoznávání EEG signálů souvisejících s pohybem s cílem vyvinout BCI. Proto jsme vyvinuli systém pro zpracování signálu v reálném čase [2], který nám umožňuje provádět experimenty se zpětnou vazbou.

Systém sestává z nezávislých modulů, které spolu komunikují přes počítačovou síť (Obr. 1). Nezávislé moduly umožňují snadné rozšíření systému o nové funkce (je jenom potřeba implementovat nové moduly). Díky propojení pomocí počítačové sítě není systém omezený pouze na jeden počítač. To je užitečné, pokud je potřeba provádět výpočty náročné na zdroje nebo pokud je potřeba samostatný počítač pro vlastní EEG nahrávání. Další výhodou distribuovaného systému je jednodušší ovládání a nastavení. Data mohou být nahrávána na počítači blízko EEG přístroje, monitorování a ovládání může být prováděno na dalším počítači, zatímco subjekt sleduje zpětnou vazbu na třetím (Obr. 1). Takový systém nám umožňuje snadno navrhovat různé experimenty, které vyžadují zpracování EEG v reálném čase.



Obrázek 1: Architektura systému.

Moduly jsou implementovány v jazyce Java, takže jsou nezávislé na použitém hardwaru a operačním systému. Komunikační protokol je založen na Real-time Transport Protocol (RTP). RTP byl navržen pro účely přenosu videa a audia a RTP pakety bývají prioritizovány na síťových prvcích. Přenos EEG pro BCI má velmi podobné požadavky jako audio a video. Nejdůležitějším parametrem je odezva. V [3] jsme odzkoušeli, že náš systém má dostatečně nízkou odezvu pro aplikace se zpětnou vazbou. Použitý komunikační protokol umožňuje kromě přenosu vlastních EEG signálů přenos dodatečných parametrů. Logovací systém je navržen tak aby ukládal, všechna přenášená data a změny stavů.

Naše první experimenty používaly jednoduchou klasifikaci, která neprováděla natrénování na subjekt. I při tomto jednoduchém přístupu jsme narazili na rozdíly mezi subjekty. Naším dalším logickým krokem bude použití pokročilejších klasifikačních algoritmů, které většinou musí být natrénovány. Protože EEG se pro jednoduché subjekty velmi liší, lepší výsledky by mělo dávat natrénování na konkrétní subjekt [4].

Aby se mohly vybrat správné koeficienty, klasifikační algoritmus musí mít informaci o subjektu během trénovací i testovací fáze.

Naším cílem je implementovat identifikaci subjektu do našeho BCI systému. Identifikační algoritmus by měl poskytnout informaci o aktuálním subjektu zbytku systému (Obr. 2). Systém by potom měl být schopen podle této informace vybrat správné parametry.

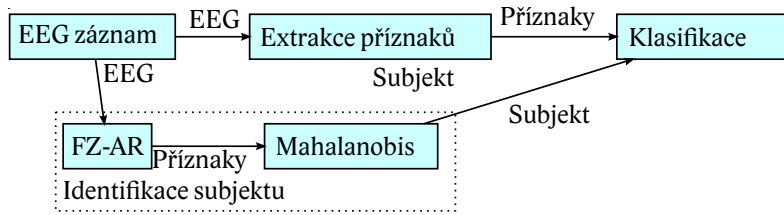
2. Metoda

2.1. Zpracování ze záznamu

Existuje několik různých příznaků, které je možné použít pro identifikaci. Náš BCI systém je založený na změnách v μ rytmu. Nahrávání EEG je nastaveno tak, aby se zachytil právě tento rytmus. Jedna z charakteristik, která je závislá na subjektu, je právě frekvence μ rytmu [5]. Proto je výhodné použít jako příznak pro klasifikaci frekvence v pásmu kolem μ rytmu. Metoda využívající právě tyto příznaky pro identifikaci subjektu je popsána v [6].

V [6] je μ rytmus modelován pomocí autoregresního modelu s frekvenční lupou (frequency zoomed auto-regressive, FZ-AR). Vlastní klasifikace je provedena pomocí regularizované Mahalanobisovy vzdálenosti. Jako příznaky pro klasifikaci byly odzkoušeny frekvence pólů FZ-AR, koeficienty FZ-AR a výkonová spektrální hustota počítaná z FZ-AR.

Klasifikace byla odzkoušena na již nahraném EEG signálu z 41 kanály. Odzkoušeny byly různé parametry a výsledky klasifikace byly porovnány. Výsledné navrhované nastavení klasifikace používá jako příznaky frekvence pólů počítané na 15 sekundových úsecích.



Obrázek 2: Integrace do BCI systému.

2.2. Zpracování v reálném čase

Pro využití výhod identifikace subjektu je potřeba aby identifikace byla součástí BCI systému. Proto jsme upravili offline metodu z [6] pro použití v reálném čase a zabudovali ji do našeho BCI systému. Protože náš systém pracuje v reálném čase a zpracování probíhá na souvislém proudu dat bez segmentace, chtěli jsme použít stejný přístup i pro identifikaci. Náš BCI systém je rozdělen do nezávislých modulů, které spolu komunikují po počítačové síti (Obr. 1). Identifikační algoritmus byl implementován jako dva takové moduly (Obr. 2).

První z nich počítá FZ-AR pomocí NLMS algoritmu. Na vstupu má EEG data a vytváří odhady frekvencí pólů FZ-AR pro každý kanál. Výstupní signál je převzorkován na frekvence paketů, takže každý paket má v sobě jeden vzorek příznaků. Protože příznaky neobsahují vysoké frekvence, není pro převzorkování potřeba žádný filtr.

Druhý modul je klasifikátor založený na Mahalanobisově vzdálenosti. Používá jakýkoliv vstup jako příznaky a vytváří výsledek klasifikace. Přepínání klasifikátoru z a do trénovacího stavu je ovládané systémovými příkazy, které mohou být poslány například z již existujícího ovládacího modulu.

Moduly mohou být snadno použity spolu se zbytkem systému. Informace o aktuálním subjektu může být poslána do jakéhokoliv modulu a použita jako dodatečná informace. Rozdělení na dva moduly bylo provedeno, aby se umožnilo i jiné použití modulů. FZ-AR modul může poskytovat příznaky jinému klasifikátoru a na druhou stranu modul Mahalanobis může klasifikovat jiné příznaky.

Průběžný odhad FZ-AR parametrů je použit pro výpočet příznaků. Vlastní frekvenční lupa je provedena pomocí modulace signálu na střední frekvenci

$$x_m[n] = x[n]e^{j\Omega_m n}, \quad (1)$$

kde x je EEG signál, x_m je modulovaný EEG signál, $\Omega_m = 2\pi \frac{f_m}{f_s}$, f_s je vzorkovací frekvence (v našem případě 200 Hz) a f_m je modulační frekvence. Průběžný výpočet v reálném čase není problém. Poté je signál decimován poměrem k . Pro decimaci je použit filtr typu FIR.

Dalším krokem je výpočet parametrů AR modelu. AR model odhaduje příští hodnotu signálu jako

$$\tilde{x}[n+1] = X[n]^T W[n], \quad (2)$$

kde \tilde{x} je signál odhadnutý modelem, p je řád AR modelu (my používáme 2), W je vektor AR koeficientů a $X[n] = (x[n], x[n-1], \dots, x[n-p])^T$ je vektor minulých hodnot signálu.

AR koeficienty je možné odhadnout několika způsoby. Zvolili jsme komplexní verzi LMS algoritmu [7], protože je jednoduchá a nepotřebuje segmentaci. AR koeficienty W jsou rekurzivně počítány pro každý nový vzorek

$$W[n+1] = W[n] + \mu \bar{X}[n] (x[n] - \tilde{x}[n]), \quad (3)$$

kde μ je koeficient učení.

Koeficienty modelu nejsou vhodné jako příznaky. Rozdíl v EEG mezi subjekty je ve frekvenci μ rytmu, proto by měly být vhodné příznaky založené na frekvencích [5]. Frekvence jsou počítané z kořenů r_i polynomu

$$P(x) = x^p + w_p x^{p-1} + w_{p-1} x^{p-2} + \dots + w_2 x + w_1, \quad (4)$$

kde w jsou AR koeficienty z vektoru W . Argumenty kořenů odpovídají frekvencím

$$f_i = \frac{\arg(r_i)}{2\pi} * k + f_m. \quad (5)$$

Takže počet příznaků odpovídá řádu modelu vynásobeném počtem kanálů. Protože kořeny polynomu nemají žádné konkrétní pořadí jsou výsledné frekvence seřazené podle velikosti.

Protože příznaky počítáme průběžně, rozptyl signálu je neznámý, takže není možné odhadnout rychlost konvergence. To lze řešit použitím normalizovaného LMS – NLMS [8]. NLMS nepoužívá konstantní μ koeficient ale definovaný jako

$$\mu = \frac{\mu_0}{X^T X + \Psi}, \quad (6)$$

kde μ_0 je koeficient učení a Ψ je koeficient, co zabraňuje dělení malým číslem.

Klasifikace je prováděna pomocí regularizované Mahalanobisovy vzdálenosti [9]. Ta je definována jako

$$d = (f - \bar{f}) \left((1 - \lambda) (S + \varepsilon I)^{-1} + \lambda I \right) (f - \bar{f})^T, \quad (7)$$

kde f je vektor příznaků, \bar{f} je vektor středních hodnot příznaků, S je kovarianční matice vektoru příznaků a λ s ε jsou parametry regularizace.

Střední hodnota a kovarianční matice jsou počítány průběžně během trénování. Průběžné hodnoty mohou být uloženy a později načteny, takže trénování může být rozděleno do několika sezení. Když je klasifikátor přepnut do identifikačního režimu, pro každý subjekt je spočítána matice

$$S' = (1 - \lambda) (S + \varepsilon I)^{-1} + \lambda I, \quad (8)$$

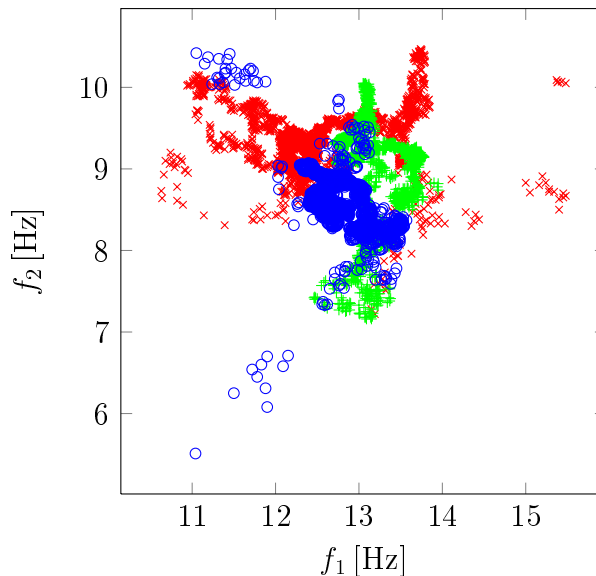
Která se následně používá pro výpočet Mahalanobisovy vzdálenosti. Na výpočet inverze je použita Choleskyho dekompozice [10].

Veškeré zpracování je prováděno na proudu dat v reálném čase, žádná segmentace není potřeba, což nám usnadnilo integraci do našeho systému pro zpracování EEG v reálném čase.

3. Výsledky

Pro testování identifikačního systému jsme použili data z experimentu popsáném v [11] a [12]. Subjekty si představovaly pohyby, takže nahrávání bylo prováděno tak, aby zachytilo jevy související s pohybem. Záznam sestával jenom ze dvou EEG kanálů. Ty byly nahrány v bipolárním zapojení na levé a pravé straně hlavy (elektrodové páry FC3 – CP3 a FC4 – CP4). μ rytmus by zde měl být přítomen. Vzorovací frekvence byla 200 Hz a signál byl v EEG zařízení filtrován pásmovou propustí 0.5 – 35 Hz. Konfigurace experimentu odpovídá plánovanému použití identifikačního algoritmu jako součásti BCI.

Experiment zahrnoval tři subjekty. Každý subjekt se zúčastnil tří nahrávacích sezení rozložených do dvou měsíců. Každé sezení sestávalo až z devíti měření. Každé měření trvalo přibližně čtyři minuty. Výsledná databáze z tohoto experimentu obsahovala 71 měření.



Obrázek 3: Projekce dvou příznaků tří různých subjektů.

Použili jsme modulační frekvenci $f_m = 10.5$ Hz, a decimační poměr $k = 20$. Decimační filtr byl typu FIR řádu 100 se zlomovou frekvencí 2.5 Hz. Pomocí frekvenční lupy byl tedy vybrán rozsah 8 – 13 Hz, kde je přítomen μ rytmus.

Pakety posílané systémem obsahovaly 5 vzorků signálu. Příznakový signál byl převzorkován na jeden vzorek na paket, takže vzorkovací frekvence příznaků byla 40 Hz.

Učící koeficienty NLMS byly nastaveny na $\mu_0 = 0.001$ a $\Psi = 100$. μ_0 koeficient byl nastaven tak, aby algoritmus konvergoval v řádu sekund. Potom byl upraven, aby výsledné skóre bylo co nejvyšší. Hodnota koeficientu Ψ nemá na výsledek moc velký vliv. Koeficienty regularizované Mahalanobisovy vzdálenosti byly zvoleny podle [6]: $\varepsilon = 0.2$ a $\lambda = 0.4$.

Pro testování byl použit skutečný systém pro zpracování EEG v reálném čase. Předem nahraná data byla použita jako zdroj signálu místo živého EEG. Aby se zkrátilo vyhodnocení, byla data posílána desetkrát rychleji než v reálném čase. To ukázalo, že systém je efektivní a v reálném čase by zvládl mnohem náročnější vyhodnocení.

Nejllepších výsledků bylo dosaženo s AR modelem druhého řádu. S dvěma EEG kanály to dává čtyři příznaky (Obr. 3).

Prvních 700 vzorků z každého měření bylo vynecháno z klasifikace, aby měl LMS algoritmus čas na prvotní konvergenci.

Byla provedena deseti násobná křížová validace na 71 měřeních. Skóre pro každé měření bylo počítáno jako poměr mezi správně klasifikovanými vzorky a všemi vzorky. Průměrné skóre křížové validace bylo $87 \pm 4\%$.

4. Závěr

Vyvinuli jsme algoritmus, který dokáže identifikovat subjekt z EEG signálu. Algoritmus byl zabudován do již existujícího BCI systému. Algoritmus je prováděn v reálném čase.

Algoritmus identifikuje subjekt, který používá BCI, což umožňuje automatické zvolení nastavení, které je závislé na subjektu. To je krok k lepší použitelnosti reálného BCI systému.

Popisovaný algoritmus používá pouze kanály, které používá vlastní BCI, takže není potřeba získávat žádná dodatečná data. Identifikace byla do BCI dodána pouze změnou softwaru.

Algoritmus byl testován pomocí křížové validace za podmínek velmi podobných zamýšlenému použití. Průměrné skóre pro tři subjekty bylo 87%. Dokázali jsme, že algoritmus je schopný

identifikovat subjekt dokonce pouze se dvěma EEG kanály.

Plánujeme použití tohoto algoritmu jako součást BCI systému. Informace o identitě subjektu bude použita pro přepínání parametrů specifických pro subjekt, jako například koeficienty klasifikátoru.

Poděkování

Výzkum je podporován grantovou agenturou Českého Vysokého Učení Technického v Praze grantem číslo SGS12/143/OHK3/2T/13.

Reference

- [1] POULOS, M. et al. Person identification based on parametric processing of the EEG. In *Electronics, Circuits and Systems, 1999. Proceedings of ICECS '99. The 6th IEEE International Conference on*, 1, p. 283–286 vol.1, 1999. doi: 10.1109/ICECS.1999.812278.
- [2] DOLEŽAL, J. – ČERNÝ, V. – ŠŤASTNÝ, J. Constructing a brain-computer interface. In *Applied Electronics (AE), 2011 International Conference on*, p. 1–4, Sept. 2011.
- [3] ŠŤASTNÝ, J. et al. Design of a modular brain-computer interface. In *Applied Electronics (AE), 2010 International Conference on*, p. 319–322, Sept. 2010.
- [4] VAUGHAN, T. et al. The wadsworth BCI research and development program: at home with BCI. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*. June 2006, 14, 2, p. 229–233. ISSN 1534-4320. doi: 10.1109/TNSRE.2006.875577.
- [5] TRAN, Y. – CRAIG, A. – MCISAAC, P. Extraversion–introversion and 8–13 Hz waves in frontal cortical regions. *Personality and Individual Differences*. 2001, 30, 2, p. 205–215. ISSN 0191-8869. doi: 10.1016/S0191-8869(00)00027-1.
- [6] ŠŤASTNÝ, J. – SOVKA, P. – KOSTÍLEK, M. Overcoming Inter-Subject Variability in BCI using EEG-Based Identification. Will be published, 2013.
- [7] WIDROW, B. – MCCOOL, J. – BALL, M. The complex LMS algorithm. *Proceedings of the IEEE*. April 1975, 63, 4, p. 719–720. ISSN 0018-9219. doi: 10.1109/PROC.1975.9807.
- [8] NITZBERG, R. Application of the Normalized LMS Algorithm to MSLC. *Aerospace and Electronic Systems, IEEE Transactions on*. Jan. 1985, AES-21, 1, p. 79–91. ISSN 0018-9251. doi: 10.1109/TAES.1985.310541.
- [9] ARCHAMBEAU, C. – VERLEYSSEN, M. Fully Nonparametric Probability Density Function Estimation with Finite Gaussian Mixture Models. In *Proceedings of the 5th International Conference on Advances in Pattern Recognition*, p. 81–84, Calcutta, India, December 2003.
- [10] KRISHNAMOORTHY, A. – MENON, D. Matrix Inversion Using Cholesky Decomposition. *CoRR*. 2011, abs/1111.4144.
- [11] ČERNÝ, V. EEG processing system implementation. Master's thesis, Czech Technical University in Prague, 2012.
- [12] DOLEŽAL, J. – ČERNÝ, V. – ŠŤASTNÝ, J. Online motor-imagery based BCI. In *Applied Electronics (AE), 2012 International Conference on*, p. 65–68, Sept. 2012.

Validace výběru parametrů pro filtrování podprostorů při zpracování pohybového EEG

Martin DOBIÁŠ¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů
dobiana3@fel.cvut.cz

Abstrakt: Výzkum popsáný v tomto příspěvku je určen k zrychlení rozhraní mozek-stroj pomocí navýšení počtu rozeznatelných EEG pohybových stavů. Proto jsou ve výzkumu zkoumány pohyby prováděné na jedné straně těla. Výzkum je zaměřen na použití metod slepé separace pro zlepšení odstupu signál/šum v pohybovém EEG. K tomuto účelu je použita metoda filtrování podprostorů vyvinutá v naší laboratoři. Konkrétně je zkoumáno nastavení podprostorového filtru použitého k odstranění nežádoucích složek signálu. Bylo nalezeno stabilní nastavení filtru, které je odolné vůči změnám způsobeným nestabilitou algoritmu analýzy nezávislých komponent a variabilitou EEG. Tímto také proběhla verifikace vyvinutého algoritmu. Pomocí aplikování metody filtrace podprostorů se podařilo zlepšit výsledky klasifikace v průměru o 16,7% dosahujíc klasifikačního skóre distálního vs. proximálního pohybu až $99,1 \pm 0,9\%$.

1. Úvod

Rychlost rozhraní mozek stroj je v dnešní době relativně nízká, pouze v řádu desítek bitů za minutu. Naším cílem je tuto rychlost zvýšit pomocí navýšení počtu rozeznatelných stavů v EEG. V našem výzkumu používáme pohybové EEG, jelikož pro lidské bytosti je velice přirozené ovládat své okolí pohybem. Většina výzkumu v této oblasti se zaměřuje na rozeznání pohybů na pravé a levé části těla [1, 2], ať už představovaných, nebo skutečných. Naproti tomu, náš výzkum je založený na studiu lateralizované aktivity, tedy na pohyby na jedné straně těla. Takovéto rozlišení umožňuje snazší navýšení počtu stavů v EEG.

Pro rozeznání jednotlivých typů pohybů používáme klasifikační systém založený na analýze časového vývoje pohybového EEG pomocí skrytých Markovských modelů. Přestože tento systém vykazuje lepší výsledky s lateralizovanou aktivitou, než většina běžně používaných [5], je vzhledem k rychlosti rozhraní další zlepšení více než žádoucí. Proto byla vyvinuta metoda filtrování podprostorů založena na analýze nezávislých komponent, se kterou jsme dosáhly výrazného zlepšení klasifikace [3,4].

Jedná se o parametrickou metodu pracující ve dvou krocích. V prvním kroku jsou odstraněny nepohybové komponenty a v následujícím jsou selektivně vybrány komponenty vykazující pohybovou aktivitu. Pouze tyto jsou v konečné fázi filtrace rekonstruovány zpět na skalp. Ve výzkumu identifikujeme vhodné nastavení parametrů pro výběr komponent, zejména pak těch pohybových. Parametry jsou zkoumány i z hlediska stability.

Nastavení podprostorových filtrů není v literatuře příliš probíráno. Výběr komponent je buď manuální [6], což v praxi znamená jen malý počet opakování ICA algoritmu při testování, nebo automatický. Ten může být založený na různých vlastnostech, například na topologickém rozložení IC [7] (Independent Components - nezávislých komponent) nebo vlastnostech časového vývoje signálu IC [8].

1.1. Pohybové EEG

Pohybová aktivita je v EEG doprovázena rytmickými změnami, které jsou zpravidla patrné v μ (8-13 Hz) a v β (12-30 Hz) pásmu. Tyto změny se nazývají desynchronizace (Event-Related Desynchronization - ERD), snížení aktivity, a synchronizace (Event-Related Synchronization - ERS), zvýšení aktivity [9]. Podle pásem se pak rozdělují na μ ERD, β ERD, μ ERS a β ERS. ERD začíná obvykle cca 2 s před pohybem a končí několik málo sekund po něm. ERS je spojeno s nástupem pohybu a je zpravidla silnější v β pásmu, jeho špička je zhruba 1 s po pohybu, kdy ERD stále ještě probíhá [13]. ERD je výraznější v μ pásmu.

1.2. Databáze

EEG databáze momentálně používaná pro výzkum byla původně nahrána pro studii [9]. Databáze obsahuje nahrávky volních pohybů pravého ramene (proximální pohyb) a pravého ukazováčku (distální pohyb) od sedmi pravorukých dobrovolníků. Pohyby byly prováděny v nepravidelných intervalech 12-15 s. Nahrávky byly pořízeny pro 59 EEG kanálů rozmístěných podle systému 10-10, $f_s = 50$ Hz. Elektroda Cz byla použita jako referenční.

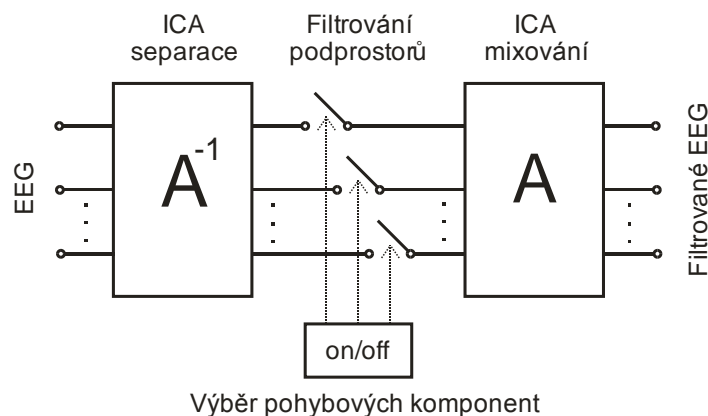
Data byla segmentována do 10 s epoch s počátkem pohybu v polovině tohoto intervalu. Počet realizací pohybů nezanešených artefakty je 52/27, 97/59, 58/40, 44/34, 101/107, 86/51, a 101/87 pro osoby 1-7 a distální/proximální pohyby.

1.3. Filtrování podprostorů

Filtrování podprostorů je technika určená ke zlepšení odstupů signál-šum v multidimenzionálním signálu. Základem metody je technika zvaná Slepá Separace Zdrojů (Blind Source Separation - BSS). Nejzákladnějším příkladem této metody je redukce dimenzionality pomocí PCA, ve které jsou data transformována do nekorelovaných podprostorů a pouze ty s vysokým výkonem jsou zachovány.

V principu se metoda skládá ze tří kroků (Obr. 1):

1. Transformace - převod do nového souřadnicového systému (PCA, ICA, nebo jiná BSS metoda)
2. Filtrace - výběr „zajímavých“ podprostorů a potlačení všech ostatních.
3. Rekonstrukce – projekce vybraných podprostorů zpět do původních souřadnic.



Obr. 1 – Filtrování podprostorů

2. Metodika

2.1. Filtrování podprostorů

2.1.1. BSS algoritmus

Pro rozklad EEG je použita vylepšená verze algoritmu FastICA [10] nazývaná EFICA [11]. Jelikož je FastICA (symetrický přístup) nedílnou součástí EFICA, ponechává si EFICA svou rychlost a robustnost a navíc pomocí jemného ladění vylepšuje nezávislé komponenty. Algoritmus rovněž obsahuje test sedlových bodů. EFICA poskytuje lepší výsledky při rozkladu pohybové EEG aktivity [3]. Na základě předchozího výzkumu [3,12] byla pro základní rozklad použita nelinearita \tanh .

2.1.2. Filtrace

Filtrace je rozdělena do dvou částí. V první části jsou na základě pěti kritérií odhaleny známé typy nepohybových komponent. Tři z nich jsou založeny na lokalizaci IC na skalpu (aktivita soustředěná na jednu elektrodu, výrazný výkon pro nepohybové oblasti, lokální maxima mimo motorickou oblast) a dvě na známých technických (krátký širokospektrální burst v době pohybu) a biologických (rostoucí výkon ve vyšších frekvenčních pásmech) artefaktech. Aplikace těchto kritérií redukuje riziko vybrání nepohybových komponent jako pohybových v další fázi, navíc jejich výpočet zabere jen malé množství výpočetních zdrojů. Podrobnější popis lze nalézt v [4]. Parametry vztahující se k těmto kritériím byly vybrány tak, aby pro žádnou osobu nebyla vyřazena žádná ručně identifikovaná pohybová komponenta.

V druhé části se ze zbylých komponent vyberou ty pohybové [3, 4]. Výběr je založen na pohybových vlastnostech EEG rozpoznávaných z časově frekvenčního vývoje nezávislé komponenty. Z něho se určuje síla ERD a ERS. Ta se počítá jako průměrná hodnota statisticky významného snížení/zvýšení EEG aktivity:

$$ERD_{ST}[i] = k \sum_{t=-2s}^{2s} \sum_{f=5Hz}^{30Hz} |ERD_i[t, f]| \quad (1)$$

$$ERS_{ST}[i] = k \sum_{t=0s}^{5s} \sum_{f=5Hz}^{35Hz} |ERS_i[t, f]|, \quad (2)$$

kde ERD a $ERS[t, f]$ je časově-frekvenční mapa statisticky významného snížení/zvýšení aktivity, t je čas, f je frekvence, i je index IC a k je normalizační konstanta. Rozsahy času a frekvence jsou podloženy fyziologickým výzkumem [9]. IC je sledována pohybovou pokud platí alespoň jedno z následujících:

$$ERD_{ST}[k] > C_{6T} \quad (3)$$

$$ERS_{ST}[k] > C_{7T} \quad (4)$$

Rozdíl možných hodnot síly ERD a ERS pro různé osoby je výrazný, proto jsou tyto parametry společně s minimálním počtem vybraných IC pomocí ERD (N_{6min}) a ERS (N_{7min}) posuzovány pro každou osobu individuálně. Minimální počet vybraných IC je nutný, aby v případě nevhodného nastavení C_{6T} či C_{7T} byly v druhé části vybrány alespoň nějaké komponenty. Ukazuje se, že mohou navíc být použity jako hlavní výběrové kritérium.

2.2. Výběr pohybových parametrů

Vzhledem k tomu, že výběr pohybových komponent má mnohem větší vliv na výsledek klasifikace než potlačení nepohybových, jsou pohybové parametry detailněji optimalizovány. Výběr se provádí z následujících hodnot v Tabulce 1.

Tab. 1: Množina parametrů

Parametr	Hodnoty
C_{6T}	0,01; 0,02; 0,03; 0,04; 0,05; 0,07; 0,08; 0,1; 0,15; 0,25
N_{6min}	2; 4; 6; 8; 10
C_{7T}	0,01; 0,03; 0,05; 0,1; 0,2
N_{7min}	0; 1; 2; 4

Tyto hodnoty byly nastaveny na základě obvyklých hodnot pro databázi. Vzhledem k nejednoznačnosti závislosti klasifikačního skóre na síle ERD/ERS zvolených komponent nebyl jiný postup příliš praktický (Obr 2). Jelikož by bylo vzhledem k časové náročnosti klasifikace pomocí HMM počítat všechny možné kombinace časově náročné, byla použita metoda rozděl a panuj, kde jsou nejdříve optimalizovány ERD parametry a potom pro 5 nejsilnějších kombinací ještě ERS parametry [3,4].

Alternativní způsoby výběru pohybových komponent nejsou příliš praktické. Jelikož s každým novým výpočtem ICA dostáváme nové komponenty je manuální výběr je nesmírně zdoluhavý, jednotvárný a náchylný k chybám. Navíc není možné ručně klasifikovat komponenty v reálném čase. Jelikož dopředu neznáme třídy, zdá se být využití pohybových vlastností EEG a finálního klasifikačního skóre jako vhodná cesta.

2.3. Validace výběru

Původní algoritmus pro výběr pohybových parametrů byl založen na principu popsaném v Sekci 2.2. Vzhledem k tomu, že hodnoty parametrů byly nastaveny pro každou osobu zvlášť a pouze na základě jediného ICA rozkladu, nastala zcela reálná možnost přeučení algoritmu pro jejich nastavení. Proto jsme se rozhodli provést validaci jejich výběru a zároveň nastavit metodu pro jejich určení, která by vykazovala stabilní výsledky přes několik běhů ICA algoritmu.

Pro validaci využíváme adaptovanou metodu křížové validace. V trénovací části jsou opakovaně pro různé sady vstupních dat hledány optimální parametry pro výběr pohybových komponent podprostorového filtru. Trénovací sady jsou složeny z vyvážené podmnožiny realizací pohybů. V experimentech používáme pětinasobnou křížovou validaci [14] ve které je každá realizace použita právě čtyřikrát.

Algoritmus může být popsán následovně:

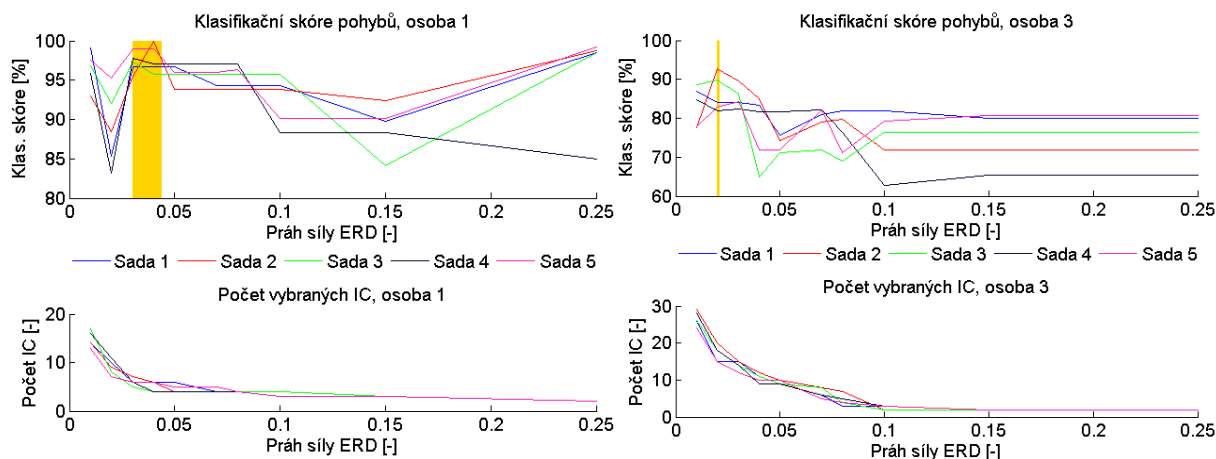
1. Vezměme 50 realizací pohybů (25 distálních, 25 proximálních) a rozdělme je do pěti vyvážených skupin. Vezměme všech pět možných kombinací čtyř skupin a nazvěme je trénovací sady. Každá tedy obsahuje 40 realizací (40 je dostačující a umožňuje provádět výpočty EFICA na slabších 32bit PC).
2. Pro každou trénovací sadu nalezneme EFICA separační matici a získáme nezávislé komponenty a
 - a. aplikujeme potlačení nepohybových komponent [3]
 - b. najdeme pět kombinací C_{6T} a N_{6min} z Tabulky 1, pomocí kterých získáme nejlepší výsledky klasifikace pohybů
 - c. najdeme nejlepší kombinace parametrů C_{7T} a N_{7min} a pěti nejlepších kombinací z bodu 2.b.
3. Prozkoumejme nejlepší kombinace ze všech pěti sad a vyberme stabilní validované parametry.

3. Výsledky

Pomocí navrhnuté metody se podařilo získat optimální hodnoty pohybových parametrů. Výsledky pro osoby 1 a 3 jsou zobrazeny v Tabulce 2. V tabulce jsou uvedeny intervaly parametrů, jejichž kombinace vede ke stejnému podprostorovému filtru (stejně vybrané IC v každé příslušné sadě). Tabulka obsahuje vždy tři nejlepší kombinace (nejlepší průměrné skóre přes všechny sady) parametrů.

Tab. 2: Výsledky optimalizace včetně skóre pro jednotlivé sady

Os.	Klas. skóre nej. ele.	C_{6T} , rozsah ERD	N_{6T}	C_{7T} , rozsah ERS	N_{7T}	Nejlepší skóre				
						1	2	3	4	5
1	99.04±0.90	0.0302 - 0.0437	6	0.0097 - 0.0101	≤7	98.83	100.00	98.57	100.00	97.78
	98.94±1.06	0.0302 - 0.0437	6	0.0248 - 0.0329	2	97.85	99.76	97.56	99.76	99.76
	98.69±1.19	0.0280 - 0.0303	≤5	0.0097 - 0.0101	≤7	98.83	97.06	99.76	100.00	97.78
3	90.46±4.25	0.0199 - 0.0207	≤15	0.0085 - 0.0100	≤4	95.32	93.78	91.77	85.78	85.64
	89.10±5.10	0.0199 - 0.0207	≤15	>0.0157	4	95.32	92.70	89.95	85.78	81.73
	87.39±5.51	0.0295 - 0.0307	≤12	0.0085 - 0.0100	≤4	95.32	89.83	87.09	85.37	79.36



Obr. 2: Vývoj klasifikačního skóre a počtu IC pro různé hodnoty C_{6T}

Z tabulky je patrné, že limitující faktory mohou být pro každý případ odlišné. Například pro osobu 3, výběr 2 je pro komponenty vybírané pomocí síly ERS důležité, že se vyberou právě 4. Oproti tomu pro osobu 3, první výběr je determinující interval C_{6T} a N_{6T} musí být pouze menší než 15. Pro osobu 1, výběr 1, ERD, lze zvolit kteroukoliv z variant, například $N_{6min} = 6$ a $C_{6T} > 0,0302$, nebo $N_{6min} \leq 6$ a $C_{6T} (0,0302; 0,0437)$, obě povedou ke stejnému výsledku.

Obrázek 2 ukazuje graf závislosti klasifikačního skóre a počtu vybraných komponent na síle ERD (C_{6T}) pro všechny sady a je v něm vyznačen optimální interval z Tabulky 2.

4. Závěr

Byl navrhnut a implementován algoritmus pro validaci procesu výběru prahových hodnot ERD a ERS pro filtrování podprostorů. Bylo ukázáno, že lze najít takové nastavení filtru, které produkuje výsledky porovnatelné s [3,4], ale je stabilní přes více běhů ICA algoritmu, což přináší více realistické výsledky. Pro první osobu bylo dosaženo zlepšení klasifikace pohybů o 22,4% (v porovnání s Laplaceovským filtrem) a pro osobu 3 pak 18,4%.

V rámci budoucích vylepšení algoritmu podprostorového filtrování je třeba dále prostudovat nutnost využití potlačení nepohybových komponent, jelikož dosavadní výsledky ukazují, že jejich vliv je minimální. Zatímco parametry N_{6T} a N_{7T} jsou vhodné bezpečnostní opatření, jejich vhodnost jako výběrových kritérií by měla být více prostudována.

Vzhledem k tomu, že tato metoda byla primárně vyvíjena pro databázi distálních vs. proximálních pohybů, je třeba připravit systém jejího nastavení pro jiné databáze tak, aby její aplikace na jiná data byla co nejméně náročná měla co nejlepší účinek.

Filtrování podprostorů je odšumovací metoda s velkým potenciálem pro zpracování pohybové EEG a může pomoci zejména při rozeznávání více nelateralizovaných pohybů a při zpracování v reálném čase jako podpůrná metoda rozhraní mozek-stroj.

Poděkování

Výzkum popsany v tomto příspěvku byl podporován interním grantem ČVUT, SGS12/143/OHK3/2T/13 "Algoritmy a hardwarové realizace číslicového zpracování signálů".

Reference

- [1] Blankertz, B.; Dornhege, G.; Krauledat, M.; Müller, K.-R.; Kunzmann, V.; Losch, F.; Curio, G. The Berlin Brain-Computer Interface: EEG-based communication without subject training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2006), 147 - 152
- [2] Lehtonen, J.; Jylänki P.; Sams, M. Online classification of single EEG trials during finger movements. *IEEE Transactions in Biomedical Engineering* (2008).
- [3] Ručkay, L. EEG decomposition for simple movement classification. České vysoké učení technické v Praze, *dizertační práce*, 2010.
- [4] Ručkay, L.; Šťastný, J.; Dobiáš, M.; Sovka, P. Subspace Filtering for High-resolution Non-Lateralized Movement EEG Classification. *Biomedical Signal Processing and Control*, v recenzním řízení.
- [5] Doležal, J.; Šťastný, J.; Sovka, P. Exploiting temporal context in high-resolution movement-related EEG classification. *Radioengineering* (2011), 20, 666-676.
- [6] Vialatte, F.-B.; Solé-Casals, J.; Maurice, M.; Latchoumane, C.; Hudson, N.; Wimalaratna, S.; Jeong, J.; Cichocki, A. Advances in Neuro-Information Processing Improving the Quality of EEG Data in Patients with Alzheimer's Disease Using ICA. *Springer-Verlag* (2009), 979-986.
- [7] Hung, C.; Lee, P.; Wu, Y.; Chen, H.; Chen, L.; Yeh, T.; Hsieh, J. Recognition of motor imagery electroencephalography using independent component analysis and machine classifiers. *International Conference on Computer Graphics, Visualization and Computer Vision* (2004).
- [8] Vorobyov, S.; Cichocki, A. Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis. *Biological Cybernetics* (2002), 86, 293-303.
- [9] Stančák, A.; Feige, B.; Lücking, C. H.; Kristeva-Feige, R. Oscillatory cortical activity and movement-related potentials in proximal and distal movements. *Clinical Neurophysiology* (2000), 111, 636-650.
- [10] Hyvärinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks* (1999), 10, 626-634.
- [11] Koldovský, Z.; Tichavský, P. & Oja, E. Efficient variant of algorithm FastICA for independent component analysis attaining the Cramér-Rao lower bound. *IEEE Transactions on Neural Networks* (2006), 17, 1265-1277.
- [12] Ručkay, L.; Šťastný, J.; Sovka, P. ICA Model Order Estimation Using Clustering Method. *Radioengineering* (2007), 16, 51-57.
- [13] Durka, P. J.; Ircha, D.; Neuper, Ch.; Pfurtscheller, G. Time-frequency microstructure of event-related desynchronization and synchronization, *Medical & Biological Engineering & Computing* (2001), 39, 315-321.
- [14] Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection, *International Joint Conference on Artificial Intelligence* (1995), 1137-1145.

Klasifikace jemných pohybů z EEG signálu

Milan KOSTÍLEK¹

¹České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů
kostimil@fel.cvut.cz

Abstrakt: Tento článek se zabývá offline klasifikací pohybového elektroencefalografického signálu (EEG) s vysokým rozlišením. V první části příspěvku je popsána použitá EEG databáze, která je specifická tím že se skládá ze dvou částí naměřených s ročním odstupem. Dále je uveden princip validace EEG databáze, a zjištění ke kterým tato validace vedla. V druhé části je pak popsán samotný klasifikační algoritmus a jeho použití pro jednoduché klasifikační experimenty (2 třídy dat - pohybové EEG a odpočinek) i pro experimenty složitější (5 tříd dat - klasifikace jemných pohybů). Výsledky klasifikačních experimentů potvrzují předpoklady, že klasifikace jemných pohybů nedosahuje tak kvalitních výsledků jako jednodušší experimenty. Dále jsou v článku uvedeny výsledky experimentů provedených s využitím průměrování spektrogramů a zkracování EEG segmentů, a je ukázáno že obě tyto metody vedou ke zlepšení klasifikačních výsledků.

1. Úvod

Frekvenční spektrum EEG signálu se běžně rozděluje do několika frekvenčních pásem. Pásmo od 8 do 13 Hz přitom bývá nazýváno α , ve stejném frekvenčním pásmu se však také nacházejí tzv. μ -rytmy, které pocházejí ze sensorimotorického centra a lze je tedy naměřit nejvýrazněji přibližně na středu hlavy v oblasti mezi levým a pravým uchem. μ -rytmy jsou generovány sensorimotorickými neurony, a jejich amplituda je největší když není vykonáván žádný pohyb a naopak je utlumována během pohybu. S poměrně velkou přesností lze na jejich základě klasifikovat úseky EEG signálu do dvou skupin, a to EEG pohybové a klidové. Tuto klasifikaci lze využít v rozhraní mozek-stroj (BCI), což je takový systém, který dokáže signály lidského mozku převést na signály jiného typu, např. signály srozumitelné pro software běžného počítače. Pro efektivní využití BCI systému je však potřeba co nejvyšší datová propustnost, které lze dosáhnout například zvýšením počtu tříd pro klasifikaci. Při klasifikaci pohybového EEG lze tedy dále rozlišovat mezi pohyby na pravé a levé straně těla, mezi pohyby jednotlivých končetin, nebo také mezi konkrétními typy pohybu jako jsou například extenzní či flexní pohyby.

Výzkum týkající se klasifikace pohybového EEG a především vylepšování jejích výsledků se v současnosti ubírá mnoha různými směry. Studie [1] shrnuje jak se jednotlivé klasifikační experimenty odlišují typy použitých klasifikátorů i rozdílnými vlastnostmi EEG signálů zvolenými jako zdroje informace pro klasifikátor. Dále je také potřeba rozlišovat zda je prováděna klasifikace pohybového EEG pro vykonávané pohyby [2] [3] nebo pro pohyby představované [4] [5] [6]. Ve studii [3] bylo ukázáno, že lze pomocí EEG signálu rozlišovat mezi poměrně blízkými pohyby zápěstí a prstů s přesností přibližně 70%. V práci [7] byla představena metoda dekompozice μ -rytmů, která slouží ke zjištění užšího μ -pásma specifického pro daný subjekt, využití tohoto nalezeného frekvenčního pásma při klasifikaci pohybového EEG pak vede ke zlepšení klasifikačních výsledků. Specifická metoda klasifikace představovaného pohybového EEG sloužící k ovládní kursoru ve třech dimenzích byla navržena ve studii [8]. To že jde zjistit z pohybového

EEG signálu nejen druh vykonávaného pohybu, ale i jeho směr a kinematika, pak ukazuje souhrn výsledků v práci [9].

V experimentech popsaných v tomto příspěvku je použit klasifikátor založený na skrytých Markovovských modelech (HMM), protože poskytuje několik výhod pro klasifikaci pohybového EEG signálu. Jako příznaky pro klasifikátor jsou pak použity spektrogramy získané metodou rychle Fourierovy transformace (FFT) z jednotlivých EEG segmentů.

2. Použitá EEG databáze

Databáze EEG signálů, která byla použita pro všechny experimenty popsané v tomto článku, byla původně nahrána Ing. Doležalem [10] pro účely výzkumu pohybového EEG a jeho využití v BCI systému. Databáze obsahuje data 10 subjektů (všichni mužského pohlaví, průměrný věk 32 let s $\sigma = 11.8$). Databáze se skládá ze dvou oddělených částí, data v první části databáze byla naměřena přibližně o rok dříve než data v druhé části. Jeden ze subjektů se neúčastnil druhého měření, proto byly pro experimenty se spojenými částmi databáze použity EEG signály pouze devíti subjektů. EEG signály byly získány z celkem 53 elektrod rozmístěných dle systému 10-10, použitá vzorkovací frekvence byla 1024 Hz. Měření každého subjektu bylo rozděleno do čtyř bloků, během prvních tří bloků vykonával subjekt extenzní a flexní pohyby pravého či levého ukazováčku, během čtvrtého bloku pak bylo naměřeno klidové EEG. Všechny EEG signály byly filtrovány Laplaceovým prostorovým filtrem a rozděleny do segmentů (realizací) dlouhých 10 sekund (± 5 s kolem okamžiku pohybu pro první tři bloky, a pravidelně po 10 sekundách pro čtvrtý blok). Stejnou databázi jsem již dříve použil pro experimenty s biometrickou identifikací osob [11] [12].

2.1. Validace databáze

Vhodnost databáze pro klasifikační experimenty byla potvrzena pomocí validačního procesu, který byl navržen za účelem zpřesnění předchozí validace provedené Ing. Doležalem [10]. Nejprve byly lokalizovány a odstraněny všechny artefakty v naměřených EEG signálech, použita byla jednoduchá manuální metoda založená na vizuálním porovnání časového průběhu signálů z více elektrod. Průměrný počet segmentů, které byly označeny jako artefakty a následně odstraněny z dalšího zpracování, byl 1.84% pro první část databáze a 5.31% pro druhou část databáze. Jako další krok byla provedena analýza ERP potenciálů (potenciály navázané na událost), což jsou signály vznikající v lidském mozku a jsou vždy spojeny s výskytem nějaké konkrétní události (pro potřeby tohoto článku jsou danou událostí jednotlivé extenzní či flexní pohyby). ERP potenciály mají výrazně nižší amplitudu než zbytek EEG signálu, proto je lze nejjednodušeji získat průměrováním velkého množství EEG segmentů [13]. Analýza ERP potenciálů neukázala žádné výrazné rozdíly mezi extenzními a flexními pohyby, a potvrdila tedy, že pro klasifikaci těchto druhů pohybů je potřeba komplexnější charakteristika EEG signálu.

Počátky jednotlivých pohybových realizací byly v původních datech vyznačeny s použitím EMG signálů naměřených z pravé a levé ruky, tato metoda ovšem nemusí vždy poskytovat přesnou informaci o počátku pohybu z důvodu velkého zašumění EMG signálů. Pro zjištění přesné pozice počátku každého segmentu byla tedy použita metoda odstranění latencí [14]. Nalezené latence byly nezanedbatelné (větší než 0.1 s) pro 11.1% segmentů z první části databáze a pro 30.1% segmentů z druhé části databáze. Série klasifikačních experimentů pak ukázala, že odstranění latencí nemá výrazný vliv na klasifikační výsledky, a lze tedy usoudit, že samotný klasifikační algoritmus (viz. Kapitola 3.1) dokáže tyto latence kompenzovat.

Dále byla provedena analýza spektrogramů pro jednotlivé EEG segmenty, protože tyto spektrogramy byly dále použity jako příznaky pro samotnou klasifikaci. Spektrogramy byly an-

alyzovány jak jednoduchou vizuální metodou, tak i pomocí statické metody založené na Kruskal-Wallisově testu [15], cílem této analýzy bylo nalézt rozdíly ve spektrogramech pohybového EEG při extenzním a flexním typu pohybu. Ve spektrogramech pohybového EEG lze nalézt dva důležité fenomény - desynchronizaci spojenou s událostí (ERD) a synchronizaci spojenou s událostí (ERS) [16]. Jak ERD tak i ERS lze dále ještě rozlišovat podle frekvenčního pásma ve kterém se nacházejí, z tohoto hlediska jsou pak ve spektrogramech pohybového EEG nejvýznamnější μ ERD a β ERS (pásmo β je frekvenční oblast přibližně 13-30 Hz). Z analýzy vyplývá, že μ ERD lze poměrně snadno nalézt ve spektrogramech všech měřených subjektů a jsou v něm patrné rozdíly mezi flexními a extenzními pohyby. Frekvenční pásmo μ (8-13 Hz) je tedy z hlediska klasifikace spektrogramů pohybového EEG pro následující experimenty nejvýznamnější.

3. Klasifikační experimenty

3.1. Metody

Pro všechny klasifikační experimenty byl použit klasifikátor založený na metodě Skrytých Markovovských Modelů (HMM) [17]. Tento klasifikátor bývá častěji používán pro rozpoznávání řeči, ovšem má také některé vlastnosti vhodné pro rozpoznávání pohybového EEG [18]. Jak již bylo zmíněno dříve jako klasifikační příznaky byly použity FFT spektrogramy, a to s časovým rozlišením 200 ms (1 s dlouhé Blackmannovo okno s překryvem 800 ms) a frekvenčním rozlišením 1 Hz (ve frekvenčním pásmu 5-40 Hz). Z důvodu malého množství jednotlivých EEG segmentů (realizací) byla použita metoda křížové validace, kdy pro každý běh klasifikačního experimentu bylo 80% všech realizací použito pro trénování klasifikátoru a zbylých 20% pro testování. Celkem bylo provedeno 25 běhů každého klasifikačního experimentu, jejichž výsledky byly na závěr zprůměrovány - 5 permutací rozdělení realizací do 5 tříd a každá ze tříd použita jednou jako testovací množina a čtyřikrát jako část množiny trénovací.

Výsledkem klasifikačních experimentů tak bylo jedno klasifikační skóre pro každý subjekt, elektrodu a typ pohybu. Klasifikační skóre poskytuje kvalitní informaci o výsledcích offline klasifikačních experimentů, ovšem nehodí se pro srovnání výsledků s online BCI experimenty, kde

Tabulka 1: Klasifikační výsledky pro jednodušší experimenty (2 třídy). ITR v bitech za minutu.

Subjekt	První část		Druhá část		Spojená databáze	
	Skóre	ITR	Skóre	ITR	Skóre	ITR
1	84.5%	2.27	95.2%	4.32	93.1%	3.83
2	92.5%	3.69	93.2%	3.84	94.4%	4.14
3	90.9%	3.36	-	-	-	-
4	98.1%	5.19	100%	6.0	95.1%	4.30
5	96.4%	4.66	100%	6.0	85.9%	2.4
6	89.2%	3.04	98.0%	5.16	96.2%	4.61
7	90.6%	3.30	97.0%	4.82	89.8%	3.14
8	84.0%	2.19	98.6%	5.38	99.4%	5.69
9	79.2%	1.57	88.7%	2.94	96.1%	4.58
10	99.6%	5.77	97.4%	4.97	93.2%	3.86
Průměr	90.5%	3.50	96.5%	4.83	93.7%	4.06
	±6.6%	±1.36	±3.6%	±1.00	±3.9%	±0.93

Tabulka 2: Klasifikační výsledky pro klasifikaci jemných pohybů (5 tříd). ITR v bitech za minutu.

Subjekt	První část		Druhá část		Spojená databáze	
	Skóre	ITR	Skóre	ITR	Skóre	ITR
1	68.4%	4.74	59.2%	3.18	53.3%	2.35
2	58.9%	3.14	57.5%	2.92	48.5%	1.76
3	64.1%	3.97	-	-	-	-
4	61.1%	3.48	62.6%	3.73	50.6%	2.01
5	61.4%	3.53	57.4%	2.91	44.4%	1.32
6	64.7%	4.08	60.0%	3.31	60.0%	3.30
7	63.3%	3.84	62.0%	3.62	54.8%	2.54
8	62.7%	3.74	58.1%	3.02	43.4%	1.21
9	57.2%	2.89	59.8%	3.27	49.4%	1.86
10	61.2%	3.49	58.3%	3.05	51.0%	2.06
Průměr	62.3%	3.69	59.4%	3.22	50.6%	2.04
	±3.1%	±0.52	±1.9%	±0.29	±5.1%	±0.64

je důležitá také rychlost klasifikace a náročnost klasifikačního experimentu. Vhodným ukazatelem pro srovnávání offline a online klasifikačních experimentů se jeví Míra Přenosu Informace (Information Transfer Rate - ITR), kterou lze v bitech za minutu spočítat podle následujícího vzorce [19]:

$$ITR = K \left[\log_2 N + P \log_2 P + (1 - P) \log_2 \left(\frac{1 - P}{N - 1} \right) \right], \quad (1)$$

kde K představuje počet realizací za minutu, N je počet tříd pro klasifikaci, a P je klasifikační skóre (pravděpodobnost že bude daná realizace zařazena do správné klasifikační třídy). Všechny výsledky klasifikačních experimentů v tomto příspěvku jsou tedy hodnoceny jak samotným klasifikačním skóre tak i pomocí ITR.

Pro ověření funkce použitého klasifikátoru byly nejprve provedeny jednodušší experimenty, které měly za úkol rozlišovat mezi pohybovým EEG a odpočinkem (tedy EEG získané

Tabulka 3: Klasifikační výsledky po zprůměrování spektrogramů (5 tříd). ITR v bitech za minutu.

Subjekt	První část		Druhá část		Spojená databáze	
	Skóre	ITR	Skóre	ITR	Skóre	ITR
1	80.7%	3.68	70.2%	2.54	-	-
2	65.6%	2.11	65.9%	2.15	-	-
3	66.1%	2.16	-	-	-	-
4	71.6%	2.68	69.0%	2.42	-	-
5	73.5%	2.87	62.1%	1.82	-	-
6	73.5%	2.87	68.5%	2.38	-	-
7	73.0%	2.82	72.1%	2.73	-	-
8	67.6%	2.30	66.9%	2.23	-	-
9	68.4%	2.37	72.2%	2.74	-	-
10	72.7%	2.79	65.6%	2.11	-	-
Průměr	71.3%	2.67	68.1%	2.35	-	-
	±4.5%	±0.46	±3.3%	±0.30	±-	±-

během čtvrtého bloku měření - viz Kapitola 2). Celkem byly provedeny tři sady těchto jednodušších experimentů - pro první část databáze (10 subjektů), pro druhou část databáze (9 subjektů), a pro spojené části databáze (9 subjektů). Pro spojení databáze byl použit algoritmus založený na normalizaci výkonu EEG signálu [20].

Dále byly provedeny experimenty se samotnou klasifikací jemných pohybů, úkolem pro klasifikátor zde tedy bylo rozlišovat mezi 5 třídami EEG - extenzní pohyb levého ukazováčku, flexní pohyb levého ukazováčku, extenzní pohyb pravého ukazováčku, flexní pohyb pravého ukazováčku, a odpočinkové EEG. Opět byly provedeny tři sady experimentů - pro první a druhou část databáze odděleně a pro spojenou databázi.

Pro případné vylepšení výsledků klasifikace byly otestovány dvě metody - průměrování klasifikačních příznaků (FFT spektrogramů) a zkracování EEG realizací. Průměrování spektrogramů by mělo potenciálně vést ke zvýraznění μ -rytmu, který je nejdůležitější částí spektrogramu pro klasifikaci (viz Kapitola 2.1), a tedy ke zlepšení klasifikačního skóre. Naopak zkracování EEG segmentů nemá za úkol zvýšit klasifikační skóre, ale zvýšit maximální možný počet realizací za minutu a tím zvýšit hodnotu ITR.

3.2. Výsledky

Výsledky jednodušších klasifikačních experimentů (2 klasifikační třídy - pohybové EEG, odpočinek) jsou shrnuty v tabulce 1. Klasifikační skóre pro jednotlivé subjekty se pohybuje v rozmezí 80% a 100%, průměrné klasifikační skóre přes všechny subjekty je pak vyšší pro druhou část databáze (96.5%) než pro část první (90.5%), a dle očekávání průměrné skóre po spojení databází je zhruba mezi nimi (93.7%).

Tabulka 2 obsahuje výsledky pro druhou sadu klasifikačních experimentů (5 klasifikačních tříd - levá extenze, levá flexe, pravá extenze, pravá flexe, odpočinek). Klasifikační skóre pro všechny osoby je nižší než u předchozích jednodušších experimentů, ale jak můžeme vidět hodnoty ITR se pohybují na podobné úrovni (především pro samostatnou první část databáze). Z výsledků je také patrné, že po spojení databází dochází ke zhoršení klasifikačních výsledků (přibližně o 10%). To vede k závěru, že mezi daty z obou částí databáze jsou podstatné rozdíly, které se nepodařilo odstranit během spojování databáze a tyto rozdíly se projevují během kla-

Tabulka 4: Klasifikační výsledky po zkracování EEG segmentů (5 tříd). ITR v bitech za minutu.

Subjekt	První část		Druhá část		Spojená databáze	
	Skóre	ITR	Skóre	ITR	Skóre	ITR
1	50.7%	4.04	50.5%	3.99	-	-
2	51.6%	4.26	54.3%	4.96	-	-
3	53.9%	4.85	-	-	-	-
4	59.5%	6.46	55.0%	5.14	-	-
5	57.6%	5.88	57.3%	5.80	-	-
6	58.9%	6.28	49.4%	3.71	-	-
7	60.3%	6.70	61.0%	6.92	-	-
8	50.2%	3.91	58.3%	6.11	-	-
9	57.1%	5.73	60.3%	6.72	-	-
10	46.5%	3.06	59.6%	6.50	-	-
Průměr	54.6%	5.12	56.2%	5.54	-	-
	±4.7%	±1.26	±4.2%	±1.17	±-	±-

sifikačních experimentů.

Dále byl proveden experiment na otestování metody průměrování spektrogramů, pro průměrování byly vždy použity spektrogramy pro dvě po sobě jdoucí EEG realizace stejného typu. Výsledky pro tyto experimenty jsou v tabulce 3, jak můžeme vidět došlo ke zlepšení klasifikačního skóre, ovšem na úkor zhoršení ITR. Pokles ITR je způsoben tím, že na každý zprůměrovaný spektrogram je potřeba dva EEG segmenty a tedy maximální počet realizací za minutu se snižuje.

Tabulka 4 obsahuje výsledky experimentů provedených po zkracování EEG segmentů, kde EEG segmenty byly zkráceny z původních 10 sekund ($\pm 5s$ kolem pohybu) na 5 sekund ($\pm 2.5s$ kolem pohybu). Jak je patrné z těchto výsledků, klasifikační skóre po zkrácení segmentů klesá v průměru o 5% až 10%, ovšem dle předpokladů stoupá hodnota ITR a to v průměru o 2 bit/min. Tyto experimenty ukazují, že je možné pro klasifikaci pohybového EEG použít i kratší EEG segmenty, a bude tedy možné v online BCI experimentech dosáhnout vysokých hodnot ITR.

4. Závěr

Klasifikační experimenty, provedené na databázi obsahující flexní a extenzní pohyby, dokazují že je možné tyto pohyby poměrně dobře klasifikovat. Průměrná úspěšnost klasifikace, s použitím HHM klasifikátoru a FFT spektrogramů jako klasifikačních příznaků, dosahuje hodnoty přibližně 60%. Dále se mi podařilo zvýšit klasifikační skóre použitím průměrovaných spektrogramů, při průměrování ze dvou EEG realizací je průměrné dosažené skóre přibližně 70%. Tyto výsledky nejsou tak dobré jako při jednoduché klasifikaci mezi pohybovým EEG a klidovým EEG, kdy se průměrné skóre pohybuje kolem 93%, ovšem klasifikace jemných pohybů je z principu náročnější a dosažené výsledky jsou tedy uspokojivé. Dále bylo otestováno zkracování EEG segmentů, a bylo potvrzeno, že může vést v případě online experimentů k urychlení klasifikace a tedy ke zvýšení objemu přenesené informace pro rozhraní mozek-stroj (BCI).

Poděkování

Výzkum popsany v tomto článku byl podporován z interního grantu ČVUT SGS12/143/OHK3/2T/13 "Algoritmy a hardwarové realizace číslicového zpracování signálů".

Reference

- [1] Lotte, F.; Ccongedo, M.; Lécuyer, A.; Lamarche, F.; Arnaldi, B. A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces. *Journal of Neural Engineering* 4 (2007), R1-R13.
- [2] Quandt, F.; Reichert, C.; Hinrichs, H.; Heinze, H. J.; Knight, R. T.; Rieger, J. W. Single trial discrimination of individual finger movements on one hand: A combined MEG and EEG study. *NeuroImage* 59 (2012), 3316-3324.
- [3] Mohamed, A. K.; Marwala, T.; John, L. R. Single-trial EEG Discrimination between Wrist and Finger Movement Imagery and Execution in a Sensorimotor BCI. *33rd Annual International Conference of the IEEE EMBS* (2011), 6289-6293.
- [4] Ramoser, H.; Müller-Gerking, J.; Pfurtscheller, G. Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement. *IEEE Transactions of Rehabilitation Engineering* 8 (2000), 441-446.

- [5] Hsu, W. EEG-based motor imagery classification using enhanced active segment selection and adaptive classifier. *Computers in Biology and Medicine* 47 (2011), 633-639.
- [6] Gu, Y.; Dai, J.; Wu, B.; Zheng, N.; Chen, W.; Zheng, X. A Detective Method for Multi-class EEG-based Motor Imagery Classification Based on OCSVM. *Journal of Convergence Information Technology* 6 (2011), 257-265.
- [7] Veluvolu, K. C.; Wang, Y.; Kavuri, S. S. Adaptive estimation of EEG-rhythms for optimal band identification in BCI. *Journal of Neuroscience Methods* 203 (2012), 163-172.
- [8] Li, T.; Hong, J.; Zhang, J. Electroencephalographic (EEG) Control of Cursor Movement in Three-dimensional Scene Based on Small-world Neural Network. *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems* (2010), 587-591.
- [9] Jerbi, K.; Vidal, J. R.; Mattout, J.; Maby, E.; Lecaigard, F.; Ossandon, T.; Hamamé, C. M.; Dalal, S. S.; Bouet, R.; Lachaux, J. P.; Leahy, R. M.; Baillet, S.; Garnero, L.; Delpuech, C.; Bertrand, O. Inferring hand movement kinematics from MEG, EEG and intracranial EEG: From brain-machine interfaces to motor rehabilitation. *IRBM* 32 (2011), 8-18.
- [10] Doležal, J.; Šťastný, J.; Sovka, P. Recording and recognition of movement related EEG signal. *2009 International Conference on Applied Electronics* (2009), 95-98.
- [11] Kostílek, M. EEG signal as biometric characteristic and its long-term temporal stability. *POSTER 2012 - 16th International Student Conference on Electrical Engineering* (2012).
- [12] Kostílek, M.; Šťastný, J. EEG biometric identification: repeatability and influence of movement-related EEG. *2012 International Conference on Applied Electronics* (2012), 147-150.
- [13] Dawson, G. D. A summation technique for the detection of small evoked potentials. *Electroencephalography and Clinical Neurophysiology* 6 (1954), 65-84.
- [14] Thornton, A. R. D. Evaluation of a technique to measure latency jitter in event-related potentials. *Journal of Neuroscience Methods* 168 (2008), 248-255.
- [15] Rosner, B. *Fundamentals of Biostatistics*. 7th ed. Cengage Learning, 2011.
- [16] Pfurtscheller, G.; da Silva, F. H. L. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology* 110 (1999), 1842-1857.
- [17] Young, S. J.; Evermann, G.; Gales, M. J. F.; Hain, T.; Kershaw, D.; Moore, G.; Odell, J.; Ollason, D.; Povey, D.; Valtchev, V.; Woodland, P. C. *The HTK Book, version 3.4*. Cambridge, UK: Cambridge University Engineering Department, 2006.
- [18] Šťastný, J.; Sovka, P. High-resolution movement EEG classification. *Computational Intelligence and Neuroscience* (2007), 95-98.
- [19] Wolpaw, J. R.; Birbaumer, N.; Heetderks, W. J.; McFarland, D. J.; Peckham, P. H.; Schalk, G.; Donchin, E.; Quantrano, L. A.; Robinson, C. J.; Vaughan, T. M. Brain-Computer Interface Technology: A Review of the First International Meeting. *IEEE Transactions on Rehabilitation Engineering* 8 (2000), 164-173.
- [20] Doležal, J.; Šťastný, J.; Švadlenka, M. EEG database merging for BCI applications. *Electroscope* (2012).

Automatická detekce a prostorové klastrování interiktálních výbojů v invazivním EEG

Radek JANČA¹, Petr JEŽDÍK¹, Roman ČMEJLA¹, Pavel KRŠEK², Petr MARUSIČ³,
Přemysl JIRUŠKA^{3,4,5}

¹Fakulta Elektrotechnická, České vysoké učení technické v Praze, Praha, ČR

²Klinika dětské neurologie and ³Neurologická klinika, 2. lékařská fakulta, Karlova Univerzita, Fakultní nemocnice Motol, Praha, Česká republika

⁴Oddělení vývojové epileptologie, Fyziologický ústav, AV ČR, Praha, ČR

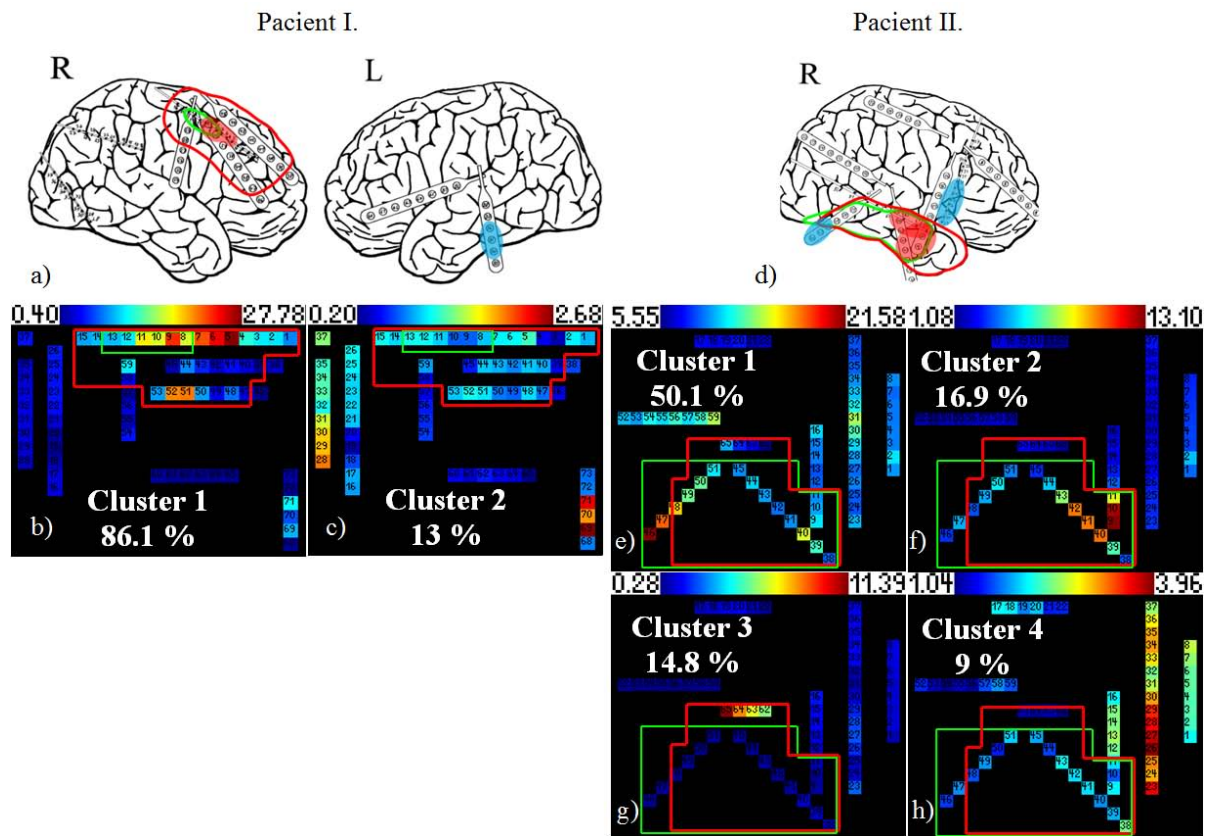
⁵Neuronal Networks Group, School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, United Kingdom

jancarad@fel.cvut.cz, jezdipl@feld.cvut.cz, cmejla@fel.cvut.cz, pavel.krsek@post.cz,
jiruskapremysl@gmail.com, Petr.Marusic@fnmotol.cz

Abstrakt: Epilepsie je nejčastějším chronickým neurologickým onemocněním v rozvojových zemích postihující okolo jednoho procenta populace. Přibližně třetina pacientů nereaguje na medikamentózní léčbu. Vybraná skupina farmakorezistentních pacientů může profitovat z chirurgické léčby. Principem je odstranění mozkové tkáně (epileptogenní zóny), která je zodpovědná za vznik záchvatů a jejímž odstraněním dojde k vymizení záchvatů. Určení této oblasti vyžaduje detailní předoperační vyšetření zahrnující elektrofyziologické a zobrazovací metody. U některých pacientů je vyžadována dlouhodobá implantace subdurálních a/nebo intra-cerebrálních elektrod. Intrakraniální záznam může poskytnout lepší prostorovou informaci o oblasti, kde vznikají záchvaty (Seizur Onset Zone - SOZ). Dodatečnou lokalizační informaci lze získat z časové a prostorové distribuce iritativní zóny, mozkové oblasti generující interiktální epileptiformní výboje.

Množství dat získaných z invazivního záznamu je enormní kvůli velkému počtu použitých elektrod (až stovky) a kontinuálního mnohadenního trvání. Subjektivní vizuální hodnocení je tedy velmi obtížné až neproveditelné. Současný trend směřuje k využívání automatických metod, které umožní: 1. analýzu dlouhotrvajících mnohakanálových záznamů a 2. extrahování relevantních informací pro zlepšení lokalizace epileptogenní zóny a tím i úspěšnosti chirurgického zákroku. Detektor hrotů je jedním z vyvíjených algoritmů, který umožňuje spolehlivou detekci výbojů. Druhý algoritmus zpracovává detekce a extrahuje časové a prostorové souvislosti mezi detekcemi, což výrazně zvyšuje výtěžnost invazivního monitoringu.

Použitím algoritmu na lidské signály prokazuje schopnost detekování nízkovoltážních výbojů v dobré shodě s expertním hodnocením (senzitivita 88,6 %). V druhém kroku klastrovací metoda využívající analýzu hlavních komponent (PCA) umožňuje identifikaci prostorové distribuce výskytu hrotů. Předběžné výsledky ukazují, že tento automatický přístup umožňuje identifikovat jasné zdroje interiktálních epileptogenních výbojů a tím potenciálně zpřesnit předoperační vyšetření pro určení iritativní zóny. Výsledky fokálního a multifokálního pacienta demonstruje obr. 1.



Obr. 1: Výsledky klastrování.

Pacient I. trpěl fokální kortikální dysplasií. Resekovaná oblast je značena červenou čarou, oblast začátku záchvatu zelenou (a). Obrázek znázorňuje prostorové rozmístění elektrod.

Frekvence hrotů za minutu je vyjádřena barevnou škálou. Klastrování ukazuje na dva dominantní klastry: klastr 1 (b) generuje 86,1 % událostí (červená oblast) a klastr 2 (c) generuje 13 % hrotů (modrá oblast). Pacient II.: Rozmístění elektrod v mozku (d), červená kontura značí chirurgicky odstraněnou oblast, zeleně je oblast počátku záchvatu. Červená plocha reprezentuje odstraněné klastry, modré byly ponechány. Barevná kortikální mapa ukazuje četnost hrotů v jednotlivých kanálech. Hlavní čtyři klastry se stejnou prostorovou distribucí jsou prezentovány. Klastr 1 (e) generuje 50,1 % událostí, klastr 2 (f) 16,9 % (odstraněno), klastr 3 (g) 14,8 % (odstraněno) a klastr 4 (h) 9 %. Oblast začátku záchvatu překrývá dva nejvýznamější klastry.

Poděkování

Tento projekt byl podporován Ministerstvem zdravotnictví České republiky IGA NT11460-4/2010, IGA NT13357-4/2012, SGS 10/272/OHK4/3T/13, Nadačním fondem Karla Janečka grantem pro podporu vědy a výzkumu 2012/10 - 00064203.

Detektor epileptiformních Vysokofrekvenčních oscilací v intrakraniálních elektroencefalografických záznamech

Tomáš HAVEL

České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

havelto3@fel.cvut.cz

Abstrakt: V článku jsou popsány principy algoritmu rozpoznávání epileptiformních Vysokofrekvenčních oscilací (HFO) od mozkové aktivity pozadí. Tento grafoelement se v poslední době ukazuje jako slibný biomarker epileptiformní tkáně. Zahrnutí oblasti v níž se HFO generuje do resekované oblasti také ukazuje na zlepšení pooperační prognózy. Aby lékaři mohli resekci plánovat efektivně, je zapotřebí nástroje, který by jim s hodnocením dlouhodobých elektroencefalografických záznamů pomohl. Tím je robustní automatický detektor HFO grafoelementů. Námi vyvíjený detektor je založen na principu nárůstu energie vysokofrekvenčních složek v průběhu HFO. Pro rozpoznání významného nárůstu se užívá dynamický práh, odvozený z parametrů modelu distribuce amplitud v jednotlivých segmentech obálky signálu. Detektor byl testován množinou intrakraniálních EEG záznamů epileptických pacientů. Odhadované výsledky, na základě neexpertně označených dat, dosahují senzitivity 80% při četnosti 4,3 falešně pozitivních detekcí za minutu.

1. Úvod

Epilepsie je skupina závažných neurologických onemocnění, které se projevují záchvatovými stavy tzv. epileptickými záchvaty. Ve většině případů se u epileptických pacientů daří zvládat jejich onemocnění pomocí léků. Přibližně u 30 % pacientů je však léčba neúčinná. Hovoří se pak o farmakorezistivní epilepsii, kdy se u pacienta i přes užívání léků nadále vyskytují záchvaty. V těchto případech se zvažuje chirurgická léčba epilepsie, která spočívá v odstranění epileptiformní tkáně. Za tímto účelem je potřeba epileptogenní oblasti mozku identifikovat a co nejpřesněji lokalizovat. Lékaři pak mohou zvážit možná rizika a rozhodnout o provedení operace.

Pacient je vyšetřován pomocí různých tomografických metod a skalpového elektroencefalogramu (EEG). Pokud jsou výsledky těchto neinvazivních metod neprůkazné a přesto se předpokládá operační zákrok, je potřeba provést další vyšetření. Na základě výsledků vyšetření a sémiologie záchvatů jsou určeny podezřelé oblasti mozku. Mozkovou elektrickou aktivitu v těchto oblastech je potřeba přesněji monitorovat a zmapovat. K tomu se využívá intrakraniálního elektroencefalogramu (iEEG). V rámci tohoto předoperačního vyšetření jsou pacientovi invazivně implantovány subdurální elektrody (umístěné na mozkové kůře) a/nebo intracerebrální hloubkové elektrody (umístěné přímo v mozkové tkáni). Mozková aktivita se standardně zaznamenává při relativně běžném režimu pacienta společně s videozáznamem (video-EEG) na specializovaném pracovišti po dobu přibližně jednoho týdne. Tento dlouhodobý EEG záznam je vyhodnocen a na základě výskytu typických epileptogenních grafoelementů

(výbojů) je lokalizována podezřelá oblast. V případě multifokální epilepsie se oblastí s epileptiformní aktivitou nachází v mozku větší počet.

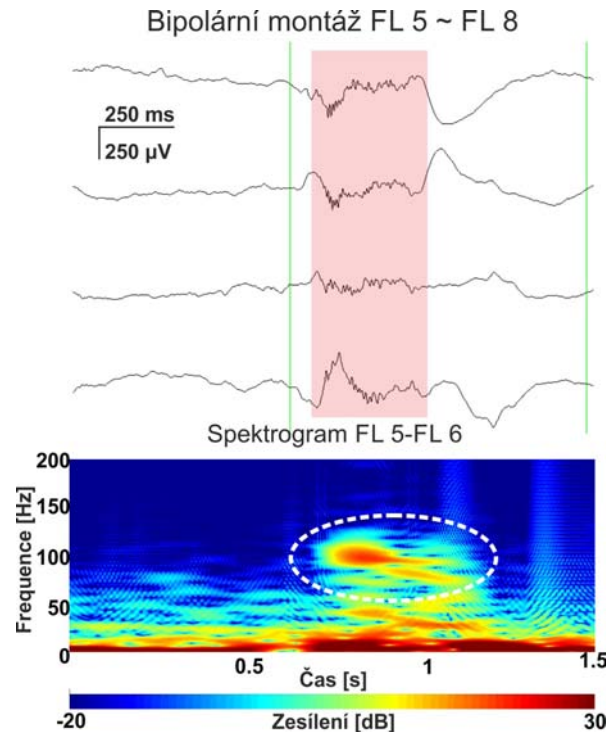
Jedním z epileptogenních grafoelementů jsou tzv. vysokofrekvenční oscilace (High-frequency oscillations, HFO). Doposud bohužel není pro tento grafoelement stanovena žádná přesná definice. Frekvenční pásmo, ve kterém se oscilace vyskytují, se nejčastěji udává 80 – 500 Hz a jejich trvání je typicky kratší než sekunda (Obr. 1). HFO byly nejprve pozorovány v záznamech z mikroelektrod [1]. Později bylo zjištěno, že se objevují i v intrakraniálních záznamech získaných pomocí makroelektrod [2]. Bylo prokázáno, že tento typ grafoelementu poskytuje cenné lokalizační údaje a pomáhá určit zónu počátku záchvatu (seizure onset zone, SOZ) [3-5]. Zóna, ve které se HFO generuje, se z pravidla neshoduje s SOZ, ale její zahrnutí do resekce má pozitivní dopad na celkový výsledek chirurgického zákroku [6,7].

Pro hodnocení dlouhodobých vícekanálových EEG záznamů je potřeba spolehlivě identifikovat HFO. Běžně se využívá pouze manuálního hodnocení záznamů lidskými experty (z pravidla neurologové). Je však známou věcí, že vzájemná i vlastní shoda hodnotitelů je silně omezená [8,9]. Současné automatické metody jsou zase zatíženy velkým počtem falešně pozitivních chyb (false positive, FP) [10].

1.1. Motivace

- Využití HFO jako relativně nového slibného biomarkeru epileptogenní tkáně.
- Usnadnění časově náročného vyhodnocování dlouhodobých záznamů lékařům v klinické praxi.
- Zpřesnění určení SOZ a iritační zóny s využitím prostorové lokalizace zdrojů HFO jako pomocných ukazatelů.
- Kvantitativní hodnocení výskytu HFO v různých interiktálních fázích iEEG záznamů (qEEG) by mělo vést k nalezení epileptických ložisek.

Cílem této studie je navrhnout a implementovat algoritmus vhodný k detekování HFO událostí v dlouhodobých iEEG záznamech. Výsledky označení událostí budou sloužit ke stanovení kvantitativních hodnocení výskytů HFO událostí u epileptických pacientů. Implementovaný algoritmus bude testován na iEEG záznamech pacientů epileptochirurgického programu.

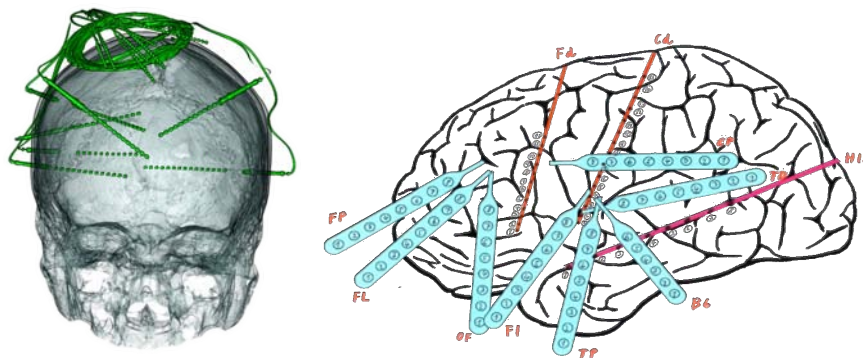


Obr. 1: Ukázka záznamu intrakraniálního EEG z několika elektrod jednoho subdurálního „stripu“ (pásku). HFO grafoelement je vyznačen červeně. Na spektrogramu z jednoho kanálu je čárkovane vyznačen typický frekvenční profil HFO, tzv. „BLOB“.

2. Materiály a metody

2.1. Množina testovacích dat

V průběhu spolupráce s Dětskou neurologickou klinikou a Neurologickou klinikou Fakultní nemocnice v Motole byly získány iEEG záznamy celkem 30 pacientů epileptochirurgického programu. Pacienti byli v rámci předoperačního vyšetření monitorováni přibližně po dobu jednoho týdne. Pacienti jsou různého věku od dětství po dospělost a trpí různými typy epileptického onemocnění. U 8 pacientů jsou k dispozici pouze iktální (záchvatové) záznamy s krátkými preiktálními (předzáchvatovými) a postiktálními (pozáchvatovými) úseky. U zbylých 22 pacientů jsou záznamy jak z iktální, tak interiktální (mezizáchvatové) fáze. Tato studie se zaměřuje na analyzování interiktálních záznamů. K analyzování tak připadá v úvahu 46 záznamů od těchto 22 pacientů. Délka záznamů se pohybuje v rozmezí 10 až 70 minut, počet elektrod je 47 až 128 a vzorkovací frekvence je 1000 Hz. Plánované umístění intrakraniálních elektrod je individuální a vychází ze sémiologie záchvatů a výsledků předcházejících vyšetření každého pacienta. Kortikální mapa se zakreslenými pozicemi elektrod, stejně jako po implantační rekonstrukce výpočetní tomografie (computed tomograph ,CT) (Obr. 2) je ve většině případů při analýzách k dispozici.



Obr. 2: Poimplantační 3D rekonstrukce výpočetní tomografie (CT), kortikální mapa se zakreslenými pozicemi elektrod

2.2. Principy detekčního algoritmu

Detekční algoritmus je založen na nárůstu energie vysokofrekvenčních komponent signálu v průběhu HFO. Pro lepší představu při čtení popisu je celý proces graficky znázorněn na (Obr. 3). Prvním krokem zpracování vícekanalového EEG záznamu je filtrace signálů. K filtraci se používá IIR filtr horní propust typu Chebyshev 2 (se zvlněným zádržným pásmem) s kompenzovaným posuvem fáze. Z filtrovaného signálu se pomocí Hilbertovy transformace stanoví obálka. Obálka se metodou posuvného okna rozsegmentuje na stejně dlouhé překrývající se úseky. V každém úseku je metodou odhadu maximální věrohodnosti (Maximum Likelihood Estimation) stanoven statistický model distribuce amplitud obálky. Distribuce amplitud je pozitivně vychýlená, nemůže být proto použita normální (Gaussova) distribuce. Pro nalezení distribuce, která nejlépe popisuje rozložení amplitud, bylo vyzkoušeno 17 různých vychýlených distribucí. Nejmenší odchylky modelu dosahuje Generalized Extreme Value nebo Alpha-stable model. Obě tyto distribuce jsou definovány čtyřmi parametry a jsou výpočetně náročné. Log-normální distribuce, definovaná dvěma parametry, poskytuje uspokojivou přesnost a má nižší výpočetní nároky. Pomocí parametrů modelu distribuce se stanoví mediány Med a střední hodnoty E . Tyto parametry distribuce byly vybrány proto, že jejich rozdíl $Med - E$ může být použit jako míra variability signálu. Získané hodnoty se interpolují na délku

signálu. Tímto postupem vznikne vektor hodnot mediánu **Med** a vektor středních hodnot **E**. Vektor prahovacích hodnot **th** se poté získá váhovanou kombinací vektorů **Med** a **E**, viz. Rovnice (1).

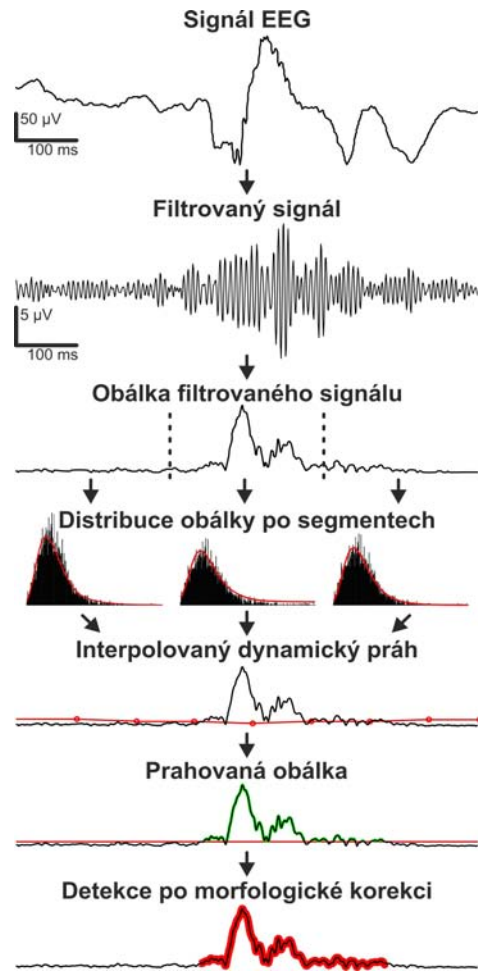
$$\mathbf{th} = c_{Med} \mathbf{Med} + c_E \mathbf{E} \quad (1)$$

Obálka filtrovaného signálu se porovnává s prahem. Ve vzorcích, ve kterých obálka práh překročí, je zaznamenána detekce. Výsledky prahování se ukládají do binární sloupcové matice, kde každý sloupec reprezentuje jeden kanál. Za účelem snížení počtu falešně pozitivních detekcí byly dodatečně implementovány post-processing metody. Pomocí binárních morfologických operací se sloučí detekce, mezi kterými je krátké přerušení (fragmentované detekce) a následně se odstraní detekce, které jsou příliš krátké. Maximální délka přerušení a minimální celková délka detekce byly zvoleny na základě publikovaných vlastností HFO [10-12]. Další metoda odstraňuje detekce úseků, ve kterých se vyskytují ostré hrany. Metoda využívá porovnání rozdílů amplitud a časů globálního minima a maxima daného detekovaného úseku. Tímto způsobem se potlačují falešné detekce filtračních, případně i technických artefaktů. Pro konkrétní nastavení parametrů detektoru se podívejte na sekci Výsledky.

3. Výsledky

Parametry detektoru byly optimalizovány na základě vizuálního hodnocení analyzovaných signálů. Snahou bylo dosáhnout nastavení, při kterém bude zachována vysoká senzitivita detektoru, ale počet falešně pozitivních detekcí bude co možná nejmenší. Mezní frekvence horní propusti byla nastavena v souladu se znalostmi HFO na 80 Hz. Délka segmentů, ze kterých se odhaduje distribuční model, byla zvolena 4 sekundy. Při této délce je zachována dostatečná dynamika. Tato délka také představuje experimentálně zjištěnou průměrnou délku kvazistacionárních úseků signálu. Překrytí segmentů se používá 50%. Byly testovány i větší hodnoty překrytí, ale nepřinášají výrazné zlepšení a jsou výpočetně náročnější. Váhovací koeficienty byly zvoleny $c_{Med} = 2$ a $c_E = 1.2$.

Implementovaná post-processing metoda spojuje detekce, které mezi sebou mají kratší úsek než 20 ms. Některé publikované detektory využívají pro detekované úseky pravidla, která udávají minimální počet vrcholů nebo period, aby byly detekce považovány za HFO [4,10]. U našeho detektoru se v post-processing metodách využívá podobných pravidel. Detekce, které mají kratší trvání, než 50 ms jsou odstraněny. Tato hodnota odpovídá čtyřem cyklům nejmolejší akceptované frekvence.



Obr. 3: Ilustrace detekčního procesu

Pro statistické hodnocení detektoru nebyly k dispozici expertem označené záznamy. Výsledky detektoru byly odhadnuty na základě záznamů označených laikem. Senzitivita byla stanovena 80% při četnosti falešně pozitivních detekcí 4,3 FP/min.

4. Diskuze

V rámci projektu byl navržen nový algoritmus pro detekování HFO událostí ve dlouhodobém EEG záznamu. Inovativním přístupem je především dynamické určení prahu, na základě statistického modelu obálky, určeného z jednotlivých časových segmentů. Detektor vycházející z popsaného algoritmu prokázal svou použitelnost při analyzování intrakraniálních EEG záznamů epileptických pacientů. I přes implementování post-processing kroků má však detektor stále relativně velkou četnost falešně pozitivních detekcí. Další vývoj detektoru bude směřovat právě ke snížení této četnosti užitím nových post-processing metod. Po získání expertně značených dat bude provedeno statistické hodnocení detektoru a optimalizace parametrů vedoucí ke zlepšení senzitivity.

Poděkování

Práce výzkumného týmu je podporována granty Ministerstva Zdravotnictví ČR IGA NT 11460-4/2010, IGA NT13357-4/2012, studentským grantem SGS 10/272/OHK4/3T/13.

Reference

- [1] Jiruska P, Bragin A. (2011) *High-frequency activity in experimental and clinical epileptic foci*. *Epilepsy Res.* 97:300-307.
- [2] Bragin A, Mody I, Wilson CL, Engel J Jr. Local generation of fast ripples in epileptic brain. *J Neurosci* 2002a; 22: 2012–21.
- [3] Bragin, A., Engel Jr., J., Wilson, C.L., Fried, I., Mathern, G.W., (1999). *Hippocampal and entorhinal cortex high-frequency oscillations (100–500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures*. *Epilepsia* 40, 127–137
- [4] Jiruska P., Finnerty G., Powell A., Lofti N., Cmejla R., and Jefferys J., (2010) *Epileptic high-frequency network activity in a model of non-lesional temporal lobe epilepsy*., *Brain: A Journal of Neurology*, vol. 133, no. 5, str. 1380-90
- [5] Jirsch JD, Urrestarazu E, LeVan P, et al. (2006) *High-frequency oscillations during human focal seizures*. *Brain* 129:1593–1608.
- [6] Jacobs J, Zijlmans M, Zelmann R, Chatillon CE, Hall J, Olivier A, Dubeau F, Gotman J. (2010) *High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery*. *Ann Neurol.* 67:209-220.
- [7] Zijlmans M, Jiruska P, Zelmann R, Leijten FS, Jefferys JG, Gotman J. (2012) *High-frequency oscillations as a new biomarker in epilepsy*. *Ann Neurol.* 71:169-78.
- [8] Blanco JA, Stead M, Krieger A, et al. (2010) *Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients*. *J Neurophysiol* 104: 2900-2912.
- [9] Gardner AB, Worrell GA, Marsh E, Dlugos D, Litt B. (2007) *Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings*. *Clin Neurophysiol* 118: 1134-1143.

- [10] Staba RJ, Wilson CL, Bragin A, Fried I, Engel J Jr. (2002) *Quantitative analysis of high-frequency oscillations (80–500 Hz) recorded in human epileptic hippocampus and entorhinal cortex.* J Neurophysiol 88:1743–1752.
- [11] Jefferys, J.G.R., Menendez de la Prida, L., Wendling, F., Bragin, A., Avoli, M., Timofeev, I., Lopes da Silva, F.H. (2012) *Mechanisms of physiological and epileptic HFO generation* Progress in Neurobiology, 98 (3), str. 250-264.
- [12] Zelmann, R., Zijlmans, M., Jacobs, J., Châtillon, C.-E., Gotman, J. (2009) *Improving the identification of High Frequency Oscillations* Clinical Neurophysiology, 120 (8), str. 1457-1464.

Využití Hilbert-Huangovy transformace v analýze epileptického EEG – pilotní studie

Jiří BALACH¹, Roman ČMEJLA¹, Petr JEŽDÍK¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

jiri.balach@fel.cvut.cz

Abstrakt: Tento článek pojednává o možnostech využití Hilbert-Huangovy transformace ve smyslu zpracování EEG dat pacientů s epilepsií. Tato metoda byla zvolena, protože přistupuje k signálu jako k nelineárnímu a nestacionárnímu procesu, jímž EEG záznam bezpochyby je. Článek je zaměřen především na empirickou modální dekompozici, která rozkládá signál na několik komponent s rozdílným frekvenčním obsahem. Diskutuje, zda se podle předpokladů v obsahu vyskytují hledané grafoelementy a zda je možné pomocí jednotlivých komponent určit jejich šíření do okolních kanálů.

1. Úvod

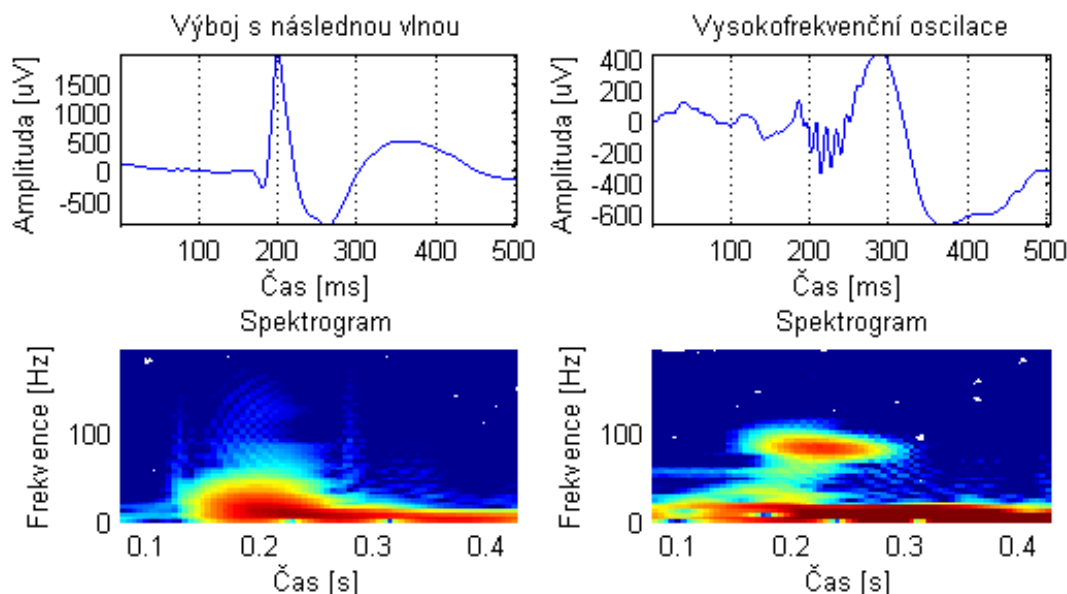
Při diagnóze epilepsie a její lokalizaci se kromě zkušeností lékaře a výsledků funkčně zobrazovacích metod využívá také několika markerů patrných v EEG (pro přesnou lokalizaci zejména intrakraniálního EEG). Do této kategorie patří výskyt typických grafoelementů. Jsou to interiktální epileptické výboje [1] nebo tzv. vysokofrekvenční oscilace [2]. Sledováním výskytu této aktivity v průběhu interiktálních fází [3] a jejich šíření ve fázi preiktální, lze přesněji lokalizovat epileptiformní tkáň, z které záchvaty vychází.

V dnešní době se při zpracování signálů předpokládá, že se jedná o lineární a stacionární signál. Nejčastěji pro tyto účely využívá Fourierova nebo vlnková transformace. K analýze nelineárních a nestacionárních procesů těmito metodami se využívá spektrogram, škálogram či Wigner-Villeho distribuce. Hilbert-Huangova transformace (HHT) [4] je metoda navržená pro analýzu nelineárních nestacionárních signálů. Jedná se o empirickou metodu, která nemá kompletní teoretický nebo matematický základ.

V této práci je využito pouze části HHT a to empirické modální dekompozice (EMD). Tato metoda rozkládá signál na jednotlivé komponenty podle jejich obsahu podobně jako banka filtrů. Ovšem při této dekompozici nejsou frekvenční pásma komponent předem známa a odvíjí se od průběhu zkoumaného signálu. Tímto způsobem jsou ze signálu postupně separovány složky od nejvyšších frekvencí až k nejnižším frekvenčním složkám. Předpokládá se, že v takto rozloženém signálu budou hledané grafoelementy separovny v jednotlivých komponentách a bude možné analyzovat nejen jejich výskyt, ale také jejich propagace mezi jednotlivými kanály.

2. Grafoelementy v epileptickém EEG

V epileptickém EEG rozeznáváme několik typů grafoelementů, které jsou generovány poškozenou tkání. Tato aktivita se pak dále šíří do ostatních mozkových center a tím může být vyvolán záchvat. Tato práce je zaměřena na dva nejpoužívanější a to na interiktální a iktální výboje a na vysokofrekvenční oscilace. Výskyt těchto grafoelementů je spojován s epileptogení tkání.



Obr. 1: Ukázka průběhu a frekvenčního rozložení grafoelementů

2.1. Epileptické výboje

Jsou to rychlé výboje naakumulované energie a obvykle mají trvání v řádu desítek milisekund. Rozlišuje se několik typů této aktivity. Jedná se buď o osamčené výboje nebo sérii několika výbojů za sebou. V průběhu epileptického záchvatu se pak může vyskytovat takzvaný komplex hrot-vlna, jenž je složen z výboje následovaného pomalou vlnou. Opět i zde se může vyskytovat několik výbojů jdoucích posobě následované vlnou. (viz obr). Pro shrnutí jsou tedy parametry tohoto grafoelementu následující:

1. **Trvání:** výboj 20 – 80 ms, vlna: 80 – 200 ms
2. **Frekvenční rozsah:** cca 10 – 50 Hz

2.2. Vysokofrekvenční oscilace

Tyto grafoelementy nemusí nutně prokazovat výskyt epilepsie. Bylo prokázáno, že vysokofrekvenční aktivita může být přirozená v hipokampu a ostatních parahipokampálních oblastech. Předpokládá se, že jsou důležité při přenosu informací a synchronizaci mozkové aktivity [5]. Převážně se jedná o aktivitu v pásmech 30-80Hz (rychlá gamma) a v pásmu 80-200Hz. [6] Vysokofrekvenční grafoelementy se dělí na dva typy v závislosti na jejich frekvenčním rozsahu.

V rozsahu 80-200Hz se nazývají Ripples a v rozsahu 200-500Hz je nazýváme Fast Ripples. O jejich přesných frekvenčních rozsazích se stále diskutuje a v různých člancích se udávají jiná pásma. [7, 8]. Nicméně je vidět, že frekvenční oblasti 30-80Hz a Ripples v určitých částech mozkové tkáně mohou korespondovat s přirozenou aktivitou mozku, při daných úkonech.

1. **Trvání:** 20 – 1000 ms
2. **Frekvenční rozsah:** 80 – 500 Hz

3. Empirická modální dekompozice

Jedná se o část procesu Hilbert-Huangovy transformace. Přesněji o rozložení signálu na několik komponent tzv. Intrinsic Mode Function (IMF). Tímto postupem jsou ze signálu postupně separovány jednotlivé složky počínaje těmi s nejvyšší frekvencí. Podrobný postup této dekompozice je následující (grafická ukázka viz. Obr. 2):

1. Nalézt maxima a minima v $x(t)$
2. Vytvořit obálky maxim a minim $X_{up}(t), X_{low}(t)$ pomocí kubické spline interpolace
3. Spočítat střední hodnotu obálek
$$m(t) = \frac{X_{up}(t) - X_{low}(t)}{2}$$
4. Odečíst střeni obálku od signálu
$$d_1(t) = x(t) - m(t)$$

$$d_i(t) = d_{i-1}(t) - m(t)$$
5. Porovnat výsledek s podmínkami
 - a) Počet lokálních extrémů je stejný, nebo se liší maximálně o 1, s počtem průchodů nulou
 - b) Porovnat koeficient s požadovanou hodnotou

$$SD_k = \frac{\sum_{i=0}^T |d_{k-1} - d_k(t)|^2}{\sum_{i=0}^T d_{k-1}^2(t)}$$

Pokud jsou obě podmínky splněny, tak:

$$c_i(t) = d(t)$$

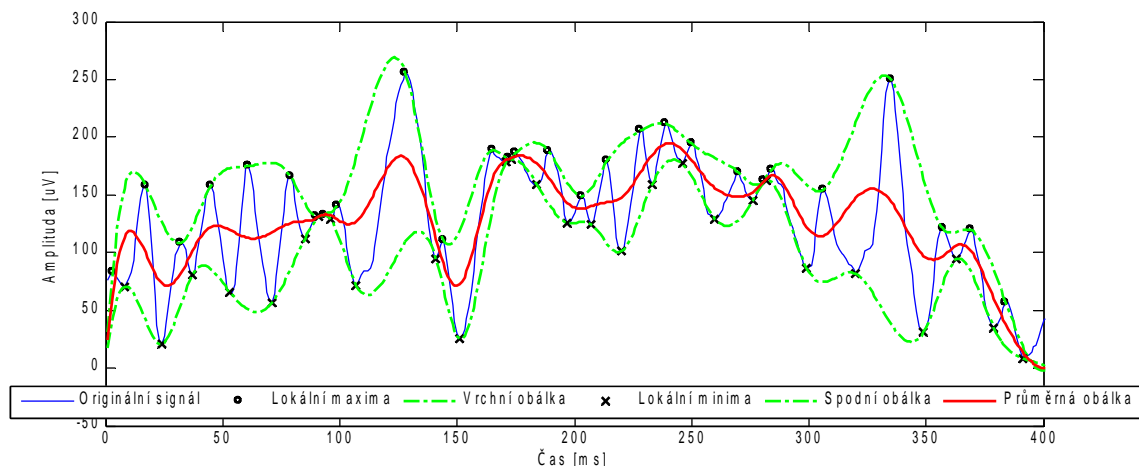
$$r_1(t) = x(t) - c_1(t)$$

$$r_i(t) = r_{i-1}(t) - c_i(t)$$

6. Opakovat krok 1. - 5. dokud není residuální funkce monotónní nebo má jen jedno maximum/minimum

Původní signál se dá složit součtem jednotlivých módů (IMF) a zbytkového residua.

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t)$$



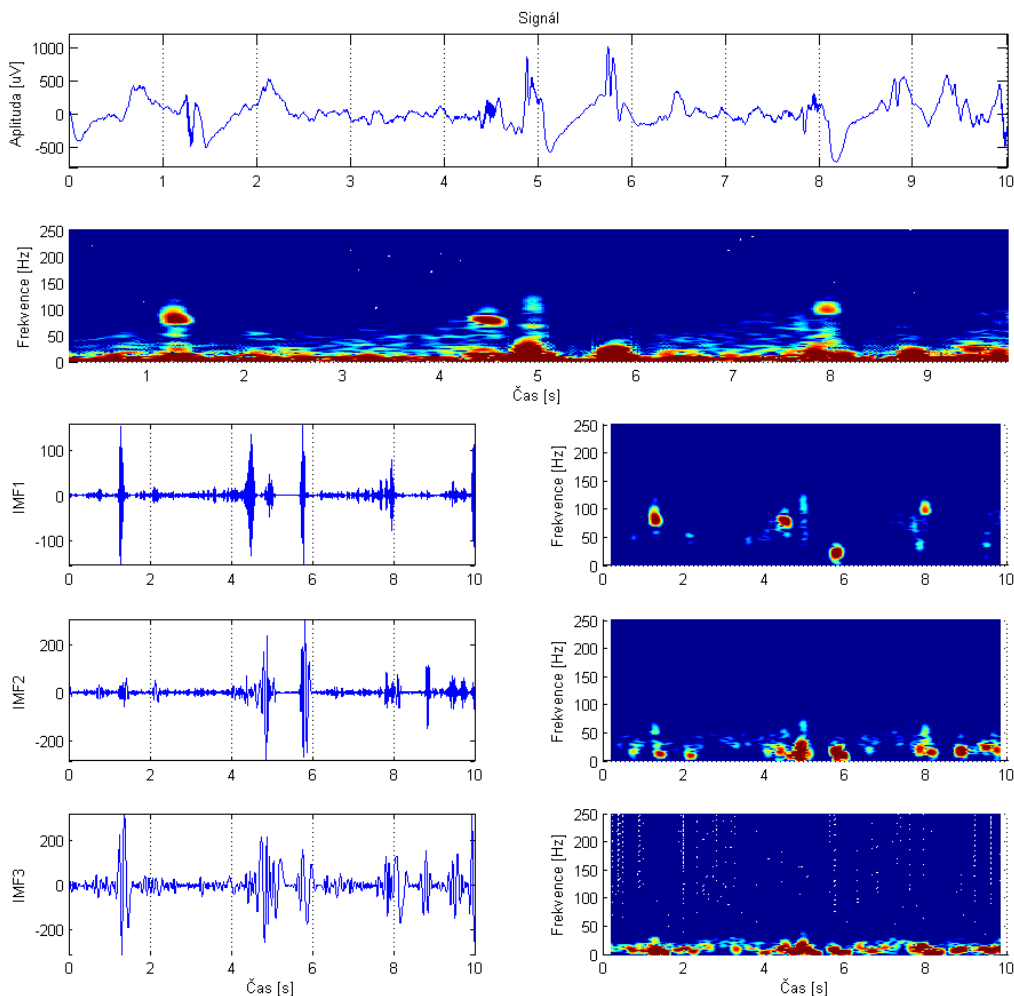
Obr. 2: Ukázka EMD postupu

4. Aplikace na EEG signál

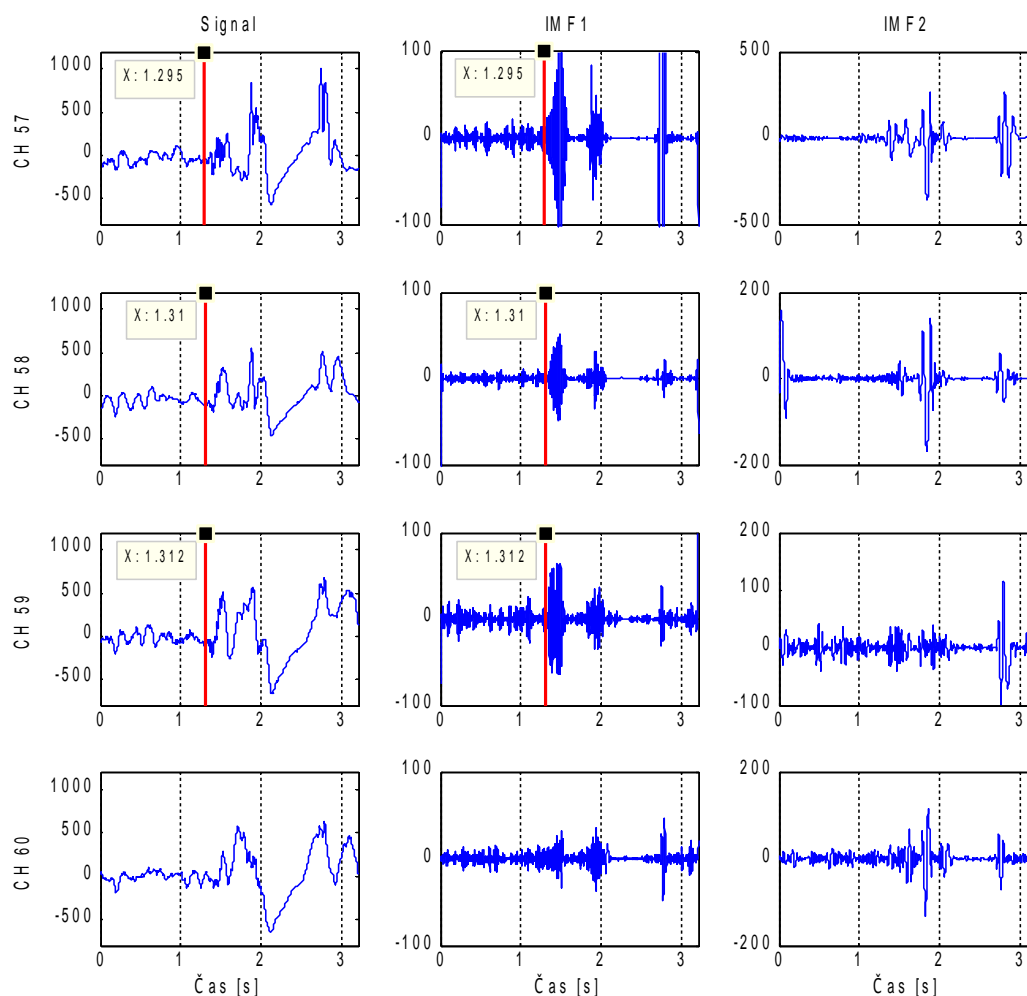
Pro účely testování byl zvolen interiktální signál EEG pacienta z dlouhodobé monitorace před chirurgickým zákrokem. Záznam byl pořízen intracerebrálními a subdurálními makroelektrodami o celkové délce 10-ti minut z celkem 97 elektrod při vzorkovací frekvenci 1000Hz. Z celého souboru dat byl následně vybrán záznam z 8 temporálních elektrod, kde se grafoelementy vyskytovali nejvíce. Na tyto data byla použita dekompozice na jednotlivé IMF složky. Takto upravený signál byl zkoumán, zda je nebo není možno separovat požadované grafoelementy a zda je pomocí této dekompozice odhalit jejich šíření mezi kanály.

4.1. Analýza grafoelementů

Podle předpokladů by se vysokofrekvenční aktivita měla objevit v prvních komponentách (IMF 1-2). Ve složkách následujících by se pak měly objevit výboje a jejich složky. Ukázka dekompozice a frekvencí obsažených v jednotlivých složkách je na Obr. 3. Zde je vidět, že pomocí první IMF se opravdu separovali vysokofrekvenční složky signálu a je následně snadné je identifikovat. Výraznější výboje (v 5 a 5,5 vteřině) jsou zase majoritně zastoupeny v následující složce. V poslení, třetí IMF jsou pak výraznější všechny nižší aktivity včetně složek z výbojů.



Obr. 3: Signál a jeho dekompozice v časové a frekvenční oblasti



Obr. 4: Ukázka propagace aktivity v jednotlivých IMFs

4.2. Analýza propagace

Předpokládá se, že mozková aktivita se postupně propaguje i do okolních oblastí. Počátek by měl ležet v hledané epileptogenní tkáni a nalezením tohoto centra by se její poloha měla zpřesnit. Pokud se tedy grafoelementy propagují do okolních kanálů a dají se separovat pomocí EMD, měla by se dát tato propagace odhalit v rámci jednotlivých IMF. Pro tyto účely byly vybrány 4 elektrody z temporální oblasti, kde se vyskytuje vysokofrekvenční aktivita.

Ukázka dekomponovaného signálu do dvou složek (IMF1 a IMF2) je znázorněna na Obr. 4. Zde je červenou čarou zvýrazněn začátek první vysokofrekvenční oscilace v jednotlivých průbězích první IMF a v odpovídajících časech také v původním signálu. Nalezené počátky vysokofrekvenčních oscilací nasnačují, že se nejspíše aktivita propaguje z kanálu 57 do sousedních kanálů 58 a 59. U těchto kanálů, ale již není zcela zřejmé kde nastala propagace dříve. Pro tyto účely je klíčové nalezení přesného počátku grafoelementu.

Propagace aktivity s nižší frekvencí není jasná a proto není zvýrazněna. Také v průběhu kanálu č. 60 (CH 60) není zaznamenán čas propagace, protože už zde není jasný začátek výraznější aktivity.

5. Závěr

Z jednotlivých analýz bylo zjištěno, že pomocí Hilbert-Huangovy transformace, repektive empirické modální dekompozice, lze separovat jednotlivé typy grafoelementů a usnadnit tak jejich detekci. Vysokofrekvenční oscilace se objevují výrazně v první složce (IMF1), interiktální výboje jsou pak zastoupeny v dalších komponentách, převážně v IMF2.

Detekce propagace aktivity v jednotlivých IMFs už nemá tak jasné výsledky, nicméně ani tento směr nelze zavrhnout. I tento směr nabízí vyhlídky do budoucna. Bude ovšem potřeba přesně detekovat počátky grafoelementů. Zatím se zdá, že nejvhodnější marker pro odhalení propagací mezi jednotlivými kanály budou vysokofrekvenční oscilace, jenž se daří separovat do první IMF a kde mají výrazný průběh.

Poděkování

Práce výzkumného týmu je podporována granty Ministerstva Zdravotnictví ČR IGA NT 13357-4/2012, IGA NT 11460-4/2010 a studentským grantem SGS13/138/OKH3/2T/13.

Reference

- [1] Stead, M.; Bower, M.; Brinkmann, B. H.; Lee, K.; Marsh, W. R.; Meyer, F. B.; Litt, B.; Van Gompel, J.; Worrell, G. A. Microseizures and the spatiotemporal scales of human partial epilepsy. *Brain*. 2010-08-27, vol. 133, issue 9, s. 2789-2797.
- [2] Zijlmans, M.; Jiruska, P.; Zelmann, R.; Leijten, F.S.S.; Jefferys, J.G.R.; Gotman, J. High-frequency oscillations as a new biomarker in epilepsy. *Annals of Neurology*. 2012, vol. 71, issue 2, s. 169-178 .
- [3] Andrade-Valenca, L.; Mari, F.; Jacobs, J.; Zijlmans, M.; Olivier, A.; Gotman, J.; Dubeau, F. Interictal high frequency oscillations (HFOs) in patients with focal epilepsy and normal MRI. *Clinical Neurophysiology*. 2012, vol. 123, issue 1, s. 100-105 .
- [4] Huang, N. E.; Shen, Z.; Long, S. R.; Wu, M. C.; Shih, H. H.; Zheng, Q.; Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995.
- [5] Engel, J. Jr.; Da Silva, F., L. High -frequency oscillations – Where we are and where we need to go., *Progress in Neurobiology* 98 (2012), s. 316-318.
- [6] Engel J. Jr.; Bragin, A.; Staba, R.; Mody, I. High-frequency oscillations: What is normal and what is not?, *Epilepsia*. 2009, roč.50, č.4, s 598-604 .
- [7] Staba, R.J.; Bragin, A. High-frequency oscillations and other electrophysiological biomarkers of epilepsy: underlying mechanisms. *Biomarkers in Medicine*. 2011, roč. 5, č. 5, s. 545-556.
- [8] Zijlmans, M.; Huiskamp, G.M.; Cremer, O.L.; Ferrier, C.H.; Van Huffelen, A.C.; Leijten, F.S.S. Epileptic high-frequency oscillations in intraoperative electrocorticography: The effect of propofol. *Epilepsia*. 2012, roč. 53, č. 10, s. 1799-1809.

Akustické modelovanie trifónov na bázi HMM

Michal Borský

České vysoké učení v Praze, Fakulta elektrotechnická, Katedra teorie obvodů
borskmic@fel.cvut.cz

Abstrakt:

Akustické modelovanie je jednou z hlavných úloh pri konštrukcii ASR systému, ktoré sa spolu s vhodnými príznakmi významne podieľa na úspešnosti rozpoznávača. Cieľom tejto práce je prezentovať výsledky z oblasti extrakcie robustných príznakov, viazania parametrov kontextovo závislých fonémov a vytvárania genderových modelov. Výsledky ukazujú, že voľbou vhodných metód odstránenia šumu z príznakov je možné znížiť chybovosť až o 50%. Natrénovanie genderových modelov prinieslo pri správnom použití zníženie asi o 1%. Voľbou správnych hodnôt konštánt pri zväzovaní pre dané množstvo tréningových dát, je možné dosiahnuť ďalšie zníženie rádovo v jednotkách percent.

1. Úvod

Automatické rozpoznávanie reči (ASR) sa v poslednom desaťročí stalo bežne používané v praktických aplikáciach ako napr. diktovacích systémoch, automatickom prepise záznamov, alebo vyhľadávania kľúčových slov. Medzi najrozšírenejšie postupy pri vytváraní akustických modelov (AM) patria skryté Markovovské Modely (HMM). Ich tréning je otázkou veľkého množstva dostupných dát a kombinácie vhodných postupov. I keď sa v dnešnej dobe môže zdať, že otázka správneho nastavenia a aplikácie jednotlivých algoritmov pri vytváraní AM je uzavrená, nie je tomu tak. V situáciach ako je napr. vysoká úroveň šumu na nahrávkach, alebo použitie ASR systému v úlohe rozpoznávania spon-tánnej reči je stále veľký priestor pre ďalšiu prácu.

2. Všeobecné nastavenie trenovania akustických modelov

Model ASR systému pozostáva z niekoľkých blokov, keď medzi najpodstatnejšie patria bloky dekodéru, akustického modelu a jazykového modelu. Akustický model je výsledkom tréningu na množine dát, z ktorých boli extrahované vhodné príznaky. V uskutočnených experimentoch boli používané príznaky prispôbené na daný problém, ale vždy bolo použité základné nastavenie, ktoré bolo jednoduché. Týmto bolo 12 mel-frekvenčných kepstrálnych koeficientov (MFCC) spolu s nulovým, doplnené o ich prvé a druhé diferencie v jednom prúde. To isté platilo aj pre AM, keď vzhľadom na dostupnú databázu boli AM natrénované v nasledujúcom nastavení: *43 monofónov + model pre ticho a krátku pauzu (sp), internal-word alebo cross-word trifóny, ľavo-pravý model bez možnosti preskokov, s výnimkou pre "sp"*.

Pre natrénovanie parametrov akustických modelov je potrebná vhodná tréningová databáza. V této práci boli AM tréningované na databázi SPEECON, ktorá obsahuje dáta od 589-tich rečníkov oboch pohlaví, nahraných štyrmi mikrofónmi (CS0-CS3). Väčšina AM vznikla z nahrávok z nízko šumového CS0 mikrofónu, avšak v prípadoch zúčastnenia možností zvýšenia robustnosti AM na šum sa pracovalo s vysoko šumovými kanálmi CS2 a

CS3. Trénovacia množina vo všetkých prípadoch obsahovala dáta z prostredí (OFFICE a ENTERTAINMENT), ktoré sa vyznačujú nízkou úrovňou šumu na pozadí. Celkovo sa jednalo o cca 55k signálov od 190-tich rečníkov v dĺžke asi 51 hodín. Ako testovacia množina bola zvolená sadu obsahujúca základné číslovky, v dĺžke asi 13.5 minúty, a množina povelov. Táto mala asi 15 minút a tvorilo ju 472 rôznych povelov používaných napr. pre GPS navigáciu alebo hlasové ovládanie prístrojov v domácnosti.

3. Varianty akustických modelov trifónov

Pri tvorbe AM pre praktickú aplikáciu je nutné zohľadniť rôzne faktory, napr. výslednú rýchlosť výpočtu pri rozpoznávaní, vplyv šumu i na úrovni modelovania, prispôbienie modelu konkrétnemu rečníkovi apod. V tejto práci sú analyzované a optimalizované akustické modely pre niekoľko vybraných špecifických situácií. V nasledujúcich častiach sú popísané výsledky experimentov spolu s krátkou všeobecnou diskusiou nad danou konkrétnou optimalizáciou.

3.1. Vplyv parametrov zväzovania stavov

V praktických systémoch je jedným z najdôležitejších parametrov rýchlosť rozpoznávania vyjadrená ako real-time (RT) faktor, na ktorý má hlavný vplyv zložitosť AM. Zložitosť AM narastá priamo úmerne s použitím kontextovo závislých fonémov (bifónov/trifónov) a počtom Gaussovských zmesí na stav. Jednou z úspešne používaných metód redukcie zložitosti je zväzovanie stavov AM na základe akustickej podobnosti, keď v súčasnosti sa preferuje zväzovania za pomoci tree-based algoritmu. Výsledný počet zviazaných stavov je v praxi kompromisom medzi veľkosťou AM a jeho schopnosťou ešte zachytiť malé rozdiely pri modelovaní rôznych alofónov. Toto často vedie na určenie optimálneho počtu Gaussoviek, pri ktorom je možné zaistiť fungovanie systému v reálnom čase. Výsledné číslo môže byť dosiahnuté dvoma rôznymi spôsobmi: *malý počet stavov po zviazaní a pridanie veľkého počtu zložiek*, alebo *veľký počet stavov po zviazaní a pridanie malého počtu zložiek*.

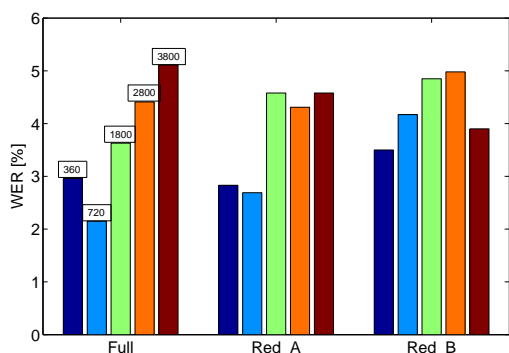
Rozdielne názory na použitie možno badať v odlišných prístupoch pri praktickom nasadení, keď napr. v práci [2] autori používajú monofóny s veľkým počtom zložiek, až 100, na stav. Opačný prístup je zase preferovaný v práci [3].

Oba tieto postupy sa môžu javiť ako ekvivalentné z hľadiska úspešnosti rozpoznávania a preto sa nasledujúci experiment zameriava na prípadné rozdiely v úspešnosti. Postup vytvorenia AM bol založený na natrénovaní multimodálnych HMM z cross-word trifónov, keď boli testované rôzne hodnoty zväzovacej kritériálnej funkcie. Počet zmesí bol následne pridaný tak, aby výsledný počet Gaussoviek bol zhruba rovnaký. Taktiež bola menená veľkosť trénovacej množiny, pretože jej veľkosť výrazne ovplyvňuje počet nájdených trifónov a množstva dát pre tréovanie. Špeciálna pozornosť bola venovaná výberu foneticky bohatých signálov (fsb) a počtu nájdených trifónov. Dovedna boli zvolené tri trénovacie množiny: **Plná** - fsb = 6498 a 15392 trifónov, **Red_A** - fsb = 6498 a 13451 trifónov a **Red_B** - fsb = 3249 a 12131 trifónov a päť prahov pre kritériálnu funkciu.

Pre všetky trénovacie množiny sa varianta väčšieho počtu stavov po zviazaní a pridanie menšieho počtu Gaussoviek javila ako lepšia. Tento prístup je však aplikovateľný

Meno	Plná		Red_A		Red_B	
	Poč. Stavov	Zmesí	Poč. Stavov	Zmesí	Poč. Stavov	Zmesí
Tri_360	6665	5	3583	7	2435	10
Tri_720	3893	7	1905	13	1303	20
Tri_1800	1926	15	923	27	630	40
Tri_2800	1410	20	665	40	465	55
Tri_3800	1138	22	555	46	376	65

Tabuľka 1: Nastavenia pre jednotlivé množiny



Obr. 1: Chyba rozpoznávania WER [%] pre asi 20k Gaussoviiek

iba v prípade, že sú si obe varianty rovnocenné z pohľadu praktickej aplikácie. Reálne systémy preto občas kombinujú obe varianty.

3.2. AM pre rozpoznávanie v zašumenom prostredí

Aplikácia väčšiny dnešných ASR systémov predpokladá použitie v podmienkach, ktoré sú na to vhodné. To znamená nízku úroveň šumu na pozadí a minimálne zkrasenie rečového signálu kanálom. Toto však nieje možné vždy zaistiť, vid' napr. GPS navigácia v idúcom aute, alebo hlasové ovládanie prístrojov. Pre oba prípady je typické, že nahrávací mikrofón je vzdialený od rečníka, keď napr. nasadenie ASR systému do inteligentného domu predpokladá zabudovanie mikrofónov do stien resp. stropu. Výsledkom môže byť značné zkrasenie rečového signálu aditívnym a konvolučným šumom. Bežne používané parametrizačné techniky sú na tieto podmienky dosť citlivé [4] a vykazujú zníženú úspešnosť rozpoznávania.

Tento experiment je preto zameraný na zvýšenie robustnosti extrakcie príznakov. Boli spracované dve metódy odstraňovania aditívneho a konvolučného šumu. Algoritmus rozšíreného spektrálneho odčítania (SS) bol použitý na potlačenie aditívneho šumu a na potlačenie konvolučného šumu bolo použité odčítanie priemerného kepra (CMS). Ich prínos bol skúmaný jak jednotlivo, tak v prípade aplikácie oboch zároveň.

V tejto úlohe pozostávala tréningová množina zo signálov z kanálov CS0, CS2 a CS3. Počas nahrávania bol pre tieto signály pomocou nahrávacieho zariadenia odhadnutý parameter SNR. Kanál CS0 obsahoval nízku úroveň šumu s hodnotou $SNR = 27.1 \pm 4.6$ dB, pre CS2 bola hodnota $SNR = 12.5 \pm 4.4$ dB a kanál CS3 mal $SNR = 6.2 \pm 4$ dB.

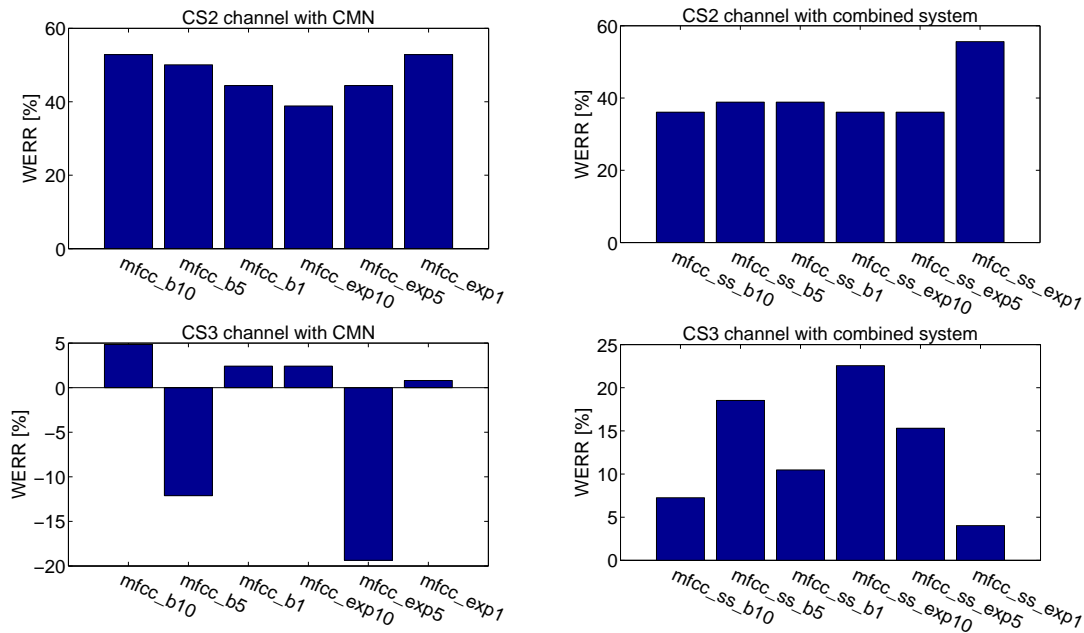
Jednotlivé odšumovacie metódy su všeobecne známe, a preto si ich čitateľ môže naštudovať v napr. [5] a [6]. Pre SS bolo použité jednotné nastavenie: *rozšírené spektrálne odčítanie, odkad spektra šumu v každom segmente zvlášť, integračná konštanta $p = 0.95$, realizované v amplitúdovej doméne, pred aplikovaním banky filtrov.*

V prípade CMN boli implementované v nástroje Ctucopy dve rozdielne prístupy, a to exponenciálne zabúdie (exp) a blokové priemerovanie (b). V oboch prípadoch bola testovaná rozličná dĺžka priemerovania a odhad sa prenášal naprieč segmentami. Zhrnutie testovaných nastavení pre extrakciu príznakov je v Tab.2.

Param.	SS	T [s]		
<i>mfcc</i>	nie	-		
<i>mfcc_ss</i>	áno	-		
<i>mfcc_b/mfcc_exp</i>	nie	1	5	10
<i>mfcc_ss_b/mfcc_ss_cms</i>	áno	1	5	10

Tabuľka 2: Zoznam použitých parametrizácií

Akustické modely boli natréňované postupom, ktorý bol opísaný v predchádzajúcej kapitole. Dosiahnuté výsledky na Obr.2 vyjadrujú zlepšenie ako redukciu chybovosti (WERR). Pre kanál CS2 bolo dosiahnuté značné zlepšenie použitím čisto algoritmu CMN, keď sa chyba znížila v priemere o 50%. Pre kanál CS3 bolo zlepšenie dosiahnuté konzistentne len pri aplikácii oboch algoritmov. Časová konštanta $T = 5s$ sa ukázala pri použití len CMN metódy nevhodná. Pre zvyšné konštanty T bolo dosiahnuté mierné zlepšenie. Zhoršenie bolo naopak pozorované, pokiaľ boli algoritmy aplikované na kanál CS0.



Obr. 2: WERR [%] pre rôzne parametrizácie

3.3. Genderové AM

V práci [7] autor porovnáva AM natrénované na celej množine o veľkosti zhruba 460k slov s modelmi vytvorenými špeciálne z mužských, resp. ženských nahrávok. Autor v práci skúma, či genderové modely natrénované na polovičnom množstve dát dokážu konkurovať všeobecným modelom na plnej množine. V závere konštatuje mierne zlepšenie (cca. 1%) pre genderové modely, pokiaľ sú aplikované na cieľovú skupinu. Ako hlavný nedostatok uvádza výrazne zhoršenie, až 8%, pokiaľ sa tieto použijú na nesprávneho rečníka. Výsledné použitie genderových modelov teda podmieňuje spoľahlivým detektorom. Akustická rozdielnosť nahrávok od mužských a ženských rečníkov vedie na rozdielne parametre AM. V tejto časti boli vytvorené genderových AM, keď tréningová množina SPEECONu bola rozdelená podľa apriornej informácie o pohlaví. Zastúpenie jednotlivých pohlaví v množinách je zhrnuté v Tab.3, keď je zrejme výchýlenie v prospech ženských rečníkov pre tréningovú množinu. Natrénované boli modely pre ženy (F) a pre mužov (M) a výsledky boli porovnané s modelmi na pohlaví nezávislými (U). Testovacia množina obsahovala 192 signálov s číslovkami, kde prevládali signály pre mužov.

	Unisex	Muži	Ženy
Signály	54240	25116	29124
slov	743	446	297

Tabuľka 3: Počet signálov a slov v množinách

Proces tréningovania prebehol bežným postupom od monofónov, ktoré boli rozšírené na internal-word trifóny, až po multimodálne HMM. Zväzovanie prebehlo pomocou tree-based algoritmu s parametrom TB 565, čo vyšlo na presne 4000 stavov po zviazaní. Modely boli ešte 2-krát pretrénované na celej množine a potom 3-krát na genderovej. Nasledne sa zvyšoval počet zložiek na stav na 8 postupne po jednej, a trikrát sa pretrénovalo po každom pridaní. Výsledky Acc [%] pre jednotlivé počty zložiek sú zhrnuté v tabuľke nižšie.

		Modely								
		4 mixtures			6 mixtures			8 Mixtures		
		U	M	F	U	M	F	U	M	F
Test	U	97.98	97.53	98.65	97.98	98.21	97.64	98.52	97.53	98.65
	M	90.98	97.53	82.19	92.33	97.96	84.18	92.33	97.98	83.84
	F	93.54	89.91	98.99	93.81	90.36	98.99	94.08	90.81	98.99

Tabuľka 4: ACC [%] pre jednotlivé modely a testovacie množiny

Použitie genderových modelov vykazuje mierne zlepšenie, keď hlavne pre ženské AM bolo dosiahnuto zvýšenie ACC až o 1%. Na druhej strane bolo pozorované výrazné zhoršenie ACC pri krížovom použití AM a testovacej množiny (F-M, M-F). Z tohoto experimentu je zrejme, že doteraz používaná tréningová množina nieje úplne vhodne zvolená vzhľadom na zastúpenie pohlaví, keď obsahuje väčšie množstvo signálov pre ženských rečníkov. Na druhej strane testovacia množina obsahuje zase väčšie množstvo signálov od mužov. Z tohoto dôvodu sa ako cieľ v najbližšej dobe stanovila revízia množín a definícia ich presného zloženia.

4. Záver

Cieľom tejto práce bolo zhrnúť dosiahnuté výsledky na poli akustického modelovania tri-fónov. Otázka počtu stavov po zviazaní versus počet pridaných Gaussovských zmesí je aj napriek doterajším záverom stále otvorená. Naše experimenty ukazujú, že pri dostatočnom množstve dát je vhodnejšie zväzovať menej reštriktívne a pridať menej zmesí. Ako protiváha však stojí fakt, že v niektorých praktických nasadeniach systému je výhodnejšia skôr druhá varianta, a preto táto téma ostáva ešte otvorená. Aplikácia algoritmov redukcie šumu dokáže významne zvýšiť úspešnosť ASR. Použitie genderových modelov vykazuje mierne zlepšenie úspešnosti, aj keď sú natrénované len na polovičnom množstve dát ako modely nezávislé. Naopak ich nesprávne použitie má za následok výrazné zhoršenie rozpoznávania. Ich praktické nasadenie je preto významne podmienené kvalitným detektorom pohlaví.

Ako ciele vytýčené v najbližšej dobe boli stanové práce na zoptimalizovaní zloženia databáz a ich rozšírenie o databázu TEMIC. V ďalších krokoch je to potom vyčistenie šum potlačujúcich algoritmov. Úplne novou, doteraz nezpracovanou kapitolou, sú možnosti nasadenia diskriminatívnych algoritmov tréningu AM a použitie dodatočných rečových príznakov. Všetky tieto časti budú spojené do výskumu v oblasti rozpoznávania reči v zhoršených podmienkach a z komprimovaných nahrávok.

PodĎakovanie

Výzkum v tejto práci bol podporený interným grantom ČVUT SGS12/143/OHK3/2T/13 “Algoritmy a hardvérové realizácie číslicového zpracovania signálov”.

Literatúra

- [1] S. Young and et al., *The HTK Book, Version 3.4.1*. Cambridge, 2009.
- [2] J. Nouza, D. Nejedlová, J. Zdánský, and J. Kolorenc, “Very large vocabulary speech recognition system for automatic transcription of Czech broadcast programs,” in *INTERSPEECH*, 2004.
- [3] P. Ircing, J. Psutka, and V. Radová, “Automatic transcription of audio archives for spoken document retrieval,” (Anaheim), pp. 448–452, ACTA Press, 2006.
- [4] J. Rajnoha and P. Pollák, “ASR systems in noisy environment: Analysis and solutions for increasing noise robustness,” *Radioengineering*, vol. 20, no. 1, pp. 74–84, 2011.
- [5] P. Sovka, P. Pollak, and J. Kybic, “Extended spectral subtraction,” in *EUSIPCO’96*, (Trieste), September 1996.
- [6] J. Droppo and A. Acero, “Environmental robustness,” in *Springer Handbook of Speech Processing*, pp. 653–680, Springer, 2008.
- [7] P. A. Olsen and S. Dharanipragada, “An efficient integrated gender detection scheme and time mediated averaging of gender dependent acoustic models.”

Zlepšení přesnosti fonetické segmentace na bázi HMM s akustickými modely trifónů

Petr Mizera

České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů
mizerpet@fel.cvut.cz

Abstrakt: Článek prezentuje výsledky analýzy přesnosti fonetické segmentace, která byla provedena na základě standardně používané metody HMM nuceného zarovnání (forced alignment). Nově byla analyzována segmentace na bázi akustických modelů trifónů. Dosažené přesnosti určení hranic hlásek jsou prezentovány pro různé složitosti akustických modelů. Dále článek shrnuje první kroky, které byly řešeny v rámci dizertační práce na téma "Rozpoznávání řeči na bázi artikulačních příznaků".

1. Úvod

Systémy automatického rozpoznávání řeči (ASR - Automatic Speech Recognition) nalézají v současné době uplatnění v mnoha aplikacích reálného života např. se jedná o systémy automatické titulkování, diktovací systémy (v soudnictví, či pro přepis lékařských zpráv) [1]. V poslední době se jedná o ASR zabudované ve vestavěných systémech např. v automobilech, GPS navigacích, inteligentních domech, či přímo v mobilních telefonech [2].

V současné době je intenzivně studována problematika rozpoznávání spontánní řeči. Rozpoznávání spontánní řeči přináší řadu problémů ve srovnání s problematikou čtené řeči. Pro případ rozpoznávání českého jazyka se jedná zejména o volnou strukturu českých vět, dále častý výskyt nespisovných, či hovorových slov, nebo neřečových událostí (nádech, váhání, zaseknutí). Častý výskyt nespisovný i hovorových slov je nutné řešit pomocí výslovnostního slovníku prostřednictvím vytváření výslovnostních variant daného slova. Při vytváření výslovnostních variant je také vhodné uvažovat o neutralizaci v češtině, kterou představuje asimilace a redukce.

Na problematiku rozpoznávání spontánní řeči je zaměřeno téma řešené dizertační práce. První kroky práce byly směřovány k vytváření výslovnostního slovníku pro český korpus spontánní řeči Czech NCCCz, který vznikl v rámci spolupráce s univerzitou Radboud University v Nijmegenu. Pro vytváření výslovnostního slovníku byl použit interní nástroj LexFix. Tento nástroj byl rozšířen o několik funkcí, které zjednodušují uživateli manuální práci s rozsáhlým korpusem. Výstupem tohoto období byl manuálně překontrolovaný výslovnostní slovník (cca. 30 tisíc slov), jenž bude při rozpoznávání spontánní řeči rozšiřovat standardní výslovnostní slovník.

Následující práce byla již směřována k rozpoznávání řeči. V počátku byla řešena vhodná volba příznaků pro parametrizaci řečového signálu, při které bylo provedeno seznámení se s nástroji HTK, CtuCopy a tvorbou jednoduchého rozpoznávače číslovek. V poslední době byly aktivity zaměřeny na analýzu přesnosti fonetické segmentace čtené řeči a v následné době bude analyzována přesnost pro spontánní řeč. Fonetická segmentace

spontánní řeči bude dále studována s ohledem na vytvoření pravidel pro redukcii výslovnosti při tvorbě výslovnostních slovníků. Výsledky analýzy přesnosti fonetické segmentaci čtené řeči budou prezentovány v následujících částech příspěvku.

2. Řečové příznaky

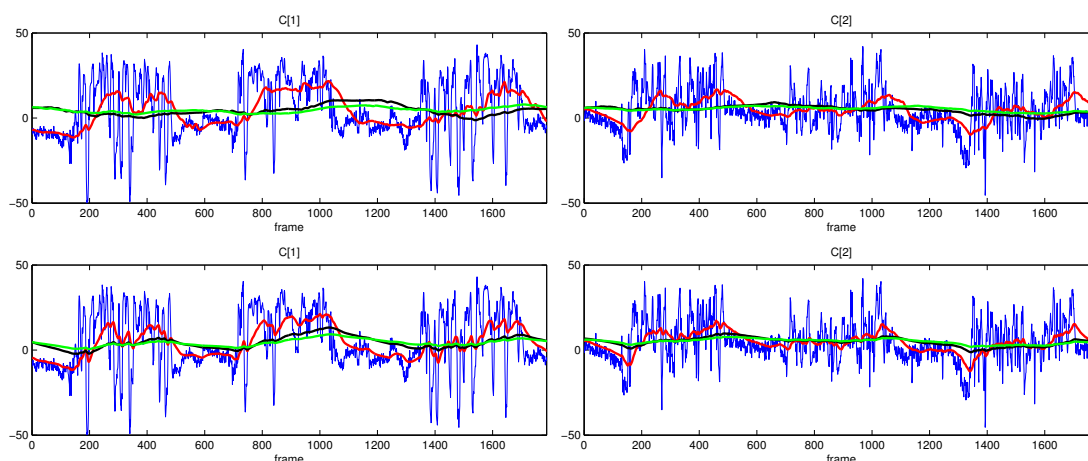
Vhodná volba příznaků pro rozpoznávání umožňuje dosáhnout vyšších úspěšností. V současné době jsou často používanými příznaky Melovské keprální koeficienty (MFCC), které jsou odhadovány v krátkých časových intervalech řečového signálu. Tyto koeficienty byly navrženy s ohledem na nelineární vnímání frekvence lidským sluchem.

Při analýze nastavení parametrů parametrizace byly využity výsledky již provedených analýz na našem pracovišti. Proto bylo pro parametrizaci zvoleno Hammingovo okno o velikost 25ms s posuvem 10ms. Melovská banka trojúhelníkových filtrů byla navržena do 22 pásem s 50 % překryvem jednotlivých pásem. Vhodné se ukázalo pracovat s 12 keprálními koeficienty plus s koeficientem C0.

Analýza byla zaměřena na nesoulad podmínek při vytváření akustických modelů pro pásmo 0-4000kHz bez převzorkování dat z různých pásem, kde nesoulad při vytváření akustických modelů byl způsoben frekvenční charakteristikou pre-empázového filtru pro různé kmitočty. Pre-empázový filtr provádí zvýraznění vyšších kmitočtů, které jsou potlačeny vlivem produkce lidské řeči. Více detailů uvedeno v [4].

Při kompenzaci vlivu pre-empáze byla také analyzována metoda odečítání průměrného kepru (CMS - Cepstral Mean Subtraction), jež je standardně používána pro kompenzaci konvolučního zkreslení (zkreslení přenosového kanálu, různé dozvuky, či ozvěny). Obecně metoda potlačuje jakékoliv konvoluční zkreslení v akustickém signálu. Pro náš případ se jednalo o kompenzaci zkreslení, které bylo způsobeno pre-empázovým filtrem.

Metoda CMS je implementována v nástroji HCOPY pod parametrem Z. Průměrné keprum, které je následně odečítáno od každého vektoru příznaků je v nástroji vypočteno v rámci celého signálu, proto je možné metodu využívat pouze při off-line zpracování. Hlavní nevýhodou této implementace je především odhad průměrného kepru v každém signálu z různého počtu vzorků.



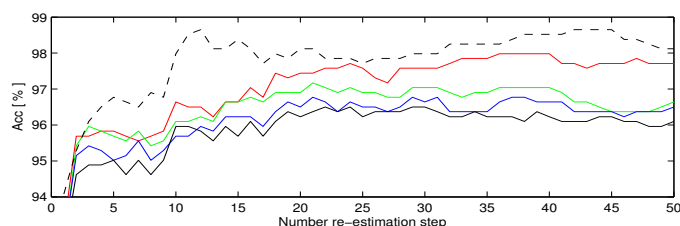
Obrázek 1: CMS s odhadem průměrného kepru EA/MA přes časové intervaly 1 s (červená), 5 s (černá), 10 s (zelená)

Proto pro naše další výzkumné aktivity bylo provedeno rozšíření interního nástroje

CtuCopy o metodu CMS. Metoda CMS byla do nástroje implementována ve dvou variantách odhadu průměrného kepru. V první variantě je průměrné keprum odhadováno na základě klouzavého průměru MA (Moving Average), druhá varianta odhadu je realizována na základě rozšíření MA o exponenciální zapomínání známá jako EA (Exponential Averaging). Průměrné keprum je tedy odhadováno z dlouhodobého časového intervalu, který je možný specifikovat v konfiguračním souboru pod parametry `fea_Z_exp` a `fea_Z_block`.

Obr. č. 1. zobrazuje průběhy odhadnutých hodnot průměrného kepru na základě výše uvedených metod přes různé dlouhodobé časové intervaly.

Analýza vlivu nesouladu způsobená pre-emfází byla vyhodnocena na základě úspěšnosti rozpoznávací českých číslovek. Rozpoznávač byl realizován pomocí nástrojů HTK. Byly použity standardní trifónové akustické modely s třemi emitujícími stavy, které byly trénovány do 12 gausovských komponent. Dosažené přesnosti jsou prezentovány na následujícím obrázku.



Obrázek 2: Přesnost pro varianty: *8k_p97_8k_p97_22b* (červená), *8k_p97_16k_p97_22b* (zelená), *8k_p97_16k_p99_22b* (černá), *8k_p97_16k_p98_22b* (modrá), *Z_8k_p97_Z_16k_p99_22b* (přerušovaná černá)

Z obr. č. 2. je vidět, že nejlepším řešením kompenzace vlivu pre-emfáze při vytváření akustických modelů pro různé vzorkovací frekvence bez nutnosti převzorkování dat je dosaženo pomocí metody CMS.

3. Fonetická segmentace

Následující část práce byla zaměřena na analýzu přesnosti fonetické segmentace, která nalézá uplatnění v mnoha aplikacích např. fonetickém výzkumu, při automatickém segmentování velkých databází, rozpoznávání či konkatenační syntéze řeči. V následující části příspěvku budou uvedeny výsledky analýzy přesnosti fonetické segmentace založené na různých akustických modelech AM.

Fonetická segmentace je realizována na základně metody HMM nuceného zarovnání hlásek (HMM-forced alignment). Přesnost této metody ovlivňují velikosti posuvu segmentačního okna při parametrizaci řečového signálu a dále kvalita akustických modelů. V následující části budou prezentovány výsledky dosažené přesnosti fonetické segmentace založené na HMM pro různé varianty akustických modelů.

Prvním krokem při realizaci algoritmu HMM nuceného zarovnání hlásek je volba řečových příznaků. Tato problematika byla již popsána v předešlé části, a proto zde bude jen přehledně shrnuto použité nastavení parametrizace v následujících bodech:

- koeficient pre-emfáze 0.97,
- Hammingovo okno s délkou 25 ms,
- posuv okna 10 ms,
- Cepstral mean normalization (CMN), průměr počítaný přes celý signál,

- banka trojúhelníkových filtrů s 22 pásy, 50 % překryvem jednotlivých pásem pro 8 kHz a 30 překrývajícími frekvenčními pásy pro 16 kHz,
- 12 MFCCs s c_0 , dynamické a akcelerační koeficienty.

3.1. Akustické modely HMM

V této práci je porovnávána přesnost fonetické segmentace založené na akustických modelech monofónů (kontextově nezávislých fonémů) a trifónů (kontextově závislých fonémů). Pro případ trifónů budou výsledné hranice fonémů lokalizovány pro střed fonému. Trifónové akustické modely přinášejí především výhodu při modelování kontextu mezi jednotlivými fonémy. Oproti tomu je nutné řešit pomocí různých svazovacích technik jejich velký počet, který je při uvažování všech variant trifónů roven třetí mocnině počtu monofónů.

V následujících bodech budou shrnuty vlastnosti vytvořených akustických modelů monofónů a trifónů:

Monofóny:

- 43 monofónů, model pro ticha a krátké pauzy (tee-model),
- levo-pravé HMM s 3 emitujícími stavy bez přeskočení,
- okolo 12–20 gausovských směrů na každý emitující stav,
- 1 nezávislý proud pro statické, dynamické a akcelerační koeficienty.

Trifóny:

- 9136 variant trifónů klonovaných z monofónů,
- levo-pravé HMM s 3 emitujícími stavy bez přeskočení,
- okolo 3–12 gausovských směrů na každý emitující stav,
- 1 nezávislý proud pro statické, dynamické a akcelerační koeficienty.

Trénování HMMs bylo realizováno pomocí nástrojů HTK. Trénovací množina obsahovala řečové signály z databáze Czech SPEECON, které byly zaznamenány v prostředí kanceláře s nízkou úrovní šumu. Celkem tyto signály představovaly 51 hodin záznamu. Trénování HMMs bylo použito Baum-Welchova algoritmu. Pro analýzu fonetické segmentace byly následně vybrány AM, které dosáhly nejvyšší úspěšnosti v jednoduché úloze rozpoznávání českých číslovek.

3.2. Vyhodnocení

Přesnost fonetické segmentace byla vyhodnocena na základě uvedených kritérií: *Shift of the Phone Beginning (SPB)*, *Shift of the Phone End (SPE)*, a *Change of the Phone Length (CPL)* [7] definovaných jako

$$SPB_{ph}[i] = beg_{ph}[i] - beg_{ph,ref}[i], \quad (1)$$

$$SPE_{ph}[i] = end_{ph}[i] - end_{ph,ref}[i], \quad (2)$$

$$CPL_{ph}[i] = end_{ph}[i] - beg_{ph,ref}[i] - end_{ph,ref}[i] + beg_{ph,ref}[i] \quad (3)$$

kde $beg_{ph}[i]$, $end_{ph}[i]$, $beg_{ph,ref}[i]$ a $end_{ph,ref}[i]$ jsou automatické a referenční hranice.

Globální vyhodnocení přesnosti určení hranic fonémů bylo realizováno pomocí střední hodnoty a směrodatné odchylky napočítané přes všechny fonémy. Detailnější pohled na dosaženou přesnost jednotlivých fonémů je z důvodů nízkých četností některých fonémů prezentováno prostřednictvím fonémových skupin, které jsou uvedeny v následující tabulce i s hodnotami četnosti. Jednotlivé fonémové skupiny byly vytvořeny na základě prací [3, 6].

skupina	fonémy	četnost
VH	<i>i, i:, u, u:</i>	143
VNH	<i>a, a:, e, e:, o, o:, o_u, a_u, e_u</i>	297
FAF	<i>f, v, s, z, S, Z, P\, Q\, x, h\, t_s, t_S, d_Z, d_z</i>	195
PLO	<i>p, b, t, d, c, J\, k, g</i>	201
NAS	<i>m, F, n, J, N</i>	92
APP	<i>r, l, j</i>	125

Tabulka 1: fonémové skupiny

3.3. Testovací množina dat

Pro testování přesnosti byla vytvořena množina signálů z české databáze SPEECON. Testovací množina byla disjunktní k množině trénovací a obsažené signály pocházely z prostředí kanceláře. Celkem testovací množinu tvořilo 32 signálů s celkovou dobou 3 minut.

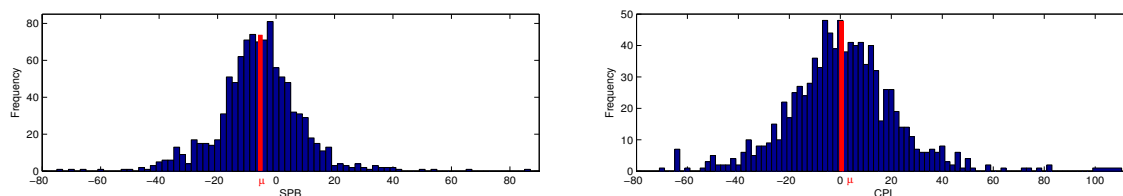
3.4. Výsledky

Nakonec byly analyzovány čtyři varianty akustických modelů, které se odlišovaly šířkou kmitočtového pásma a aplikací metody CMS. Následující tabulka prezentuje globální hodnoty jednotlivých kritérií počítaných přes všechny skupiny fonémů. Hodnoty pro jednotlivé skupiny fonémů je možné nalézt v práci [5].

	<i>8k_2510_22</i>				<i>16k_2510_30</i>			
	<i>mfcc_0_d_a</i>		<i>mfcc_0_d_a_z</i>		<i>mfcc_0_d_a</i>		<i>mfcc_0_d_a_z</i>	
	<i>mono12</i>	<i>tri4</i>	<i>mono15</i>	<i>tri3</i>	<i>mono19</i>	<i>tri12</i>	<i>mono20</i>	<i>tri5</i>
<i>SPB</i> [ms]	-10.2±15.7	-5.9±14.6	-10.9±16.3	-9.9±16.4	-10.3±15.6	-6.8±15.8	-10.9±17.2	-9.6±15.6
<i>CPL</i> [ms]	0.4±23.1	0.2±22.8	0.1±22.3	0.6±23.5	0.2±22.8	0.6±22.8	0.2±23.9	0.6±22.0

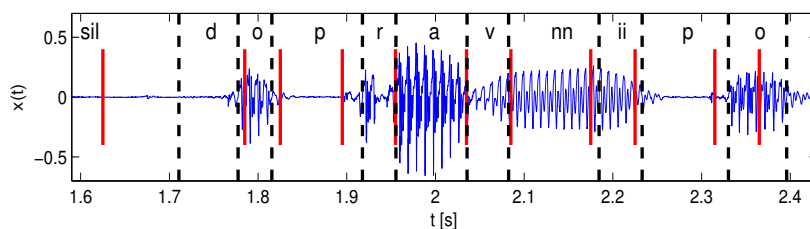
Tabulka 2: Výsledky globálních kritérií [ms]

Hodnoty kritérií SPB, CPL prezentované v tabulce dosahují pro variantu AM trifónů systematicky nižších hodnot střední hodnoty i směrodatné odchylky v porovnání s AM monofónů. Vliv metody CMS, která byla při této analýze počítána přes celou promluvu nepřinesla výrazné zlepšení při lokalizaci hranic fonémů. Podrobný náhled na hodnoty kritérií jsou zobrazeny pomocí histogramů na následujících obrázcích.



Obrázek 3: Histogramy kritérií SPB a CPL pro variantu *mfcc_0_d_a_8k_2510_22*

Nejlepší přesnosti bylo dosaženo pro variantu AM trifónů s čtyřmi gausovskými směsí bez CMS pro frekvenční pásmo 0-4kHz. Nejvyšších nepřesností určení hranic fonému dochází typicky na okraji slov. Tato situace je zachycena na obr. č. 4., z něhož je patrné, že určení hranic fonémů uprostřed slov probíhá s velkou přesností oproti určování hranic na okrajích.



Obrázek 4: Ilustrativní srovnání automatické a manuální fonetické segmentace (černá)

4. Závěr

V této práci byly realizovány první dílčí úlohy nezbytné pro rozpoznávání spontánní řeči. Z provedené analýzy fonetické segmentace na bázi HMM pro různé složitosti AM se ukázalo jako vhodné použít AM kontextově závislých fonémů. Pro nejlepší variantu AM bylo dosaženo přesnosti určení hranic 5ms pro určení začátku, či konce fonému dosaženou s poměrně malou standardní odchylkou 10ms. Při analýze se ukázalo, že metoda CMS implantovaná v nástroji HCOPY nepřinesla výrazné zlepšení přesnosti fonetické segmentace.

V následujících obdobích budou zaměřeny aktivity na analyzování přínosu příznaků na bázi časových trajektorií (TempoRAI Patterns - TRAP) v úloze fonetické segmentace. Budou provedeny první pokusy s fonetickou segmentací spontánní řeči, na jejichž základě bude následně snahou stanovit nová pravidla pro redukci výslovnosti.

Poděkování

Tato práce vznikla za podpory vnitřního grantu ČVUT SGS12/143/OHK3/2T/13 Algoritmy a hardwarové realizace číslicového zpracování signálů.

Reference

- [1] F. Jurcicek. *Speech Recognition for Live TV*. In IEEE Signal Processing Society - SLTC Newsletter, April 2009
- [2] L Heungkyu, K. Hanseok. *Intelligent Speech Interactive Agent on a Car Navigation Environment Using Embedded ASR and TTS*. In ICART, December 2004
- [3] P. Machac, R. Skarnitzl, *Fonetická segmentace hlásek*. Praha: Nakladatelství Epocha, 2009. (In Czech language, translated title: Phonetic segmentation of phones)
- [4] P. Mizera, *Mismatch Effect in Preemphasis Application within Speech Recognition Systems*. In CVUT POSTER 2013
- [5] P. Mizera, P. Pollak, *Accuracy of HMM-Based Phonetic Segmentation Using Monophone or Triphone Acoustic Model*. In APPEL 2013, Plzeň
- [6] P. Pollak, J. Volin, and R. Skarnitzl, *Phone Segmentation Tool with Integrated Pronunciation Lexicon and Czech Phonetically Labelled Reference Database*. In 6th International Conference on Language Resources and Evaluation. Marrakech (Morocco), 2008, vol. 1, p. 1-5.
- [7] P. Pollak, J. Volin, R. Skarnitzl, *Influence of HMM's Parameters on the Accuracy of Phone Segmentation - Evaluation Baseline*. In Proceedings of the 16th Conference Joined with the 15th Czech-German Workshop "Speech Processing". Prague, 2005, vol. 1, p. 302-309.

Prozodie a modelování přízvukových taktů

Jan BARTOŠEK¹, Václav HANŽL¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

bartoj11@fel.cvut.cz

Abstrakt: Studie modelování přízvukových taktů pro český jazyk zčásti navazuje a rozšiřuje příspěvek z ročníku loňského. V aktuálním příspěvku je opět nejprve představena prozodie jako suprasegmentální úroveň řečové informace (zdůrazněna je zejména souběžnost "nízkofrekvenční" modulace intonace sděleného projevu pro určení modality věty s "vysokofrekvenční" modulací přízvukových taktů) a dále je zevrubně probrána možnost detekce přízvukových taktů jen za pomoci informace o intonačním průběhu na bázi skrytých Markovových modelů (HMM). Detekce přízvukových taktů velmi úzce souvisí s umístěním hranic jednotlivých slov v promluvě a znalost takové informace lze v systému na rozpoznávání řeči často výhodně upotřebit. Ačkoliv v článku operujeme nad českými daty, užití principy jsou obecně aplikovatelné i na další jazyky.

Přízvukový takt neboli stopa (anglicky "foot") v češtině tvoří skutečnou jednotku řeči a je tak i vnímána posluchačem. Její strojové určení z promluvy je však obtížné zejména díky variabilitě s jakou jsou v češtině přízvukové takty realizovány (jak zvýšením, tak snížením intonace i dynamiky první slabiky v taktu).

Ve studii vycházíme z poznatku, že významný pokles F_0 uprostřed zkoumané množiny slabik pravděpodobně nepředstavuje realizaci intonace v rámci přízvukového taktu a že získaná posloupnost F_0 ve středech slabik bude tomuto trendu odpovídat. Dále vycházíme z předpokladu, že známe časové okamžiky slabičných středů (nejčastěji samohláska) a příslušnou hodnotu základní frekvence.

Jako nástroj pro manipulaci s HMM byl zvolen balík programů HTK. V experimentu operujeme nad datovou základnou podmnožiny českého SPEECONu (cca 10 000 promluv), kde k promluvám známe jejich textový přepis. Okamžiky slabičných středů (nukleí) získáváme force-alignmentem promluv oproti jejich přepisům za použití natrénovaných akustických modelů pro trifóny (čas prostředního stavu modelu je brán jako tato časová značka). Intonační průběh promluv je odhadnut použitím autokorelační funkce a po vhodném post-processingu je z něj extrahována intonační informace ve slabičných středech. Tato akustická informace stává prvním vstupem pro trénování/testování vzniklých modelů. Druhým vstupem je korespondující informace lexikální, kterou obdržíme vhodnou transformací textových přepisů promluv. K tomu byl navržen a implementován modul "Text2Foot converter" pracující na základě klitických pravidel pro češtinu (pohlčení jednoslabičných slov neboli "vhodné mazání mezer mezi slovy"). Pracuje v několika fázích a na jeho výstupu dostáváme přepis promluvy nikoliv ve slovech oddělených mezerami, ale v přízvukových taktech. Z obou typů informace lze poté natrénovat vhodné modely reprezentující stopy o různých délkách.

Během přípravy akustických dat bylo užito několik typů normalizace intonace: vzhledem k celkovému průměru intonace během promluvy, vzhledem k průměru intonace v rámci dané stopy (známe v trénovací množině, v reálu není k dispozici). Spojením obou přístupů lze normalizovat oproti křivce proložené získanými „intonačními body“, kdy se snažíme z promluv odfiltrout větnou intonaci.

Prozodie a modelování přízvukových taktů

- Doktorand: Ing. Jan Bartošek
 - Školitel: Ing. Václav Hanžl, CSc.
- **Rámcové téma dizertace:**
- Využití prozodie v rozpoznávání souvislé řeči

LDD 2013

1

Osnova přednášky

- Úvod – opakování: Co je prozodie
- Intonace, větná melodie
- Přízvuk, přízvukový takt
- Experiment – možnosti detekce českých přízvukových taktů pomocí intonace
- Výsledky experimentu
- Závěr

LDD 2013

2

Úvod

- náš výzkum - čeština, ale principy zobecnitelné i na ostatní jazyky

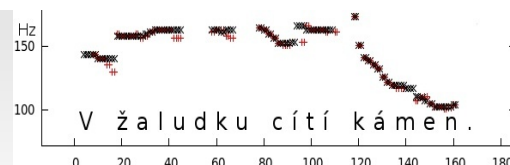
Prozodie

- označuje souhrn jevů na suprasegmentální úrovni řeči (segment=hláska,fón)
- intonace, přízvuk, rytmus a mluvní tempo

LDD 2013

3

Intonace:



- průběh výšky hlasivkového tónu v čase (řeč = znělé hl. + neznělé hl. + pauzy)
 - intonace na úrovni věty = větná melodie
 - intonace i v rámci přízvukových taktů
- F0 - základní hlasivkový tón (fyzikální)
- Pitch - vnímaná výška tónu (často log(F0))

LDD 2013

4

Přízvuková analýza

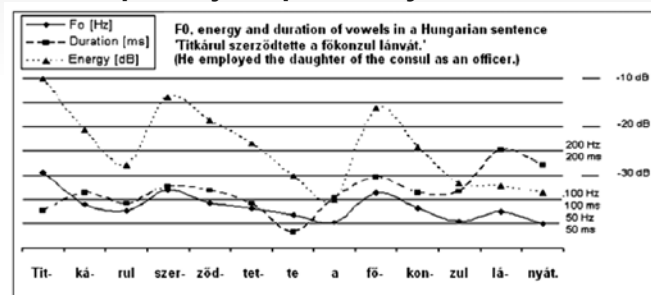
- dva typy přízvuku (prominence):
 - explicitní (větný) - určujeme ho sami → "důraz" na určitém slově ve větě
 - implicitní (lexikální) "přízvuk" → segmentace řeči na "přízvukové takty", klíčová role pro rytmus řeči
- umístění přízvuku v přízvukovém taktu (p.t.) závisí na konkrétním jazyce, v češtině je přízvuk na první slabice (fixed-stressed) vs. angličtina
- běžný mluvčí/posluchač dokáže určit hranice p.t.

LDD 2013

5

Prostředky pro realizaci prominence (intonace, intenzita, délka)

- maďarština - zvukové kvality spolu korespondují na přízvukných slabikách:



LDD 2013

6

Realizace lexikální prominence v češtině

v češtině spolu zvukové kvality takto nekorespondují, přízvuk je realizován:

- jak zvýšením, tak snížením intonace
- jak zvýšením, tak snížením intenzity
- délka slabik nehraje významnou roli
- významný pokles F0 uprostřed zkoumané množiny slabik pravděpodobně nepředstavuje realizaci intonace v rámci přízvukového taktu [Palková2003]

LDD 2013

7

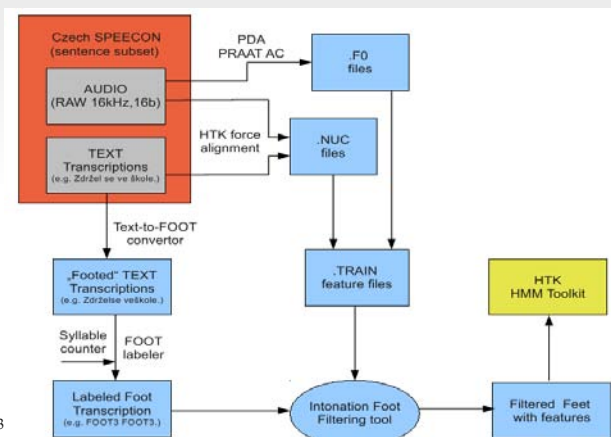
Experiment

- cíl: strojově rozdělit promluvu na přízvukové taktu pomocí prozodické informace (intonace)
- zdrojová řečová databáze: podmnožina SPEECON CZ (cca 10 000 vět)
- příprava trénovacích dat nebyla zcela jednoduchá, vytvořen lexikální modul pro převod vět na sekvenci přízvukových taktů podle mluvnických pravidel
- Realizace: HTK HMM Toolkit

LDD 2013

8

Experiment - příprava dat



LDD 2013

9

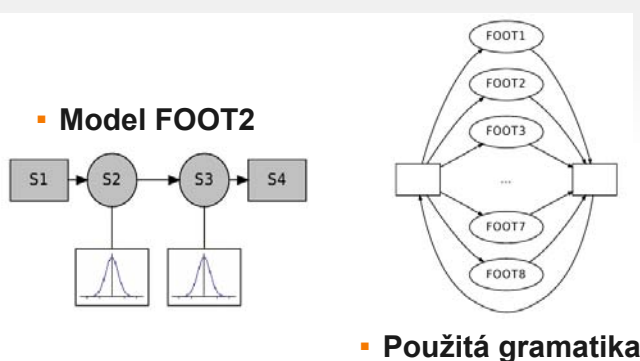
Experiment - Text2Foot converter

- Fáze 0: neslabičné předložky ('v', 'k', 's', 'z')
- Fáze 1: enklitika
 - jsem, jsme, ... + bych, bys, ...
 - si, se
 - mi, mně, ti, mu, ...
 - mě, tě, ho, ...
 - mnou, něj, něm, ... + prý, však
- Fáze 2: proklitika (u, o, ze, ke, ve, na, ...)

LDD 2013

10

Experiment - vizualizace HMM modelů



LDD 2013

11

Experiment - typy normalizací

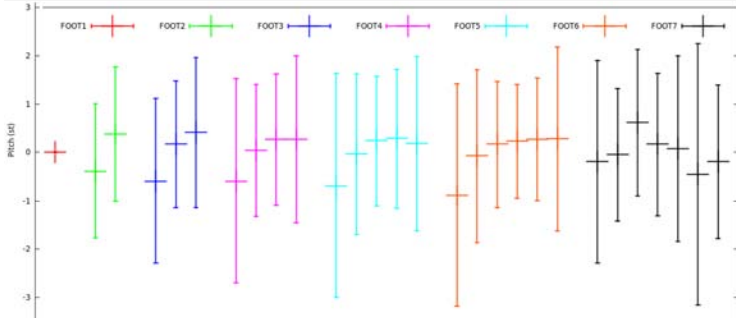
- NORM0
- NORM1
- NORM2
- NORM3

LDD 2013

12

Experiment - natr. modely

Norm1



LDD 2013

13

Experiment - výsledky

Training data	Testing data	Corr	Acc	D	S	I
norm0	norm0	34.2	18.2	2326	2718	1223
norm1	norm1	53.4	46.1	1739	1755	561
norm2	norm2	30.4	18.7	2659	2563	876
norm3	norm3	32.1	20.4	2240	2154	759
norm1	norm3	38.7	32.1	1460	3145	493

LDD 2013

14

Závěr

- Experiment – v reálném případě užití (norm1/norm3) dosahujeme úsp. 32%
- pro lepší výsledky bude zřejmě nutné pročištění dat (kolaborace s fonetiky)
- Využití informace o intenzitě/energii pravděpodobně situaci nezlepší (přesnost)
- Velká očekávání máme od využití spektrálního zesílení ve středech slabik

LDD 2013

15

Zdroje

- Z. Palková, Fonetika a fonologie češtiny [Phonetics and phonology of Czech]. Karolinum, Praha, 1994.
- K. Vicsi and G. Szaszak, "Automatic segmentation of continuous speech on word level based on supra-segmental features", 2005, International Journal of Speech Technology, vol. 8, pp. 363–370,
- Z. Palková and J. Volín "The role of f0 contours in determining foot boundaries in Czech," in In Proceedings of the 15th ICPHS, Barcelona, vol. 2, 2003, pp. 1783–1786.
- P. Hauser, Základy skladby češtiny. Brno: Masarykova univerzita, 2003.

LDD 2013

16

Děkuji za pozornost

LDD 2013

17

Význam difúzní magnetické rezonance u dětí s vývojovou dysfázií

Josef VAVŘINA¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

vavrijos@fel.cvut.cz

Abstrakt: Tým LANNA spolupracuje na grantu s neurologickou klinikou při diagnostice vývojové dysfázie. Naším hlavním cílem je pomoci při diagnostice nemoci metodami založenými na zpracování řeči. Tyto metody jsou neinvazivní a jejich náklady nejsou vysoké. Cílem tohoto článku je představit naopak metodu velmi nákladnou založenou na difúzní magnetické rezonanci (DTI - diffusion tensor imaging), kdy je cílem hledat dráhy neuronů spojující řečová centra v mozku a z parametrů tohoto propojení určit míru onemocnění. Článek se také zabývá metodikou získávání parametrů spojení řečových center a představuje metody pro analýzu těchto parametrů. Zatím nejsou k dispozici referenční data zdravých dětí, které by významně přispěly k analýze.

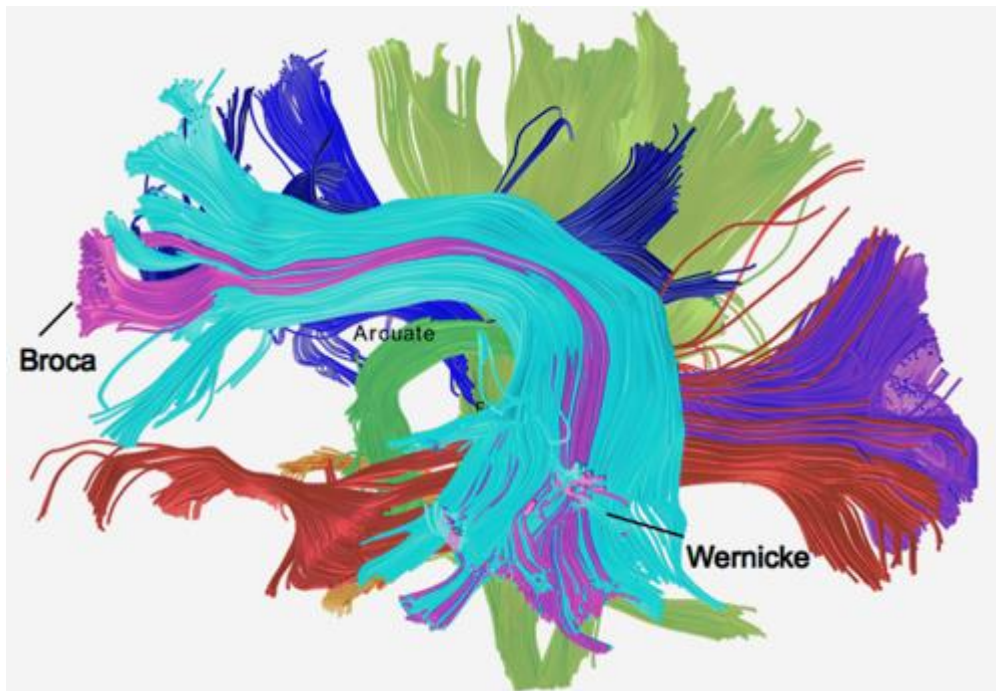
1. Úvod

Náš tým LANNA spolupracuje na grantu s dětskou neurologickou klinou v Motole řadu let. Cílem našeho grantu je diagnostika vývojové dysfázie u dětí ve věku 6 - 12 let. Vývojová dysfázie je vrozené onemocnění mozku, kde jeden z hlavních důsledků nemoci je opožděný vývoj řeči. Nemoc je často komplexnější porucha mozku a není možné ji jednoduše diagnostikovat, protože porucha řeči může nastat v mozku z mnoha důvodů. Nemoc se také projevuje například problémy s jemnou motorikou, prostorovou představivostí atd. Pro naši skupinu je důležitý vývoj řeči. Léčba nemoci je do jisté míry individuální, kdy velice záleží na míře onemocnění, její příčině a na případných dalších poruchách mozku [1].

Spolupracujeme tedy s logopedy, psychology, neurology, EEG specialisty a nyní i s výsledky z MR traktografie. Cílem je korelovat data jednotlivých oddělení a zjistit jejich vztah k míře onemocnění. Cílem grantu je nalézt metodiku, která by dokázala nemoc levně, časově nenáročně a příjemně pro všechny strany diagnostikovat. Tento článek se zabývá naopak metodou velice nákladnou, nepříjemnou ale velice objektivní a to mapováním mozkových drah pomocí difúzní magnetické rezonance. Výsledky z tohoto vyšetření by měly

sloužit jako reference pro jiné metody a pro hledání odpovědi na vývojovou dysfázii v samotném mozku.

Metoda DTI je poměrně nová metoda, která byla poprvé popsána v roce 1994 [2]. Metoda je založena na difúzi látek v mozku, které více difundují ve směru nervových vláken [3]. Nejběžnější metoda nalezení nervových vláken je pomocí algoritmu line propagation, kdy je v počátečním objemu vytvořen určitý počet počátečních bodů, z kterých algoritmus po směru nejvyšší difúze rekonstruuje nervová vlákna. Metoda se používá pro diagnostiku nebo i pomoc při operaci lézí, ale také pro hledání zrakových svazků nebo také pro hledání fascilus arcuatus, který právě spojuje Brocovo a Wernickeovo centrum (centra řeči - Obr. 1) [4].



Obr. 1: Fascilus arcuatus spojující Wernicke a Broca řečová centra v mozku[5]

2. Příprava dat

Pomocí MR traktografie bylo vyšetřeno 42 dětí s vývojovou dysfází. Získávání dat probíhá v těchto krocích:

1. Provedení vyšetření dítěte pomocí MR traktografie
2. Vytvoření mapy mozkových drah celého mozku
3. Nalezení fasciculus arcuatus pro obě hemisféry
4. Získání parametrů fasciculu arcuatu

Počítačové zpracování bylo provedeno pomocí software MedINRIA.

Fascilus artuatus má pak tyto naměřené parametry [6]:

1. Volume - objem svazku
2. Fibres - počet vláken svazku
3. Délka vláken
4. FA - frakční anizotropie určuje směrovost svazku

5. ADC (apparent diffusion coefficient) - celková difuzivita tkáně

U parametrů délky, FA a ADC je určena průměrná, maximální, minimální hodnota a směrodatná odchylka svazku. Všechny hodnoty jsou určeny pro obě hemisféry mozku, protože Brocovo a Wernickeovo centrum se nachází v obou hemisférách. V dospělém věku se u praváků dominantní centra nacházejí v levé hemisféře v cca 95% případů [1]. U dětí s vývojovou dysfázií nemusí docházet k tomuto předpokladu.

Z dat 42 dětí se jen u 38 z nich podařilo najít fasciculus artuatus alespoň v jedné hemisféře a u 34 dětí se podařilo nalézt tento trakt v obou hemisférách. Pro každé z těchto dětí je k dispozici 28 parametrů z MR traktografie.

3. Metody zpracování

Metody jsou založené na statistických metodách[7], které dají základní informace o datech a na metodách klasifikace bez učitele hledající shluky dat s podobnými parametry. V současné době jsou pořizována data MR traktografie zdravých dětí, které výrazně rozšíří možnosti zpracování dat.

Statistické zpracování:

Toto zpracování slouží k základnímu pochopení dat. V jakých hodnotách se hodnoty pohybují, zda jsou v datech velké výchyly apod. Hlavní důraz byl kladen zejména na zjištění dominance levé nebo pravé hemisféry, který byl prováděn srovnáním parametru objemu levé a pravé hemisféry případně počtu vláken levé a pravé hemisféry. Statistické zpracování zkoumá vliv věku na jednotlivé parametry a korelaci parametrů mezi sebou. Hlavní přínos této metody bude zřetelný až při srovnání s daty zdravých dětí.

Shluková analýza:

Tato analýza je výhodná pro data, o kterých nemáme mnoho informací a nemůžeme použít klasifikační metody, které potřebují referenční skupinu dat pro trénink. Cílem diagnostiky je nemocné děti rozdělit do skupin podle míry jejich onemocnění. Analýza by měla data rozdělit do jednotlivých skupin. Použili jsme dvě metody pro shlukovou analýzu s dvěma sadami vstupních parametrů.

1. **PCA**- Pomocí metody hlavních komponent (PCA [8]) je možné efektivně zredukovat velké množství dimenzí vstupních dat do 3D prostoru, pro které je možné provést shlukovou analýzu.
2. **SOM**- (Self-organizing-maps [9]) využívají projekci do 2D a následně metodou na základě k-means provádí shlukovou analýzu.

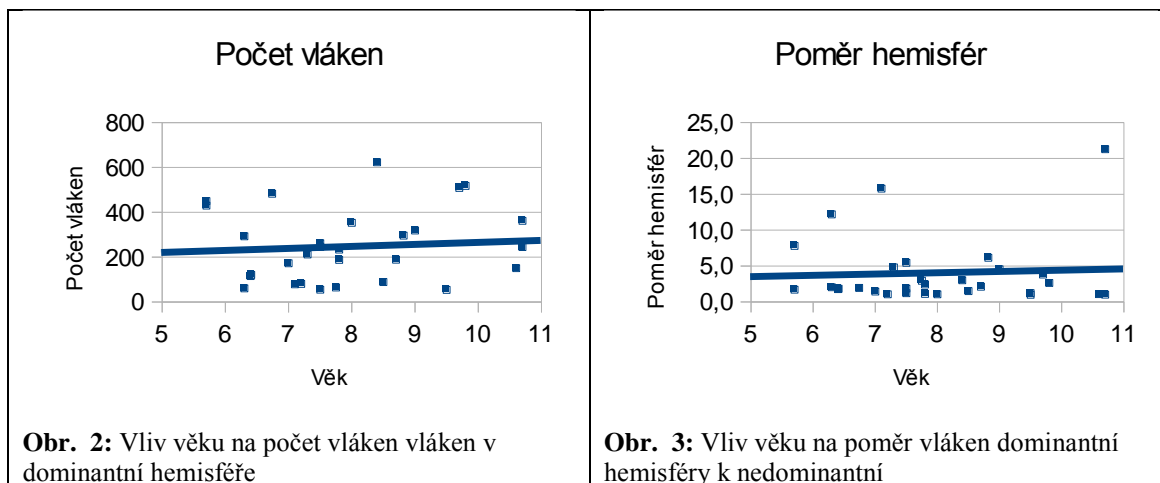
Dvě skupiny vstupních dat:

1. **Všechny parametry** - pro vstup do obou shlukových analýz bylo použito všech 28 parametrů z MR traktografie.
2. **Zajímavé parametry** - po statistické analýze a doporučení neurologů se ukázalo, že počet vláken a objem svazku pro obě hemisféry jsou vhodnými parametry pro shlukovou analýzu.

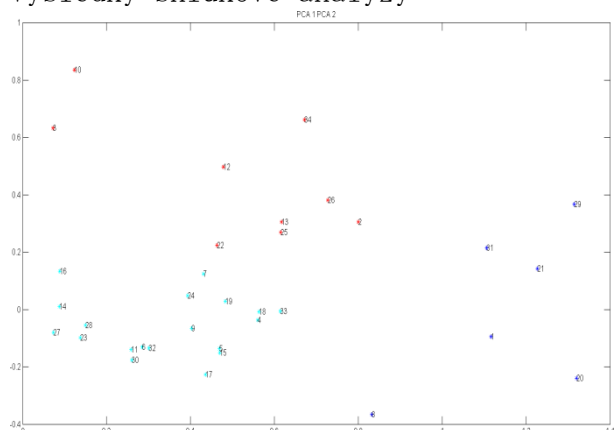
4. Výsledky

Zajímavé výsledky ze statistického zpracování jsou:

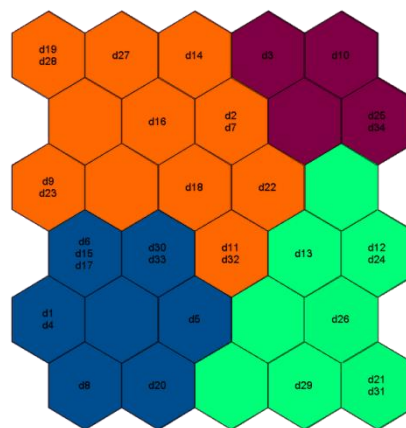
- cca 70% dětí má dominantní levou hemisféru pokud jde o poměr počtů vláken levé a pravé hemisféry.
- průměrný počet vláken levé hemisféry je 218 vláken, pravé hemisféry 131 vláken (poměr 1,66) naproti tomu poměr objemu svazku je v poměru 1,34. Vláken v levé hemisféře je více, ale nemají takový objem. Poměry ostatních parametrů jsou pro levou i pravou hemisféru stejné. Například u délky vláken se ale nic jiného neočekávalo.
- Počet vláken v levé i pravé hemisféře je u dětí velice různý. Některé děti mají v jedné hemisféře například jen několik jednotek vláken nebo se trakt vůbec nepodařilo najít, jiné děti mají v každé hemisféře přes 200 vláken. Řádové rozdíly jsou i v počtu vláken dominantní hemisféry, kdy nejmenší hodnoty se pohybují kolem 50 vláken a nejvyšší přes 500.
- Počet vláken dominantní hemisféry je rostoucí parametr. Poměr vláken v hemisférách (dominantní vs. Nedominantní) nelze dobře předpovídat s věkem dítěte. Některé poměry vláken jsou velmi vysoké a lineární regresi velice vychylují (obr. 2 a obr. 3)
- Korelace počtu vláken k objemu svazku je v pravé hemisféře 0,59 a v levé hemisféře 0,82.



Výsledky shlukové analýzy:



Obr. 4: Projekce do PC1 a PC2 s metodou k-means. Čísla odpovídají pořadí dítěte ve vstupním vektoru.



Obr. 5: Shluková analýza pomocí SOM. Čísla odpovídají pořadí dítěte ve vstupní vektoru.

Obě metody shlukují data velice podobně (obr. 4 a 5), děti se stejným pořadovým číslem jsou v obou metodách vždy podobně vzdáleny. Další zajímavý poznatek je, že pokud použijeme všechny parametry svazku na vstupu nebo jen zajímavé parametry, tak obě metody vytvářejí shluky také velice obdobně. V neposlední řadě je nutné zmínit, že ani jedna metoda nevytváří ohraničené shluky dat. Dalo by se říci, že je vytvořen jeden hlavní shluk a pak individuálně rozprostřená data kolem tohoto shluku (viz. Obr. 4.).

Míra onemocnění vývojovou dysfázií nicméně nemusí korelovat s parametry svazku mezi řečovými centry a také nemusí odpovídat rozdělením do ohraničených skupin.

5. Závěr

V článku byl představen význam MR traktografie ve vztahu k onemocnění vývojovou dysfázií. Byla popsána metoda získávání dat, kde základem metody je vypočítání parametrů fascilus arcuatus, který spojuje řečová centra v mozku. Vyšetření MR traktografie bylo podrobena 42 dětí, kde pro 34 dětí byla data kompletní. Data byla prozkoumána statistickými metodami a metodami shlukové analýzy. Statistické metody provádí základní analýzu dat se zaměřením na poměr levé a pravé hemisféry a význam jednotlivých parametrů. Zajímavé parametry jak z pohledu neurologů, tak z výsledků statistické analýzy ukazují, že velký význam hraje počet vláken ve fascilus arcuatus. Velkou roli hraje také objem tohoto svazku, ale ten silně koreluje s počtem vláken zejména pak v levé hemisféře. Vliv věku na parametr počtu vláken v dominantní hemisféře nebo i poměr vláken v dominantní a nedominantní hemisféře nebyl prokázán.

V blízké době by mělo být provedeno vyšetření MR traktografie i u zdravých dětí, které získá referenční data, která umožní další analýzu.

Cílem metody vyšetření MR traktografií je přinést referenční data pro méně náročné metody, kdy by bylo možné například pomocí EEG naměřit korelovatelné hodnoty. Onemocnění vývojovou dysfázií je nicméně velice

komplexní onemocnění mozku a jediná metoda MR traktografie nezíská jednoznačnou diagnózu, pro kterou je zapotřebí spolupráce více vyšetření.

Poděkování

Tato práce byla podpořena jako součást Grant IGA MH CR agency, No. NT11443-5/2010 a No. SGS12/185/OHK4/3T/13.

Reference

- [1] Russell J. Love; Web, Wanda G.: *Mozek a řeč*. Portál. 2009. ISBN:978-80-7367-464-9
- [2] Basser PJ, Mattiello J, Le Bihan D.: *MR diffusion tensor spectroscopy and imaging*. Biophys. J. 1994; 66: 259-267.
- [3] A. Zolal a kolektiv: *Použití DTI traktografie v neuronavigaci při operacích mozkových nádorů: kazuistiky*. Česká a slovenská neurologie a neurochirurgie. 2008, roč. 71/104, č. 3, s. 352-357. ISSN: 1210-7859; 1802-4041 (elektronická verze).
- [4] Mori S, van Zijl PC: *Fiber tracking: principles and strategies - a technical review*. NMR Biomed. 2002 Nov-Dec;15(7-8):468-80.
- [5] Fernandez-Miranda JC, Pathak S, Schneider W.: *High-definition fiber tractography and language*. J Neurosurg. 2010 Jul;113(1):156-7; author reply 157-8. doi: 10.3171/2009.10.JNS091460.
- [6] M. Keřkovský, A. Šprláková-Puková, T. Kašpárek, P. Fadrus, M. Mechl, V. Válek: *Diffusion tensor imaging - současné možnosti MR zobrazení bílé hmoty mozku*. Česká a slovenská neurologie a neurochirurgie. 2010; 73/106(2): 136-142
- [7] Zvárová, J.: *Základy statistiky pro biomedicínské obory*, Karolinum - nakladatelství Univerzity Karlovy, Praha, 2001
- [8] I. T. Jolliffe, *Principal Component Analysis*. Springer, ISBN: 978-0387954424 (2002)
- [9] T. Kohonen: *Self-Organizing Maps*. Springer-Verlag, 3rd edition (2001)

Klasifikace promluv pacientů s vývojovou dysfázií

Martina NEJEPSOVÁ¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

nejepmar@fel.cvut.cz

Abstrakt: Vady hlasu a řeči u dětí vedou k redukci jejich mluveného projevu. Vývojová dysfázie je jednou z nejrozšířenějších poruch u dětí. V případě včasného rozpoznání a zahájení terapie je možné dosáhnout začlenění dětí mezi běžnou populaci.

K návrhu metody pro rozpoznávání vady řeči u dětí byla pořízena databáze promluv od 271 dětí s různou mírou závažnosti vývojové dysfázie. Výběr 100 zdravých dětí z databáze pro hodnocení věkové závislosti byl užit jako reference.

Z potřeby roztrždit tyto promluvy byl připraven subjektivní poslechový test zaměřený na odhad závažnosti vývojové dysfázie na základě vyslechnuté promluvy (říkanky a spontánního popisu) a znalosti biologického věku mluvčího. Testu, obsahujícího 371 promluv, se zúčastnili 4 logopedi a 3 zkušení dobrovolníci. Na základě jejich hodnocení byly promluvy rozděleny do 4 kategorií dle závažnosti (0 – zdravý mluvčí, 1 – mírná, 2 – střední, 3 – závažná porucha vývojové dysfázie).

Tyto kategorie promluv byly analyzovány a hledány takové charakteristické parametry pro sledování věkové závislosti a závažnosti vývojové dysfázie. Ve shodě se subjektivním poslechovým testem byla analyzována pouze samovolně pořízená říkanka (235 promluv) a spontánní popis (290 promluv). Mezi tyto charakteristiky patří například míra srozumitelnosti říkanky metodou DTW, rychlost a délka promluvy u říkanky a i spontánního popisu, poměr řeč-pauza, počet slov ve spontánním popisu.

Kromě objektivních parametrů byly subjektivně sledovány i ukazatele vyplývající z pozorování při pořizování databáze. Mezi sledované ukazatele patří například četnost užitých slovních druhů.

Vzájemným porovnáním výsledků analýz a subjektivního hodnocení byl navržen postup, který bude sloužit pro automatickou klasifikaci promluv a tím i urychlením rozpoznání poruchy řeči.

KLASIFIKACE PROMLUV PACIENTŮ S VÝVOJOVOU DYSFÁZIÍ

Doktorand: Martina Nejepsová
Školitel: Doc. Ing. Roman Čmejla, CSc.

24.5.2013

OBSAH PREZENTACE

- Motivace
- Vývojová dysfázie
- Databáze promluv pacientů
- Subjektivní poslechový test
- Akustické analýzy promluv
- Shrnutí
- Další cíle

MOTIVACE

- Pořízení databáze promluv pacientů s vývojovou dysfázií
- Porovnání objektivního a subjektivního hodnocení projevu pacientů s vývojovou dysfázií
- Zjednodušení diagnostiky vývojové dysfázie

VÝVOJOVÁ DYSFÁZIE

- Specificky narušený vývoj řeči – porucha zpracování řečového signálu
- Ztížená schopnost nebo neschopnost naučit se verbálně komunikovat, ačkoliv podmínky pro rozvoj řeči jsou přiměřené
- Symptomatologie v mluveném projevu
 - Opožděný či nerovnoměrný vývoj řeči
 - Artikulační obtíže – srozumitelnost projevu
 - Prosodické problémy – rytmus a intonace projevu
 - Neodpovídající slovní zásoba, záměna či chyby ve slovosledu
 - Rozdíl mezi verbálními a neverbálními schopnostmi
- Diagnostika
 - Na základě hodnocení několika specialistů

DATABÁZE

- Obsah databáze: říkanka

En ten týky,
dva špalíky,
čert vyletěl z elektriky.
Bez klobouku bos,
Natloukl si nos.

spontánní popis



Databáze pacientů s vývojovou dysfázií	188 promluv (127 samovolně nahraných)	261 promluv (190 samovolně nahraných)
Referenční databáze zdravých dětí	98 promluv	100 promluv

SUBJEKTIVNÍ POSLECHOVÝ TEST

- Odhad závažnosti vývojové dysfázie na základě vyslechnuté promluvy a znalosti biologického věku mluvčího
- Hodnocené subjekty
 - Promluvy – říkanka a spontánní popis
 - 271 promluv dětí s vývojovou dysfázií a 100 promluv zdravých dětí
 - Informace o hodnoceném subjektu – biologický věk
- Test
 - Klasifikační subjektivní poslechový test s nominální škálou
 - Kategorizace do 4 tříd: 0 – 1 – 2 – 3
- Hodnotitelé
 - 4 specialisté, 3 laici

SUBJEKTIVNÍ POSLECHOVÝ TEST

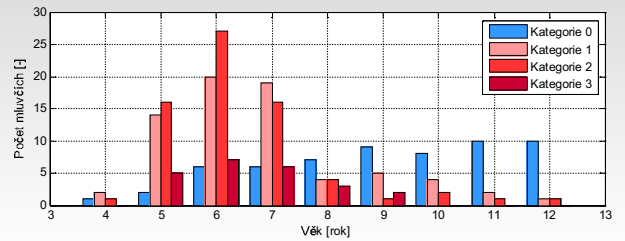
- Kategorizace promluv pacientů s vývojovou dysfázií

Věk mluvčího	4	5	6	6	8	9	10	11	12	celkem
Kategorie 0	2	2	6	6	6	9	8	10	10	60
Kategorie 1	3	16	24	21	4	5	4	2	1	81
Kategorie 2	9	38	50	25	6	1	2	2	1	135
Kategorie 3	10	26	35	18	3	2	0	0	1	95
Celkem	24	83	115	60	21	16	14	14	13	371

ANALÝZA ŘÍKANKY

- Projevy vývojové dysfázie
 - Poruchy paměti – problémy se zapamatováním a reprodukcí říkanky
 - Špatná artikulace
 - Změna tempa během promluvy

- Kategorie a počty promluv

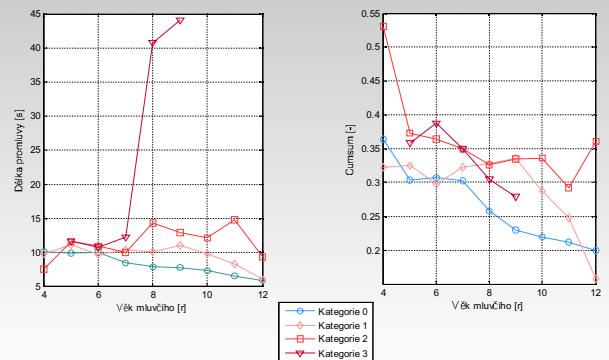


ANALÝZA ŘÍKANKY

- Objektivní analýza
 - Délka celé promluvy
 - F0 celé promluvy
 - Porovnávání celé promluvy metodou DTW
 - Poměr řeči (respektive pauzy) v promluvě
 - Počet a délka pauz v promluvě
 - Počet slov v promluvě
 - Rychlost artikulace celé promluvy
 - Odišnost rychlostí artikulace jednotlivých veršů
 - Změna rytmu říkanky
- Ostatní hodnocení
 - Dodržování veršů
 - Shoda s rýmem a veršem

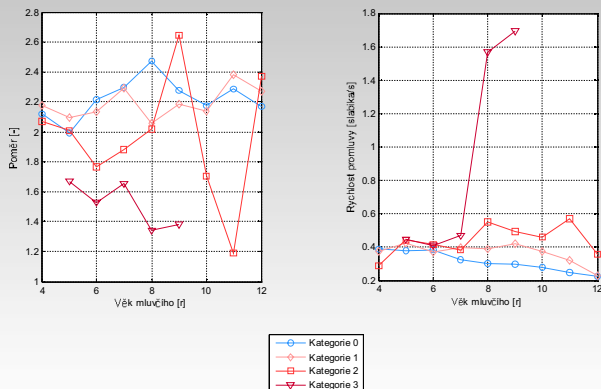
ANALÝZA ŘÍKANKY

- Délka promluvy
- Míra srozumitelnosti promluvy metodou DTW



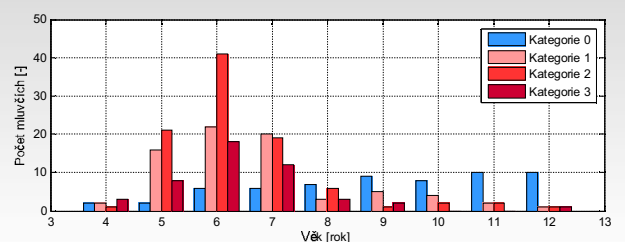
ANALÝZA ŘÍKANKY

- Poměr řeč-pauza
- Rychlost promluvy



ANALÝZA SPONTÁNNÍHO PROJEVU

- Projevy vývojové dysfázie
 - Malá slovní zásoba
 - Špatná artikulace
 - Změna tempa během promluvy
- Kategorie a počty promluv

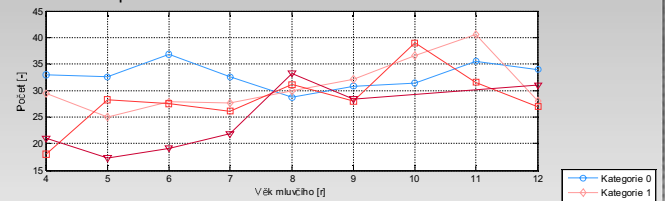


ANALÝZA SPONTÁNNÍHO PROJEVU

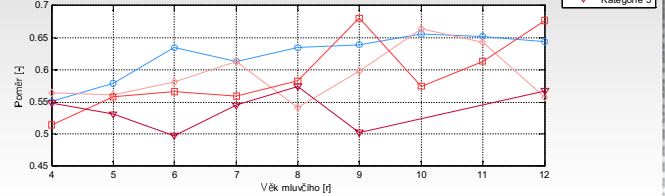
- Objektivní analýza
 - Délka promluvy
 - F0 celé promluvy
 - Poměr řeči (respektive pauz) v promluvě
 - Rychlost artikulace celé promluvy
 - Počet a délka pauz v promluvě
 - Počet slov v promluvě
 - Rychlost promluvy
- Ostatní hodnocení
 - Počet užitých slovních druhů v promluvě

ANALÝZA SPONTÁNNÍHO PROJEVU

- Poměr řeči v promluvě

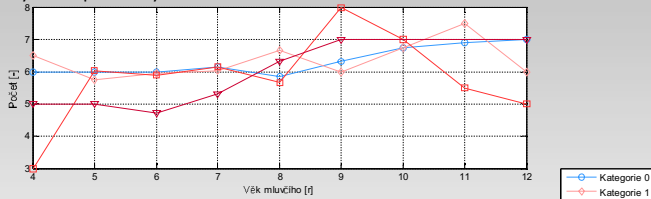


- Počet slov v promluvě

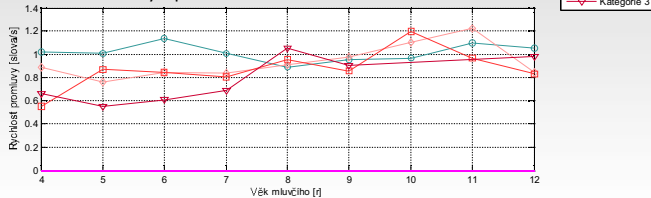


ANALÝZA SPONTÁNNÍHO PROJEVU

- Rychlost promluvy



- Užité slovní druhy v promluvě



SHRNUTÍ

- Kategorizace promluv a porovnání výsledků analýz na základě subjektivního hodnocení
- Akustické analýzy promluv
- Volba vhodných akustických parametrů pro automatickou klasifikaci závažnosti poruchy

DALŠÍ CÍLE

- Rozšíření počtu parametrizací promluv
- Automatizace hodnocení a kategorizace nahrávek
- Sledování trendu poruchy řeči v průběhu léčebné terapie

LITERATURA

- [1] Balkó I.: K výzkumu tempa řeči a tempa artikulace v různých řečových úlohách: *Bohemistika*, Nr. 3, 2005.
- [2] Flipsen P. jr.: Syllables per word in typical and delayed speech acquisition Intonace. *Clinical Linguistics & Phonetics*, Volume 20, No. 4, 2006.
- [3] Janda J.: *Posuzování logopedického věku dítěte*. Disertační práce. Praha: ČVUT v Praze, 2012.
- [4] Melka A.: *Základy experimentální akustiky*, AMU, Praha 2005.
- [5] Potamianos A., Narayanan S.: A review of the acoustic and linguistic properties of children's speech. *Multimedia Signal Processing*, IEEE 9th Workshop on Volume, 2007.
- [6] Scotta S.K., Clegg F., Rudgec P., Burgess P.: Foreign accent syndrome, speech rhythm and the functional neuroanatomy of speech production. *Journal of Neurolinguistics*, Volume 19, Issue 5, 2006.
- [7] Shriberg L.D., Fourakis M., Hall S.D., Karlsson H.B., Lohmeier H.L., McSweeney J.L., Potter N.L., Scheer-Cohen A.R., Strand E.A., Tilkens C.M., Wilson D.L.: Perceptual and acoustic reliability estimates for the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics*, Volume 24, No. 10, 2010.
- [8] Škodová E., Jedlička I. a kolektiv: *Klinická logopedie – 2. aktualizované vydání*, PORTAL, Praha, 2007.
- [9] Vick J., Moore C.A., Campbell T., Shriberg L., Green J., Truemper K.: Multivariate Classification of Children with Speech Delay of Unknown Origin. *The Annual Convention of the American Speech-Language-Hearing Association*, New Orleans, 2009.

Hodnocení promluv pacientů s dysfonií - Subjektivní testy

Adam Stráník¹, Roman ČMEJLA¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

stranada@fel.cvut.cz

Abstrakt: Chraptivost je obecné označení pro přítomnost patologických příměsí v hlase a je také jedním z prvních příznaků onemocnění hlasového ústrojí. Je provázána různými jevy jako například chraptivostí, chraplavostí, dyšností nebo hlasovým napětím, které při běžné komunikaci vnímá nejen školený specialista, ale i běžný laik. Tyto jevy nemusí být pro posluchače vždy nepříjemné, nicméně, jak je zmíněno výše, mohou být předzvěstí počínajícího vážnějšího onemocnění hlasového ústrojí.

Pro určení míry chraptivosti odborná veřejnost vypracovala několik postupů, jak od sebe rozlišovat různé druhy a tíže chraptivosti. Tyto metody jsou jak subjektivní, tak objektivní. Mezi subjektivní metody se řadí různé poslechové testy, při kterých odborník poslouchá hlasový projev pacienta a na základě vlastních zkušeností vyplňuje určité formuláře ve kterých je popsáno, na jaké aspekty řečového projevu by se měl zaměřovat. Objektivních metod existuje celá řada, některé jsou invazivní (laryngoskopie, laryngostroboskopie,...) a jiné neinvazivní (elektroglottografie, analýza akustického signálu). Z hlediska ceny vyšetření a komfortu z pohledu pacienta je vyšetření pomocí analýzy akustického signálu nejpříjemnější. Metody akustické analýzy hlasu mají také výhodu v tom, že zpětně neovlivňují pacienta, jako je tomu například u laryngoskopie.

Naším cílem je vyvinout automatický systém, který by odborníkům umožnil objektivizovat míru chraptivosti v hlase na základě akustické analýzy hlasového signálu.

Pro tento účel je nutné vytvořit vhodnou databázi promluv, kterou odborníci subjektivně ohodnotí vhodně zvolenou metodou. Na základě výsledků subjektivního hodnocení je možné hledat vhodně akustické parametry, které nejlépe popisují jednotlivé sledované parametry chraptivého hlasu.

Byly vytvořeny a poté ohodnoceny dvě databáze promluv, které vycházejí z archivní databáze nahrávek Foniatrické kliniky 1. LF UK a VFN. První databáze obsahuje celkem 469 nahrávek prodloužené fonace hlásky /a/, z toho 397 nahrávek pochází od pacientů s chraptivostí a 72 od pacientů z kontrolní skupiny. Druhá databáze obsahuje 593 nahrávek čteného standardizovaného textu "Podzim na Starém bělidle" a skládá se z 356 chraptivých a 237 kontrolních promluv. Obě databáze byly nezávisle na sobě dvakrát ohodnoceny pěti odborníky z Foniatrické kliniky. Opakované hodnocení proběhlo vždy nejdříve po čtrnáctidenní prodlevě a s obměněným pořadím nahrávek.

K hodnocení obou databází byla použita metoda GRBT, která popisuje všechny aspekty chraptivého hlasu: G (grade - celkový dojem z hlasu), R (roughness - chraplavost, míra nepravidelnosti kmitání hlasivek), B (breathiness - dyšnost, míra šumové příměsí v hlase) a T (tension - dojem spojený s hlasovou slabostí, nebo naopak s přemáháním hlasu).

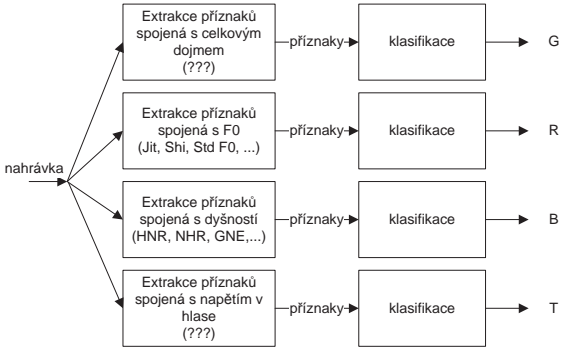
V prezentaci jsou předvedeny hlavní metody hodnocení subjektivních testů pro hodnocení míry shody skupiny hodnotitelů (tzv. inter-rater variability) a shody hodnotitele při opakovaném hodnocení (tzv. intra-rater variability). Dosažené výsledky ukazují, že vyšší shody hodnotitelé dosahují při hodnocení čteného textu než při hodnocení prodloužené fonace hlásky /a/.

Hodnocení promluv pacientů s dysfonií – subjektivní testy Adam Stráník, Roman Čmejlá Úvod Databáze Hodnocení Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot Výsledky Základní náhled na hodnocení Intra- rater variability Inter- rater variability Diskuze Závěr Reference 42-1	<h2>Hodnocení promluv pacientů s dysfonií – subjektivní testy</h2> <p>Adam Stráník, Roman Čmejlá</p> <p>K13131 – Katedra teorie obvodů Fakulta elektrotechnická České vysoké učení technické v Praze</p> <p>Letní Doktorandské Dny 2013</p> 
--	---

Hodnocení promluv pacientů s dysfonií – subjektivní testy Adam Stráník, Roman Čmejlá Úvod Databáze Hodnocení Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot Výsledky Základní náhled na hodnocení Intra- rater variability Inter- rater variability Diskuze Závěr Reference 42-2	<ol style="list-style-type: none"> 1 Úvod 2 Databáze 3 Hodnocení <ul style="list-style-type: none"> Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot 4 Výsledky <ul style="list-style-type: none"> Základní náhled na hodnocení Intra- rater variability Inter- rater variability 5 Diskuze 6 Závěr
--	--

Hodnocení promluv pacientů s dysfonií – subjektivní testy Adam Stráník, Roman Čmejlá Úvod Databáze Hodnocení Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot Výsledky Základní náhled na hodnocení Intra- rater variability Inter- rater variability Diskuze Závěr Reference 42-3	<h2>Úvod</h2> <ul style="list-style-type: none"> • chraptivost je obecné označení pro přítomnost patologických příměsí v hlase • je jedním z prvních příznaků onemocnění hlasového ústrojí • subjektivní hodnocení chraptivosti metodou GRBT <ul style="list-style-type: none"> • G ... celkový dojem z hlasu • R ... chraplavost – míra nepravidlosti kmitání hlasivek • B ... dyšnost – míra šumové příměsí • T ... napětí v hlase, nebo naopak hlasová slabost • parametry G, R, B hodnocené 0 (bez příznaku) až 3 (nejvíce přítomná) body • parametr T hodnocen -3 až 3 body, kde <ul style="list-style-type: none"> • <-2;0) bodů označuje hlasovou slabost • (0;2> body označuje přemáhání hlasu • 0 bodů je normální hlas • k hodnocení je možné používat půlbody
--	---

Hodnocení promluv pacientů s dysfonií – subjektivní testy Adam Stráník, Roman Čmejlá Úvod Databáze Hodnocení Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot Výsledky Základní náhled na hodnocení Intra- rater variability Inter- rater variability Diskuze Závěr Reference 42-4	<h2>Ukázky chraptivých hlasů</h2> <p>kontrolní skupina $G_0 R_0 B_0 T_0$</p> <p>psychogenní dysfonie $G_3 R_2 B_3 T_{-2}$</p> <p>spastická dysfonie $G_3 R_2 B_0 T_2$</p> <p>spastická dysfonie $G_3 R_3 B_1 T_2$</p> <p>psychogenní afonie $G_3 R_0 B_3 T_2$</p>
--	--

Hodnocení promluv pacientů s dysfonií – subjektivní testy Adam Stráník, Roman Čmejlá Úvod Databáze Hodnocení Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot Výsledky Základní náhled na hodnocení Intra- rater variability Inter- rater variability Diskuze Závěr Reference 42-5	<h2>Motivace</h2> 
--	---

Hodnocení promluv pacientů s dysfonií – subjektivní testy Adam Stráník, Roman Čmejlá Úvod Databáze Hodnocení Metodika Intra- rater variability Inter- rater variability Kvantifikace hodnot Výsledky Základní náhled na hodnocení Intra- rater variability Inter- rater variability Diskuze Závěr Reference 42-7	<h2>Databáze</h2> <ul style="list-style-type: none"> • databáze Foniatrcké kliniky 1. LF UK a VFN • prodloužené fonace hlásky /a/ <ul style="list-style-type: none"> • 469 nahrávek <ul style="list-style-type: none"> • 397 chraptivých • 72 kontrolní skupina • čtený text „Podzim na Starém běldle“ <ul style="list-style-type: none"> • 593 nahrávek <ul style="list-style-type: none"> • 356 chraptivých • 237 kontrolní skupina (z toho 44 nahrávek po půlhodinové zátěži hlasitým čtením)
--	---

Hodnocení promluv pacientů s dysfonií – subjektivní testy

Adam Stráník, Roman Čmejla

Hodnocení chraptivých promluv

pořadí	G	R	B	T	pořadí	G	R	B	T	pořadí	G	R	B	T
1					43					85				
2					44					86				
3					45					87				
4					46					88				
5					47					89				
6					48					90				
7					49					91				
8					50					92				
9					51					93				
10					52					94				

- slepé hodnocení – hodnotitel vidí jen pořadové číslo nahrávky
- opakované hodnocení je přečíslované (fce. randperm v MATLABu)
- ⇒ sloučení patologických promluv a kontrolní skupiny
- opakované hodnocení nejdříve po 14 dnech
- prodloužená fonace a čtený text byly hodnoceny zvlášť

42-10

Hodnocení promluv pacientů s dysfonií – subjektivní testy

Adam Stráník, Roman Čmejla

Intra- rater variability

- také nazýváno **Test – retest agreement**
- ukazuje míru shody v hodnocení při opakovaných testech
- odhalí náhodné vyplnění testu (dotazníku)
- používá se **Cronbachovo alfa** – α_C CRONBACH [1951]

$$\alpha_C = \frac{K\bar{r}}{1 + (K-1)\bar{r}}; \quad (1)$$

kde K je počet hodnocení (pro jednou opakované hodnocení $K=2$, pro dvakrát opakované hodnocení $K=3$ atd.) a \bar{r} je průměr horního trojúhelníku korelační matice

- hodnoty $\alpha_C \in (-1; 1 >$
 - teoreticky i záporné, ale počítají se pouze kladné
 - čím vyšší, tím vyšší shoda

42-12

Hodnocení promluv pacientů s dysfonií – subjektivní testy

Adam Stráník, Roman Čmejla

Inter- rater variability

- někdy nazýváno **reprodukovatelnost**
- ukazují
 - míru shody skupiny hodnotitelů
 - jak dobře je definovaný test
 - jak dobře hodnotitelé pochopili zadání
 - jakou mají zkušenost s hodnocením
- používají se tzv. κ statistiky
 - Cohenovo κ_C
 - Fleissovo κ_F
- hodnoty opět teoreticky i záporné, maximum 1

42-14

Hodnocení promluv pacientů s dysfonií – subjektivní testy

Adam Stráník, Roman Čmejla

Cohenovo κ_C

- vzájemná shoda **dvou hodnotitelů**
- definované jako COHNEN [1960]

$$\kappa_C = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}, \quad (2)$$

kde $Pr(a)$ je relativní shoda mezi hodnotiteli a $Pr(e)$ je hypotetická pravděpodobnost náhodné shody

- viz příklad

42-15

Hodnocení promluv pacientů s dysfonií – subjektivní testy

Adam Stráník, Roman Čmejla

Cohenovo κ_C – příklad

- dvojice hodnotitelů Marcelová – Vokřál, 1. test, G

		Marcelová			
		0	1	2	3
Vokřál	0	35	41	4	0
	1	39	130	54	3
	2	0	23	71	8
	3	0	1	24	36

- relativní shoda $Pr(a) = (35 + 130 + 71 + 36)/469 = 0,58$
- možnost náhodné shody $Pr(e) = 0,03 + 0,20 + 0,07 + 0,01 = 0,31$ určuje, s jakou pravděpodobností se mohli shodnout u jednotlivých hodnocení
 - hodnocení 0 použil Vokřál u $(35 + 41 + 4 + 0)/469 = 0,17$ nahrávek, Marcelová u $(35 + 39 + 0 + 0)/469 = 0,16$
 - shodnout se tedy mohli u $0,17 * 0,16 = 0,03$ nahrávek

$$\kappa_C = \frac{0,58 - 0,31}{1 - 0,31} = 0,39$$

42-16

Hodnocení promluv pacientů s dysfonií – subjektivní testy

Adam Stráník, Roman Čmejla

Fleissovo κ_F

- vzájemná shoda **n hodnotitelů**
- definované jako FLEISS [1971]

$$\kappa_F = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}, \quad (3)$$

kde $1 - \bar{P}_e$ je míra shody dosažitelná náhodou, $\bar{P} - \bar{P}_e$ je míra skutečně dosažené shody

- viz příklad

42-17

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Fleissovo κ_F – příklad

nahrávka i	0	1	2	3	P_i
1	0	0	1	4	0,6
2	0	0	5	0	1
3	0	1	3	1	0,3
4	1	4	0	0	0,6
5	0	0	4	1	0,6
...
469	5	0	0	0	1
total	541	924	623	257	
p_j	0,23	0,39	0,27	0,11	

- zjistit podíl všech přiřazení do j -té kategorie – parametr p_j
 - např.: $p_0 = 541/2345 = 0,23$, kde 2345 je počet všech hodnocení (469*5)
- určit $\bar{P}_e = \sum_{j=1}^k p_j^2$

42-18

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Fleissovo κ_F – příklad

nahrávka i	0	1	2	3	P_i
1	0	0	1	4	0,6
2	0	0	5	0	1
3	0	1	3	1	0,3
4	1	4	0	0	0,6
5	0	0	4	1	0,6
...
469	5	0	0	0	1
total	541	924	623	257	
p_j	0,23	0,39	0,27	0,11	

- určit míru shody u i -té nahrávky (tzn. spočítat, kolik párů hodnotitelů hodnotilo do stejné kategorie ku všem možným pářům hodnotitelů)

$$P_i = \frac{1}{n(n-1)} \left[\left(\sum_{j=1}^k n_{ij}^2 \right) - (n) \right]$$
- např.: $P_1 = \frac{1}{5(5-1)}(0^2 + 0^2 + 1^2 + 4^2 - 5) = 0,6$

42-19

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Fleissovo κ_F – příklad

nahrávka i	0	1	2	3	P_i
1	0	0	1	4	0,6
2	0	0	5	0	1
3	0	1	3	1	0,3
4	1	4	0	0	0,6
5	0	0	4	1	0,6
...
469	5	0	0	0	1
total	541	924	623	257	
p_j	0,23	0,39	0,27	0,11	

- $\bar{P} = \frac{1}{N} \sum_{i=1}^N P_i$
- po určení ostatních P_i a p_j dopočítáme

$$\bar{P} = 0,50$$

$$\bar{P}_e = 0,29$$
- takže lze určit

$$\kappa_F = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e} = \frac{0,50 - 0,29}{1 - 0,29} = 0,29$$

42-20

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Kvantifikace hodnot α_C

- obvykle se používá následující tabulka

α	slovní popis
$< 0,9; 1 >$	vynikající
$< 0,8; 0,9 >$	dobrá
$< 0,7; 0,8 >$	přijatelná
$< 0,6; 0,7 >$	diskutabilní
$< 0,5; 0,6 >$	slabá
$(-1; 0,5)$	nepřijatelná

- nutné brát s rezervou
- obecně platí že čím více, tím lépe

42-22

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Kvantifikace hodnot κ

- obvykle se používá následující tabulka

κ	slovní popis
$< 0,8; 1 >$	vynikající
$< 0,6; 0,8 >$	značná
$< 0,4; 0,6 >$	průměrná, mírná
$< 0,2; 0,4 >$	přiměřená, ucházející
$< 0; 0,2 >$	nepatrná
$(-1; 0)$	bezvýznamná

- nutné brát s rezervou
- obecně platí že čím více, tím lépe

42-23

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Základní náhled na hodnocení - Spearmanův korelační koeficient

Heatmaps showing Spearman correlation coefficients for /a/ and čtený text across different parameters (G, R, B, T).

parametr	/a/				Podzim			
	G	R	B	T	G	R	B	T
$\bar{O}r$	0,65	0,48	0,66	0,12	0,82	0,71	0,75	0,17
max r	0,75	0,66	0,82	0,49	0,90	0,83	0,85	0,61
min r	0,52	0,25	0,53	-0,18	0,77	0,61	0,65	-0,25

42-26

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Opakované hodnocení

Na základě Wilcoxonova rank-sum testu ověření, zda hodnotitelé nebyli při opakovaném hodnocení hodnější nebo přísnější (uvádí De BODT et al. [1997] a KARNELL et al. [2007])

hodnotitel	/a/				Podzim			
	G	R	B	T	G	R	B	T
Dezort	↓	↓	↓
Mercelová	.	.	.	↑*
Mrkvičková	.	.	.	↑	↑	↑	↑	.
Richterová	↑	.	↑	.
Vokřál	.	.	↓
všichni	↓	.	↓	↓	↑	↑	↑	.

* změna znaménka hodnocení parametru

Úvod
Data
Metodika
Výsledky
Závěr
Reference

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Intra- rater variability

Shoda při opakovaném testu

hodnotitel	Cronbachovo α_C							
	/a/				čtený text			
	G	R	B	T	G	R	B	T
Dezort	0.80	0.74	0.81	0.50	0.90	0.85	0.88	0.57
Mercelová	0.84	0.67	0.85	0.53	0.94	0.90	0.83	0.63
Mrkvičková	0.72	0.64	0.67	0.10	0.90	0.88	0.85	0.50
Richterová	0.85	0.82	0.81	0.65	0.88	0.81	0.86	0.54
Vokřál	0.91	0.83	0.89	0.48	0.98	0.96	0.96	0.62
všichni	0.95	0.91	0.95	0.56	0.98	0.96	0.96	0.62

vynikající dobrá přijatelná diskutabilní slabá nepřijatelná

Úvod
Data
Metodika
Výsledky
Závěr
Reference

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Cohenovo κ_C - /a/

	1. kolo				2. kolo			
	G	R	B	T	G	R	B	T
Me – Vo	0.39	0.24	0.30	0.05	0.38	0.24	0.40	0.03
Me – Mr	0.25	0.24	0.28	0.16	0.32	0.14	0.27	0.12
Me – Ri	0.38	0.29	0.34	0.23	0.41	0.20	0.43	0.21
Me – De	0.29	0.25	0.40	0.21	0.20	0.18	0.41	0.12
Vo – Mr	0.29	0.24	0.25	0.09	0.34	0.26	0.32	0.10
Vo – Ri	0.39	0.26	0.35	0.03	0.33	0.20	0.33	0.04
Vo – De	0.28	0.32	0.41	0.09	0.17	0.18	0.40	0.15
Mr – Ri	0.27	0.27	0.17	0.08	0.31	0.17	0.28	0.05
Mr – De	0.25	0.25	0.21	0.08	0.20	0.11	0.32	0.16
Ri – De	0.18	0.22	0.38	0.23	0.19	0.21	0.38	0.18
max	0.39	0.32	0.41	0.23	0.41	0.26	0.43	0.21
min	0.18	0.22	0.17	0.03	0.17	0.11	0.27	0.03
mean	0.30	0.26	0.31	0.13	0.29	0.19	0.35	0.12

vynikající značná průměrná přiměřená nepatrná bezvýznamná

Úvod
Data
Metodika
Výsledky
Závěr
Reference

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Cohenovo κ_C - čtený text

	1. kolo				2. kolo			
	G	R	B	T	G	R	B	T
Me – Vo	0.64	0.54	0.52	0.22	0.66	0.56	0.41	0.22
Me – Mr	0.62	0.57	0.52	0.44	0.22	0.32	0.28	0.26
Me – Ri	0.34	0.28	0.26	0.18	0.14	0.21	0.15	0.13
Me – De	0.55	0.47	0.43	0.43	0.60	0.48	0.44	0.46
Vo – Mr	0.65	0.58	0.57	0.14	0.26	0.27	0.35	0.05
Vo – Ri	0.33	0.22	0.32	0.03	0.16	0.16	0.16	-0.01
Vo – De	0.57	0.53	0.55	0.22	0.57	0.45	0.46	0.17
Mr – Ri	0.41	0.34	0.23	0.09	0.49	0.29	0.38	0.25
Mr – De	0.47	0.39	0.54	0.43	0.16	0.23	0.43	0.33
Ri – De	0.20	0.16	0.27	0.16	0.09	0.19	0.29	0.17
max	0.65	0.58	0.57	0.46	0.66	0.56	0.47	0.46
min	0.20	0.16	0.23	0.03	0.09	0.16	0.15	-0.01
mean	0.48	0.41	0.42	0.24	0.34	0.32	0.34	0.20

vynikající značná průměrná přiměřená nepatrná bezvýznamná

Úvod
Data
Metodika
Výsledky
Závěr
Reference

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Fleissovo κ_F

Shoda skupiny

	/a/				čtený text			
	G	R	B	T	G	R	B	T
1. kolo	0,29	0,26	0,30	0,09	0,48	0,38	0,36	0,18
2. kolo	0,27	0,18	0,35	0,08	0,41	0,36	0,37	0,19
1. a 2. kolo	0,29	0,23	0,33	0,12	0,45	0,39	0,37	0,22

vynikající značná průměrná přiměřená nepatrná bezvýznamná

Úvod
Data
Metodika
Výsledky
Závěr
Reference

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Diskuze – Test-retest

Porovnání s článkem „Test-Retest Study of the GRBAS Scale: Influence of Experience and Professional Background on Perceptual Rating of voice Quality“ De BODT et al. [1997]

- 12 nahrávek
 - 3 ukázka
 - vyhodnocení na 9 nahrávkách
- prodloužená /a/ a /i/, foneticky balancovaný text 115 slov
- 23 hodnotitelů
 - 13 ENT (otolaryngologové)
 - 10 S/LP (speech-language pathologists – foniatři)
 - toho
 - E ... experienced, pokud se zabývali s hlasovou patologií alespoň třetinu profesního života, ale alespoň 3 roky
 - I ... inexperienced
- pouze patologické hlasy

Úvod
Data
Metodika
Výsledky
Diskuze
Závěr
Reference

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Diskuze – Test-retest

- porovnání s článkem „Test-Retest Study of the GRBAS Scale: Influence of Experience and Professional Background on Perceptual Rating of voice Quality“ De BODT et al. [1997]
- Test-retest

K_C	I	E	ENT	S/LP	All	/a/	čtený
G	0,50	0,70	0,62	0,58	0,60	0,41	0,54
R	0,29	0,40	0,27	0,43	0,35	0,33	0,52
B	0,29	0,45	0,30	0,46	0,38	0,43	0,49

vynikající
značná
průměrná
přiměřená
nepatrná
bezvýznamná

42-36

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Diskuze – Inter-rater

- porovnání s článkem „Test-Retest Study of the GRBAS Scale: Influence of Experience and Professional Background on Perceptual Rating of voice Quality“ De BODT et al. [1997]
- Agreement between observers

K_F	I	E	ENT	S/LP	All	/a/	čtený
G	0,39	0,49	0,45	0,46	0,44	0,29	0,45
R	0,16	0,20	0,14	0,25	0,17	0,23	0,39
B	0,21	0,20	0,19	0,31	0,21	0,33	0,37

vynikající
značná
průměrná
přiměřená
nepatrná
bezvýznamná

42-37

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Závěr

- všichni hodnotitelé dosahují dobrých test – retest výsledků (kromě parametru T) jak pro čtený text, tak pro prodlouženo fonaci
- nepotvrdilo se, že při opakovaném hodnocení jsou hodnotitelé přesnější
- konzistence hodnocení skupiny
 - vyšší pro čtený text
 - v porovnání s článkem máme srovnatelné, pro čtený text lehce lepší výsledky
- provedené subjektivní **hodnocení je použitelné** pro další experimenty

42-39

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Díky za pozornost

Dotazy?

Tato práce je podporována z grantů **GACR 102/12/2230** a **SGS12/185/OHK4/3T/13**.

42-40

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Bibliography I

Lee J. CRONBACH. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:297–334, 1951.

Jacob COHNEN. A coefficient of agreement for nominal scales. *Educational and Physiological Measurement*, 20:37–46, 1960.

Joseph L. FLEISS. Measuring nominal scale agreement among many raters. *Physiological Bulletin*, 76(5):378–382, 1971.

Marc S. De BODT, Floris L. WUYTS, Van de HEYNING Paul H., and Christophe CROUX. Test-retest study of the grbas scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11(1):74–80, 1997.

42-41

Hodnocení promluv pacientů s dysfonií – subjektivní testy
Adam Stránek, Roman Čmejla

Bibliography II

Michael P. KARNELL, Sarah D. MELTON, Jana M. CHILDES, Todd C. COLEMAN, Scott A. DAILEY, and Henry T. HOFFMAN. Reliability of clinician-based (grbas and cape-v) and patient-based (v-rqol and ipvi) documentation of voice disorders. *Journal of Voice*, 21(5):576–590, 2007. doi: DOI:10.1016/j.jvoice.2006.05.001.

Jungo SAWA and Toshihiko MORIKAWA. Interrater reliability for multiple raters in clinical trial of ordinal scale. *Drug Information Journal*, 41:595–605, 2007.

42-42

Hodnocení koktavosti pomocí automatických algoritmů ve čtených promluvách

Tomáš Lustyk¹, Petr Bergl, Roman ČMEJLA¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

lustytom@fel.cvut

Abstrakt: Tento příspěvek je zaměřen na objektivní hodnocení neplynulosti řeči pomocí automatických algoritmů ve čtených promluvách. Takovéto algoritmy by našly využití v oblasti logopedie a poruch řeči zejména ve spojitosti s koktavostí. Mohli by být nápomocny při objektivním určování tíže poruchy, sledování vývoje nemoci a léčby, nebo volbě vhodného léčebného postupu.

Mezi projevy charakterizující koktavost, které se objevují v řeči, se řadí repeticce, prolongace, tonoklony, četné pauzy a přerušení slov. V příspěvku je popsáno sedm algoritmů, které analyzují například pravidelnost výdeje energie v řeči, množství ticha nebo počet náhlých spektrálních změn v promluvě. Úspěšnost algoritmů je hodnocena na databázi promluv obsahující široké spektrum mluvčích (od mluvčích bez projevů koktavosti a po mluvčí s vysokým stupněm koktavosti). Pro vyhodnocení schopnosti algoritmu popsat neplynulost řeči jsou všechny algoritmy srovnávány se subjektivním hodnocením. Tato kontrolní data byla vytvořena pomocí Kondášovy stupnice a stupnice LBDL. Pearsonův korelační koeficient a ANOVA analýza slouží pro vyhodnocení úspěšnosti algoritmů.

Výsledky algoritmů naznačují, že automatické hodnocení neplynulosti by bylo možné použít pro čtené promluvy. Hlavními závěry z experimentů jsou: algoritmy jsou schopny popsat celkový dojem z promluvy; algoritmy jsou schopny rozlišit mezi jednotlivými stupni koktavosti; některé algoritmy umožňují popsat jednotlivé projevy neplynulosti zejména pak fixované projevy bez slyšitelného zvuku (pauzy). Velkou výhodou algoritmů je, že lze metody použité pro jejich výpočet lze zaměnit za jiné, například detektor náhlých změn ve spektru za jiný.

Hodnocení koktavosti pomocí automatických algoritmů ve čtených promluvách

Tomáš Lustyk, Petr Bergl, Roman Čmejla

Letní doktorandské dny 2013



Obsah

- Koktavost
- Databáze promluv, hodnocení tíže poruchy
- Algoritmy
- Srovnání hodnocení a automatických měření
- Závěr a další vývoj

2

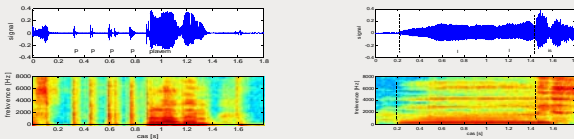
Koktavost

• Příznaky vnější

repetice (opakování hlásek nebo slabik), obrázek vlevo

prolongace (prodlužování hlásek), obrázek vpravo

tonoklony, četné pauzy, přerušení slov



• Příznaky vnitřní

3

Motivace

- Subjektivní hodnocení lékařů
- Automatický systém hodnocení tíže poruchy
- Nástroj, který by mohl pomoci při diagnóze, sledování vývoje nemoci a léčby, výběru léčebného postupu

4

Databáze promluv, hodnocení tíže poruchy

- Vznik na Foniatrické klinice 1. LF UK a VFN v Praze
- Obsahuje 160 mluvčích různého věku a různé tíže poruch
 - čtené promluvy 118 signálů
 - spontánní promluvy 134 signálů
- Hodnocení promluv (kontrolní data pro algoritmy):
 - modifikovaná Kondášova stupnice (0, 1, 2, 3, 4)
 - hodnocení LBDL (the Lidcombe Behavioral Data Language of stuttering)

5

LBDL

- uvažuje sedm charakteristik projevů koktavosti:

repetice slabik (sr) - „pa...pa...papí“

repetice hlásek (isr) - „s...s...strom“

opakování vícetlakových výrazů (msur) - „je ...je to...je to...“, „stro...stro...stromy“

fixovaný stav se slyšitelným zvukem (prolongace) (fpwaa) - „Illlíst“

fixovaný stav bez slyšitelného zvuku (pauzy, přerušení) (fpwoaa) - „lis...(bez zvuku) tí“

nadbytečné řečové projevy (svb) - „na e-e-e-na e-e-na e-obrázku“

nadbytečné neřečové projevy (video) - pohyby končetin, grimasy

kombinované:

opakované = pouze opakované projevy

fixované = pouze fixované projevy

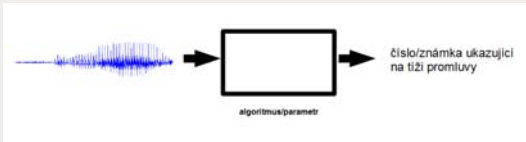
celková známka = všechny projevy

- umožňuje nalézt vhodné parametry pro určitý projev (např. který parametr popisuje nejlépe repetice), snadné pro použití

6

Algoritmy

Sedm parametrů, snaha popsat různé projevy koktavosti



- pravidelnost energie (RSE, snaží se zaznamenat nepravidelnost výdeje energie při promluvách koktavých)
- průměrná délka ticha (ALS, popisující velké množství ticha v promluvách koktavých)
- počet maxim Bayesova detektoru (ESF, využívá detektoru spektrálních změn BACD, vychází z předpokladu, že mluvčí s poruchou plynulosti řeči v promluvě za určitou dobu vysloví méně slov/hlásek než mluvčí z kontrolní skupiny)

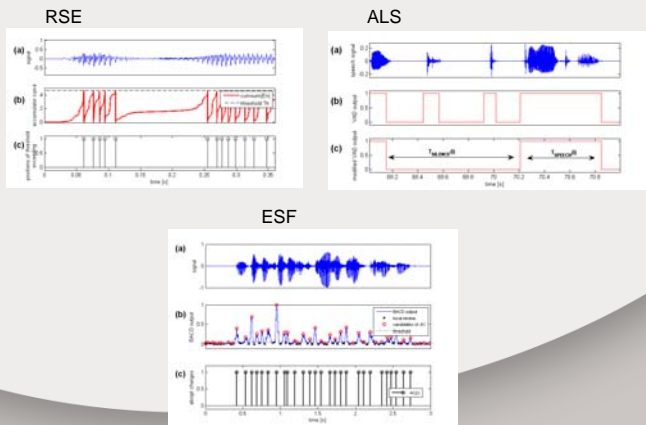
7

Algoritmy

- rozestupy při překročení prahu (SET, využívá detektor spektrálních změn, snaží se postihnout prolongace a pauzy v řeči koktavých)
- směrodatná odchylka z 11 po sobě jdoucích intervalů (SDI11, opět využívá BACD, snaží se postihnout celkovou plynulost/dojem z promluvy);
- počet spektrálních změn v krátkém intervalu (SCSI, využívá Bayesova detektoru, cílem návrhu parametru bylo postihnout častá přerušení a pauzy v řeči)
- počet spektrálních změn v úsecích řeči (NSI, využívá VAD i BACD, snaží se popisovat pauzy a přerušení v promluvě).

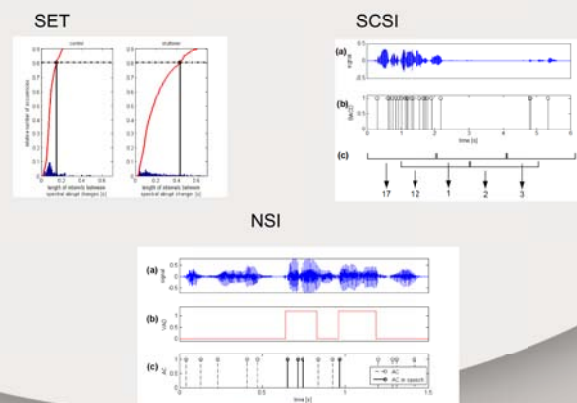
8

Algoritmy (ukázka)



9

Algoritmy (ukázka)



10

Výsledky

- Srovnání hodnot algoritmů a hodnocení lékařů a LBDL stupnice
- Čtené promluvy
- Pearsonův korelační koeficient, analýza ANOVA
- Najdou parametry rozdíl mezi jednotlivými skupinami neplynulosti?
- Je možné najít algoritmy vhodné pro popis určitých projevů koktavosti?

11

Výsledky

- Rozdíl mezi jednotlivými skupinami

	RSE	ALS	ESF	SET	SDI11	SCSI	NSI
ANOVA F(4,117)	25,72*	40,54*	42,84*	23,48*	38,65*	42,54*	42,21*
0 vs. 1	NS	NS	NS	NS	NS	NS	NS
1 vs. 2	NS	p<0,001	p<0,001	p<0,05	p<0,001	p<0,001	p<0,001
2 vs. 3	p<0,001	p<0,001	p<0,001	p<0,001	p<0,001	p<0,001	p<0,001
3 vs. 4	p<0,001	p<0,001	NS	NS	NS	NS	p<0,001

NS – nevýznamné, * p < 0,001

12

Výsledky

- algoritmy x všechny projevy kocktavosti LBDL

charakteristika	RSE	ALS	ESF	SET	SDH1	SCSI	NSI
sr	0.43	0.48	-0.49	0.45	0.46	-0.48	-0.48
isr	0.66	0.53	-0.51	0.49	0.53	-0.54	-0.50
msur	0.42	0.46	-0.54	0.56	0.56	-0.57	-0.47
fpwaa	0.39	0.25	-0.46	0.47	0.49	-0.48	-0.33
fpwoaa	0.65	0.84	-0.67	0.74	0.63	-0.72	-0.84
svb	0.37	0.31	-0.31	0.28	0.37	-0.32	-0.26
opakované	0.67	0.63	-0.63	0.61	0.64	-0.65	-0.61
fixované	0.69	0.81	-0.73	0.80	0.71	-0.78	-0.84
celková známka	0.76	0.80	-0.76	0.78	0.76	-0.80	-0.79

13

Závěr, další směr vývoje

- Algoritmy jsou schopny popsat celkový celkovou úroveň plynulosti řeči ve čtené promluvě
- Některé algoritmy umí popisat jednotlivé projevy kocktavosti, zejména pak fixované projevy bez slyšitelného zvuku (pauzy)
 - + možnost záměny jednoho nástroje za jiný
 - popis projevů opakujících se projevů a prolongací pouze v omezeném rozsahu
- Další směr práce (spontánní promluvy, systém prajucící v reálném čase)

Tato práce byla podporována z grantu GAČR P102/12/2230

14

Děkuji za pozornost

15

Literatura

Cmejla, R., Ruzs, J., Bergl, P., and Vokral, J. (2012). "Bayesian changepoint detection for the automatic assessment of fluency and articulatory disorders", *Speech Communication*, In press, available online 16 August 2012.

Cucchiari, C., Strik, H., and Boves, L. (2000). "Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology", *J. Acoust. Soc. Am.* 107, 989-999.

Kalinowski, J. (2003). "Self-reported ecacy of an all in-the-ear-canal prosthetic device to inhibit stuttering during one hundred hours of university teaching: an autobiographical clinical commentary", *Disability and Rehabilitation* 25, 107-111.

Lechta, V. and collective (2004). *Diagnoza narusene komunikacni schopnosti*. (Portal).

Teesson, K., Packman, A., and Onslow, M. (2003). "The Lidcombe bahavioral data language of stuttering", *J. Speech Lang. Hear. Res.* 46, 1009-1015.

Bloodstein, O. and Bernstein Ratner, N. (2008). *A handbook on Stuttering*, sixth edition (Delmar, Cengage Learning).

16

Charakteristiky promluv pacientů s Parkinsonovou nemocí extrahované z řečové diadochokinetické úlohy

Michal NOVOTNÝ¹, Jan RUSZ¹, Roman ČMEJLA¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

novotm26@fel.cvut.cz

Abstrakt: Přestože je Parkinsonova nemoc (PN) v současné době jedním z nejčastějších neurodegenerativních onemocnění, existuje pouze symptomatická léčba a studie ukazují, že k diagnóze může docházet až 10 let po výskytu prvních projevů. Jedním z prvních symptomů PN je hypokinetická dysartrie. Práce zaměřené na diagnostikování PN skrze hypokinetickou dysartrii využívají zejména prodlouženou fonaci, která není schopna postihnout artikulační aspekty promluvy, nebo na souvislých promluvách (čtený text, monolog), které jsou příliš složité pro automatické zpracování. Jako řešení tohoto problému se jeví řečová diadochokinetická (DDK) úloha, založená na rychlém opakování slabik /pa/-/ta/-/ka/. Tato úloha díky rychlým opakovaným pohybům artikulačního ústrojí odráží nedostatky v načasování a přesnosti artikulačních pohybů a zároveň díky pevně danému rámci úlohy poskytuje dostatečné zjednodušení pro automatické zpracování. Výsledky této práce potvrzují možnost hodnocení různých aspektů PN artikulace získaných z DDK úlohy. Ze 12 navržených biomarkerů rozříděných do 6 skupin zachycujících různé dimenze řeči, ukázalo statisticky signifikantní rozdíly mezi PN a kontrolní skupinou 8 biomarkerů ze všech skupin. Navíc výsledek klasifikace (celková úspěšnost detekce odpovídající 88.3 ± 8.1 %) SVM klasifikátoru založeného právě na těchto biomarkerech prokazuje použitelnost úlohy v případě detekce PN promluv.

1. Úvod

Parkinsonova nemoc (PN) je degenerativní poškození mozku, které postihuje zejména lidi po padesátém roce života [1]. PN poškozuje dopaminové cesty v substantia nigra, což vede k postupnému ubývání dopaminu v této oblasti [2]. Dopaminový deficit následně vede zejména k motorickým, ale i senzorickým obtížím, poruchám spánku, nálad a paměti [2].

Incidence Parkinsonovy nemoci (PN) rovná 100 případů na 100 000 obyvatel řadí tuto nemoc na druhé místo mezi neurodegenerativními onemocněními hned za Alzheimerovu nemoc [1]. Toto onemocnění navíc spadá do rozsáhlé skupiny označované pojmem parkinsonismus. Choroby řazené do této oblasti se vyznačují čtyřmi základními společnými symptomy (bradikynézie či akinézie, svalová rigidita, tremor, nestabilní postoj) [2]. Parkinsonova nemoc navíc reaguje na léčbu prekurzorem dopaminu tzv. L-dopou. Nicméně podávání L-dopy poskytuje pouze symptomatickou léčbu, která nemá vliv na progresi onemocnění [3]. Navíc dřívější studie prokázaly, že k výskytu motorických symptomů využívaných pro diagnózu dochází při poklesu dopaminu o přibližně 80 % původní hodnoty, což znamená, že diagnóza může být zpožděna až 10 let za propuknutím nemoci [2], [4].

Studie [4] ukázala, že existují subklinické symptomy, které se projevují již před manifestací základních motorických symptomů. Mezi tyto symptomy patří podle [postumna] zejména poruchy spánku a řeči. Poruchy řeči označované jako hypokinetická dysartrie shrnují řadu symptomů spojených se sníženou pohyblivostí artikulačního ústrojí. Mezi tyto symptomy patří chraptivost, monopitch, snížená hlasitost, palilálie, snížená srozumitelnost, hypernasalita, bradykinézie, rigidita a nepřesná artikulace.

Mezi nejčastěji používané úlohy zaměřené na hodnocení dysarthrických promluv patří prodloužená fonace a souvislé promluvy (čtený text, monolog). Nevýhodami těchto úloh je přílišná složitost zpracování v případě souvislých promluv a zaměření pouze na symptomy spojené s dysfonií v případě prodloužené fonace. Jednou z dalších možností je tzv. diadochokinetická (DDK) řečová úloha založená na rychlém opakování slabik /pa-/ta-/ka/. Výhodou této úlohy je, že spojuje jednoduchost zpracování pevně dané úlohy s měřením artikulačních schopností podporovaných charakterem úlohy. Díky těmto vlastnostem byl navržen algoritmus automatické segmentace jednotlivých promluv, který v každé slabice označil pozice exploze (E), počátky vokálu (V) a konce vokálu (O) a tím umožnil efektivně zpracovat velké množství promluv [5]. Pro potřeby následného zpracování segmentovaných promluv bylo třeba navrhnout biomarkery pomocí kterých by bylo možné hodnotit různé aspekty dysarthrické řeči. Ve studii je navrženo dvanáct biomarkerů rozdělených do šesti skupin spojených s různými fyziologickými oblastmi řeči.

V této práci jsou představeny jednotlivé biomarkery a je ohodnocena jejich statistická významnost. Navíc pro potřeby určení použitelnosti automatického zpracování byl realizován *Support Vector Machine* (SVM) klasifikátor, který má za úkol automaticky rozdělit promluvy z PN a kontrolní skupiny (KS).

2. Metodika

Účastníci

Data byla nahrána od 46 rodilých českých mluvčích, ze kterých bylo ve skupině 24 (20 mužů a 4 ženy) diagnostikována PN. Všichni mluvčí s PN byly vyšetřeny krátce po diagnóze, ještě před nasazením léčby. Střední věk PN skupiny činí 60.9 (\pm směrodatná odchylka 12.6 let), střední doba trvání onemocnění 31.3 \pm 22.3 měsíce; stádium nemoci bylo na stupnici Hoehna a Yahrové rovné 2.2 \pm 0.5, podle Unified Parkinson's Disease Rating Scale (UPDRS) III 17.4 \pm 7.1.

Kontrolní skupina zahrnuje 22 dobrovolníků (15 mužů and 7 žen) se středním věkem 8.7 \pm 14.6 let, dobrovolníci nemají žádnou předchozí neurologickou historii. Obě jsou věkově vyvážené a neexistují žádné větší rozdíly věkového rozložení. Nikdo z účastníků (PN ani KS) neprodělal žádnou řečovou terapii. Studie byla schválena etickou komisí Všeobecné fakultní nemocnice v Praze a každý účastník poskytl informovaný souhlas.

Nahrávky

Nahrávání bylo prováděno v tiché místnosti pomocí kondenzátorového mikrofonu ve vzdálenosti 15 cm od úst. Data byla následně ukládána do osobního počítače se vzorkovací frekvencí 48 KHz a kvantována na 16 bit. Každá promluva byla nahrávána řečovým terapeutem na jednotlivých sezeních, kde byli všichni účastníci požádáni, aby rychle opakovali sérii slabik /pa-/ta-/ka/ tak rychle, pravidelně a dlouho jak dokáží. Každý účastník nahrál dvě nahrávky.

Celkově bylo nahráno 1644 slabik, z nichž 753 bylo ze skupiny PN and 891 ze skupiny KS. Tento zdánlivý nepoměr je způsoben sníženou schopností PN účastníků provádět dlouhé DDK promluvy a je kompenzován tím, že jsou biomarkery počítány pouze z prvních pěti opakování.

Biomarkery

Některé biomarkery nebylo možné hodnotit celkově pro všechny slabiky dohromady, a proto byla provedena měření pro každou slabiku zvlášť, v jiných měřeních pak byla k celkovému hodnocení doplněna měření jednotlivých slabik jako rozšiřující. Výsledkem je tak 24 měření pro 12 biomarkerů. Výstupy jednotlivých měření jsou uvedeny v tabulce Tab. 1, která shrnuje střední hodnoty a směrodatné odchylky pro KS a PN skupinu. Navíc je pro každý biomarker pomocí Wilcoxonova testu stanovena hodnota statistické významnosti.

Kvalita hlasu

První skupina je zaměřená na artikulaci vokálu a sestává se ze tří parametrů. Prvním parametrem je kvocient podobnosti vokálu (Vowel Similarity Quotient - VSQ) hodnotící pomocí autokorelace celé samohlásky kvalitu kmitání hlasivek a přesnost nastavení laryngeálních artikulátorů. Podobným parametrem je kvocient podobnosti prvních 30ms vokálu (VSQ₃₀), který také využívá autokorelace, tento parametr vychází z předchozí studie [6], kde je 30ms okno doporučeno pro velkou citlivost na periodicitu. Načasování přenastavení artikulátorů v průběhu fonace hodnotí kvocient variability vokálu (Vowel Variability Quotient - VVQ), který pomocí směrodatné odchylky popisuje pravidelnost délky jednotlivých vokálů.

Koordinace laryngeální a supralaryngeální aktivity

Druhá skupina zahrnuje biomarkery hodnotící koordinaci laryngeálních (hlasivky, laryngeální svaly) a supralaryngeálních (čelist, jazyk, rty) artikulátorů. Parametr voice onset time (VOT) je definovaný jako vzdálenost mezi počátkem konsonanty a počátkem znělosti. Normalizovaný VOT (VOT_n) je pak hodnotou VOT normovanou délkou celé slabiky. Oba parametry VOT a VOT_n popisují přesnost načasování počátečního uvolnění konsonanty (supralaryngeální artikulátory) a nástupu znělosti (laryngeální artikulátory).

Přesnost artikulace konsonanty

Nepřesnost artikulace konsonanty se může projevit jako nedostatky v polohování supralaryngeálních artikulátorů a může vést ke zvýšení šumové složky ve spektru konsonanty. Z tohoto důvodu jsou parametry hodnotící přesnost artikulace zaměřeny na spektrální oblast: jedná se o sklon spektra (Consonant Spectral Trend - CST) a první spektrální moment (Consonant Spectral Momentum - CSM). Parametr CSM je počítán jako těžiště celého spektra, kdežto parametr CST je definován jako sklon přímky proložené částí spektra definovanou v [7] (/p/ [2750 - 3250], /t/ [2250 - 2750], /k/ [1750 - 2250]).

Pohyby jazyka

Čtvrtá kategorie popisuje pohyby jazyka a je spojena s druhým formantem. Sklon druhého formantu (Second Formant Trend - SFT) je počítán jako úhel regrese druhé formantové frekvence vokálu.

Slábnutí okluze

Slábnutí okluze se projevuje jako nedostatečné uzavření artikulátorů v okamžiku pauzy, jež vede k turbulentnímu proudění vzduchu. Turbulence poté produkují zvýšenou šumovou složku, která se projevuje snížením poměru signál šum a tedy celkovým snížením parametru SNR (Signal-to-Noise Ratio).

Celkové načasování řeči

Rychlost promluvy je hodnocena pomocí parametru DDK rate udávajícího počet za sekundu vyřčených slabik. Druhým parametrem hodnotícím rychlost z pohledu pauz je DDK pace, popisující průměrnou délku pauzy mezi jednotlivými slabikami. Parametrem hodnotícím pravidelnost promluvy je DDK variability založená na směrodatné odchylce délek pauz.

Klasifikátor

Pro určení další využitelnosti navržených biomarkerů byl realizován SVM klasifikátor s radiálním kernelovským jádrem, který umožňuje pomocí dvou parametrů (penalizační σ a kernelovský C) nastavit flexibilitu (C) a přímou (σ) rozhodovací roviny. V našem případě byly oba parametry rovny dvěma, což poskytovalo přiměřenou flexibilitu a přitom ještě nedocházelo k přeučení. Poté byly hledány možné kombinace jednotlivých biomarkerů, poskytující co možná nejlepší poměr správně klasifikovaných promluv.

3. Výsledky

Biomarkery

Výsledky pro skupinu zaměřenou na *kvalitu hlasu* ukazují statisticky signifikantní pokles kvality PN hlasu pouze v případě parametru VSQ₃₀. V kombinaci se shodnými výsledky VSQ a VSQ₃₀ pro kontrolní skupinu, naznačuje signifikantní VSQ₃₀ silnější dopad PN na artikulační pohyby při nástupu znělosti než na artikulaci ukončení znělosti. Parametr VVQ ukázal zvýšenou variabilitu délek vokálu, nicméně tyto rozdíly nedosáhly dostačující úrovně signifikance, což může být způsobeno například vyšší citlivostí na rozdíly mezi jednotlivými mluvčími, která zastíní rozdíly mezi KS a PN skupinou.

Všechny parametry ze skupiny *koordinace laryngeálních a supralaryngeálních svalů* ukazují významné rozdíly. Prodloužení délky VOT u parkinsoniků ukazuje problémy s rychlou koordinací různých skupin svalů. Výsledky navíc ukazují, že klasický parametr VOT poskytuje lepší výstupy, než délkou slabiky normalizovaný VOTn, navržený v práci [8]. Z toho je možné usuzovat, že jednoslabičný charakter DDK úlohy poskytuje dostatečnou normalizaci.

Tab.1 Výsledky měření jednotlivých biomarkerů

Biomarkery	KS		PN		Wilcoxonův test
	Mean	STD	Mean	STD	
Kvalita hlasu					
VSQ	0.45	0.10	0.40	0.14	p = 0.135
VSQ30	0.45	0.10	0.36	0.11	p < 0.050
VVQ (ms)	50.62	6.82	55.37	11.05	p = 0.136
Koordinace laryngeální and supralaryngeální aktivity					
VOT: all (ms)	20.68	6.23	34.15	6.88	p < 0.001
VOT: pa (ms)	14.32	4.42	28.20	7.02	p < 0.001
VOT: ta (ms)	22.85	8.81	35.57	10.46	p < 0.001
VOT: ka (ms)	24.86	9.55	38.78	7.59	p < 0.001
VOTn: all (%)	27.87	6.66	28.78	23.33	p < 0.050
VOTn: pa (%)	21.72	5.54	14.45	64.21	p < 0.010
VOTn: ta (%)	29.69	8.15	34.07	15.57	p < 0.050
VOTn: ka (%)	32.19	9.36	38.00	10.31	p < 0.050
Přesnost artikulace konsonanty					
CST: pa (rad *10 ⁻³)	-0.91	0.27	-0.90	0.49	p = 0.913
CST: ta (rad *10 ⁻³)	-1.67	1.89	-1.22	1.05	p < 0.050
CST: ka (rad *10 ⁻³)	-27.03	20.06	-19.68	21.23	p = 0.074
CSM: pa (kHz)	1.858	0.276	1.843	0.379	p = 0.504
CSM: ta (kHz)	2.116	0.566	2.459	1.002	p = 0.362
CSM: ka (kHz)	2.208	0.368	2.194	0.514	p = 0.724
Pohyby jazyka					
SFT: pa (rad)	-0.05	0.27	-0.03	0.25	p = 0.913
SFT: ta (rad)	0.55	0.23	0.28	0.25	p < 0.010
SFT: ka (rad)	-0.54	0.19	-0.46	0.18	p = 0.091
Slábnutí okluze					
SNR (dB)	16.00	3.77	13.86	4.49	p < 0.050
Celkové načasování řeči					
DDK rate (slabik/s)	7.68	0.66	6.68	0.88	p < 0.001
DDK pace (ms)	64.26	10.87	67.68	18.66	p = 0.799
DDK variability(ms)	0.14	0.20	0.62	0.99	p < 0.010

Spektrální charakteristiky popisující *přesnost artikulace konsonanty* potvrzují nepřesné nastavení artikulačního ústrojí. Parametr CSM, díky snížené citlivosti na méně energetické změny ve vyšších částech spektra, sice nepotvrdil prokazatelný rozdíl, nicméně měření CST: ta prokázalo signifikantní rozdíl v nastavení artikulačního ústrojí mezi KS a PN skupinou. Navíc měření CST: ka ukazuje také zhoršení PN artikulace, a ač nedosahuje hladiny statistické významnosti, poukazuje tento fakt na ovlivnění pohybu jazyka.

Výsledky skupiny zaměřené na *přesnost artikulace konsonanty* jsou ve shodě s výsledky získanými z měření cílených na *pohyby jazyka*. Výsledky této skupiny potvrzují signifikantní snížení pohybů jazyka v případě slabiky /ta/ a nesignifikantní snížení pohybů jazyka v případě slabiky /ka/ u skupiny PN.

Měření týkající se *Slábnutí okluze* ukázalo u parkinsoniků signifikantní snížení parametru SNR. Tento výsledek potvrzuje zvýšení šumové složky v pauzách mezi jednotlivými slabikami a tento nárůst šumu naznačuje existenci turbulentního proudění způsobeného slábnutím okluze.

Celkové načasování řeči u parkinsoniků ukazuje, díky sníženému DDK rate, na snížení rychlosti, navíc nezměněný DDK pace a zvýšená DKK variability odhalují na efekt popisovaný jako krátké uspěchané řečové segmenty (short rushes of speech [9]), kdy jsou části řeči prováděny v rychlejším tempu, ale zároveň je mezi jednotlivými segmenty delší mezera. Uvážíme-li navíc výsledky skupin zaměřených na *přesnost artikulace konsonanty* a *pohyby jazyka* můžeme předpokládat, že tyto zrychlené segmenty jsou jednotlivé sekvence /pa/-/ta/-/ka/.

Klasifikace

Tabulka 2 ukazuje vysokou míru úspěšnosti detekce (88 %), která je srovnatelná s jinými studii dosahujícími (81 %) pro kombinaci různých měření a databázi založenou na brzkých stádiích PN [10], a pro úlohy založené na prodloužené fonaci a databázi obsahující také pozdější stadia, u kterých je možné předpokládat více patrné postižení řeči (97 %) [11], (89%) [12]. Nejlepší výsledek klasifikace pro jediné měření VOT: all rovný 81 % navíc ukazuje, že v brzkých stádiích má PN spíše vliv na artikulaci než na fonaci což je v rozporu s pracemi [13], [14]. To může být způsobeno zejména zkoumáním brzkých stádií PN ve kterých nedošlo k tak výrazné progresi, a tedy ani tak pokročilému poškození řeči.

4. Závěr

Prezentované výsledky ukazují, že řečová DDK úloha poskytuje dostatečné zjednodušení pro automatické zpracování a zároveň zachovává potřebné informace pro klasifikaci s vysokou mírou úspěšnosti. Parametry použité pro nejlepší klasifikátor navíc potvrzují relevanci z fyziologického hlediska, které je pro autory jedním z hlavních kritérií. Výsledky dosažené pouze na jedné úloze jsou srovnatelné s dalšími pracemi, a navíc poskytují možnost zlepšení výsledků kombinací s dalšími detektory založenými na vyšetřování fonace, či na měření faciokineze z pořízených videí.

Poděkování

Tato práce je podporována z grantů SGS12/185/OHK4/3T/13, GACR 102/12/2230.

Tab.2 Hodnocení biomarkerů

Použité biomarkery	Popis	Úspěšnost detekce (%)
VSS, VOT: All, CST: ta, SFT: ta, SNR, DDK rate	Nejlepší	88.3 ± 8.1
Všechny	Nejlepší pro všechny biomarkery	84.6 ± 8.8
VOT: all	Nejlepší pro jeden biomarker	80.9 ± 8.9

Reference

- [1] Van Den Eeden S. K.; Tanner C. M.; Bernstein A. L.; Fross R. D.; Leimpeter A.; Bloch D. A.; Nelson L. M. Incidence of Parkinson's disease: Variation by age, gender, and race/ethnicity. *Am. J. Epidemiol.* (2003) 157, 1015-1022.
- [2] Rodriguez-Oroz M. C.; Jahanshahi M.; Krack P.; Litvan I.; Macias R.; Bezard E.; Obeso J. A. Initial clinical manifestations of Parkinson's disease: features and pathophysiological mechanisms. *Lancet Neurol.* (2009) 8, 1128-1139.
- [3] Djaldetti, R., & Melamed, E. Management of response fluctuations: Practical guidelines. *Neurology* (1998) 51(Suppl. 2), S36–S40.
- [4] Postuma R. B.; Lang A. E.; Gagnon J. F.; Pelletier A.; Montplaisir J. Y. How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. *Brain* (2012) 135, 1860-1870.
- [5] M. Novotný Automatická segmentace hlásek při rychlém opakování slabik (/pa/ – /ta/ – /ka/) u hypokinetické dysartrie Letní doktorandské dny Praha 2012.
- [6] Sapir S.; Ramig L. O.; Spielman J. L.; Fox C. Formant centralization ratio: a proposal for a new acoustic measure of dysarthric speech. *J. Speech Lang. Hear. Res.* (2010) 53, 114–125.
- [7] Hansen J. H. L.; Gray S. S.; Kim W. Automatic voice onset time detection for unvoiced stops (/p/, /t/, /k/) with application to accent classification. *Speech Communication* (2010) 52, 777-789.
- [8] Fischer E.; Goberman A. M. Voice onset time in Parkinson's disease. *J. of Communication Disord.* (2010) 43, 21-34.
- [9] Darley F. L.; Aronson A. E.; Brown J. R. Motor speech disorders. *Philadelphia: W. B. Saunders* (1975).
- [10] Rusz J.; Čmejla R.; Ružičková H.; Klempíř J.; Majerová V.; Picmausová J.; Roth J.; Růžicka E. Acoustic assessment of voice and speech disorders in Parkinson's disease through quick vocal test. *Movement Disorders*, (2011) 26(10): 1951-1952.
- [11] Tsanas A.; Little M. A.; McSharry P. E.; Spielman J.; Ramig L. O. Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Transactions on Biomedical Engineering* (2012) 59, 1264-1271.
- [12] Little M. A.; McSharry P. E.; Hunter E. J.; Spielman J.; Ramig L. O. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE Transactions on Biomedical Engineering* (2009) 56(4), 1015-1022.
- [13] Logemann J. A.; Fisher H. B.; Boshes B.; Blonsky E. R. Frequency and cooccurrence of vocal tract dysfunction in the speech of a large sample of Parkinson patients. *J. Speech Hear. Disord.* (1978) 43, 47–57.
- [14] Zwirner P.; Barnes G. Vocal tract steadiness: A measure of phonatory stability and upper airway motor control during phonation in dysarthria. *J. Speech Hear. Res.* (1992) 35, 761–768.

Objektivní metody hodnocení důrazu u Parkinsonovy nemoci

Tereza TYKALOVÁ¹

¹ České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů

tykalter@fel.cvut.cz

Abstrakt:

Cíle: Ačkoli snížená schopnost vyjádřit důraz je považována za jednu z nejvíce postižených řečových dimenzí u Parkinsonovy nemoci (PN), mechanismus vyjádření důrazu nebyl doposud prozkoumán s použitím objektivních akustických metod. Hlavním cílem této práce je kvantifikovat efekt Parkinsonovy nemoci na prosodické charakteristiky a popsat mechanismus tvorby důrazu u lidí s PN.

Metody: Schopnost vyjádřit větný důraz byla zkoumána u skupiny 20 mužů s PN v brzké fázi a porovnána s 20 zdravými jedinci stejného věku a pohlaví. Každý účastník studie byl požádán o přečtení krátkého textu s pěti vyznačenými slovy. Cílem úlohy bylo zdůraznit těchto pět vyznačených slov v kontextu jednotlivých vět. Akustické analýzy se skládaly z měření základní hlasivkové frekvence (F0), intenzity a délky trvání slov. Dále bylo navrženo nové měření nazvané Stress Pattern Index (SPI), které odráží efekt všech tří základních prosodických charakteristik využívaných při vyjádření důrazu.

Výsledky: Ačkoli pacienti s PN vykazovali sníženou schopnost vyjádřit důraz, stále byli schopní významně zvýšit F0, intenzitu a také výrazně prodloužit délku slov. Ve skutečnosti, nebyly nalezeny žádné statisticky významné rozdíly mezi PN a zdravými jedinci s využitím měření F0, intenzity, délky trvání slov a rozsahu intenzity. Nicméně jedinci s PN vykazovali významně snížené hodnoty u měření SPI a rozsahu F0.

Závěr: Snížená schopnost vyjádřit důraz byla pozorována dokonce již u pacientů v brzkých stádiích PN. Schopnost zlepšení akustických ukazatelů při cíleném důrazu naznačuje důležitost zvážení terapie řeči u osob s PN.

Objektivní metody hodnocení důrazu u Parkinsonovy nemoci (PN)

T. Tykalová

školitel: Doc. Ing. Roman Čmejla, CSc.

Letní doktorandské dny 2013

01/16

Parkinsonova nemoc

- neurodegenerativní onemocnění
- **výskyt:** 1-2% lidí nad 60 let
- **úbytek mozkových buněk** v substantia nigra → pokles hladiny dopaminu → špatný přenos signálů mezi neurony
- **hlavní příznaky:** poruchy hybnosti (třes, ztuhlost svalů, neschopnost započít pohyb)
- **další příznaky:** poruchy řeči u 90% pacientů, často jeden z prvních příznaků nemoci

02/16

Motivace

OBECNÁ motivace:

- porozumění problematice
- stanovení včasné a přesné diagnózy
- monitorování progresu nemoci – biomarker
- monitorování efektu léčby
- zpětná vazba při terapii řeči

03/16

Motivace

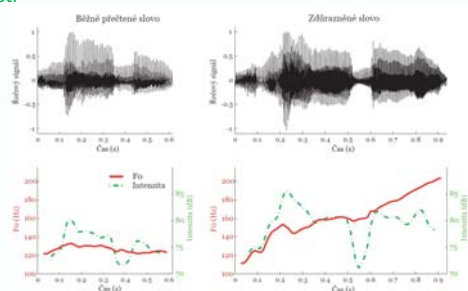
MOTIVACE pro výběr tématu:

- osoby s PN mají sníženou schopnost vyjádřit důraz
- jedna z nejvíce postižených řečových dimenzí u PN (Darley, 1969 – Differential diagnostic patterns of dysarthria)
- přesto téměř žádné akustické studie (Ma, 2010 - Intonation contrast in Cantonese speakers with hypokinetic dysarthria associated with Parkinson's disease; Cheang, 2007 - An acoustic investigation of Parkinsonian speech in linguistic and emotional contexts)

04/16

Princip vyjádření důrazu

- prodloužení slova
- **zvýšení FO**
- **zvýšení hlasitosti**



05/16

Princip vyjádření důrazu - závěr

Parametr	při DŮRAZU	v důsledku NEMOCI (PN vs. KS)
délka slova	↑	↓
F0 (základní hlasivková frekvence)	↑	↓
Intenzita (hlasitost)	↑	↓

06/16

Databáze – pacienti

- 36 osob mužů → 20 osob s diagnózou PD
- brzká stádia
 - bez farmakoterapeutické léčby
- 16 zdravých osob = kontrolní skupina (KS)
- bez historie neurolog. onemocnění
 - bez historie komunikačních problémů

07/16

Databáze – řečová úloha

- cíl úlohy: během čtení nepřírozně zdůraznit slova vyznačená v textu

Dnes jsme to již nestihli,
možná **ZÍTRA** navštívíme všechny své známé.

Příbuzné jsme již navštívili,
možná zítra navštívíme všechny své **ZNÁMÉ**.

08/16

Databáze – řečová úloha

- výsledek: pro 5 slov - 2 x verze přečtení

zdůrazněná

běžně přečtená

ZNÁMÉ **SVÉ**
ZÍTRA **VŠECHNY**
NAVŠTÍVÍME

známé své
zítra všechny
navštívíme

09/16

Metodika – akustická měření

- Stanovení 5 standardně používaných parametrů

- střední hodnota F0
- rozsah F0
- střední hodnota intenzity
- rozsah intenzity
- délka promluvy

+ návrh nového **parametr SPI** (Stress Pattern Index), který odráží efekt všech základních parametrů (tj. F0, intenzity a délky promluvy)

$$SPI = (1 + \ln(F0_{\max} / F0_{\min})) \sum E_n$$

10/16

Metodika – statistické vyhodnocení

- 1) Hodnoty parametrů rozděleny do 2 skupin dle způsobu přečtení:

- skupina DURAZ – 1 x zdůrazněná realizace
- skupina NORMA – průměr z 4 x běžně čtená realizace

- 2) Statisticky významné rozdíly mezi PN a KS

- víceměrná analýza rozptylu (ANOVA)

11/16

Interpretace výsledků ANOVA

	F0	Intenzita	Délka slova	F0 - rozsah	Intenzita - rozsah	SPI
SPEAKER GROUP	-	<0.001	<0.001	<0.001	<0.001	<0.001
STRESS CONDITION	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
SPEAKER GROUP × STRESS CONDITION	-	-	-	0.0014	-	<0.001

Daný efekt nebo interakce bude statisticky významná jestliže:

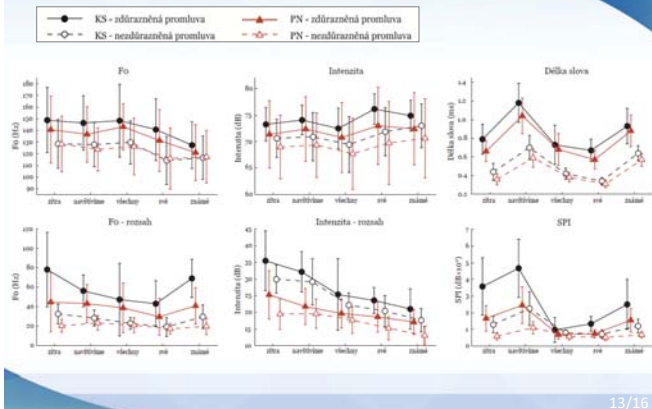
SPEAKER GROUP (PN × KS) – budou nalezeny rozdíly mezi skupinou PN a KS
-> osoby s PN budou mít poškozenou řeč

STRESS CONDITION (DURAZ × NORMA) – budou nalezeny rozdíly mezi zdůrazněnou a běžně čtenou verzí -> parametr bude využíván při tvorbě důrazu

SPEAKERS GROUP × STRESS CONDITION – budou nalezeny rozdíly mezi skupinami v závislosti na typu čtení -> osoby s PN nebudou schopni vyjádřit důraz tak efektivně jako KS

12/16

Výsledky



Výsledky - shrnutí

- pacienti s PN mají **sníženou schopnost** vyjádření důrazu
- **!nicméně jsou stále schopni ho vyjádřit!**
- pacienti s PN jsou schopni **cíleně zlepšit řečový výkon** z hlediska všech základních parametrů (F0, intenzita, délka promluvy)
- nově navržený **parametr SPI je vhodný** pro detekci důrazu u PN

14/16

Závěr

- zlepšení akustických ukazatelů při cíleném důrazu naznačuje důležitost **zvážení terapie řeči** u PN
- nově navržený parametr SPI může být vhodným ukazatelem pro hodnocení **poruch řeči** a monitorování **efektů léčby**

15/16

Děkuji za pozornost

Reference:

- Darley FL, Aronson AE, Brown JR. Differential diagnostic patterns of dysarthria. *J Speech Hear Res.* 1969;12: 246–269.
- Ma JKY, Whitehill TL, So SYS. Intonation contrast in Cantonese speakers with hypokinetic dysarthria associated with Parkinson's disease. *J Speech Lang Hear R.* 2010;53: 836–849.
- Cheang HS, Pell MD. An acoustic investigation of Parkinsonian speech in linguistic and emotional contexts. *J Neurolinguist.* 2007;20: 221–241.

16/16