

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Analyza poruch hlasu u Parkinsonovy nemoci pomocí základní hlasivkové frekvence
Jméno autora:	Vojtech Illner
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra radioelektroniky
Oponent práce:	Ing. Tomáš Bořil, Ph.D
Pracoviště oponenta práce:	Fonetický ústav FF UK

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání vyžadovalo nastudovat určité teoretické partie odpovídající spíše magisterské úrovni.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	

Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Kromě překlepů, kterým se práce takového rozsahu pochopitelně jen těžko vyhne (např. vocal chords místo cords či folds, hypokinetická apod.) je však přítomen častý problém s interpunkcí a problém s dělením slov (mater-iálů, ne-jpravděpodobnější, defin-ujeme, ne-jlepší). Rušivě působil překlad „semitóny“, když existuje ustálené pojmenování půltóny.	
Úvodní kapitoly nejsou logicky členěny. Sekce 1.2 začíná odkazem 23 let staré přehledové publikace, ze které je citováno shrnutí, že existuje velké množství estimátorů F0 a žádný přístup se neukázal jako objektivně nejlepší. Následuje nijak nezdůvodněný výběr 3 metod pro jejich práci a jejich podrobnější popis. Teprve kapitola 2 začíná sekcí 2.1 Výběr estimátorů, která odkazuje na aktuální studii z roku 2014 porovnávající 10 estimátorů, ze kterých se autor bakalářské práce rozhodl vybrat právě 3 již dříve zmiňované metody. Tato sekce měla být uvedena na začátku části 1.2.	

Výběr zdrojů, korektnost citací	A - výborně
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	
V několika případech nejsou citovány původní zdroje, které by byly pro daný účel vhodnější (např. stupnice ERB).	

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Str. 3 – není zřejmé, zda estimátor F_0 pracuje se signálem $U(t)$, o kterém se ze začátku mluví jako o zdrojovém signálu hlasivek (viz též obr. 1.2), anebo o výsledném řečovém zvuku, který prošel celým vokálním aparátem. Text odstavce 1.1.2 působí na čtenáře, jako by tím bylo myšleno podle potřeby nějak tak vše dohromady.

Podrobnější popis 3 metod v sekci 2.1 se snaží ve čtenáři vytvořit pocit, že mu princip metod objasňuje. Speciálně u metody SWIPE dochází fakticky spíše k opaku, postup odvození působí až magicky:

„Hannovo okno bylo zvoleno jako vhodný kompromis mezi jeho spektrálními vlastnostmi ... a výpočetní náročností.“

– v tuto chvíli mi není zřejmé, čím je Hannovo okno méně výpočetně náročné než jiná běžně používaná okna.

Nebo (parafrázuji): „Ze spektra je vypočten modul, který je vhodné logaritmovat, protože takový postup je použit např. u výpočtu kepstra; pro náš případ však tento přístup není vhodný, jelikož logaritmus způsobuje velká údolí. Byl tedy zvolen alternativní přístup, kdy je modul spektra umocněn na $1/2$. Tato volba přinesla nejlepší výsledky oproti jiným možnostem, jako např. umocněním na druhou nebo neupravovat modul spektra vůbec.“

U sekce kapitoly 2.2.1 „Gold standard“ mohu pro případný další výzkum doporučit zvážení volně dostupných databází řeči obsahující zdrojový elektroglografický (EGG) signál, z nichž vypočítaný odhad F_0 by měl být mnohem blíže „ground truth“ než hodnoty odhadované (a vizuálně a percepčně korigované) z výsledného řečového signálu.

Str. 44 – Pearsonův chí-kvadrát test pro testování normality dat: jako vhodnější metodu bych doporučil specializovaný Shapiro-Wilkův test (nebo pro obecná rozdělení Kolmogorov-Smirnovův test), které nevyžadují umělé rozdělení spojitých hodnot do intervalů, tedy v podstatě vytváření „histogramu“.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Práce se zabývá aktuálním tématem a přináší závěry, které mohou dílčím způsobem zkvalitnit výpočty koeficientů vhodných pro vyhodnocování poruch hlasu u Parkinsonovy nemoci. Výběr posuzovaných metod je rozumný a způsob posuzování jejich úspěšnosti je adekvátní úrovni bakalářské práce. Oceňuji sekci 2.4.2 a následující, které se snaží objevit a pochopit úskalí hodnocení jednotlivých metod, přitom autor navrhuje mechanismy filtrování chyb.

Přestože formální úroveň práce není zcela dokonalá, důležitější je odborná úroveň, kde autor nejen musel nastudovat náročnější partie, ale zároveň práci obohotali i vlastními úvahami a nápady na zlepšení, které skutečně vedly ke stabilnějším výsledkům.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Otázky k obhajobě

Str. 4: Uvádíte, že referenční hodnotu pro pultóny určíte dostatečně nízkou, abychom nemohli dostávat záporná čísla. Proč by vadila záporná čísla?

Proč je ideální hodnotou následně zvoleno 60 Hz? Na str. 5 a 21 uvádíte jako minimum F_0 50 Hz.

Str. 20, sekce 2.2.1: „vizuální a poslechová kontrola F_0 pro vytvoření gold standardu“ – jak často byla v promluvách přítomna třepená fonace? Jak jste se s ní vypořádali, a to zejména na koncích nádechových úseků? Vzhledem k typické neperiodičnosti takových úseků estimátory často odhadují víceméně náhodné hodnoty, mnohdy je též detekován stoupající trend F_0 , přestože poslechově se jedná o klesající melodém a náhodná třepenost je spíše „zabarvením“ hlasu a signalizací, že se mluvčí chtěl dostat s hodnotou F_0 pod svoje fyziologické minimum. Upravovali jste tyto úseky podle zamýšleného (a slyšeného) směru intonace, nebo jste jim ponechali náhodný charakter, nebo je z množiny gold standardu úplně vyřadili?

Str. 21, vstupní parametry: i když asi tuším důvod, mohl byste vysvětlit, proč F_{0max} bylo zvoleno 500 Hz, když v úvodu (str. 5) zmiňujete jako maximum pro F_0 hodnotu 800 Hz?

Str. 21, zvolený timestep = 0.01 sekund jako kompromis mezi přesností a výpočetní náročností („dostatečná jemnost“ vs. velký počet vzorků). Dovedete odborněji zdůvodnit volbu kroku 10 ms? Jak přesné rozlišení pro analýzu Parkinsonovy potřebujete? Jak tato hodnota souvisí s typickým trváním hlásek, jedné periody a množstvím událostí, které potřebujete pro vyhodnocení zachytit?

Str. 22, Spearmanův korelační koeficient a příslušná p-hodnota, jak uvádíte: „značící pravděpodobnost nulové korelace za předpokladu nenulové“ a „tedy např. hodnota $p = 0.05$ značí, že náš předpoklad existence korelace by měl být v 5 % případů mylný“. V pozdějších kapitolách máte již definici p-hodnoty správně. Jak je to tedy s hypotézou u korelačního koeficientu? Co říká obdržená p-hodnota? A jaký je rozdíl mezi interpretací konkrétní obdržené p-hodnoty a prahové hladiny významnosti $\alpha = 0.05$ pro určení pravděpodobnosti chyby 1. druhu ve frekventistickém přístupu vyhodnocování hypotéz? Lze z jedné p-hodnoty přiřadit hypotézám pravděpodobnosti?

Str. 50, závěr: zcela nesouhlasím s tvrzením, parafrázuji: „Pomocí směrodatné odchylky F_0 odhadnuté metodou SWIPE lze úspěšně rozlišit mezi hodnotami od zdravého člověka a od nemocného s Parkinsonovou nemocí (PN) i v hlučnějším prostředí do SNR 6 dB“.

To, že byl pomocí t-testu objeven statisticky významný rozdíl mezi středními hodnotami nezávislých skupin zdravých a nemocných, je dobré zjištění, neznamená však, že z jedné hodnoty mohou úspěšně klasifikovat. Na obr. 3.1 je zřejmé, že rozdělení se překrývají v podstatě v celém rozsahu hodnot. Upřesněte tedy prosím, jak by bylo možné koeficient pro dané účely užitečně využít, tedy jak lépe postavit experiment (než jedno měření s odpovědí ano/ne), aby u konkrétního člověka mohlo dojít k vyhodnocení rizika PN.

Datum: 5.6.2018

Podpis: