

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	<b>Implementace keprstrálního detektoru řečové aktivity při výpočtu řečových příznaků</b>
<b>Jméno autora:</b>	<b>Michal Kosek</b>
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra teorie obvodů
<b>Oponent práce:</b>	Ing. Jan Bartošek, PhD.
<b>Pracoviště oponenta práce:</b>	IBM ČR

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	<b>náročnější</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Jelikož se jedná o práci bakalářskou, jejíž zadání a realizace obnáší nejen porozumění pokročilejším technikám zpracování číslicových signálů, ale i zorientování se v cizím zdrojovém kódu a jeho vhodné rozšíření, považuji zadání zcela jistě za náročnější.	

<b>Splnění zadání</b>	<b>splněno</b>
<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>	
Autor splnil všechny body zadání beze zbytku a bez jakýchkoliv výhrad. Implementační a zejména experimentální část práce je díky své komplexnosti dle mého názoru vypracována nad rámec zadání.	

<b>Zvolený postup řešení</b>	<b>vynikající</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Kromě metod řešení, které jsou již součástí zadání práce, student vhodně navrhl a zvolil všechny části práce, které nebyly pevně nadefinovány. To se týká jak konkrétního začlenění implementovaného detektoru do existujícího nástroje CtuCopy, tak například návrhu dvou různých metod pro výpočet adaptivních prahů kritériální funkce či vyhlazování mediánovou filtrace. Výborně je uchopena experimentální část práce.	

<b>Odborná úroveň</b>	<b>A - výborně</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Autorovi práce nelze upřít studiem (či jinou praxí) získaný přehled, ale i nadhled v oblasti jak číslicového zpracování signálů, tak v oblasti analytických a programátorských dovedností, které účelně využil v implementační části práce.	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>A - výborně</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Po formální stránce je práce na vysoké úrovni. Z práce je zřejmé, že nebyla šita horkou jehlou a nad větami autor přemýšlel a zachytil v nich velmi výstižně podstatu popisovaných skutečností, přitom v rozsahu přiměřeném práci bakalářské. Až na jednotky drobných překlepů a chybek (např. v popisu obrázku 5.2 má být zřejmě uvedena zkoumaná PLP parametrizace místo MFCC) je jazykově na velmi dobré úrovni. Místy bych uvítal o něco obsáhlejší popis méně standardních jevů či triků, které mohou být pro čtenáře na první pohled obtížněji uchopitelné (např. jednorůchodový výpočet směrodatné odchylky s použitím nikde nepopsané proměnné na levé straně rovnice 3.10).	

<b>Výběr zdrojů, korektnost citací</b>	<b>B - velmi dobře</b>
<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</i>	

*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.*

Citované zdroje práce jsou vybrány vhodně. Chvillemi je obtížné v textu odlišit autorovy vlastní úvahy od těch převzatých (viz např. vztahy pro adaptivní prahy na stranách 14 a 15). Dále mi v práci chybí např. původ existujícího nástroje 'vadcrit' a zdroj tvrzení o rozdílnosti dvou zvuků v kapitole 2.3. Autorovi bych doporučil do budoucna využívat nabízených citačních možností o něco více. K porušení citační etiky však zcela jistě v práci nedošlo.

#### **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.*

Dosažené výsledky práce jsou dle mého názoru nad rámec typických výsledků bakalářských prací a to hned z několika důvodů: V první řadě to je obsáhlá experimentální část opírající se o akademicky uznávanou kombinaci řečově orientovaných databází, jejíž části byly jistě podkladem pro publikační činnost autora na konferenci POSTER 2017, a ve které mi chybí pouze jakékoliv (byť hrubé) srovnání implementovaného VAD s některými dalšími existujícími. Autorův požadavek na 2s neřečovou inicializaci testovaných signálů mi přijde validní i pro případ reálného nasazení. V řadě druhé je to komplexně uchopená implementace rozšiřující existující nástroj CtuCopy s velkou volbou konfiguračních parametrů a s plným interním využitím potenciálu existujícího nástroje. Věřím, že takto obohacený nástroj najde budoucí použití. Zároveň nelze než pochválit autorovo vlastní přispění do v práci „konzumovaného“ git repozitáře QUT-NOISE v podobě nové větve pro volitelný počátek řeči v generovaných signálech.

### **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.*

**Předkládaná závěrečná práce se zabývá vhodně zvoleným tématem detekce řečových/neřečových úseků v signálu a elegantně rozšiřuje existující nástroj parametrizace řečového signálu. Ve všech ohledech naplňuje své zadání a zároveň splňuje všechny požadované náležitosti. Až na výše zmíněné drobné poznámky jsem na práci nenalezl jakékoliv nedostatky a považuji ji za vynikající.**

Otázky ke zodpovězení autorovi při obhajobě práce:

- V kapitole 3.1.4. („Post-processing výstupu detektoru řečové aktivity“) autor uvádí vhodnou délku mediánového filtru s odkazem na literaturu jako  $M=5$ . Je tato délka nějak závislá na vzorkovací frekvenci analyzovaného signálu?
- V kapitole 5.4.2 autor uvádí výsledky vyhlazování median-filtrem pro různý řád/délku tohoto filtru (konkrétně 5, 9, 11 a 17). V závěru práce pak autor reportuje pozitivní vliv takové filtrace na úspěšnost detekce zejména v případech nižší chybovosti nefiltrované detekce až do řádu 17 tohoto filtru. Testoval autor i ještě vyšší řády median-filtru? Pokud ne, jaké chování by od takového experimentu očekával?
- Lze porovnat dosažené výsledky s existující prací podobného zaměření (např. s citacemi [23] nebo [24])?
- Jakých úspěšností dosahují současné state-of-the-art VAD a na jakých metodách strojového učení jsou založeny?

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

Datum: 21.1.2018

Podpis: