

Posudek doktorské disertační práce

Ing. Ivana Míči

OPTIMALIZOVANÁ DETEKCE ŘEČOVÉ AKTIVITY V PROSTŘEDÍ S PROMĚNNÝMI VLASTNOSTMI

Doktorská disertační práce Ing. Ivana Míči má rozsah 86 stran textu včetně 2 příloh. Pečlivě sestavený seznam literatury obsahuje 40 citací z oblasti tématu disertační práce, statistiky a číslicového zpracování signálů, převážně z posledních 10 let. Jako přílohu k disertační práci jsem obdržel výpis 16 publikací disertanta, včetně kopie článku v *Advances in Communications, Computers, Circuits and Devices* z roku 2010, kde je Ing. Míča hlavním autorem a bodové hodnocení doktoranda. Všechny publikace doktoranda jsou se spoluautory, některé byly předneseny na zahraničních konferencích. Práce obsahuje seznam obrázků, tabulek a seznam symbolů, veličin a zkratk. Téma práce odpovídá studijnímu oboru *Teleinformatika* a je vysoce aktuální z hlediska automatické detekce řeči v komunikačních systémech, mobilních komunikacích a také pro analýzu a rozpoznávání řeči. Disertace byla vypracována v rámci aktivní spolupráce autora na projektu SIX operačního programu Výzkum a vývoj pro inovace, dále na projektu Ministerstva školství, mládeže a tělovýchovy a také na projektu Ministerstva průmyslu.

Hlavním cílem disertační práce je, jak je zřejmé už z názvu práce, automatická detekce řečové aktivních úseků signálu. V práci jsou shrnuty teoretické přístupy hlavních detekčních metod, které jsou analyzovány pomocí laboratorní i aplikačně specifické databáze s proměnnými pracovními podmínkami. Vlastnosti detekčních algoritmů jsou hodnoceny pomocí kontingenčních tabulek a prezentovány pomocí tabulek a grafů. Na základě jejich funkce a spolehlivosti je na základě podrobné analýzy navržen a ověřen optimalizovaný detektor řečové aktivity pro prostředí s proměnnými vlastnostmi s různými druhy hluku, který vykazuje podstatně lepší vlastnosti než detektor použitý jako referenční.

Vlastní práce je rozdělena do 5 kapitol. Stručný úvod do problematiky v rozsahu asi 2 stran shrnuje historii a současný stav a principy funkce detektorů automatické detekce řeči. V první kapitole je na 10 stránkách shrnuta problematika detektorů řeči včetně hodnocení jejich kvality a vlivu šumu na jejich funkci. V následující kapitole jsou na 13 stranách popsány laboratorní a aplikačně specifická databáze pro testování výkonnosti detektorů řečové aktivity. V kapitole třetí jsou na 13 stranách shrnuty a analyzovány vybrané techniky klasifikace signálových segmentů řeči. Jako referenční detektor pro porovnání výkonnosti detekčních algoritmů je použit standard G.729B v revizi z roku 2012. Je stručně popsán vlastní algoritmus, vyhodnocena jeho výkonnost pomocí laboratorní i telefonní databáze a je proveden rozbor nepříznivých případů. Jako základ pro optimalizovaný systém byl zvolen detektor učící se bez učitele, který je podrobně popsán. Kapitola čtvrtá obsahuje hlavní přínos práce daný návrhem optimalizovaného detektoru řečové aktivity. Má rozsah 10 stran a je dokumentován řadou tabulek a grafů, ze kterých vyplývá jeho vysoká účinnost v porovnání s referenčním algoritmem. Poslední kapitola pak obsahuje stručné shrnutí výsledků disertační práce a nástin dalšího výzkumu. V první příloze jsou shrnuty vztahy pro výpočet segmentálních parametrů řeči. Druhá příloha je věnována konceptům strojového učení a shrnuje základní statistické metody.

Disertační práce je psána přehledně, s hojnými citacemi literatury, s dobrou grafickou úpravou a je dobře čitelná. Práce je psána na dobré matematické úrovni a dokládá velkou znalost problematiky a statistiky. S tím kontrastuje poměrně velké množství překlepů, chybějících nebo přebývajících slov. V některých grafech nejsou popsány osy, jejich význam

je však zřejmý z textu. V seznamu literatury jsou neúplné citace [Míč+10], [Pre08] a [TTN01]. Dále nejsou citace většinou psány v běžně používané formě, u časopisů chybí vol. a no., „In“ se používá u sborníků z konferencí. Co také znamená plus (+) v některých citacích? Cenný je systematický a chronologický popis jednotlivých přístupů detekce řečových úseků. Návrh optimalizovaného algoritmu vychází z podrobné analýzy vlastností referenčního detektoru G.729B verze z roku 2012 a to jak pomocí laboratorní databáze s různými druhy šumů, tak pomocí aplikačně specifické databáze záznamů telefonních hovorů z prostředí kontaktního centra. Optimalizovaný detektor řečové aktivity považuji za hlavní a původní přínos práce.

Uvedené připomínky nesnižují význam práce, jsou míněny jako poznámka pro přípravu publikací o nové optimalizované metodě detekce řečové aktivity v prostředí s proměnnými parametry. Podle mého názoru je téma disertační práce aktuální z hlediska současného stavu vědy i potřeb praxe a cíl navrhnout a ověřit účinný detektor řečové aktivity byl v plné míře splněn. Jako hlavní výsledek práce považuji systematické hodnocení stávajících algoritmů pomocí laboratorních i provozních signálů a především návrh optimalizovaného detektoru řečové aktivity umožněný důkladnou analýzou příčin nedostatků těchto algoritmů. Podle mého názoru lze očekávat další zajímavé výsledky při nasazení navrženého optimalizovaného detektoru do reálného provozu.

K práci a do diskuse během obhajoby mám následující otázky:

- Při obhajobě by bylo vhodné předvést, jako ilustraci ke grafům, také akustické ukázky řeči s různými typy rušení.
- Jak se chová optimalizovaný detektor v místech neznělých exploziv v řeči?
- Jaké mediánové filtry, prosté nerekurzivní nebo rekurzivní, byly v práci použity?

Závěr

Autor disertační práci prokázal schopnost vědecky pracovat a navrhnout, realizovat a vyhodnotit nový optimalizovaný detektor řečové aktivity. Práce svědčí o velkém přehledu autora ve vědecké literatuře, o jeho zkušenostech při aplikaci metod číslicového zpracování signálů a dokumentuje jeho experimentální přístup k řešení aktuální a perspektivní úlohy. Vytčené cíle autor v plné míře splnil. Většina původních výsledků autora byla publikována i když zatím většinou na konferencích. Jádro disertační práce bylo publikováno v renomovaném časopise. Po pečlivém prostudování disertační práce Ing. Ivana Míči jsem dospěl k názoru, že autor má potřebnou vědeckou erudici, že práce splňuje podmínky tvůrčí vědecké práce a odpovídá obecně uznávaným požadavkům k udělení vědecké hodnosti doktora. Disertační práci proto doporučuji k obhajobě.

V Praze, dne 5. června 2014

Ing. Robert Vích, DrSc., dr. h. c. mult.

Ústav fotoniky a elektroniky AVČR, v. v. i.