

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Hlasová identifikace pohlaví na bázi GMM a její aplikace při rozpoznávání řeči
Jméno autora:	Bc. Tereza Pytelová
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Ing. Josef Rajnoha, Ph.D.
Pracoviště oponenta práce:	SAP ČR

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání závěrečné práce spadá do náročnějších. Student se musí a zorientovat v oblasti rozpoznávání řeči, které v sobě zahrnuje problematiku z více oblastí a vyžaduje pro úspěšnou realizaci dobré porozumění tématu. Na druhou stranu je možné s využitím stávajících velmi dobře popsaných nástrojů a průměrnou programátorskou znalostí celkem rychle sestavit základní systém realizující úlohu rozpoznávání a vytvoření takového systému tak není netriviální úlohou, jejíž úspěšné splnění by bylo nad rámec náročnosti magisterské práce.	

Splnění zadání	splněno s většími výhradami
<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>	
Zadání práce je rozděleno na část seznámení s problematikou rozpoznávání a následně na dílčí body týkající se výběru metod a realizace vlastního rozpoznávače. Bohužel už v prvním bodu je zřejmé, že ke splnění nedošlo v dostatečné míře. Teoretická část obsahuje popis získávání řečových příznaků a základní princip rozpoznávání řeči, ale výrazně zde chybí návaznosti na zkoumanou problematiku rozpoznávání pohlaví řečníka. Stejně tak postrádám alespoň základní popis již existujícího výzkumu v dané oblasti, což vyvolává dojem, že se studentka s aktuálním stavem problematiky neseznámila. Poslední bod zadání - implementace rozpoznávače řeči, který by využíval výsledky vlastního rozpoznání pohlaví řečníka - nebyl vůbec realizován. A bohužel právě tento bod je, podle mě, základním kamenem pro zhodnocení přínosu metody pro její možné využití v procesech rozpoznávání řeči. Výsledky práce tak působí dojmem nedokončeného výzkumu.	

Zvolený postup řešení	částečně vhodný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Využití GMM pro rozpoznávání řeči je standardně používanou metodou a experimenty ukazují, že se tímto způsobem dosahuje zajímavých úspěšností. Zároveň jsou v práci analyzovány vybrané aspekty, které mohou výslednou úspěšnost ovlivnit, což ukazuje na správný směr, kudy by se měl výzkum v oblasti ubírat. Bohužel chybí alespoň mírný vhled do aktuálního stavu problematiky s ohledem na metody a nastavení používané v rámci jiných systémů či pracovišť, například analyzovaný počet keprálních koeficientů pro získání vhodných řečových příznaků, jejichž rozmezí 20 až 40 by mohlo být na základě takových znalostí korigován ještě před vlastními experimenty. Stejně tak v teoretické části chybí diskuse nad metodami, které se v dané oblasti standardně používají a mohou být případnou alternativou k rozpoznávání pomocí GMM.	

Odborná úroveň	E - dostatečně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce využívá především již běžně známých poznatků z oblasti rozpoznávání řeči. Bohužel opět chybí srovnání s aktuálně dosahovanými výsledky jiných pracovišť zabývajících se obdobnou problematikou v odpovídajících podmínkách, což odbornou úroveň výrazně snižuje. Jak již bylo zmíněno, text působí v některých místech pouze jako výčet metod bez větší návaznosti na jejich vliv při rozpoznávání pohlaví řečníka. Například keprální koeficienty jsou uvozeny pouze větou „Cepstral analysis allows to describe frequency spectrum by some coefficients.“ bez další diskuse nad přínosem využití těchto parametrů. Zde se nelze ubránit dojmu, že se studentka o pochopení významu koeficientů ani nepokusila.	

Formální a jazyková úroveň, rozsah práce

D - uspokojivě

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Po formální stránce se jedná o standardně zpracovanou závěrečnou práci, kde lze kladně hodnotit výběr anglického jazyka a tím možnost oslovit širší spektrum odborné veřejnosti. Na škodu čitelnosti je ale v práci možné najít gramatické chyby, překlepy a nedostatky v úpravě textu. Čitelnost textu snižuje i chybějící popis os, různý formát a velikost grafů, případně chybějící popis významu hodnot v závorce v Příloze A. Bohužel také chybí některé aspekty výběru dat a parametrů. Například informace, který z nahrávacích kanálů byl použit na experimenty nebo do jaké míry bylo sledováno zastoupení jednotlivých mluvčích v trénovacích a testovacích datech (tabulky uvádí pouze počet vět v jednotlivých sadách dat).

Výběr zdrojů, korektnost citací

E - dostatečně

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.

Pro popis standardních technik v oblasti rozpoznávání řeči bylo využito relevantních zdrojů, ačkoliv je s podivem, že v práci nejsou použity zdroje navržené v rámci zadání práce. V některých případech se ale v rámci textu nedá určit, zda se jedná o myšlenku autorky, nebo zda jde o princip, jehož detail byl popsán na základě nějakého zdroje. Zároveň, jak bylo zmíněno, chybí citace zdroje popisující aktuální stav problematiky. Také první zmínka o nástroji CtuCopy by měla být doplněna odkazem na zdroj v Referencích, aby odkaz nemusel čtenář dohledávat až v některé z následných kapitol. Mezi zdroji bohužel opět výrazně postrádám zdroj k tématu rozpoznávání pohlaví mluvčího, ačkoliv přímo k oblasti detekce řečové aktivity, což je pro tuto práci část okrajová, jsou uvedeny hned čtyři zdroje.

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 ukazují, že výběr metody byl zvolen správně, ale celkové pojetí by vyžadovalo mnohem detailnější zaměření na danou problematiku, jak z teoretického, tak z praktického pohledu. Předpokládal bych, že aktuálně zpracovaný rozsah práce by měl sloužit pouze jako jeden ze základních stavebních kamenů pro výslednou práci.

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.

Teoretická část práce vzbuzuje dojem, že se studentka se zadanou tematikou neseznámila na dostatečné úrovni. V rámci praktické části pak byla realizována jen část původního zadání. Jelikož však bylo dosaženo dílčích výsledků, které ukazují na správné nakročení k dalšímu možnému postupu v rámci analýzy tématu rozpoznávání pohlaví mluvčího a jeho využití pro rozpoznávání řeči, navrhuji ohodnotit práci jako dostatečnou s následujícími upřesňujícími dotazy, kterými lze doplnit hlavní zmíněné nedostatky práce:

- Proč nebyly, navzdory zadání, dosažené výsledky zpracovány do úlohy rozpoznávání řeči?
- Lze alespoň stručně shrnout aktuální stav problematiky rozpoznávání pohlaví mluvčího? Standardně používané metody a dosahované výsledky.
- Jaký význam má použití kepstrálních koeficientů v rámci úloh rozpoznávání řeči?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **E - dostatečně**.

Datum: 3.6.2016

Podpis: