

Oponentský posudok

habilitačnej práce Mgr. Ondreja Šucha, PhD.

“Obvody na rozpoznávanie reči založené na min a max operáciách”

Habilitant, pán Mgr. Ondrej Šuch, PhD. vypracoval habilitačnú prácu na stále aktuálnu tému, týkajúcu sa rozpoznania reči so zameraním na rozpoznanie reči založené na min a max operáciách. Práca má spolu 102 strán textu, vzťahov, tabuliek a obrázkov. Práca je rozdelená spolu s úvodom a zoznamom publikácií do 9 kapitol, pričom kapitoly 6, 7 a 8 predstavujú nosnú časť práce a predstavujú vlastný vedecký prínos autora k riešeniu problematiky, pôvodné pedagogické spracovanie problematiky je spomenuté v kapitole 1-5. Tým, že práca obsahuje obe zložky vedeckopedagogickej publikácie, môžem konštatovať, že autor splnil požiadavky kladené na práce tohto druhu.

Mgr. Šuch v svojej habilitačnej práci analyzuje možnosti rozpoznania reči s využitím a implementáciou v nanotechnologických obvodoch. Práca predstavuje autorovu snahu o príspevok do analýzy signálov so zameraním na multimodálnu komunikáciu človek-stroj ako aj na metodiku sémantického prezerania veľkých multimediálnych databáz. Autor ukázal hlboké vedomosti nielen z oblasti analýzy signálov ale aj využitia týchto poznatkov pri definovaní pravidiel pre nový prístup a filozofiu rozpoznania reči.

Práca má veľmi široký záber a na jej dôkladné čítanie je potrebný široký vedecký základ. To je však miestami na úkor presného, uceleného a logického vysvetlenia jednotlivých oblastí a v nich použitých metód (napr. opis MFCC strana 42, klasifikatory, atď.)

V práci som postrádal prehľad klasifikátorov, čo to je, aké triedy poznáme, čo sa pri nich hodnotí, druhy chýb, teoretické hranice, teda ucelená teória ako aj praktické realizácie klasifikátorov s popisom ich vlastností (NN, rozhodovacie stromy, KNN, atď.). Z takéhoto úvodu môže potom vzniknúť požiadavka novej teórie, ktorú autor neskôr predkladá.

V práci nie sú vôbec vysvetlené, resp. prezentované súčasné základné princípy rozpoznávania reči, t.j. DTW a HMM, resp. HMM a neurónové siete, HMM a SVM, atď. Až analýza súčasného stavu, resp. analýza nedostatkov dáva predpoklad k vytýčeniu nového vedeckého cieľa.

V rozpoznaní reči pomocou min a max operáciách sa robí vždy len klasifikácia párov foném (vybraných hlások, ktorých vzorky sú v spektre lineárne separovateľné!- až na jeden pár), čo so skutočným rozpoznávaním

reči nemá veľa spoločného. Teda názov práce nezodpovedá celkom obsahu (názov mal obsahovať klasifikácia foném, resp. ich rozpoznávanie).

Niekoľko poznámok a otázok:

-Nejasný opis a záver ohľadom problematiky konverencie Fourierových radov i napriek tomu, že sú známe podmienky konverencie FR.

-Stačí vzorkovať $\frac{1}{2} f_{\max}$, to ale neplatí pre periodické signály, resp. výkonové, str. 12.

-Str 23. nie sú vysvetlené h , h^* , g , g^* , i napriek tomu, že tieto koeficienty majú v blasti waveletovej transformácie a bánk filtrov veľký význam?

-Str 23 f_i a ψ_i sú distribúcie? Nie sú to funkcie kolmé na f_i a ψ_i funkcie. Aký je rozdiel medzi funkciou a distribúciou?

-Čo predstavuje $\arg(x(t))$ pre reálne signály, a potom ako sa počíta $d/dt \arg(x(t))$ a aký to má význam? Str. 26

-Čo znamená $x^*(f)$, str. 27.

-str. 51 Z pravdepodobnosti $P(C1/x) + P(C2/x) = 1$ (2 triedy), čo pre 2 triedy a sigmoidálnu funkciu vo všeobecnosti nenastane. Takže tá priamo nemože vyjadrovať $P(C1/x)$. Ako skutočne vyzerá $P(C/x)$ pre 2 Gaussove rozdelenia s rovnakými kovariančnými maticami?

-Prečo GMM a kernel SVM nie sú zaradené do nelineárnych klasifikátorov?

-str 63 $f_1 < c$, čo to znamená, nie je f_1 funkcia, nechýba jej argument?

-str 63 nedostatočná definícia „supp“. Je definovaná pre premennú s_i , ale už nie pre funkciu $\text{supp}(f(s_i)) = ?$

-str 73. referencia na Váš článok je len vo forme url? Ten teda nebol oficiálne vydaný a recenzovaný žiadnou vedeckou komunitou? Ak áno treba uviesť konferenciu, časopis a všetky relevantné údaje.

-Str 75 Čo je d vo vzorci?

-Klasifikácia a navrhnuté klasifikátory sú vždy robené formou fonéma 1 vs. fonéma 2 prečo? To je pre skutočnú úlohu rozpoznávania foném dosť nepraktické, ak vôbec použiteľné berúc do úvahy ich skutočný počet!

-Výsledky sú uvedené ako simulácie analógových obvodov na digitálnom počítači, t.j. nie sú realizované na žiadnych obvodoch. Navyše pri praktických realizáciach sa zdajú byť navrhované „memristors“ veľmi pomalé, obr. 7.8 (switching probability), resp. pre aké časy t konvergujú obvody na obr. 7.5 ku min, resp. max hodnotám? Ako sa zrealizuje tréning a vyhodnotenie v analógovej oblasti? Ako vôbec prakticky prebehne celý proces rozpoznávania foném v analógovej oblasti? Aký to má praktický význam pre rozpoznávanie foném resp. reči? Detekcie istých dejov v signáloch by boli oveľa vhodnejšou aplikáciou pre tieto obvody ako „rozpoznávanie foném“, napr. náhrada lineárnych spriahnutých filtrov atď.

-Ako je možné, že optimalizáciou logistickej regresie sa dosiahli výrazné chyby pri klasifikácii, a to aj napriek tomu že dáta boli lineárne separovateľné? Akú optimalizáciu ste použili?

Niektoré chyby vo vzorcoch:

Str. 14 chýba normalizačný člen 2π

Str. 24 $C_x(t, f)$

Str. 25 $R_x(\tau)$

Str. 27 $t_x(f)$

Str. 27 $Fid(.,.)$

Str. 27 $ch()cx()$ - kovariancia ci $xcorr$?

Str. 47 dve definicie iy asi jedno ma označovat inú fonému!

Str. 73 $Fl(Q)=Fl$, čo to znamená?

Štruktúra práce je prehľadná a logická. Z hľadiska pedagogického sa jedná o materiál nielen zrozumiteľný ale aj atraktívny pre študentov vysokoškolského štúdia.

Pozitívne hodnotím formálnu prepracovanosť práce, keďže autor venoval pozornosť všetkým častiam diela. Avšak chýba zoznam skratiek a symbolov i keď sa hojne používali. Symboly často krát neboli vôbec vysvetlené, resp. na rôznych miestach mohli mať rôzny význam. Zoznam použitej literatúry je dostatočne obsiahly, a prehľadný. Autor používa vhodnú terminológiu. V práci som našiel iba niekoľko tlačových chýb, vo vzorcoch bolo viac chýb a keďže neboli číslované je obtiažne ich všetky vypisovať. Obrázky sú prehľadné a obsahujú výstižný slovný opis.

So záujmom som sa zoznámil aj s celkovou vedeckovýskumnou ako aj pedagogickou prácou habilitanta. Pedagogické výsledky ako aj publikačná činnosť odpovedá úrovni docenta. Vysoko oceňujem najmä citačný index autora, ako aj prijaté patenty. V spise sú uvedené 4, ale žiadne bližšie informácie. Aj keď niektoré parametre pre habilitáciu sú splnené len tesne, tento fakt je bohato vyvážený nadpriemernou aktivitou v ostatných parametroch celkového hodnotenia.

Habilitant svojou odbornou aktivitou preukázal schopnosť samostatnej tvorivej práce. Z predložených materiálov je jasná pedagogická aktivita vrátane prednášok a vedenia študentov ako aj schopnosť prezentovať výsledky na medzinárodných konferenciách.

Záverom môžem s potešením skonštatovať, že habilitant splnil všetky predpoklady pre udelenie titulu docent. Navrhujem preto pokračovať v začatom habilitačnom konaní a

odporúčam vymenovať Mgr. Ondreja Šucha, PhD. za docenta.



V Bratislave, 14. februára 2014