

POSUDOK DIZERTAČNEJ PRÁCE

Ing. Marek Hruz: Extrakce příznaků pro automatickou analýzu gest znakového jazyka

Predložená dizertačná práca sa zaoberá aktuálnou problematikou rozpoznávania posunkovej reči, čo je časťou širšieho cieľa, ktorý je aj v pozornosti Európskej komisie v oblasti podpory európskeho výskumu - eInclusion. Konkrétne ciele práce boli nasledovné:

Navrhnuť a implementovať systém pre sledovanie pohybu častí tela s explicitným modelom ich prekryvania. Vstup dát je z vizuálneho systému (vrátane farby), pričom sa nepoužívajú značky, systém je pravdepodobnostný, získané charakteristiky sa používajú pre rozpoznávanie a kategorizáciu, nájsť minimálnu množinu charakteristík s použitím metód redukcie dimenzie a dekorelácie.

Doktorand navrhol a implementoval systém schopný sledovať pohyb hlavy a rúk vo video sekvenciách, ktorý pomáha expertovi pre posunkovú reč pri anotácii, konkrétne výpočte charakteristických črt (features) pre rozpoznávanie, pričom použil vlastný originálny algoritmus ako spracovať prekryvajúce sa časti tela hovoriacej osoby. Túto časť možno považovať za najväčší prínos k výskumu tejto problematiky nakoľko podobná metódy nebola doposiaľ publikovaná a experimentálne výsledky sú lepšie ako pri inej publikovanej metóde. Z tohto dôvodu predstavuje predložená dizertačná práca vedecký prínos v danej oblasti výskumu.

Metódy a implementované subsystemy v rámci výskumu sú systematicky a podrobne dokumentované, dobre štruktúrované s dostatočným množstvom obrazových ilustrácií, tabuliek a grafov. Práca je napísaná v Angličtine, čo umožňuje jej čitateľnosť aj v zahraničí a nevznikajú problémy s korektnými prekladmi špecializovaných odborných výrazov. V texte som nenašiel podstatné štylistické alebo gramatické chyby. Písomná časť je rozdelená do ôsmich kapitol a dvoch príloh. Kapitoly 1, 2, 4 a 5 obsahujú motivácie a všeobecné informácie ako aj stručný výklad relevantných problémov rozpoznávania znakovkej reči. Kapitola 3 sa opisuje rôzne korpusy znakovkej, vrátane tých s ktorými pracoval aj autor práce. Výsledky vlastného výskumu sú obsiahnuté v kapitolách 6, 7 a 8 ako aj prílohách.

Vzhľadom k rozsahu práce vyskytli sa aj niektoré formálne metodické a formálne nedostatky:

Vyberám nasledovné: Autor používa v texte opakovane skratky, pred tým ako ich definuje pričom niektoré nie sú v slovníku skratiek napríklad EM (s. 25) LDA, HLDA, FA (na s. 53 definované na s. 73). Nakoľko metódy LDA, HLDA a FA sú použité v jadre práce, mali by byť stručne vysvetlené a ich výber zdôvodnený vo všeobecnej teoretickej časti (napr. kapitole 3 a 5).

V niektorých častiach teoretickej časti chýbajú odkazy na zdroje informácií napríklad v opise nemeckého a európskeho korpusu (s. 12-13).

V niektorých formulách chýbajú opisy premenných napríklad vo vzorcoch pre fourierove, deskripty, momenty, region growing feature ako aj domena B splajnov (rozsah hodnôt t), v Gausovom klasifikátore čo je T , na s. 40 čo je to k , v rovnici 6.3 čo je p_m , v rovnici entropie (6.5) ako sa určí pravdepodobnosť $p()$. Formule 6.20 a 6.21 sú zrejme chybné.

Pre pochopenie významu -čo dokazujú by bolo treba vysvetliť interpretáciu diagramov na obr. 6.5 a 6.6.

Tak isto aj na obr. 6.12. Prečo dôjde k náhlemu poklesu počtu pixlov po ich náraste (asi spojenie oblastí dvoch častí tela) pod priemerné množstvo (plocha). Čiastočne sa to týka aj obr. 6.18.

Explicitné porovnania výsledkov výskumu doktoranda s publikovanými výsledkami je len v prípade sledovania pohybu na základe výsledkov chybovosti testov na s. 78 (chýba index odkazu na literatúru, uvedené len meno autora).

Otázky:

Bolo by možné globálne kvantitatívne porovnať ktorá postupnosť metód a dátových množín je jednoznačne najlepšia pre sledovanie pohybu pri rozpoznávaní posunkovej reči, aby sa výsledky výskumu v tejto oblasti mohli „merať“ a exaktne porovnávať (niečo ako benchmark testy).

V práci sa experimentálne vyhodnotila chybovosť jednotlivých metód, nebola zmienka o rýchlosti výpočtov, čo by bolo zaujímavé z toho dôvodu, že výsledný systém bude musieť pracovať s ďalšími metódami a v reálnom čase. Autor sa nimi nezaobrá preto lebo súčasné počítače s tým nemajú problémy alebo cieľom bolo zamerať sa len na otázku spoľahlivosti rozpoznávania ?

Záver

Téma dizertačnej práce je aktuálna, má pomerne široký záber vyžadujúci si vedomosti z viac príbuzných disciplín a má praktické využitie. Doktorand musel naštudovať a analyzovať mnohé metódy, vybrať a implementovať najvhodnejšie experimentovať a nájsť vlastné originálne riešenia.

Doktorand splnil ciele dizertačnej práce -- stanovil si hypotézy, ktoré teoreticky a experimentálne dokázal a výsledky publikoval na medzinárodných vedeckých fórach (vrátane časopisov a konferencií ACM a IEEE) čím dokázal schopnosti samostatnej vedeckej činnosti a preto navrhujem aby mu bola po úspešnej obhajobe písomnej časti udelená vedecká hodnosť Philosophiae Doctor (PhD.).

V Bratislave 10.5. 2013

doc. Ing. Martin Šperka, PhD.
oponent



Oponentní posudek disertační práce

Název práce: Extrakce příznaků pro automatickou analýzu gest znakového jazyka

Student: Ing. Marek Hruž

Předkládaná práce je zaměřena na vytvoření systému pro extrakci příznaků vhodných pro automatický systém rozpoznávání znakového jazyka z video záznamu. Téma práce je aktuální a odráží potřeby oblasti komunikace člověka a stroje pro specifický případ komunikace neslyšících. Vytvořený systém má ovšem přesah i do jiných oblastí interakce člověk-stroj. Práce přináší nové poznatky především v oblasti řešení sledování překrývajících se objektů a využití apriorní znalosti o vzájemných pozicích rukou a proto je pro daný obor významná.

Vytyčené cíle disertační práce jsou shrnuty v šesti bodech. Mezi nejvýznamnější cíle patří založení systému na pravděpodobnostním přístupu, snaha o obecnost neomezením se na použití vizuálních značek a průzkum vlivu metod redukce příznaků na celkovou úspěšnost systému. Další cíle jsou zaměřeny na nutné úkoly potřebné pro vytvoření celého systému, jako vytvoření databáze, či nástroje pro polo-automatickou anotaci dat.

Všechny vytyčené cíle lze považovat za splněné a dostatečně dokumentované v samotné disertační práci. Student prokázal schopnost samostatné vědecké práce, tím že realizoval její jednotlivé kroky (studium literatury, definování problému, vytvoření vlastního řešení, experimentální ověření, srovnání řešení s ostatními přístupy a rozbor výsledků). Student vytvořil automatický systém pro extrakci příznaků použitelných pro automatickou analýzu gest znakového jazyka. Systém byl ověřen pro laboratorní podmínky a dosahuje úspěšnosti sledování 90% - 95%. Všechny kroky práce jsou dostatečně popsány a analyzovány. Vyzdvihují především množství experimentů provedených pro objasnění vlivu různých metod redukce dimenze na úspěšnost sledování a provedení experimentů pro samotné rozpoznávání izolovaných znaků. Vlastní přínos studenta spatřuji především ve vytvoření metody pro analýzu překryvů založené na pravděpodobnostním přístupu. Tento postup umožňuje snadné rozšiřování metody například o využití apriorní znalosti o vzájemné poloze rukou, která byla v práci představena. Největší nedostatek práce spatřuji v porovnání metody se state of the art. Problematika rozpoznávání řeči je obecně zatížena tímto problémem, neboť metody jsou vyvíjeny většinou pro specifický jazyk. Tyto komplikace vznikají například z důvodů odlišných jazykových modelů nutných pro rozpoznávání plynoucích z odlišností jednotlivých řečí. Práce je v tomto případě zaměřena na sledování rukou a hlavy pro generování příznaků. Porovnání metody s ostatními přístupy brání tedy pouze uzpůsobení metody specifické databázi. Evropský projekt SignSpeak, který sdružuje výzkum v oblasti rozpoznávání znakové řeči a nabízí dostatečné množství anotovaných databází, z nichž by bylo jistě možné vybrat jednu vhodnou pro porovnání metody. Výsledky metody student srovnává s vlastní implementací metody pro sledování objektů podle článku [67] a proto je obtížné zhodnotit kvalitu předkládané metody.

Formální a jazyková úprava je kvalitní. Text práce obsahuje minimální množství chyb. Členění práce je logické a přehledné. Jednotlivé kapitoly dostatečně popisují danou problematiku.

Západočeská univerzita v Plzni

Doručeno: 21.10.2013

ZCU 033667/2013

listy: 5

druh:

přílohy:



zcupesc70282

Student předkládá v práci přehled 23 publikací. U sedmi publikací je student uveden jako hlavní autor. Většina publikací jsou publikace ve sborníku z konference, kdy některé z nich patří k významným v oboru. Dva spolu-autorské články jsou publikovány v impaktovaných časopisech. Pouze část publikací je přímo vztažena k řešení a cílům disertační práce a proto je disertační práce hlavním zdrojem informací.

K práci mám následující připomínky:

- Hlavním cílem práce je vytvoření systému pro extrakci příznaků vhodných pro rozpoznávání gest znakového jazyka. Tímto problémem se zabývá několik výzkumných týmů s uspokojivými výsledky pro dané omezení úlohy. V práci mi tedy chybí přímé srovnání vytvořené metody s podobnými metodami.
- Práce neobsahuje žádnou zmínku o výpočetní náročnosti metody. Je zde pouze zmíněno, že metoda byla implementována pomocí jazyka C++. Výpočetní náročnost je jedním z hlavních kritérií při reálném nasazení metody.
- Kapitola 1.3 – V kapitole je uvedeno, že práce existují zařízení (Microsoft Kinect), pro získávání dat vhodnější než standardní kamera ovšem jejich dostupnost je nižší. S tímto bodem není možné souhlasit, neboť cenová dostupnost standardních kamer je srovnatelná s cenou těchto zařízení.
- Kapitola 3.1 obsahuje popis databáze využití pro testování extrahovaných příznaků pro rozpoznávání gest. Databáze podle popisu obsahuje tři odlišné pohledy na řečníka. Z práce je patrné, že informace o 3D poloze rukou a hlavy by napomohla jak správnosti sledování, tak samotnému rozpoznávání gest. Metoda však využívá pouze jeden pohled na řečníka. Není mi tedy zřejmé, z jakého důvodu byly různé pohledy zaznamenávány, případně proč nejsou využity.
- Kapitola 6.5 – Použitá metoda předzpracování by měla odstraňovat z obrazu všechny malé a velké objekty. V tomto případě nevím, co je velkými objekty myšleno.
- Tabulka 6.4 – Tabulka uvádí výsledky rozpoznávání izolovaných znaků pro různé příznaky. Není zde ovšem vysvětleno, zda dochází k nějaké redukci dimenze příznaků. Dimenze příznaků Comb-2 je 226. Předpokládám, že rozpoznávání nebylo prováděno v prostoru s takto vysokou dimenzí.
- Kapitola 7.1 uvádí chybovost metody sledování. Z textu není zřejmé, proč byla pro prezentaci výsledků zvolena právě tolerance 25 bodů.

Celkově hodnotím práci jako kvalitní. Práce obsahuje vlastní přínos studenta k danému oboru. Představená metoda by měla být otestována i na jiné vhodné databázi. Toto testování by pomohlo srovnat úspěšnost metody s ostatními přístupy.

Autor ve své disertační práci **prokázal** schopnost samostatné vědecké práce v daném oboru. Z tohoto důvodu **doporučuji** disertační práci k obhajobě.

V Nových Hradech, dne 14.10.2013

..........
Ing. Petr Císař Ph.D.