

Posudek habilitační práce

Název: Adaptation of speech recognition systems to selected real-world deployment conditions

Habilitation: Ing. Petr Červa, Ph.D.

Oponent: Prof. Ing. Josef Psutka, CSc.

Předložená habilitační práce s názvem „Adaptace systémů rozpoznávání řeči na vybrané reálné podmínky nasazení“ je tvořena souborem 12 vybraných publikací autora, které jsou doplněny poměrně detailním komentářem. Celá práce je rozdělena tematicky do tří částí, které řeší různé aspekty adaptace systému rozpoznávání řeči na rušivé jevy či reálné podmínky nasazení, s cílem zabezpečit co nejpřesnější výstup systému. Vybrané publikace zaměřené na zvolené téma považují za velmi aktuální, čehož dokladem je i to, že výzkumné práce byly podpořeny několika výzkumnými projekty, na jejichž řešení se habilitant různou formou podílel.

První část habilitační práce je věnována problematice neřízené adaptace na hlas řečníka. Přílohou této části jsou 4 publikace. Řešeny jsou problémy adaptace na hlas v případě využití systému rozpoznávání na on-line přepis přednášek pro sluchově postižené studenty (1 článek) a další 3 články řeší problematiku adaptace na hlas pro off-line přepisy pořadů uložených v rozhlasovém archivu apod. Články jsou vesměs publikovány v kvalitních recenzovaných zdrojích (3x konf. Interspeech, 1x článek v časopisu Speech Communication), přičemž využívají nejmodernější metody zpracování (odpovídající dané době publikování), které jsou efektivně přizpůsobovány vždy dané úloze. V časopiseckém článku je zmíněno využití Bayesova informačního kritéria (BIC) při vytváření Gaussových směsí. V této souvislosti bych rád položil dotaz do diskuse: „Jaké zkušenosti máte s využitím diagonálních vs. plných kovariančních matic v Gaussových směsích v úlohách rozpoznávání/diarizace“.

Druhá část habilitační práce řeší problematiku neřečových událostí, které mohou negativně ovlivnit kvalitu fungování systému rozpoznávání řeči. Jsou předloženy 4 články, které jsou publikovány opět na kvalitních recenzovaných konferencích (3x ICASSP a 1x Interspeech). První dva články se zabývají problematikou detekce aktivity řečníka (Speaker activity detection, SAD) a jsou cíleny na využití v úlohách přepisu televizních/rozhlasových pořadů. Další 2 články řeší problematiku rozpoznávání řeči pro případy, že se na pozadí akustického kanálu objevuje hudba. Publikované výsledky jsou získány pro češtinu a další slovanské jazyky a jsou plně srovnatelné s výsledky dosahovanými pro obdobné úlohy ve světových laboratořích. K této tematické oblasti mám opět dotaz do odborné diskuse: „V práci je kladen zvýšený důraz na kvalitu rozpoznávání řeči v případě, že na pozadí zpracovávané řeči se objevuje hudba. To je jistě velmi frekventovaný případ, který se objevuje při zpracování

rozhlasového vysílání. Nicméně, v této souvislosti by mne zajímal Váš názor na řešení situace, kdy na pozadí aktivní řeči je slyšet vzdálený hovor?"

Třetí část práce se zabývá jednak problematikou adaptace systému na rozpoznávání nového jazyku, kde cílem je vývoj akustického modelu pro nový jazyk (zde jde zřejmě o chorvatštinu), a to s využitím navržené cross-jazykové adaptace z akustického modelu češtiny. Další práce se týkají problematiky identifikace jazyka z akustického signálu s konkrétním zaměřením na slovanské a skandinávské jazyky. Tato část práce je podpořena opět 4 publikacemi (1x článek v imp. časopisu, 1x konf. Interspeech, 2x mezin. konf. TSD a Specom). Články prošly mezinárodním recenzním řízením a považuji je opět za velmi kvalitní.

Závěrem bych chtěl položit habilitantovi do odborné diskuse poslední dotaz: „Můžete vysvětlit, na jakém softwarovém nástroji probíhaly vývojové práce, a to jak v případě dřívějších prací s GMM-HMM, tak zejména v případě současného vývoje systémů DNN-HMM?"

Závěrem lze konstatovat, že celkově práce velmi dobře dokumentuje jak výzkumné schopnosti uchazeče, tak i rozsah a výbornou kvalitu jeho publikačních výsledků. Mohu proto rád konstatovat, že habilitační práce ing. Petra Červy, Ph.D. splňuje po odborné stránce podmínky odst. (3) §72 Zákona o vysokých školách č. 111/1998 Sb., a proto doporučuji práci k obhajobě před vědeckou radou Fakulty mechatroniky, informatiky a mezioborových studií TUL v Liberci a po jejím úspěšném průběhu jmenovat ing. Petra Červu, Ph.D. docentem, v oboru Technická kybernetika.

V Plzni 24. 8. 2021

prof. Ing. Josef Psutka, CSc.