

Challenges and Techniques for Dialectal Arabic Speech Recognition and Machine Translation

Authors **Mohamed Elmahdy, Mark Hasegawa-Johnson, Eiman Mustafawi, Rehab Duwairi, Wolfgang Minker**

Institutions Qatar University, Doha, Qatar
University of Illinois, Urbana, USA
Ulm University, Ulm, Germany

e mail melmahdy@ieee.org

In this research, we propose novel techniques to improve automatic speech recognition (ASR) and statistical machine translation (SMT) for dialectal Arabic. Since dialectal Arabic speech resources are very sparse, we describe how existing Modern Standard Arabic (MSA) speech data can be applied to dialectal Arabic acoustic modeling. Our assumption is that MSA is always a second language for all Arabic speakers, and in most cases we can identify the original dialect of a speaker even though he is speaking MSA. Hence, an acoustic model trained with sufficient number of MSA speakers will implicitly model the acoustic features for the different Arabic dialects. Since, MSA and dialectal Arabic do not share the same phoneme set, we propose phoneme sets normalization in order to crosslingually use MSA in dialectal Arabic ASR. After normalization, we applied state-of-the-art acoustic model adaptation techniques to adapt MSA acoustic models with little amount of dialectal speech. Results indicate significant decrease in word error rate (WER). Since it is hard to phonetically transcribe large amounts of dialectal Arabic speech, we studied the use of graphemic acoustic models where phonetic transcription is approximated to be word letters instead of phonemes. A large number of Gaussians in the Gaussian mixture model is used to model missing vowels. In the case of graphemic adaptation, significant decrease in WER was also observed. The approaches were applied with Egyptian Arabic and Levantine Arabic. The reported experimental work was performed while the first author was at the German University in Cairo in collaboration with Ulm University. This work will be extended at Qatar University in collaboration with the University of Illinois to cover ASR and SMT for Qatari broadcast TV. We propose novel algorithms for learning the similarities and differences between Qatari Arabic (QA) and MSA, for purposes of automatic speech translation and speech-to-text machine translation, building on our own definitive research in the relative phonological, morphological, and syntactic systems of QA and MSA, and in the application of translation to interlingual semantic parse. Furthermore, we propose a novel efficient and accurate speech-to-text translation system, building on our research in landmark-based and segment-based ASR.