Enhanced Amharic Speech Recognition Systems

Woubie, Abraham

Abstract:

Pronunciation variation is one of the main factors that degrade the performance of Amharic ASRS. It is caused either by intra-speaker or inter-speaker variability. This paper describes how the performance of a speaker dependent continuous Amharic speech recognizer is enhanced by modeling pronunciation variation. It uses three methods to design Amharic pronunciation dictionaries. The first method is a grapheme based canonical pronunciation dictionary that contains a single pronunciation for each word in the lexicon. The second method is a grapheme based multiple pronunciation dictionary that contains alternate pronunciations for some of the words in the lexicon. The pronunciation variants in the second method are generated using knowledge based approach. The third method is a grapheme based multiple pronunciation dictionary where the pronunciation variants are generated using data-derived approach. Using the second and third methods has led to a larger improvement in SER compared to the benchmark first method. The SER rates measured for the first method are 39%, 41%, 42% and 44% for speaker1, speaker2, speaker3 and speaker4 respectively. The SER rates measured for the second method are 31%, 33%, 35% and 38% for speaker1, speaker2, speaker3 and speaker4 respectively. Compared to the first method, a statistically significant decrement of 8%, 8%, 7% and 6% SER is measured in the second method for speaker1, speaker2, speaker3 and speaker4 respectively. Using the third method for only one of the four speakers has led to a 6% SER which is a further decrement of 25% SER compared to the second method. Using the acoustic evidence transcription of this speaker to the other three speakers has led to 12%, 17% and 19% SER for speaker2, speaker3 and speaker4 respectively. Compared to the second method, a statistically significant decrement of 21%, 18% and 19% SER is measured in the third method for speaker2, speaker3 and speaker4 respectively. Key words: Automatic Speech Recognition Systems, Pronunciation Dictionary, Pronunciation Variation, Pronunciation Variation Modeling.