

Karunananda, Viraj
Institute of Human Resource Development, University of Colombo

Poster

Phonemes assemble algorithm for automatic speech recognition in Sinhala language

Speech recognition is the process of converting a speech signal to a sequence of words in the form of digital data, by means of an algorithm implemented as a computer program. Since speech recognition is probabilistic, the most probable decoding of the audio signal is output as the recognized text, but multiple hypotheses are considered during the process. One of the major challenges of speech recognition is the variability of speech. The speech recognition problem was soon considered a pattern recognition problem.

The Sinhala language's phoneme to grapheme relationship is quite straightforward. There is one to one mapping between phonemes and graphemes, without any exceptions. The task we chose for the evaluation of our method is phoneme to grapheme transcription in the Sinhala language, which can also be regarded as straight speech recognition in which the acoustic model is perfect. In order to input speech to a computer, the user sends phoneme sequences to the computer. Then the computer indexes the grapheme values to the input phoneme or phoneme combination according to P2G matching algorithm. The output is the most appropriate character for the input phoneme or phoneme combination.

The implementation of phonemes assemble algorithm for the Sinhala language is performed on digitized phoneme database for phoneme to grapheme model. Traditional methodologies such as end point detection, zero crossing level and sound wave energy statistics were used for isolated phoneme recognition. The algorithm relied on 'search a grapheme' but although not seen in phoneme stack, it can also make it to print. This poster will outline some of the previous attempts to construct such rule sets and will describe new and successful approaches to the construction of phonetic assemble algorithm for the Sinhala language.