Spoken Language Identification Using Prosody, Phonotactics, and Acoustics: A Review

Irshad Ahmad Thukroo, Rumaan Bashir, and Kaiser J. Giri

https://doi.org/10.1142/S0219649222500575 Cited by: 0 (Source: Crossref)

< Previous

Abstract

Spoken language identification (LID) is the identification of language present in a speech segment despite its size (duration and speed), ambiance (topic and emotion), and moderator (gender, age, demographic region). Information Technology has touched new vistas for a couple of decades mostly to simplify the day-to-day life of humans. One of the key contributions of Information Technology is the application of Artificial Intelligence to achieve better results. The advent of artificial intelligence has given rise to a new branch of Natural Language Processing (NLP) called Computational Linguistics, which generates frameworks for intelligently manipulating spoken language knowledge and has brought human-machine into a new stage. In this context, speech has arisen to be one of the imperative forms of interfaces, which is the basic mode of communication for us, and generally the most preferred one. Recognition of the spoken language is a frontend for several technologies, like multiple languages conversation systems, expressed translation software, multilingual speech recognition, spoken word extraction, speech production systems. This paper reviews and summarises the different levels of information that can be used for language identification. A broad study of acoustic, phonetic, and prosody features has been provided and various classifiers have been used for spoken language identification specifically for Indian languages. This paper has investigated various existing spoken language identification models implemented using prosodic, phonotactic, acoustic, and deep learning approaches, the datasets used, and performance measures utilized for their analysis. It also highlights the main features and challenges faced by these models. Moreover, this review analyses the efficiency of the spoken language models that can help the researchers to propose new language identification models for speech signals.

Next >