Feature Extraction of Automatic Speaker Recognition, Analysis and Evaluation in Real Environment

Edward L. Campbell
Gabriel Hernández
José Ramón Calvo

Images and Signals Group. CENATAV Research Division. DATYS Enterprise. Cuba.
{ecampbell,gsierra,jcalvo}@cenatav.co.cu

Abstract

An Automatic Speaker Recognition is a biometric system that allows you to identify and verify people, using voice as a discriminatory feature. The purpose of this paper is the feature extraction stage, performing an analysis of effectiveness in real environment. The features extraction has as objective to capture the associated characteristic space of the speaker, being the Mel features and its linear variant the most used methods. In real conditions, the environment over which the speech signal is processed tends not to be ideal, nor is the duration of the speech, so it's necessary to use robust techniques for assuring a lower degradation grade of system effectiveness; techniques such as Power Normalization, Hilbert Envelope and Modulation of Mean Duration are described, analyzed and evaluated.

Keywords: *Automatic Speaker Recognition, Feature extraction Robustness*

Disponible en https://www.springer.com/gp/book/9783030011314