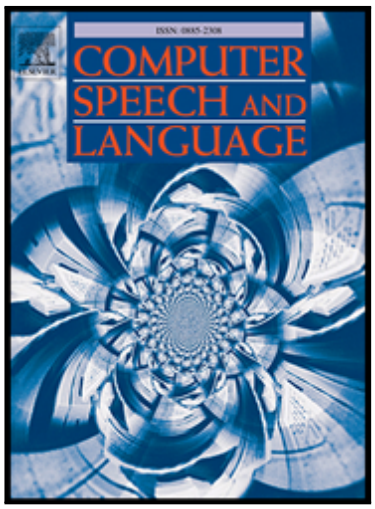

Efficient Estimation and Model Generalization for the Total Variability Model

Ruchir Travadi, Shrikanth Narayanan

PII: S0885-2308(18)30014-7
DOI: 10.1016/j.csl.2018.07.003
Reference: YCSLA 937

To appear in: *Computer Speech & Language*

Received date: 8 January 2018
Revised date: 26 June 2018
Accepted date: 9 July 2018

Please cite this article as: Ruchir Travadi, Shrikanth Narayanan, Efficient Estimation and Model Generalization for the Total Variability Model, *Computer Speech & Language* (2018), doi: 10.1016/j.csl.2018.07.003

Efficient Estimation and Model Generalization for the Total Variability Model

Ruchir Travadi, Shrikanth Narayanan

Signal Analysis and Interpretation Lab, University of Southern California

Abstract

A number of audio signal processing applications characterize different properties of the source underlying an audio signal by analyzing the distribution of a sequence of feature vectors obtained from the signal. The Total Variability Model has been widely used for this purpose as a mechanism for capturing the variability in the feature vector distribution across different signals within a low dimensional representation. In order to arrive at a compact representation, a number of assumptions are made within the model regarding the properties of this distribution. In this paper, we first present an analysis of a parameter estimation method for the model which offers a computationally efficient alternative to the widely used Expectation Maximization (EM) algorithm, but relies on the validity of the model assumptions, using experiments on speaker and language identification tasks. To explain some of the results obtained using this method, we present an extensive statistical analysis aimed at verifying the validity of some of the model assumptions. We show that many of these model assumptions are not valid for the observed data, and propose model generalizations to replace these assumptions. The proposed generalizations lead to a better performance while also opening up possibilities for discriminative training of the model.

Keywords: Total Variability Model; i-vector; Speaker Identification; Language Identification