Voice verification using i-vectors and neural networks with limited training data

ABSTRACT

This study proposes an approach to voice identification based on neural networks (DNN) for i- Vector. Modern voice identification systems based on DNN use large amounts of labeled training data. Using the LRE i-Vector Machine Learning Challenge restricts access to ready-to-use i-Vector for learning and testing the voice identification system. This poses unique challenges in developing DNN-based voice identification systems, since optimized external interfaces and network architectures can no longer be used. We propose to use the training i-Vectors to train the initial DNN to identify the voice. Next, we present a novel strategy for using this initial DNN to strip the language labels of the inappropriate set from the development data. The final DNN for voice identification is trained using the original training data and the estimated out-of-set language data. We show that augmenting the training set with out-of- set labels leads to a significant improvement in voice identification performance. In this paper, we studied the possibility of using neural networks for speech identification. In particular, standard approaches to speech recognition were considered, the concept of an artificial neuron as an object used in speech identification was defined. A speech recognition option using a neural network was investigated, and steps were presented to perform this task. Accuracy using neural networks with limited learning data and a higher i-vector dimension is superior to others with a score of 92.1%. From this study, we can conclude that the size of the UBM and the dimension of the i-vector affect the accuracy of voice identification based on the i- vector.