Speaker Verification Using Adapted Gaussian Mixture Models

Final Term Paper Proposal

Eduardo Martins Barros de Albuquerque Tenório (*student*)
Tsang Ing Ren (*advisor*)
{embat,tir}@cin.ufpe.br

Centro de Informática,
Universidade Federal de Pernambuco

October 30, 2014

Abstract

The proposed project is a reproduction of MIT Lincoln Laboratory's Gaussian Mixture Model (GMM)-based speaker verification system used in several NIST Speaker Recognition Evaluations (SREs) implemented by Douglas A. Reynolds. The system is built around the likelihood ratio test for verification, using simple but effective GMMs for likelihood functions, an Universal Background Model (UBM)-GMM for alternative speaker representation, and a form of Bayesian adaptation to derive speaker models from the UBM-GMM. Additionally, Fractional Covariance Matrix (FCM) is used in an attempt to improve the performance.

***keywords**: speaker recognition; Gaussian Mixture Models; likelihood ratio test; Universal Background Model; Fractional Covariance Matrix.*

1 Introduction

Voice is a very convenient tool for human beings. We are used to communicate and command throught speech and be identified by our voices, so it is natural that the technology follow the path and become more "vocal". From orders to a computer game to a system authentication, the knowledge of specific techniques and which is more appropriate is imperative. The first demands *speech recognition* to extract information about what is said while the second uses *speaker recognition* techniques to identify who is speaking through the analysis of vocal characteristics. Speaker recognition, in turn, is a major area of Computer Engineering, divided in two subfields: *speaker verification*, when a speaker's identity is verified by voice, and *speaker identification*, when a speaker is recognized in a group.

The proposed project is focused on speaker verification and reproduces the idea presented in [1]: a likelihood test using UBM-GMM and adapted GMMs to classify a speaker as enrolled or imposter, through rigorous training and testing. A secondary objective is use the theory of FCM [2] to try to improve the results.

2 Objectives

The project main objectives are:

1. Development of a state-of-the-art speaker verification system based on [1], trained with noise free utterances and tested in a noisy environment (utterances recorded with natural background noise) to evaluate robustness.
2. Explore the theory of FCM to try to improve the system's performance, repeating the experiment from the first objective.

3 Schedule

The schedule is divided in months, from October (when this proposal is written) to February (when the paper is delivered).

October 2014

- Literature research.
- Report writing.

November 2014

- Literature research.
- UBM-GMM development and training.
- Simple speaker's GMM development and testing.
- Adapted speaker's GMM development and testing.

- FCM development.
- Report writing.

December 2014

- FCM testing.
- Report writing.

January 2015

- Experiments and analysis of results.
- Report writing.

February 2015

- Project review.
- Delivery and presentation.

4 Evaluators

The requested evaluator is Professor George Darmiton da Cunha Cavalcanti (`gdcc@cin.ufpe.br`). If unavailable, the presence of Professor Carlos Alexandre Barros de Melo (`cabm@cin.ufpe.br`) is requested.

References

[1] D. A. Reynolds et al., "Speaker verification using adapted gaussian mixture models," *Digital Signal Processing*, vol. 10, (1-3) pp. 19-41, 2000.

[2] C. Gao et al., "Theory of fractional covariance matrix and its applications in PCA and 2D-PCA," *Expert Systems with Applications*, vol. 40, (1-3) pp. 5395-5401, 2013.

Signatures

Eduardo Martins Barros de Albuquerque Tenório
(*student*)

Tsang Ing Ren
(*advisor*)

Recife, October 30, 2014.