Text-Independent Speaker Recognition Using 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

June 11, 2015

Abstract

The proposed project is an implementation of speaker recognition systems, both identification and verification. The systems are built using Gaussian Mixture Models, as proposed in several papers from Douglas A. Reynolds. The use of Fractional Covariance Matrix is studied as an possible increase for the traditional recognition systems.

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 claimed identity is verified, and speaker identification, when a speaker is recognized in a group, both using the speaker's voice.

The proposed project is focused on speaker recognition and reproduces the ideas presented in [1] and [2]. A secondary objective is use the theory of FCM [3] to try to improve the results.

2 Objectives

The project main objectives are:

- 1. Development of a state-of-the-art speaker recognition system based on [1] and [2].
- 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 March (when this proposal is written) to June (when the paper is delivered).

March 2015

• Literature research.

April 2015

- Simple speaker GMM development.
- UBM-GMM development.
- Adapted speaker GMM development.

May 2015

• FCM development.

June 2015

- Experiments and analysis of results.
- 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

- Douglas A. Reynolds, "A Gaussian Mixture Modeling Approach to Text-Independent Speaker Identification," *PhD thesis, Georgia Institute of Technology*, 1992
- [2] Douglas A. Reynolds ., "Speaker verification using adapted gaussian mixture models," *Digital Signal Processing*, vol. 10, (1-3) pp. 19-41, 2000.
- [3] Chaobang Gao, Jiliu Zhou and Qiang Pu, "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, June 11, 2015.