Universidade Federal de Pernambuco

Centro de Informática

Improvements in a Gaussian Mixture Models based Speaker Verification System using Fractional Covariance Matrix

Eduardo Martins Barros de Albuquerque Tenório

January 12, 2015

Abstract

TODO EDITAR Abstract goes here

Dedication

TODO EDITAR To mum and dad

Declaration

TODO EDITAR I declare that..

Acknowledgements

I am thankful to my parents, for the support and patience during the graduation,
To my adviser, Tsang Ing Ren, for the guidance,
To Cleice Souza, for the previous readings and help.

Contents

Chapter 1

Introduction

Chapter 2

Speaker Recognition System

Chapter 3

Feature Extraction

The feature extraction process transforms the speech signal in a sequence of vectors representing the unique characteristics of the speaker's vocal tract. According to [1], an ideal characteristic must be:

- of high inter-speaker and low intra-speaker variability;
- robust in the presence of noise and distortion;
- frequent and natural in the speech;
- easy to measure and extract;
- difficult to be artificially produced;
- not affected by health issues and long term vocal variations.

Being an acoustic wave in time, the raw speech signal in difficulty to process

3.1 The Mel Scale

3.2 Mel Frequency Cepstral Coefficient

TODO referenciar Davis and Mermelstein [2], mostrando que seus experimentos colocam o MFCC como uma técnica de representação de características melhor que as demais (LFCC, LPC, RC e LPCC).

3.3 Energy

Chapter 4

Gaussian Mixture Models

Chapter 5

Fractional Covariance Matrix

Chapter 6

Experiments

Chapter 7

Conclusion

Appendix A

Codes

Bibliography

[1] Jared J. Wolf. "Efficient acoustic parameters for speaker recognition". In: *Journal of the Acoustical Society of America* 51 (1972), pp. 2044–2056.

[2] Steven B. Davis and Paul Mermelstein. "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* ASSP-28.4 (1980), pp. 357–366.