A Study of Speaker Verification System

黃威勛、李立民

ABSTRACT

The main purpose of speaker verification is to identify the speaker according to the related information of voice signals, and it requires a lot of steps to catch the differences between these signals by computers. In this thesis, Mel-Frequency Cepstrum Coefficients, MFCCs, are used as voice characteristic coefficients to match the characteristics of human pronunciation and hearing. Gaussian mixture model is widely used in the field of text independent speaker verification. However, differences of various speakers' voice are not only caused by different oral cavity shapes and vocal cords, but also the articulation speed. Because Gaussian mixture model does not consider the difference of articulation speed, high-order ergodic Gaussian model is adopted in this thesis to implement the text independent speaker verification system. These two models are tested under the same condition, and the results show that high-order ergodic Gaussian model can improve the performance. Equal Error Rate reduces 3.8 percentages.

Keywords: gaussian mixture model; high-order ergodic gaussian model; speaker verification

Table of Contents

第一章 緒論
第二章 語音特徵參數擷取
第三章 向量量化
第四章 建立語者模型
第五章 實驗結果
第六章 結論及未來研究方向

REFERENCES

鍾偉仁, "語者辨認與驗證之初步研究", 國立台灣大學碩士論文, 民國 90 年 6 月。

吳金池, "語者辨識系統之研究", 國立中央大學碩士論文, 民國 90 年 6 月。

王小川, "語音訊號處理", 全華科技圖書有限公司, 民國 94 年 6 月。

李佳蒨, "以高階隱藏式馬可夫模型作語音辨認之研究", 大葉大學碩士論文, 民國 95 年 6 月。