A Study of Speaker Verification System

E-mail: 9707896@ mail.dyu.edu.tw

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 apeakers' 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

REFERENCES

[1] D. A. Reynolds, " Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models", IEEE Trans. Speech and Audio Processing, Vol. 3, No. 1, pp. 72-83, 1995.

[2] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, " Speaker Verification Using Adapted Gaussian Mixture Models", Digital Signal Processing, Vol. 10, pp. 19-41, 2000.

[3] S. Davis, and P. Mermelstein, " Comparing of Parametric Representations for Monosyllable Word Recognition in Continuously Spoken Sentence", IEEE Trans. On Acoustic, Speech and Signal Processing, pp. 357-366, 1980.

[4] L. R. Rabiner, and M. R. Sambur, " An algorithm for determining the endpoints of isolated utterances", The Bell System Technique Journal, Vol. 54, pp. 297 - 315, 1975.

[5] T. K. Moon, " The Expectation-Maximization Algorithm", IEEE, Signal Processing Magazine, Nov. 1996.

[6] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M.Przybocki, " The DET curve in assessment of detection task performance", In

Proceedings of the European Conference on Speech Communication and Technology, pp. 1895-1898, 1997.

[7] S. Young, The HTK book, Version 3.4, Cambridge University Engineering Department, 2002.

[8]   "   ”   90   6

[9]   "   ”   90   6

[10]   "   ”   2005

[11]   "   ”   95   6