

{tag}

{/tag}

Databases and Expert Systems  
2012 by IJCA Journal

IJCA Special Issue on Software Engineering,

©

SEDEX - Number 2

Year of Publication: 2012

Authors:

El Bachir Tazi

Abderrahim Benabbou

Mostafa Harti

{bibtex}sedex1016.bib{/bibtex}

## Abstract

The performances of Speaker Identification Systems (SIS) are strongly influenced by the quality of the speech signal. Most of these systems are based on Gaussian Mixture Models (GMM) that is trained using a training speech database. The mismatch between the training conditions and the testing conditions has a deep impact on the accuracy of these systems and represents a barrier for their operation in real conditions generally affected by noises disturbances. The Voice Activity Detection (VAD) is a very useful technique for improving the performance of these

systems working in these scenarios. In this paper we have used within the feature extraction process, a robust VAD module, that yield high speech/non-speech discrimination accuracy and improve the performance of the SIS in noisy environments. A set of experiments which we have conducted on our proper database containing 37 Arabic speaker in order to evaluate the performances of our SIS based on gammatone frequency cepstral coefficients (GFCC) front-end combined to VAD algorithm show 7.84% average improvement of Identification Rate (IR) performance of our SIS based on GFCC robust method compared to a baseline MFCC method. 2.13% average improvement accuracy as a benefit of VAD technique is observed when the Signal per Noise Ratio (SNR) changes from 40 dB to 0dB.

## Refer

## ences

- J. P. Campbell, "Speaker identification: A tutorial," Proc. IEEE, vol. 85, pp. 1437-1462, 1997.
- S. Furui, Digital speech processing, synthesis, and identification. New York: Marcel Dekker, 2001.
- D. A. Reynolds, et al., "The SuperSID project: exploiting high-level information for high-accuracy speaker identification," in Proc. ICASSP, pp. 784-787, 2003.
- D. A. Reynolds, "Speaker identification and verification using Gaussian mixture speaker models," Speech Comm., vol. 17, pp. 91-108, 1995.
- Y. Shao and D. L. Wang, "Robust speaker identification using binary time-frequency masks," in Proc. ICASSP, vol. 1, pp. 645-648, 2006.
- Sohn, J., Sung, W., 1998. A voice activity detector employing soft decision based noise spectrum adaptation. In: Internat. Conf. on Acoust. Speech Signal Process., Vol. 1, pp. 365-368
- J. A. Haigh and J. S. Mason, "Robust voice activity detection using cepstral features," in IEEE TEN-CON, 1993, pp. 321-324
- D. K. Freeman, G. Cosier, C. B. Southcott, and I. Boyd, "The voice activity detector for the pan European digital cellular mobile telephone service," in Proc. Int. Conf. Acoustics, Speech, Signal Processing, May 1989, pp. 369-372.
- W. Abdulla, "Auditory based feature vectors for speech recognition systems," Advances in Communications and Software Technologies, N. E. Mastorakis & V. V. Kluev, Editor. WSEAS Press. pp 231-236, 2002.
- M. Kleinschmidt, J. Tchorz and B. Kollmeier, Combining speech enhancement and auditory feature extraction for robust speech recognition, Speech Communication, Vol. 34, Issues 1-2, pp. 75-91, 2001.
- B. Tazi, A. Benabbou, M. Harti, "Improved Feature Extraction for Text independent Automatic Speaker Identification System," in CMT'2012, EST USMBA Fez 22,23 and 24 Mars 2012
- Douglas A. Reynolds et Richard C. Rose, "Robust text-independent speaker identification using gaussian mixture speaker models," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol 3, N° 1 pp: 72-83, January 1995.
- Reynolds, Douglas A. Thomas F. Quatieri, and Robert B. Dunn. Speaker Verification Using Adapted Gaussian Mixture Models. Digital Signal Processing. vol. 10, pp. 19-41, 2000.

- Dempster, A. P. , Laird, N. M. , and Rubin, D. B. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society, B, 39, 1–38. December 1976.
- <http://www.speech.kth.se/wavesurfer/>
- S. Furui, An Overview of speaker recognition technology In Proceedings of the ESCA Workshop on Automatic Speaker Recognition, Identification and Verification, pages 1-9, Martigny, Switzerland, April 1994.

Computer Science

### Index Terms

Pattern Recognition

### Keywords

Gaussian Mixture Models (gmm) Mel Frequency Cepstral Coefficients (mfcc)  
Gammatone Frequency Cepstral Coefficients (gfcc)

Speaker Identification System (sis)

Voice Activity Detection (vad)