Abstract

The performance of speaker recognition system considerably degrades if the sample used for speaker recognition task has voices from different speakers in the close vicinity. Solutions to these problems are needed, especially for signals collected in a practical environment, such as in a room with background noise and reverberation. This paper presents a method of determining number of speakers in multi-speaker condition using excitation source information. Speech in a multi-speaker environment are collected using two spatially separated microphones which results in time delay of arrival of speech signals with respect to a given speaker. This time delay is estimated from the cross correlation function of Hilbert envelopes of LP Residual signals. Thus by estimating the difference in time delay for different speakers the number of speakers can be determined. The performance of the proposed method is evaluated by adding different types of noise to the clean speech signal which illustrates the robustness of the proposed method.

References

Index Terms

Computer Science

Power Systems

Keywords

Excitation Source Information

Instants Of Glottal Closure (gcis)

Linear Prediction (lp)

Residual

Envelop (he)

Time Delay Estimation

Different Types Of Noises.