IMPLEMENTATION OF THE HIDDEN MARKOV MODEL IN THE ROBUST SPEECH RECOGNITION

Mr. Vikas G. Mulik[1] , Mr.Vikram A. Mane[2], Mr. Imran M. Jamadar[3]

[1,2] E&TC Department, [3] Automobile Department ,
Annasaheb Dange College of Engineering & Technology, Ashta, (India)

ABSTRACT

Automatic Speech Recognition applications such as voice command and control, audio indexing, speech-to-speech translation, do not usually work well in noisy environments. This paper proposes a new model for a noise robust Automatic Speech Recognition (ASR) based on Hidden Markov Model (HMM) structure with a novel approach for robust speech recognition. A robust and practical speech recognition and Hidden Markov Model (HMM) was proposed aiming at improving speech recognition rate in noise environmental conditions. The system is comprised of three main sections, a pre-processing section, a feature extracting section and a HMM processing section. The proposed estimation methods are applied in combination with oracle masks, which provide an upper performance bound, as well as masks derived from speech presence probability, which represent a more realistic scenario.

Keywords: Automatic speech recognition (ASR), Hidden Markov Model (HMM), robust, feature extracting, oracle masks, speech presence probability.

I. INTRODUCTION

In real-world, in automatic speech recognition (ASR) applications, speech signals generally suffers, degradation by acoustic noise, resulting in decreased system performance. Traditionally the problem of noise robust ASR has been approached in front end feature extraction by reducing variability due to noise while retaining important discriminative information. As an alternative to the previous studies, proposed method explores the missing features approach to speech recognition, in which unreliable spectral components are detected and compensated for accordingly.

Proposed Hidden Markov Model [HMM] based approach is for the reconstruction of spectral speech components degraded by acoustic noise. The basic HMM theory was published in a set of papers by Baum et al. around 1967. TheHMMs derive from the (simpler) Markov chains. Markov chain is a discrete (discrete-time) random process with the Markov property. It is a discrete-time random process because it is in a discrete state (from among a finite number of possible states) at each "step". The Markov property states that the probability of being in any particular state only depends on the previous state it was at before.

By implicitly quantizing spectrographic data, feature trajectories can be interpreted as transitioning through a HMM-defined trellis, with respect to time or with respect to frequency. The proposed method utilizes observed speech together with a local noise estimate to compute observation statistics. With the mentioned transitional and observation information, the proposed HMM-based missing data algorithm uses the traditional forward–backward algorithm to obtain optimal spectral estimates in the MMSE sense.

A major component of missing feature approaches for robust speech recognition systems is mask estimation, which detects the spectral location of reliable features. Many missing feature studies include results based upon oracle masks, for which exact knowledge of a clean .version of the input speech signal is known. Knowledge of oracle masks provides an upper performance bound for data imputation techniques.

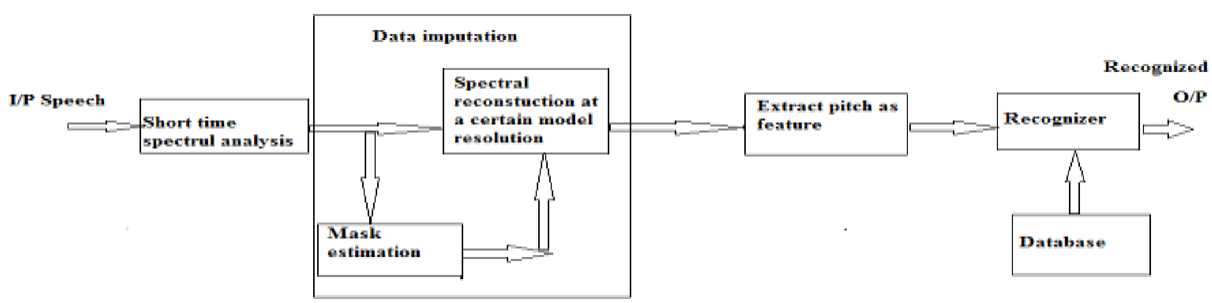
In Order to convey realistic results, mask estimation must be performed. It is proposed to develop spectral masks based on speech presence probability (SPP), in which the probability of active speech is based upon the statistical distribution of speech and noise spectral magnitudes. It is proposed to obtain clean speech & identify the speaker.

II LIMITATION OF EXISTING SYSTEM

a) In the previous studies of ASR system. The MMSE-based techniques can be meaningfully improved by introducing some source a priori knowledge in their formulation by means of a source model. This modeling makes it possible to exploit the high degree of correlation among consecutive signal samples. Speech and images are usually modeled as first-order Markov processes. The main drawback of the HMM-based MMSE techniques is the large computational burden that they involve. HMM-based MMSE estimation involves a recursive probability computation that requires large operation.[3][4]

b) In previous methods of speech recognition uses the front end feature extraction of speech but due to noise it lost the important discriminative information. So it is directly effect on the system performance.[1]

III.FIGURE



In the first stage with the help of microphone we are getting speech and use this speech signal as a input signal. This signal is processed using short time spectral analysis to making a frame[1]. This frequency frame is given to the data imputation block. Imputation means the process of replacing substituted value. After mask estimation extracting feature of speech signal here we can use pitch as a feature. This pitch is compared with the database.

IV.CONCLUSION

In this paper, we have presented a novel HMM-based framework for estimation of unreliable spectrographic data. We utilize hidden Markov models to reconstruct corrupted spectral components based on reliable information, unreliable observations model, to improve noise robust speech recognition.

REFERENCES

Journal papers

[1]Bengt J. Borgström*"*HMM-Based Reconstruction of Unreliable Spectrographic Data for Noise Robust Speech Recognition*"IEEE transaction on audio, speech, and language processing*, vol. 18, no. 6, August 2010

[2] W. Kim and R. Stern, "Band-independent mask estimation for missingfeature reconstruction in the presence of unknown background noise"*in Proc. ICASSP*, 2006, pp. 305–308.

[3] B. Raj and R. Stern, "Missing feature approaches in speech recognition," *IEEE Signal Process.*Mag., vol. 22, no. 5, pp. 101–116, Sep.2005.

[4] B. Raj, M. L. Seltzer, and R. M. Stern, "Reconstruction of missing features for robust speech recognition,"*SpeechCommun*., vol. 43, pp.275–296, 2004.

[5] R. Martin, "Noise power spectral density estimation based on op-timal smoothing and minimum statistics," *IEEE Trans. Speech Audio Process*., vol. 9, no. 5, pp. 504–512, Jul. 2001.

[6] B. Strope and A. Alwan, "A model of dynamic auditory perception and its application to robust word recognition,"*IEEE Trans. Speech Audio Process*., vol. 5, no. 5, pp. 451–464, Sep. 1997.

[7] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition,"*Proc. IEEE*, vol. 77, no. 2, pp. 257–286,Feb. 1989.

Books

[1]Chris Rowden "Speech Processing" ,1991