IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

Speech and Speaker Recognition System Using Artificial Neural Network and Hidden Markov Model

Pallavi Chandratre[1], Shweta Chile[2], Sarvesh Gawali[3], Asmita Jadhav[4]

Associate Professor, Computer Department, SSJCOE College, Dombivli, India [1]

Student, Computer Department, SSJCOE College, Dombivli, India [2, 3, 4]

Abstract: Speaker recognition is becoming an increasingly popular technology in today's society. Besides being cost effective and producing a strong return on investment in all the defined business cases, speaker recognition lends itself well to a variety of uses and implementations. These implementations can range from corridor security to safer driving to increased productivity. By focusing on the technology and companies that drive voice recognition and identification systems, we can learn current implementations and predict future trends.

Keywords: Hidden Markov Model, Artificial Neural Network, Spectogram, Zero Cross Ratio.

I. INTRODUCTION

Speech is the most efficient way to train a machine or communicate with a machine. This work focuses on the objective to recognize the word or the phase spoken by human, keywords in high speed. Speaker recognition is categorized to speaker verification and speaker identification. The goal of automatic speaker recognition is to extract, characterize and identify the information about speaker identity to identify a speaker by his or her voice.

Speaker Recognition approaches can be subdivided into two approaches: text-dependent and text-independent approaches. In the first approach, the speaker is asked to utter a specific phrase pin-code, password, etc.; while in the second approach, the speaker identification engine should catch the characteristics of the uttered speech irrespective of the spoken text. Recognition systems based on hidden Markov models are effective under particular circumstances, but have some major limitations that limit applicability of ASR technology in real-world environments.

II. PROPOSED SYSTEM

Here we propose a methodology to identify speaker and detection of speech. In this approach the input speech signal is acquired at first. The algorithm involves the acquisition of speech signal, processing, matching with the trainingdata and also speaker identification. The algorithm is divided into following parts:

- Acquisition of speech signals
- Reduction of Noise & silence
- Feature Extraction
- Recognition of speech using ANN
- Recognition of speaker using HMM

III. ARTIFICIAL NEURAL NETWORK:

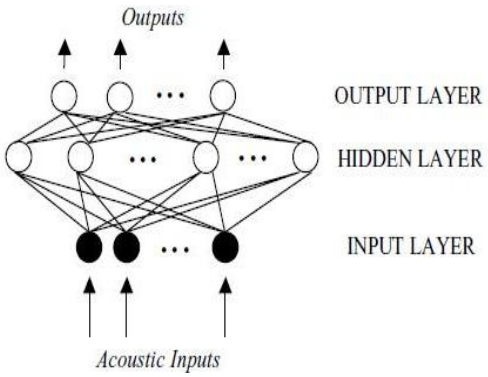

Fig1. Artificial Neural Network

An Artificial Neural Network is a collection of simple processing element, called units or Nodes, which are connected to each other and organised in layers. The processing ability of the network is stored in the inter unit connections, or weights, which are tuned in the learning process. In the learning process, a set of training patterns is presented to the network, and the weights are adjusted to minimize the error between the outputs of the net and the true target values. Thiupdate algorithm of the weights is called back-propagation

IV. HIDDEN MARKOV MODEL

A well-known and widely used statistical method that describes the spectral properties of speech frames is based on the HMM approach. The underlying assumption is that the speech signal can be depicted as a parametric random process, and that the parameters of this process can be estimated in a well-defined manner. A HMM is typically defined as a stochastic finite automaton, usually with a Bakis topology when used for speech. Hidden Markov model (HMM) is the most popular parametric model at the acoustic level.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

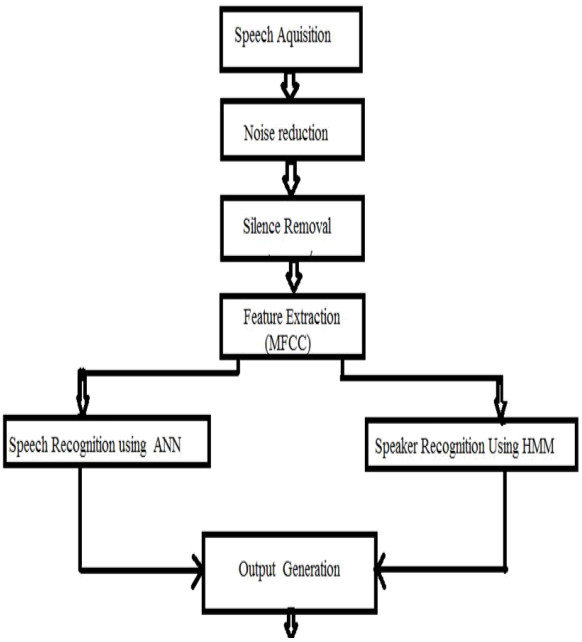

Fig2. Hidden Markov Model

The system can be applied for word recognition as well as for recognition of other speech segments like vowels, phones, words etc. We have a vocabulary of M words to be recognized and each word is to be modelled by a distinct HMM as Shown in Fig.3

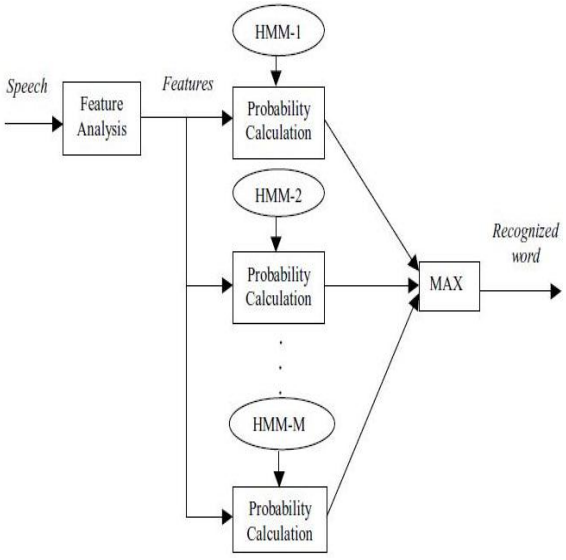

Fig3. HMM Word Recognition

V. EXPECTED RESULTS

Fig4. Acquisition of speech signal

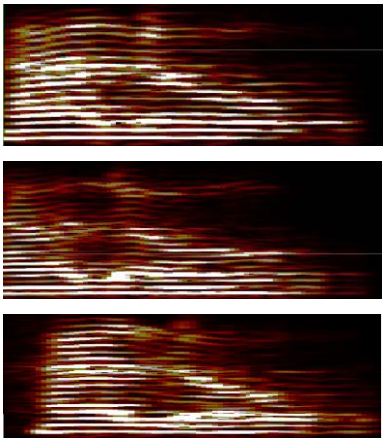

Fig5. Analysis of spectrogram

REFERENCES

[1] –C.T Chen W.D. Chang<<A Feed forward Neural Network With Function Shape Auto tuning>>.Neural Networks Vol9 No 4.pp.627-641.June 1996

[2] J.N Hawang, S-R Lay, M Maechler, R.D. Martin, J. Schimert Regression Modelling in Back-Propogation and Projection Pursunt Learning>>IESE Transaction on Neural Networks ,5(2),342-353

[3] L. Vecci, P .Campolucci, F. Piazza, A. Unicini<<Approximation Capabilities of Adaptive Spline Neural Networks>>In Proceedings of ICNN'97 Huston TX.