VOICE AND SPEECH RECOGNITION IN TAMIL LANGUAGE

KIRAN.R
Department of information Technology,
Sri Sai Ram Engineering College,
Chennai-44, Tamil Nadu.
kiranramasamy95@gmail.com

NIVEDHA.K
Department of information Technology,
Sri Sai Ram Engineering College,
Chennai-44, Tamil Nadu.
niveakalathi26@gmail.com

PAVITHRA DEVI.S
Department of information Technology,
Sri Sai Ram Engineering College,
Chennai-44, Tamil Nadu.
pavithu29@gmail.com

SUBHA.T
Assistant Professor,
Department of information Technology,
Sri Sai Ram Engineering College,
Chennai-44, Tamil Nadu.
subha.it@sairam.edu.in

ABSTRACT:

In our project, our intention is to create a voice and speech recognition system in smart phones that recognizes voice and captures the speech data in Tamil and stores and converts the captured speech as text in Tamil language itself. This can be used in voice dialing, sending SMS by saying out the message and the captured message is sent to the recipient in Tamil. There has not been much consideration for Tamil language to be used in voice and speech recognition in smart phones. For native users Tamil voice recognitions and speech would provide more flexibility in smart phone experience. Also people who have only been used to their native language Tamil, would feel easier to use the speech recognition system in their smart phones if provided in Tamil. There will be no more difficulty in usage of phones for local users and there is no need for any learning to use the smart phone. Automatic Speech Recognition (ASR) system have achieved a great success in many applications. Among them, Template Matching techniques like Dynamic Time Warping (DTW), Statistical Pattern Matching techniques such as Hidden Markov Model (HMM) and Gaussian Mixture Models (GMM), Machine Learning techniques such as Neural Networks (NN), Support Vector Machine (SVM), and Decision Trees (DT) are most popular. For this system, highest word recognition accuracy is achieved with HMM technique. It offered 100% accuracy during training process and approximately 98% for testing process.

Keywords-**Voice, Speech, Recognition, Tamil, HMM, smartphone.**

1. INTRODUCTION:

In computer science and electrical engineering, **speech recognition** (SR) is the translation of spoken words into text. It is also known as "automatic speech recognition" (ASR), "computer speech recognition", or just "speech to text" (STT).The term voice recognition or speaker identification refers to identifying the speaker, rather than what they are saying. Recognizing the speaker can simplify the task of translating speech in systems that have been trained on a specific person's voice or it can be used to authenticate or verify the identity of a speaker as part of a security process. Speech recognition applications include voice user interfaces such as voice dialing (e.g. "Call home"), call routing (e.g. "I would like to make a collect call"), search (e.g. find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), speech-to-text processing (e.g., word processors or emails), and aircraft (usually termed Direct Voice Input).ASR applications are becoming common and useful in this day and age as many of the modern devices are designed and produced user-friendly for the convenience of general public. Speaking/communicating directly with the machine to achieve desired objectives make usage of modern devices easier and convenient. Though there are many interactive software applications, the use of these applications are limited due to language barriers. Hence development of speech recognition systems in local languages will help anyone to make use of this technological advancement.

2. EXISTING SYSTEM:

There are many existing systems for voice and speech recognitions that are used worldwide in form of software and mobile applications. Some of the popular speech recognition systems are Dragon natural speaking, speech logger, talk text, etc. Most commonly used voice and speech recognition software in smart phones is Google now. These existing systems are available only in few languages and have not been used in any of the Indian languages. Thus this hinders the native users to make use of the technical advancement of ASR systems.

3. DISADVANTAGES OF THE EXISTING SYSTEMS:

i.   The use of these applications are limited due to language barriers, that is there is no flexibility for native users.

ii.   Lack of development in speech recognition systems in local languages has hindered Indian smart phone users to make use of this technological advancement who feel difficulty to use these systems in a foreign language rather than their own language.

4. PROPOSED SYSTEM:

Our intention is to create a voice and speech recognition system in smart phone that recognizes voice and captures the speech data in Tamil and stores them and converts the captured speech as text in Tamil language itself. This can be used in voice dialing, sending SMS by saying out the message and the captured message is sent to the recipient in Tamil. There has not been much consideration for Tamil language to be used in voice and speech recognition in smart phones. For native users Tamil voice and speech

recognition would provide more flexibility to use smart phone. Tamil is one of the most widely spoken languages of the world with more than 77 million speakers. Hence, there is an urgent need for the system, so that people can easily interact in Tamil language.

5. HMM - A SPEECH RECOGNITION TECHNIQUE:

HMM is a statistical pattern matching approach for speech recognition. It is the dominant technique for speech recognition based on statistical acoustic and language model is HMM. It uses the automatic learning procedures to model the variations of speech. HMMs is simple networks that can generate speech, using a number of states for each model. The parameters of the model are,

> State Transition Probabilities
>
> Means
>
> Variances
>
> Mixture Weights

that characterize the state output distributions. Each word, or each phoneme, will have a different output distribution. A HMM for a sequence of words or phonemes is made by concatenating the individual trained HMM for the separate words and phonemes. It works by starting at upper left corner of trellis and generate observations according to permissible transitions and output probabilities. The output and transition probabilities define a HMM. This algorithm not only can compute likelihood of single path and can compute overall likelihood of observation string as sum of overall paths in trellis. The HMM addresses three problems to perform ASR and the solutions are given by the following three algorithms.

i. Computing the overall likelihood generating strings of observations from HMM by using the *Forward algorithm.*

ii. Decoding the most likely state sequence/best path from HMM by using the *Viterbi algorithm.*

iii. Learning parameters (output and transition probabilities) of HMM from data by using Baum-Welch also called *Forward-Backward algorithm.*

6. HMM TECHNIQUE - ADVANTAGES:

i. Extendable, because each HMM uses only positive data so they scale well.

ii. Extremely reduce the time and complexity of recognition process for training large vocabulary.

iii. Continuous mixture density HMM is the prevalent model structure that achieves best results in accuracy.

iv. Applications of HMM have been broadened to keyword spotting, speech understanding, and machine translation.

v. Non-speech related applications are emerging as well.

v. complexity of recognition process for training large vocabulary.
.

7. ADVANTAGES OF PROPOSED ASR SYSTEM:

i. Easier, faster, better usage of phones for Tamil smart phone users.

ii. Improves the standards of Tamil language in the smart phone world.

iii. Wide usage of Tamil in speech recognition systems increases the prevalence and creates awareness of the language among people.

iv. Easier for both sender and receiver in case of text messages and the recognition system in mobile phone converts the voice data in Tamil in to Tamil text itself.

v. There will be no more difficulty in usage of phones for local users and there is no need for any learning to use the smart phone.

9. STRENGTH OF HIDDEN MARKOV MODEL:

The strength of the HMM method lies in two broad areas:
(1) Mathematical Framework
(2) Implementation Structure.

Mathematical Framework:HMM uses the statistical methodology and it provides straightforward solutions to related problems.

Implementation Structure: HMM is flexible,so it helps in dealing with various sophisticated speech-recognition tasks and it is easy to implement.

10. SYSTEM FLOW MODEL:
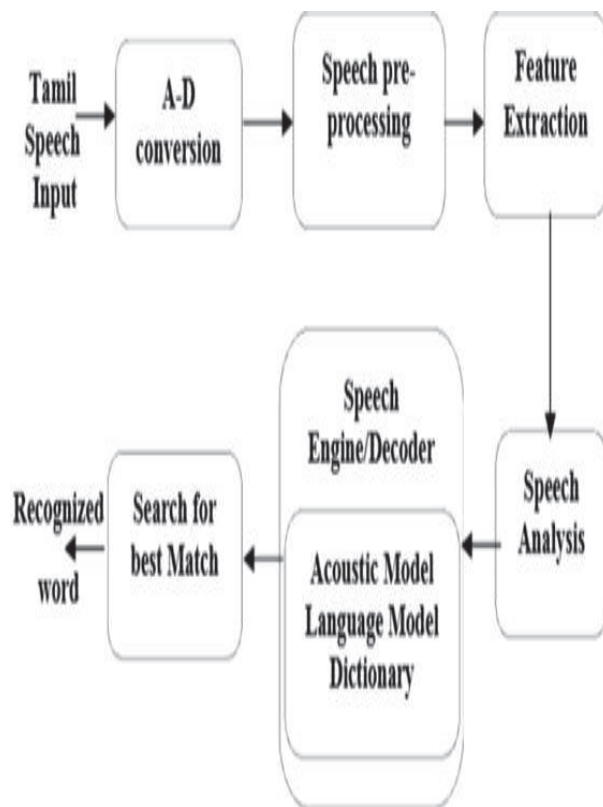
11. SYSTEM ARCHITECTURE(CMU SPHINX):
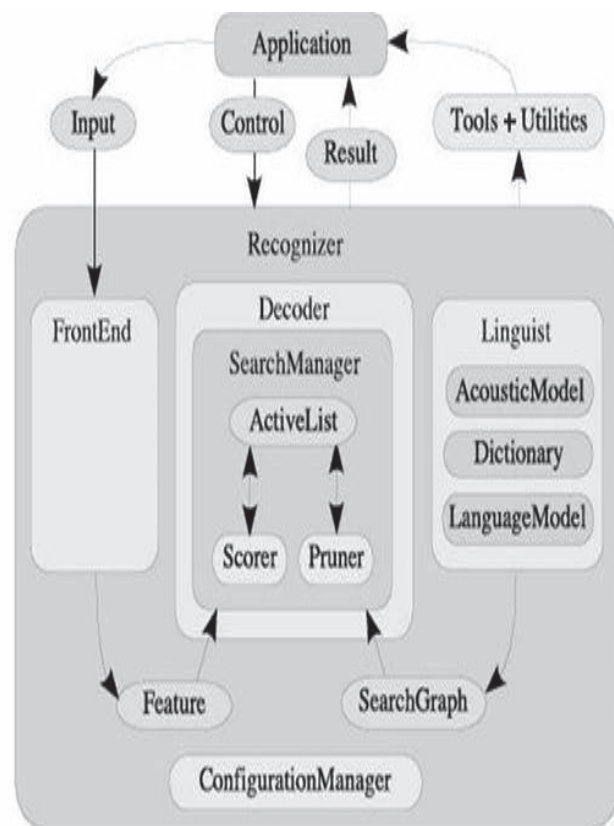
There are three primary modules in the Sphinx-4 framework:

Front End
Linguist
Decoder

Front End - The Front End takes one or more input signals and parameterizes them into a sequence of *Features.*

Linguist - The Linguist translates any type of standard language model, along with pronunciation information from the *Dictionary* and structural information from one or more sets of *Acoustic Models*, into a *Search Graph*.

Decoder - The *Search Manager* in the Decoder uses the Features from the Front End and the Search Graph from the Linguist to perform the actual decoding and generates *Results*.

11. CONCLUSION:

Thus the voice and speech recognition system in Tamil language when implemented would be of great use to native users who feel difficulty with Foreign languages that are default in their smart phones. This enables all types of smart phone users to easily use the technical advancements in their phones in their own indigenous language.

REFERENCES:

[1]   Ms. Vimala C., V. Radha, Speaker Independent Isolated Speech Recognition System for Tamil Language using HMM, Procedia Engineering, Volume 30, 2012, Pages 1097-1102,
[2]   V.S.Dharun, M.Karnan, 2012, Voice and Speech Recognition for Words and Numerals, IJMER, Vol. 2, 5, pp 3406-3414.
[3]   Cini Kurian, A Survey on Speech Recogntion in Indian Languages , *Department of Computer Science ,* Al-Ameen College, Edathala, Aluva, Kerala-683561
[4]   Dr.Thenarasu.S, Tamil pronouncing dictionary: A promising tend in language technology developmentDepartment of linguistics, Central University of Kerala.
[5] Willie Walker, Paul Lamere, Philip Kwok, Bhiksha Raj, Evandro Gouven, Peter Wolf,Sphinx-4: A Flexible Open Source Framework for speech recognition.
[6]   S.Karpagavalli, E.Chandra, A Hierarchical approach in Tamil phoneme classification, *Department of Computer science*, PSGR college for women, Coimbatore.
[7]   http://cmusphinx.source forge.net/wiki/phoneme recognition