Scientific Journal of Impact Factor (SJIF): 5.71
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 03, March -2018

AUTO-SPEECH CASTING SYSTEM FOR PROGRAMING IDE PLATFROM

Pooja Barade [1], Mahesh Chande [2],Neha Debadwar[3],
Varsha Mahajan[4],Prof.Sonali Patil[5]

[1,2,3,4]*B.E. Student, Department of Computer Engineering, Siddhant College of Engineering Pune, Maharashtra, India*
[5]*Associate Professor, Department of Computer Engineering, Siddhant College of Engineering Pune, Maharashtra, India*

ABSTRACT--*Speech Recognition is the process of automatically recognizing a certain word spoken by a particular speaker based on individual information included in speech waves. This technique makes it possible to use the speaker"s voice to verify his/her identity and provide controlled access to services like voice based biometrics, database access services, voice based dialing, voice mail and remote access to computers. Voice can be a powerful tool for use in human computer interaction because it is the fundamental means of human communication. With the rapid growth of wireless communications, the need for voice recognition techniques has increased greatly. Portability and wearability, which are necessary items for being computationally powerful computer devices, will be reinforced by attaching voice applications, since voice can support invisible communication with a computer device as a natural way of communicating. Signal processing front end for extracting the feature set is an important stage in any speech recognition system. The optimum feature set is still not yet decided though the vast efforts of researchers. There are many types of features, which are derived differently and have good impact on the recognition rate. This project presents one of the techniques to extract the feature set from a speech signal, which can be used in speech recognition systems.*

Keywords*- HCI, Voice Identification, Text To Speech, Speech To Text, Speaker recognition, Voice casting, Voice similarity, Voice applications, Voice-based interfaces.*

I. INTRODUCTION

With advances in new technologies, computer devices have grown in popularity to become one of the most common consumer devices. Even as these devices are shrinking in size, however, their capability and content are changing into more complex and diverse functionalities to meet user requests. Now, it is common for many computer devices to include a phone, a personal directory, a memo capability, an alarm clock, a scheduler, a camera, games, and several applications which were working in Personal Digital Assistants (PDAs) before, so there is no boundary between computer devices and PDAs.

However computer devices in which designers have worked more and more to decrease their size are likely to have small keypads and screens, whereas they should have more complex and diverse functions for users. Their functionality and ease-ofuse are greatly limited, and thus many researchers look to find alternative communication channels when interacting with these devices. In recent years, many researchers in the area of human computer interaction (HCI) have attempted to enhance the effectiveness and efficiency with which work and other activities are performed using voice based interfaces. Even if voice technology has been explored for use in desktop computer and telephone information system, the role of voice in interfaces has received little attention because of its difficulty of use and tiresomeness of recognition.

Actually, in the past, the accuracy of voice recognition was unacceptably low, and its a role in a system was questionable because of ambiguity and error. However, voice technology has reached the point of commercial viability and reliability now, and also many computer devices adapt voice applications for providing better services to users. Using voice allows the interface size to be scaled down because voice interaction requires only audio I/O devices such as a microphone and speaker, which are already quite small and inexpensive. Currently, in a computer device voice interfaces need only small space and power consumption, but are able to provide every user with a friendly interface by adding a feeling of natural interaction. voice interfaces are sufficient to replace graphical user interfaces for accessing all information and content without using keyboards, buttons, and touch screens, since voice is the fundamental means of human communication .

II. RELATED WORK

In[1] Voice technology has been explored for use in desktop computers and telephone information systems, so that multiple studies have been focused on voice recognition systems or applications for general computer systems. Roni Rosenfeld and others considered how to build voice application interfaces, especially achieving reliable and accurate speech recognition, and presented their thoughts about the future of speech-based interaction with at least three fundamental advantages for speech:

Speech[2] is an ambient medium rather than an intentional one. Visual activity requires our focused attention while speech allows us to do something else.

Speech[2] is descriptive rather than referential. When we speak we describe objects in terms of their roles and attributes. In visual situations we point to or grasp the objects of interest. For this reason, speech and pointing are to a large extent complementary, and can often be combined to great effect.
Speech[8] requires more modest physical resources. Speech-based interaction can be scaled down to much smaller and much cheaper form-factors than visual or manual modalities.

III. EXISTING SYSTEM

A voice application based on voice interfaces is also useful as a form of input especially when someone's hands, eyes, or ear use computer devices. Voice interfaces and voice recognition technology allow people working in active environments to use them without any holding or touching devices. According to Rick Beasley and others, voice user interfaces (VUIs) are a new concept to many who now have the task of doing everything it takes to develop a voice XML application. Major differences between VUIs Other studies have focused on voice applications especially a speech interface for handheld devices that allow user to capture and randomly access voice notes, which are segments of digitized speech containing thoughts and ideas. To improve user friendliness and dialogue success rate for third generation computer communication systems, multimodal interfaces based on speech displayed a platform which has restricted the functionality to speech centric multimodal interfaces with two input modes: speech and touch, and two output modes: audio and vision. A voice assisted simulation animation architecture (VAS Arch) described voice assist technology by providing software architecture integrating with speech input and output. It supports simulation-animation environments by providing input through spoken commands, mouse manipulation, and keyboard entry so that it can provide more user friendliness to end users.

IV. PROPOSED SYSTEM

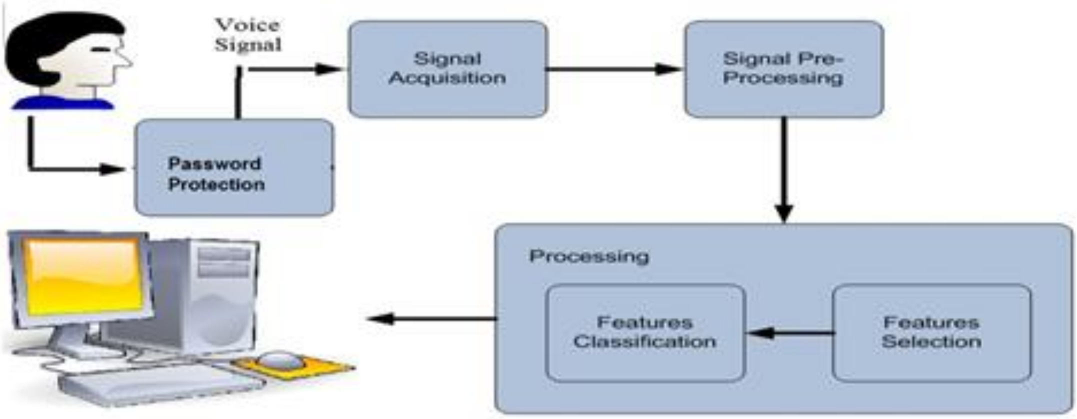


Figure 1: Proposed System

According to B. Tognazzini, voice can be used three ways: to command the computer, to enter information, and to communicate with other people. In this part, we discuss the general components of building a voice application. As seen in Figure 1, a fundamental voice application consists of four basic parts: end user, front-end interfaces, voice recognition system, and dictionary-and-text file database. Each component is explained as follows:

End Users:-
Generally end users mean device users. They can use devices to communicate and make voice feedback with the application, and especially end users are the users who currently use computer devices.

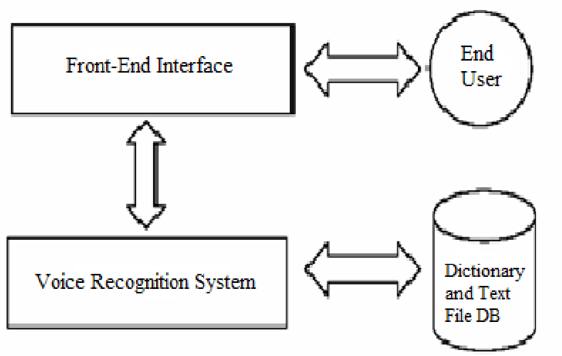


Figure 2: General Voice Application

V. IMPLEMENTATION

5.1. Technology

- Java :Java platform is used for developing the applications.
  - jdk 1.7: Java development kit is used for running the program it has its own jre.
  - jre 1.7: It is java runtime environment used for compilation.
- JVM : Java virtual machine is used for running the java program.Jvm has 3 notions i.e. specification,implementation and instance.
- JMF : Java media framework is used to manage and configure external and internal hardware devices in an application on java platform.
- SVM : Support vector machine is used to store the frames and it overrides the previous frames.

5.2. Algorithm

5.2.1. Interval Search algorithm :-

Algorithm 1: Single interval intersection counter

Input : Sorted interval starts and ends B_S and B_E , query interval a
Output : Number of intervals c intersecting a
Function ICOUNT(BS,BE,a)**begin**

```
    first ⟵ BINARYSEARCH(B_S,a.end)
    last  ⟵ BINARYSEARCH(B_E,a.start)
    c     ⟵ first – last      /*=|B| - (last +(|B|- first)) */
    Return c
```

5.2.2. Token Passing Algorithm :-

Initialisation:

Each model initial state holds a token with value 0;
All other states hold a token with value ∞

Algorithm 2:

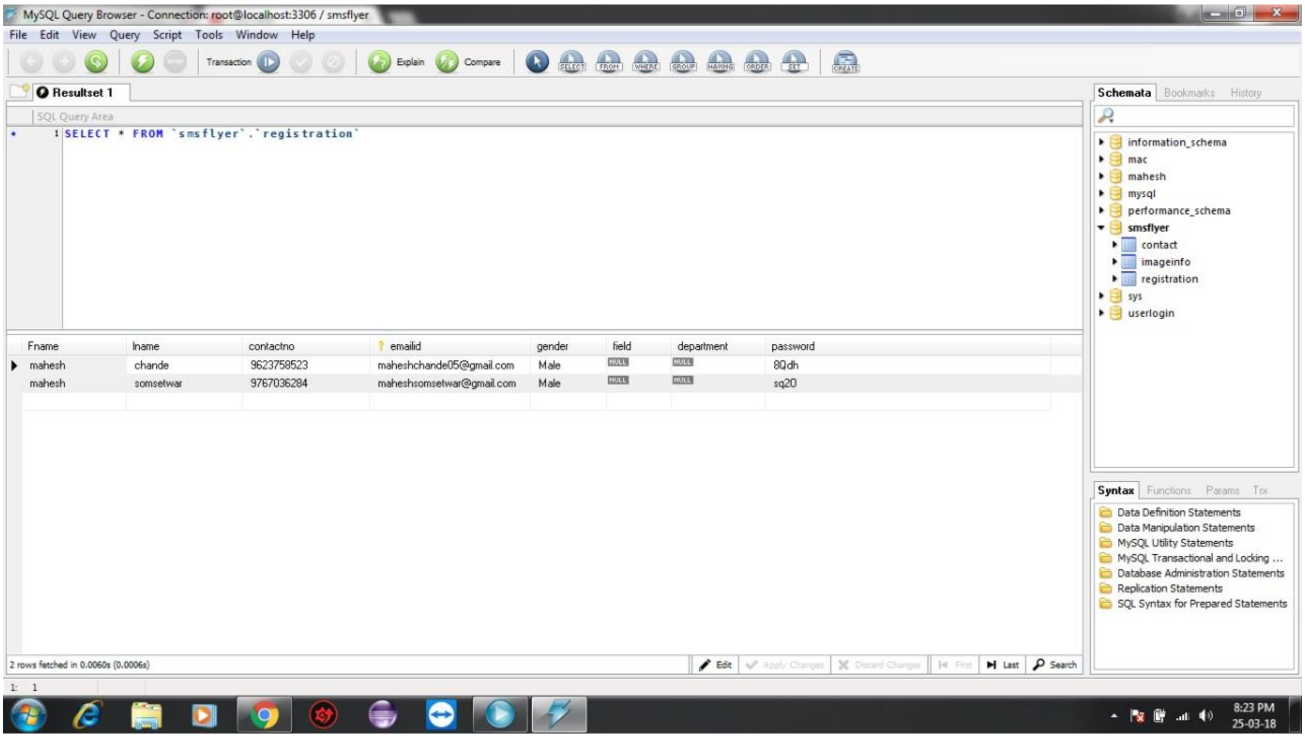
for t := 1 **to** T **do**
  for each state *i* **do**
    Pass a copy of the token in state *i* to all connecting
    States *j*, incrementing its s value by $p_{ij} + d_j(t)$;
  end;
  Discard the original tokens;
  for each states *i* **do**
    Find the token in state *i* with the smallest *s*
    Value and discard the rest
  end;
end;

5.2.3. Synchronous Viterbi Beam Search:

Algorithm 3 : Phone Synchronous Viterbi Beam Search(S,E,Q,T).

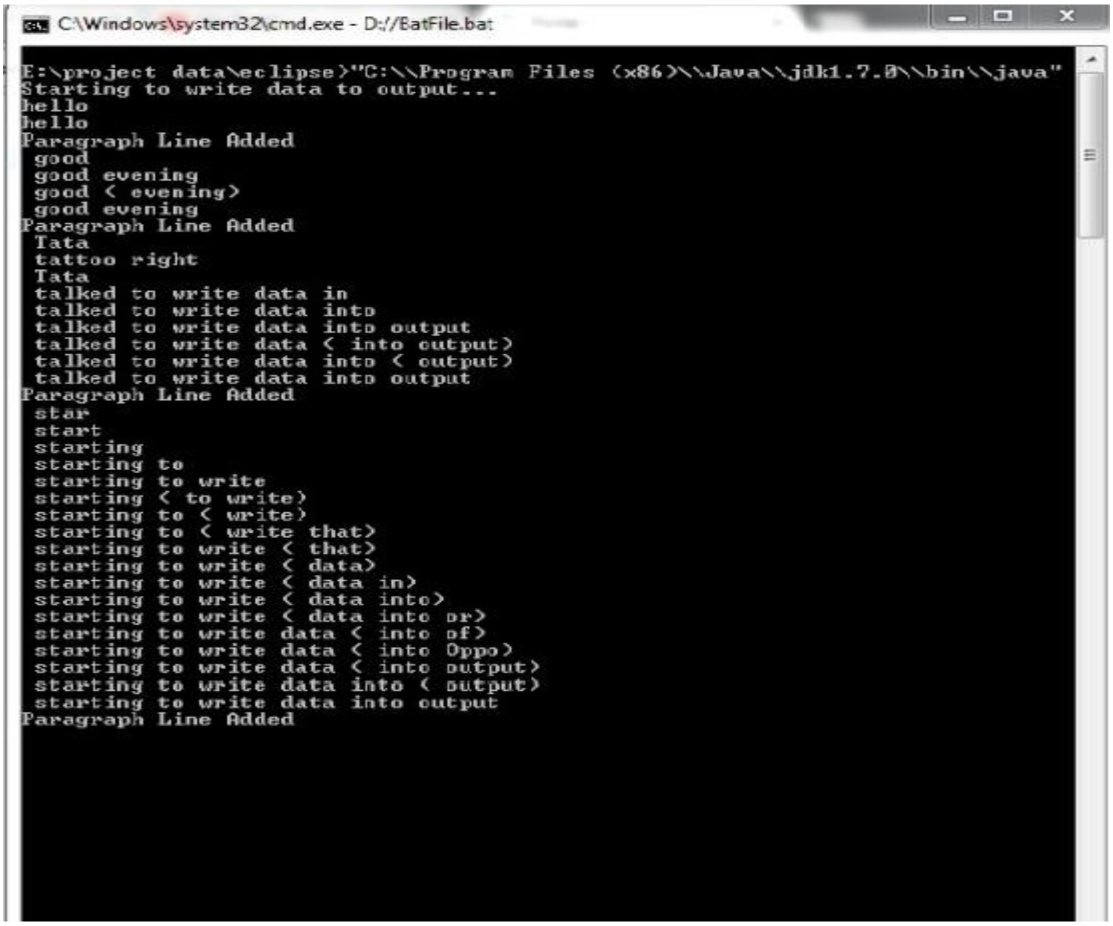
Q ⟵ S   >initialization with start node
For each t ϵ [1,T] **do**   > frame-wise *NN* Propagation
  F ⟵ *NN* Propagate (t)
  if !is BlankFrame(F) **then**   >Phone-wise WFST search
    Q ⟵ viterbi BeamSearch(F,Q)
  End if
End for
B ⟵ finalTransition(E,S,Q)   >to reach end node
Backtrace(B)

5.3. **Database :-**



VI. EXPERIMENTAL RESULT

6.1. Process Window :-



6.2. Result Window :-

VII. CONCLUSION

For fast program execution & construction process. In Existing system there is a problem of time consumption. To overcome this problem we use signals & system techniques. And we also use voice commands.

REFERENCES

1. Zhehuai Chen,YimengZhuang, YanminQian, And Kai Yu, "Phone Synchronous Speech Recognition With CTC Lattices" IEEE/Acm Transactions On Audio, Speech, And Language Processing, Vol. 25, No. 1, January 2017.

2. Nicolas Obin,Axel Roebel, "Similarity Search Of Acted Voices For Automatic Voice Casting" IEEE/Acm Transactions On Audio, Speech, And Language Processing, Vol. 24, No. 9, September 2016

3. H.KidoAnd H.Kasuya,"Everyday Expressions Associated With Voice QualityOf Normal UtteranceextractionBy Perceptual Evaluation", J. Acoust. Soc. Jpn., Vol. 57, No. 5, Pp. 337344, 2001.

4. F. Nolan, P. French, K. Mcdougall, L. Stevens, And T. Hudson, "The Role Of Voice Quality Settings In Perceived Voice Similarity", Presented At The InternationalAssociation Forensic Phonetics Acoustics, Vienna, Austria, 2011.

5. Z. Karam,W. M. Campbell, And N. Dehak, "Graph Relational Features For Speaker Recognition And Mining", In Proc. IEEE Statist. Signal Process. Workshop,2011, Pp. 525528.

6. D. A. Reynolds, T. F. Quatieri, And R. B. Dunn, "Speaker Verification Using Adapted Gaussian Mixture Models, Digital Signal Process", Vol. 10, Nos. 13, Pp. 1941, 2000.

7. W. M. Campbell, D. E. Sturim, And D. A. Reynolds, "Support Vector Machines Using Gmmsupervectors For Speaker Verification", IEEE Signal Process. Lett.,Vol. 13, No. 5, Pp. 308311, May 2006.

8. N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, And P. Ouellet, “Front-End Factor Analysis For Speaker Verification”, IEEE Trans. Audio, Speech, Lang. Process., Vol. 19, No. 4, Pp. 788798, May 2011.

9. B. SchullerEt Al., “The INTERSPEECH 2012 Speaker Trait Challenge”,Presented At The Interspeech, Portland, Oregon, USA, 2012.

10. N. Obin, A. Roebel, And G. Bachman, “On Automatic Voice Casting For Expressive Speech: Speaker Recognition Vs. Speech Classification”, Presented At The International Conf.Acoustics, Speech, Signal Processing.,Florence, Italy,2014.

11. X. Huang, A. Acero, AndH.-W. Hon, “Spoken Language Processing: A GuideTo Theory, Algorithm, And System Evelopment”, 1st Ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

12. F. Jelinek, “Statistical Methods For Speech Recognition”. Cambridge, MA,USA:MIT Press, 1997.

13. T. Hori And A. Nakamura,”SpeechRecognition Algorithms Using Weighted Finitestate Transducers”, Synthesis Lectures Speech Audio Process., Vol.9, No. 1,Pp. 1162,2013.

