International Journal of Pure and Applied Mathematics
Volume 117 No. 20 2017, 311-317
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

Mobile Transaction for Visually Challenged People using Voice Biometric system

RajaKumar.B, Kalaichelvi V, Manimozhi K, Meenakshi P, Vimala Devi P

Assistant Professor, SASTRA UNIVERSITY, KUMBAKONAM ,INDIA
Email :brk@src.sastra.edu

Abstract: Now- a-days, all the transaction are done through the mobile device. It can provide a foundation for improving the online shopping, buying and selling online environments. However, not all potential users have the capabilities that allow them to use the existing methodologies. To improve the search, authentication and navigation of various online transaction and features of mobile devices, we have proposed a model that entail the voice navigation system will be helpful for the visually challenged people for secured mobile transaction.

Keywords: Voice enabled search, biometric system, Visually-disabled users,

I. INTRODUCTION

Mobile voice search is a new research area. It provides users an easier way to search for information using voice from mobile devices. Recent advances in mobile technology have enabled users of voice recognition products to achieve desirable results. Voice Biometric systems can provide knowledge about "liveliness" of the sample being entered by applying liveliness detection techniques. This makes them capable to detect and handle spoofing.

Visually challenged people are depend upon the others for their access of mobile devices so speaking a query is much easier than typing on small devices as well as searching the contacts to make a call. iPhone devices uses voice-enabled local search system is Speak4it.

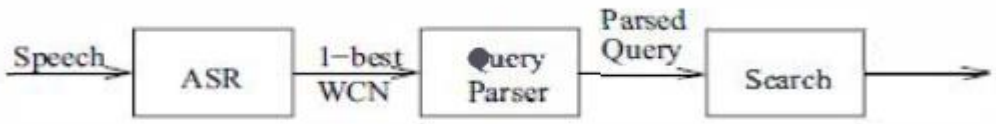

Figure 1: A schematic diagram of major technology components in Speak4it

The available system is based on the natural language understanding (NLU) component is one of the key technology modules. Fig.1. illustrates a schematic diagram of Speak4it with three major technology components such as Automatic Speech Recognizer (ASR), Natural Language Understanding (NLU) and Search. The ASR and Search components perform

speech recognition and search tasks. I Speak4it, the speech recognizer is the AT&T Watson ASR engine [1].

This model focuses only on visually challenged people. When they want to make a call, need to search the address book of the mobile devices. Most of the speech recognition system uses a language model to determine the probability of different recognition hypotheses. In this paper, we present a new approach which is based on the query that leads to appropriate search content and related information. In our approach we have considered three factors that 1.Vocabulary of the language is a set of familiar words within a person's language. A vocabulary, usually developed with age, serves as a useful and fundamental tool for communication and acquiring knowledge. Acquiring an extensive vocabulary is one of the largest challenges in learning a second language. 2. Language model is used in speech recognition, machine translation, part-of-speech tagging, parsing, handwriting recognition, information retrieval and other applications and 3.computational complexity of applying the language model.

II. Literature Survey

To assist the visually disabled individuals, in terms of their mode of interactions with mobile devices for their communication, Junlan Feng et al., proposed an assistive system utilizing voice recognition input technology. In the application, the voice commands are generated to assist the disabled people to operate basic electronic equipment's without necessarily touching the devices. The recent trends and technology provides the various features for visually disabled people to enjoy part of the technological innovations, propose a framework forcontrolling computer desktop with issuance of Voice commands [3].

There are two ends one with a language model that simply used a list of the most frequent queries in their entirely would have the lowest coverage, but would provide the best predictive power within the covered queries (have the lowest per-query perplexity), and would be the least computationally expensive and the other is with sub-word n-gram language models, which have very high coverage, but rather low predictive power (high per query perplexity) [4].

In the fields of computational linguistics and probability, an **n-gram** is a contiguous sequence of **n** items from a given sequence of text or speech. The items can be phonemes, syllables, letters, words or base pairs according to the application. The **n-grams** typically are collected from a text or speech corpus.

Research Challenges

A voice enabled interface faces so many challenges in processing the query because the voice query may vary from person to person in-terms of following reasons: - Voice modulation, time delay between the queries -A auditory environment contains multiple concurrent sources - Named entities are entirely different form the application concepts- User expects the searched contact must perform effective communication with the contacts because it dealt with cost and time.

These challenges can be rectified by introducing a new system called Intelligent Voice Response System (IVRS) [10].

III . Proposed Methodology

Phase I Voice enabled interface

Visually impaired people generally have difficulties in making phone calls; so our approach will be useful for their communication. Voice/Speech recognition is a field of computer science that deals with designing of voice enabled search engine systems that recognize spoken words. It is a technology that allows a device to identify the words that a person speaks into a microphone or telephone. The voice enabled system has at least five different phases: Parameter Extraction, Segmentation, Sound Classification, Sentence analysis and Word boundary determination, and Word recognition.

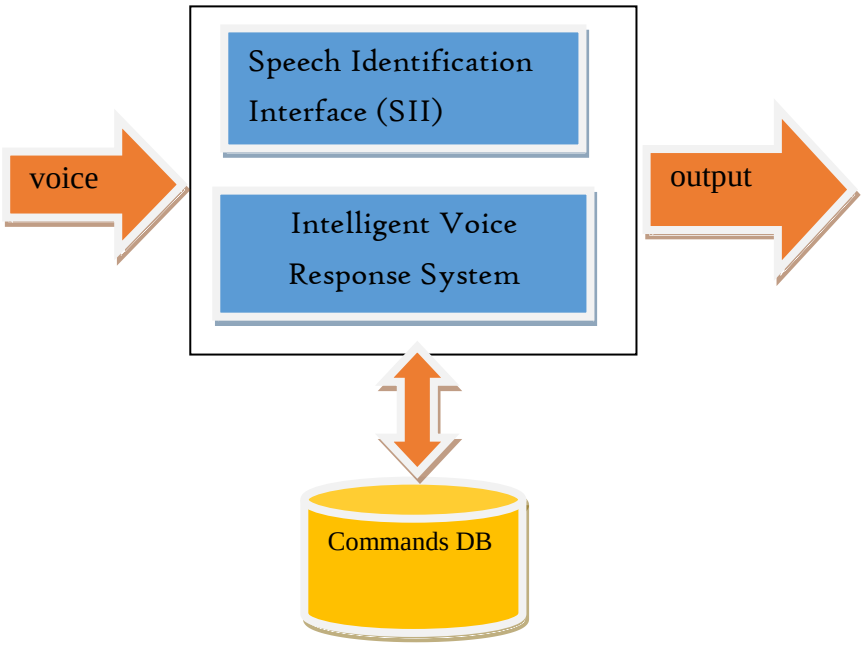

Fig.2.Voice Enabled Interface

Fig.2. shows the proposed approach for voice enabled interface, which consists of various modules such as –Speech Identification Interface (SII) and Intelligent Voice Response System (IVRS).

A. **Speech Identification Interface (SII):**

This module gets the input from the user and processes it and the processed voice is a command, can be used for contact search in any mobile devices. The given command or query can be segmented and associate each segment with needed concepts in the application / address book. A command is a part of a speech which can be used to open or search an application in our context it is a contact in the address book. The set of possible commands are stored into the database. The SII receives the command and compare with the existing commands.If it exists then it leads to the next operation. Named Entity Extraction (NEE), it tempts to identity entities of input voice [11]. It is used to well train the machine learning approaches, annotate the required data, provides the prior probability and segmentation of query / speech.

B. **Intelligent Voice Response System (IVRS):**

This module can be invoked in every output produced by the various states of the entire process. So the user can confirm themselves that whether they have given correct input and received the correct output because this system should assist the visually challenged people without the help of others.

The commands of the proposed model defined in a mobile database which can be compared with the given input while processing the commands. These commands are defined in the Table I. The voice command vector gets the values from the command database.

Table I . Commands and Actions

Voice Commands	Actions
Search	Ready to search the content
Open	It opens the interface and set into the ready state to receive the nextcommand
Cancel	Cancels the process of making call
Call	Makes the call
Hold	Ends the communication

Phase II – Authentication for Mobile transaction

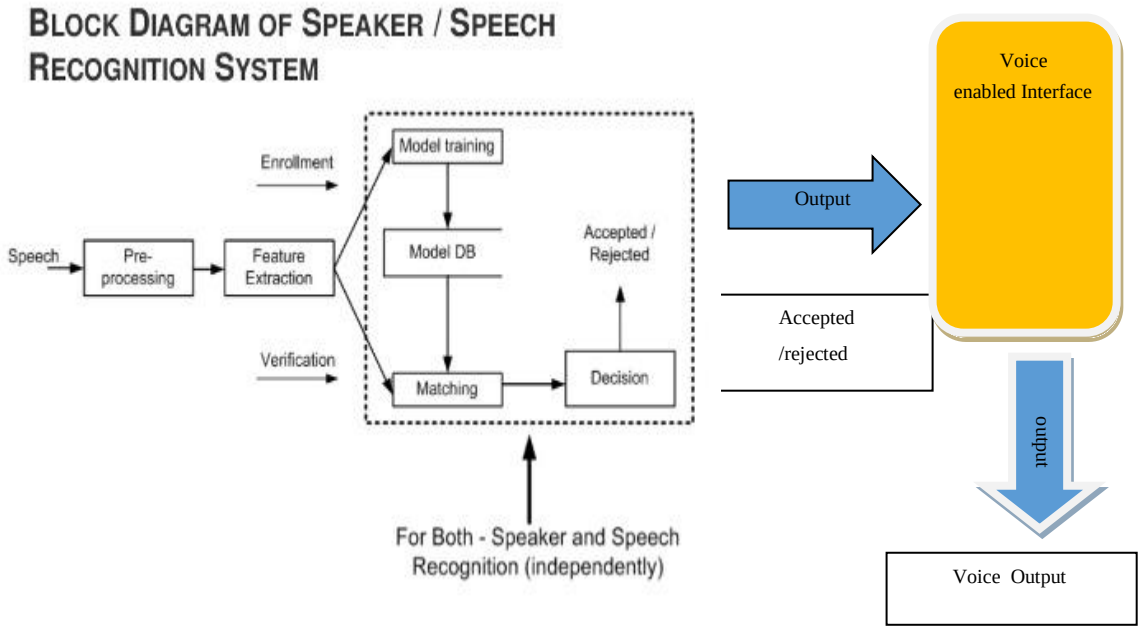

Fig.3. Biometric authentication with Voice Enabled Interface

In this phase, authentication for e-commerce transaction is done by enrolment and authentication phase. The accepted or rejected information as voice output. The voice enabled interface module check the relevant output from biometric authentication systems by binary values either 0 or 1. The output is converted to voice output which is mapped by 0 for

rejected and 1 for accepted and the appropriated message is read out by mobile built-in tools .(ie google speech).

Procedure Voice(input voicecmd)
Begin
If VoiceCmd(1)="Open" then
Open the application
Accept(Read) the Next Cmd
Else
End
If VoiceCmd(0)="Search" then
List ofContacts
Refine the search
Repeat the above step
Accept(Read) the Next Cmd
I VoiceCmd(3)= "SearchContact"
Process the Call/Cancel Operation
Else
Repeat the above step
Repeat the step from begin

IV. CONCLUSION

The proposed model is a new approach for the visually challenged peoples because they can select the different commands and navigate the mobile and make transaction. A typical circumstance contains multiple parallel sources that are also reflected by surfaces. While human listeners are able to segregate andrecognize a target signal under such adverse conditions, ASR remains a challenging problem .SII are trained on clear speech and problem of mismatch when tested in noisy and reverberant conditions. In our future work we are going to enhance this approach for biometric we can use different traits and recognition methods.

REFERENCES:

[I] V. Gotn, G. Riccardi C. Allauzen D. Hakkani A. Ljolje, S.Parthasarathy M. Rahim, and M. Saraclar., "The at&tWatson speech recognizer," ICASSP, 2005.

[2] Gerr F., Marcus M., Ensor c., Kleinbaum D., Cohen S., Edwards A., Gentry E., Ortiz D., Monteilh C." A prospective study of computer users: Study design and incidence of Musculo skeletal symptoms and disorders", American Joural of Industrial Medicine 41, pp.221-235,2002.

[3] Junlan Feng. , AT&T Labs Research, "Query Parsing in Mobile Voice Search", April 26-30, 2010, Raleigh, North Carolina, USA.

[4] R. Lau, "Subword Lexical Modelling for Speech Recognition", Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, 1998.

[5] http://en.wikipedia.org/wiki/Language_modei

[6] http://en.wikipedia.org/wiki/Bigram

[7] A. Stolcke, "Entropy-based pruning of backoff language models" , In Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop, pages 270-

274, Lansdowne, VA, 1998.
[8] Scan sof, "Embeded speech soloutions retrieved", January 25, 2005 http://www.speechworks.com/
[9] Robertson, J., Wong, Y.T., Chung, c., and Kim, O.K., "Automatic Speech Recognition for Generalised Time Based Media Retrieval and Indexing", Proceedings of 6th ACM International Conference on Multimedia(pp 241-246) Bristol, England, 1998.
[10] RajaKumar.B, Michael Raj.T.F, "Sensor Based Medical Monitoring Model with Intelligent Response System", International Conference on Computer Communication and Informatics, Coimbatore, 2013.
[11] B. T. F. Pengo, "Unsupervised query segmentation using generative language models and Wikipedia", In Proceedings of WWW, 2008.
[12] Manas Kumar Choudhury and Y.Srinivas, "A Novel Framework for Multimodal Biometrics based on Dimensionality Reduction Technique for Reducing Overlapping Features", International Journal of Current Engineering and Technology , E-ISSN 2277 – 4106, P-ISSN 2347 – 5161 Vol.5, No.1 (Feb 2015)
[13] Itorobong A. Inam, Ambrose A. Azeta, Olawande Daramola, "Comparative analysis and review of interactive voice response systems", Information Communication Technology and Society (ICTAS), May 2017
[14] Mohan Dholvan , Dr. Anitha Sheela Kancharla, "Speech enabled Voice Response System– Review", International Journal of Computer Engineering and Applications, Volume 9, Issue 10, Oct. 2015