VOICE COMMANDS CONTROL RECOGNITION ANDROID APPS

Renu Tarneja[1], Huma Khan[2], Prof. R. A. Agrawal[3] Prof. Dinesh. D. Patil[4]

[1,2] Bachelor of Engineering Students, Department of Computer Science & Engineering

[3,] Assistant Professor, Department of Computer Science & Engineering

[4,] Associate Professor & Head, Computer Science & Engineering Department

Hindi Seva Mandal's, SHRI SANT GADGE BABA COLLEGE OF ENGINEERING & TECHNOLOGY,

BHUSAWAL – 425203, DIST – JALGAON, MAHARASHTRA, INDIA

E-mail –tarneja.renu77@gmail.com

Contact-+91-9404051266

ABSTRACT- Smart mobile phones have become modern-day Swiss Army knives like for an example an Android smart phone is a handheld smart phone, a music player, a notepad, a GPS navigation unit and more functionalities, all rolled into one sleek device that fits in your pocket. Today's smart phones do so many things for us that sometimes we don't even think about how we do them. Even though Smart phones do all these new things, the most natural way of interacting with a phone remains the same as what it always has been traditionally. The mobile phone users are increase day by day and everyone wants to develop new things which are improved version of existing one. Voice Actions are a series of spoken commands that let you control your phone using your voice. Voice Recognition is a fascinating field spanning several areas of computer science and mathematics. Speech recognition technology, which is able to recognize human speech and change to text, or to perform a command, has emerged as the 'Next Big Thing' of the IT industry. Speech recognition is technology that uses desired equipment and a service which can be controlled through voice without touching the screen of the android smart phone. Reliable speech recognition is a hard problem, requiring a combination of many techniques; however modern methods have been able to achieve an impressive degree of accuracy. At the end, we're pleased to introduce some new set of features for Voice Actions for Android platform smart mobile phones. The device proposed here is an interactive android smart phone, which is capable of recognizing spoken words. We propose to develop interactive application which can run on the tablet or any android based phone. The application helps the user to open any application as well as call any contact through voice. Users can command a mobile device to do something via speech. These commands are then immediately executed.

Keywords: Smart Phone, Android, Speech Recognition, Voice, Mobile Phones, Application, Call.

INTRODUCTION

People love their mobile phones because they can actually stay in touch wherever they are. That means not just for talking, but e-mailing, texting, and so on. We are constructing with the growth of mobile phone technology. As per the users are increase day by day, facilities are also increasing. Let's start with simple regular handsets which were used just for making calls, Smart Phone have affected our lives and have become a part of it. Now a day's smart phones are not use for making calls but they have innumerable uses and can be used as a Camera , Music player, Tablet PC, T.V. , Web browser etc . New application and operating systems are required with the new technologies. In recent years, smart phones have placed an increasing emphasis on bringing speech technologies into limelight usage. This focus has led to products such as Speech server. Now let's limit our focus towards voice message system it is a service component of phone, using standardized communications protocols.

As we all know cell phones are very important part of modern life. Many of us need to make a urgent call or message at anytime from anywhere. Many of them needs their cell phones when they can't do so example at the time of driving, cooking accidents may occur because of this activity ,a voice recognition application for mobile device is being develop to avoid harmful incidents. Android is the one of the best suitable operating system for this kind of system. It is the open source operating system that is use to develop application for mobile users.

Speech recognition also appeared as part of ongoing research in progress in 1950s, but was not popularized until the mid-2000s, with low voice recognition. Presently, related speech recognition technologies, which have been previously used limitedly for special-purposes, have been rapidly evolving because of the proliferation of portable computing terminals such as smart phones interconnected with the expansion of the cloud infrastructure . Siri is one of the most prominent examples of a mobile voice interface, latest iPhone have built in voice activated personal assistant . Also Android, the Windows Phone platform, and most other mobile systems have voice functionality as well as many other applications. While these interfaces still have considerable constraint, we are inching closer to machine interfaces we can actually talk to.

LITERATURE REVIEW

Hae-Duck J. Jeong, Sang-Kug Ye, Jiyoung Lim, Ilsun You and Woo Seok Hyun[1] had proposed a computer remote control system using voice recognition technologies of mobile devices and wireless communication technologies for the blind and physically

disabled population as assistive technology. These people experience difficulty and inconvenience using computers through a keyboard and/or mouse. The purpose of this system is to provide a way that the blind and physically disabled population can easily control many functions of a computer via voice. The configuration of the system consists of a mobile device such as a smartphone, a PC server, and a Google server that are connected to each other. Users can command a mobile device to do something via voice; such as writing emails, checking the weather forecast, or managing a schedule. These commands are then immediately executed. The proposed system also provided blind people with a function via TTS(Text To Voice) of the Google server if they want to receive contents of a document stored in a computer. Halimah, B.Z. Azlina, A. ; Behrang, P. ; Choo, W.O. [2] had proposed the Mg Sys Visi system that has the capability of access to World Wide Web by browsing in the Internet, checking, sending and receiving email, searching in the Internet, and listening to the content of the search only by giving a voice command to the system. In addition, the system was built with a translator that has the functionality to convert html codes to voice; voice to Braille and then to text again. The system comprised of five modules namely: Automatic Speech Recognition (ASR), Text-to-Speech (TTS), Search engine, Print (Text-Braille) and Translator (Text-to-Braille and Braille-to -Text) module, was originally designed and developed for the visually impaired learners, can be used for other users of specially needs like the elderly, and the physically impaired learners. Initial testing of the system indicates very positive results.Md. Sipon Miah, and Tapan Kumar Godder [3] had proposed a voice Control Keyboard Systems, fully controlled by a computer, and display output on the display device with predefined time. So the project will work as a helping system for those person who has small knowledge about computer system even those person who are illiterate they can operate computer system. They implemented the system in other system for example voice control car system.

DESIGN

Our Android App which we are going to design will have these functionality: Selection of any App and Calling a Specific Contact from Phone book. We may use incremental model to implement our android app with above functionalities. Because if we would like to make any changes in the android app in future, it would be easily accommodated. Incremental model is an enhancement of waterfall model. The product is designed, implemented, integrated and tested as a series of incremental builds. Many commercial software companies and system vendor uses popular model software evolution. Incremental software development model may be applicable to projects where:
- ✓ Software Requirements are well defined, but realization may be delayed.
- ✓ The basic software functionality are required early

Initially software requirements are broken down into multiple standalone modules under incremental model. Under the software project these modules are drafted according to the level of priority they have. According to the implementation and progression of the project every module as a standalone function has to be developed. Although they can be inter related but they sure are able to exists without having the need of other modules and functionality. Incremental Model is one of the most followed model by large number of project which requires implementing individual functions and adding standalone models in the long run. There are three phases in each increment: design, implementation and analysis. During the phase of design of the first increment, the functionality with topmost priority is selected and the design is prepared. In the phase of implementation, the design and testing is implemented. In the phase of analysis, the functional capability of the partially developed product is analyzed. The development process is repeated until all the functions of the project are implemented.

METHODOLOGY

There are several well defined steps for the process of voice recognition. Different systems vary on the nature of these steps, as well as how each step is implemented, but similar methodology is followed by most successful systems.
1. Firstly the sound wave is divided into evenly spaced blocks.
2. Each block is processed for important characteristics, for example strength across various frequency ranges, total energy and number of zero crossings.
3. Attempt to associate each block with a phone using this characteristic vector, this is the most basic unit of speech, producing a string of phones.
4. Search the word whose model is the most likely similar to the string of phones which was produced.

A spectrum analysis of the block is typically being performed in step2. This can be done with a bank of frequency filters or with Fast Fourier Transform (FFT), but Linear Precidive Coding is one of the most successful technique to date. Analyzing the total energy, the number of zero crossings, and the change in the features over time are the additional important features. Step 3 is often done through a decision tree. The field of consideration is narrowed by phone which often has very prominent characteristics. Similar sounding phones are separated using additional characteristics. The final decisions are often mistaken, and these mistakes must be accounted for later. With a high degree of success using Hidden Markov Models (HMM's) step 4 has been implemented. To determine which model is the most likely match a Hidden Markov Model is constructed for each word in the vocabulary, and then the string of phones is compared against each HMM.
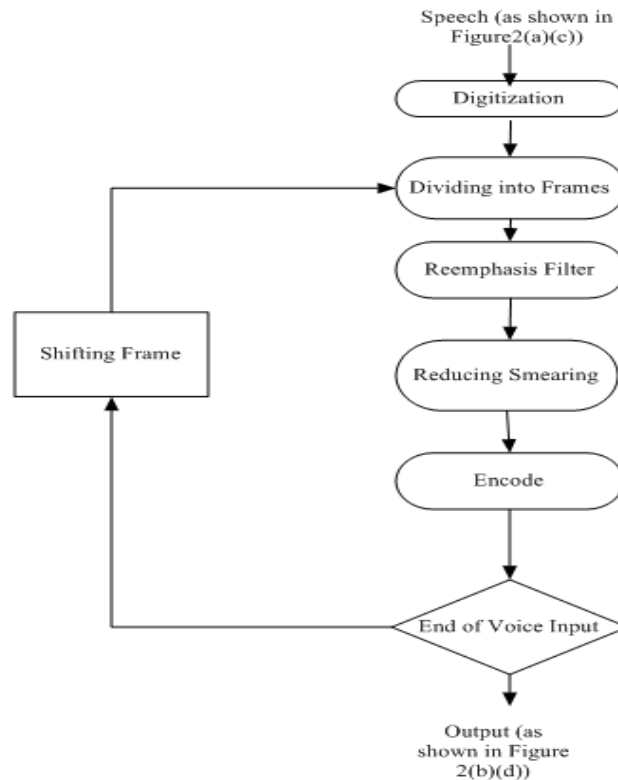
IMPLEMENTATION

Figure 1: Flowchart - Converting Speech signal into a set parameters suitable for ASR

Figure 1 shows the steps involve in converting speech signal into a set of parameters are: The main purpose of the digitization process is to produce a sampled data representation of the speech signal with as high a Signal to Noise ratio (SNR) as possible. The process of grouping digitalized speech into a set of samples, called frame. A one coefficient digital filter, known as a Reemphasis filter. This stage spectrally flattens the frame using a first order filter. Windows are functions defined across the time record which are periodic in the time record. They start and stop at zero and are smooth functions in between. When the time record is windowed, its points are multiplied by the window function, time bin by time bin, and the resulting time record is by definition periodic. It may not be identical from record to record, but it will be periodic (zero at each end). In the frequency domain, a window acts like a filter. The net result of windowing is to reduce the amount of smearing in the spectrum from signals not exactly periodic with the time record. To encode the speech signal into a suitable set of parameters three basic classes of techniques are being used:

- Fourier transformations
- Filtering through digital filter-banks
- Linear prediction

Figure 2 shows the results obtained after giving voice commands as input in part (a)The voice input is given to open an app and the corresponding app is opened in part (b).In part (c) the voice is input for calling a contact and in part (d) the corresponding contact is called.
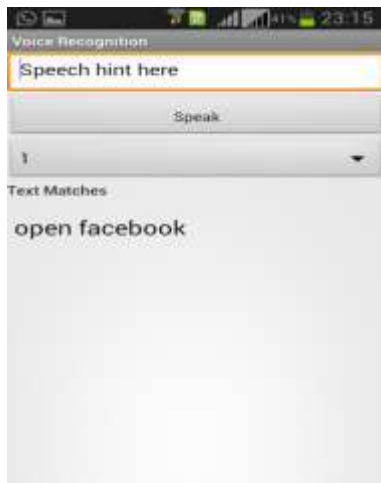
(a)

(b)

(c)

(d)

Figure 2: Results Obtained after giving voice commands as input.

Table 1: shows the Comparative Study of our voice recognition system on different Android Version..The performance of good, fair, best and excellent is being developed on the basis of number of seconds our voice recognition system takes for recognising   a voice input. it is found that on Jellybeans the performance of our voice recognition system proved to be best, and on Kitkat version it proves to be excellent and in the older version éclair and froyo the performance appear to be good but in the honeycomb and ice cream sandwich version it appears to be fair.

TABLE 1: Comparative Study of our voice recognition system on different Android Version

Android Version	Number of seconds - recognizing Voice Input	Performance
Eclair (2.0–2.1)	8sec	Good
Froyo (2.2–2.2.3)	8sec	Good
Gingerbread (2.3–2.3.7)	8sec	Good
Honeycomb (3.0–3.2.6)	7sec	Fair
IceCream Sandwich (4.0–4.0.4)	7sec	Fair
Jelly Bean (4.1–4.3.1)	6sec	Best
KitKat (4.4–4.4.4)	5sec	Excellent

Note: The voice recognition system will not work on android versions below eclair.

TARGET AUDIENCE & APPLICATIONS

The goal is to provide a system that can easily help the blind, physically disabled personalities to operate and control a smart mobile phone through their voice commands as input. The scope of this system is not just limited to such personalities but even more it will acquiesce excellent outcome for others too.

TABLE 2: Application of Voice Recognition System

Applicable Areas	Details
Education Sector	Speech to text processing, to correct pronunciation of vocabulary on foreign languages. Use keyboard to enter text verbally for handicapped students.
Medical Sector	Precision surgery, Automatic wheel chair, Medical transcription (digital speech to text).
Military Sector	Automatic aircraft control, helicopter, Training air traffic controller, Automatic ammunition control
Communication Sector	Voice dialling, Telephone directory inquiry without operator assistance.
Domestic Sector	Ovens, refrigerators, washing machine, home appliances control etc.
General	Use for security purposes at highly secure places, Dictation system on market. To translate data from one language to another, video gaming and ATM (data entry).

ACKNOWLEDGEMENT

CONCLUSION

A smart phone based voice recognition system can be developed to operate an app with simple voice commands. This technology can be implemented into a user friendly and compact device. By creating online voice-controlled applications, people are free to choose which device better suits them. This paper proposed the capability of modern voice recognition software to increase independence for persons with disabilities. Major purpose of this system was to provide a system so that the blind and physically disabled population can easily control many functions of a smart phone via voice. The system is very useful for the general population as well. Users command a mobile device to do something via voice such as directly controlling smart phones. These commands are then immediately executed.

REFERENCES:

[1] Hae-Duck J. Jeong, Sang-Kug Ye, Jiyoung Lim, Ilsun You, and WooSeok Hyun ," A Computer Remote Control System Based on Speech Recognition Technologies of Mobile Devices and Wireless Communication Technologies", 'IEEE Conference Publication',2013,page no. 595-600

[2] Halimah, B.Z. Azlina, A. ; Behrang, P. ; Choo, W.O.,"Voice recognition system for the visually impaired: Virtual cognitive approach "IEEE Conference Publications Volume: 2 ,DOI: 10.1109/ITSIM.2008.4631738, Publication Year: 2008 , Page(s): 1 - 6 .

[3] Md. Sipon Miah, and Tapan Kumar Godder , "Design Voice Control Keyboard System using Speech Application Programming Interface "IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 6, November 2010 ISSN (Online): 1694-0814 www.IJCSI.org 269 To 277.

[4] Jarng, S.S.: "Analysis of HMM Voice Recognition Algorithm." Journal of Advanced Engineering and Technology , publication year-2010 volume no. 3(3), page no. 241–249.

[5] Jeong, H.D., Ye, S.K., Lim, J., You, I., Hyun, W., Song, H.K." A Remote Computer Control System Using Speech Recognition Technologies of Mobile Devices." The Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing: Future Internet and Next Generation Networks (FINGNet-2013).publication year-2013 pp. 595–600.

[6] Ms. Anuja Jadhav Prof. Arvind Patil." Android Speech to Text Converter for SMS Application" IOSR Journal of Engineering Mar. 2012, Vol. 2(3) pp: 420-423.

[7] Mao, Q.R., Zhan, Y.Z." A Novel Hierarchical Speech Emotion Recognition Method Based on Improved DDAGSVM." Computer Science and Information Systems (ComSIS) publication year 2010,vol 7(1), page no-211–221.

[8] Torrente, J., ´A.d. Blanco, ´A. Serrano-Laguna, Vallejo-Pinto, J., Moreno-Ger, P., Fern´andez- Manj´on, B.''Towards a Low Cost Adaptation of Educational Games for People with Disabilities". Computer Science and Information Systems (ComSIS)publication year 2014, vol 11(1), page no 369–391.

[9] Nakano, S. ; Tsubaki, T. ; Hironaka, S.," Applying a voice_recognition_system for substation inspection/maintenance services",IET   Conference   Publications,   Volume: 5 DOI:10.1049/cp:19990940 Publication Year: 1999 , Page(s): 280 - 283 vol.5.

[10] Ferrando, F. ; Nouveau, G. ; Philip, B. ; Pradeilles, P. ; Soulenq, V. ; Van-Staen, G. ; Courmontagn," A voice_recognition system for a submarine piloting", OCEANS 2009 - EUROPE, IEEE Conference Publications DOI:10.1109/OCEANSE.2009.5278189 Publication Year: 2009 , Page(s): 1 – 6.

[11] Marieta Gâta," speech recognition application using voice xml ", Proceedings of the International Conference on Theory and Applications of Mathematics and Informatics - ICTAMI ,publication year 2004, Thessaloniki, Greece 146.

[12] Bodruzzaman, M. ; Kuah, K. ; Jamil, T. ; Wang, C. ; "Parametric feature-based voice_recognition_system using artificial neural network "IEEE Conference Publications , Publication Year: 1993 DOI:10.1109/SECON.1993.465673,