

Analysis of Speech Recognition with Voice Assistant technology

Mr. Yashraj R Marwah¹, Ms. Swapnali D. Mahadik¹

¹(Auxillary Tester, BNP Paribas, Mumbai, India)

²(Assistant Professor, Navinchandra Mehta Institute of Technology and Development, Mumbai, India)

Abstract: Digitization brings new possibilities to ease the daily life activities by the means of Voice Assistant Technology(VAT). Due to this VAT managing all the daily work can be performed in one go. An activity which need to be done by a human manually is now done by a voice assistant. The main objective of this paper is to know whether it will be able to identify the voice is of the same person or someone else is imitating it. Security is the most important factors of this paper. This would not only increase the security but also reliability of the system.

Keywords: Digitization, Voice Assistant Technology (VAT), Security

I. Introduction

In this digital world people want their task to be completed fast and in quick time. People want an assistant who can perform their task by just saying it. Voice Assistant Technology (VAT) is where it comes into picture. Several consumer-level products developed in the last few years have brought inexpensive voice assistants into everyday use, and more features and platforms are being added all the time. [1]. VAT is a technology that a person uses to do the daily work/task in no time like Playing the music, to know what are the latest headlines in news, to set alarm and reminders etc. Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistant are the most popular voice assistants and are embedded in smartphones or dedicated home speakers. Users can ask their assistants questions, control home automation devices and media playback via voice, and manage other basic tasks such as email, to-do lists, and calendars with verbal commands [1].

Recent studies show that these systems are increasingly becoming part of many people's everyday lives. The worldwide number of assistant users is expected to grow from 390 million in 2015 to 1.8 billion in 2021 which will result in a total revenue increase from \$1.6 billion to \$15.8 billion. [3] But as it is said with something good and nice there is always a problem attached. Like that only VAT also has disadvantages. One of the main disadvantages related to this voice-assistant technology is security.

Comparing between the original voice of a person and a dummy voice of the person to check whether the VAT can make the difference is discussed in this paper. Comparing of voice can be done using MATLAB R2018b software which can record the voice and show it in a graphical format.

II. Literature View

Knote, R.; Janson, A.; Eigenbrod, L. & Söllner, M. (2018) has presented an exhaustive, integrative literature review to build a solid basis for future research. Identified five functional principles and three research domains which appear promising for future research, especially in the information systems field. Hence, contributing by providing a consolidated, integrated view on prior research and lay the foundation for an SPA classification scheme. [3]

Imrie P, and Bednar P. (2013) has reviewed and discussed the ways in which new technology could be harnessed to create an intelligent Virtual Personal Assistant (VPA) with a focus on user-based information. [4]

Dr. Kshama V. Kulkarni, Dr. Kotrappa Sirbi, Mr. Abhijit J. Patankar has proposed a system that has capability to work with and without Internet Connectivity. It is named as Personal Assistant with Voice Recognition Intelligence, which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user. [5]

Ms. Ayushi Y. Vadwala, Ms. Krina A. Suthar, Ms. Yesha A. Karmakar, Prof. Nirali Pandya has developed an Intelligent Android Voice Assistant system using speech recognition technologies of mobile devices for every common user who is interested in AI personal assistant. The purpose of voice assistant systems is the exchange of information in a more interactive approach using speech to communicate. [6]

III. Process Framework

This section represents the actual implementation of model in a Matlab which records the voice and plots the graph which can be easily analyzed.

For Creating graph from recording the voice using MATLAB application:

Open MATLAB R2018b Application. Create method audiorecorder () which contains 3 parameters, first it contains the sampling frequency then the number of bits and then the channel Id and assign it to variable say a1. Now Record the voice using the method record which contains the variable of the audiorecorder () and for how much time the recording should be done. Now get the audio using the method getaudiodata () which contain the recorded object and assign it to a new variable say b1. Use method sound () to listen the recording which is stored in b1. Use Plot () method to view the graphical representation of the recorded voice.

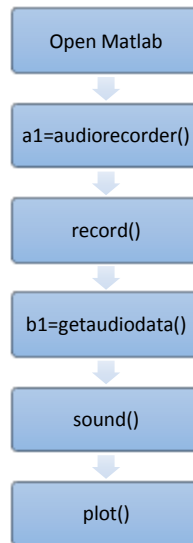


Figure no 1: Process of recording voice and plotting a graph

```
Command Window
>> a1 = audiorecorder(8000, 8, 1);
>> record(a1, 5);
>> b1 = getaudiodata(a1);
>> sound(b1)
>> plot(b1)
fx >> |
```

Figure no 2: Steps for recording the voice and plotting the graph

Graphical Representation:

1) Graph On Laptop Speaker:

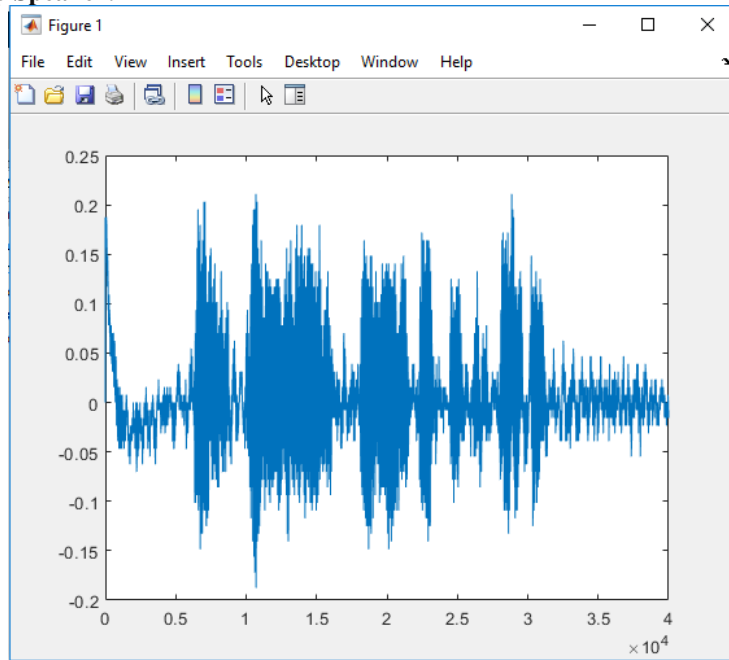


Figure no 3: Graph of voice recorded on speaker of laptop

Fig 3 shows that on recording the voice on speaker of laptop the frequency width is larger and it changes according to the pitch of the voice.

2)Graph using Earphone attached to laptop:

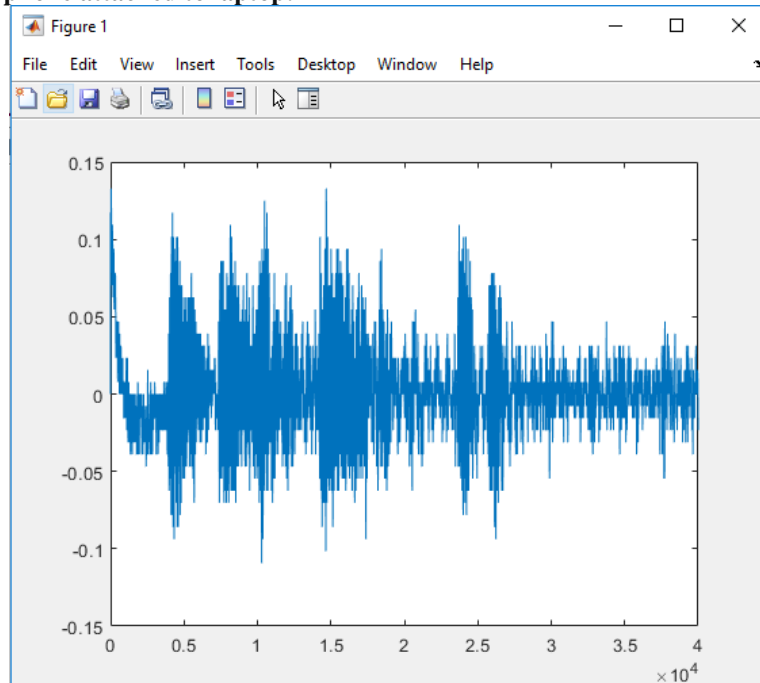


Figure no5: Graph of voice recorded using headphone attached to laptop

Figure 4 shows that on recording the voice using headphone attached to laptop the frequency width is thin as compared to voice recorded on speaker and it changes according to the pitch of the voice. As shown in Figure 3 and Figure 4 it is clear that due to change in frequency the voice width changes and it can be different every time according to the pitch of the voice. It can be seen in the graph that in different platforms the frequency changes, due to which identifying the original voice will be difficult.

For Example: A parrot owner was in for a great surprise, when he realized his African parrot was already in Christmas spirits before him. Rocco, an African grey parrot, used his owner's Amazon Alexa device

to order a range of products, BBC reported. According to the report, Rocco ordered his favorite snacks including watermelons, raisins, broccoli and ice cream, along with a light bulb and a kite. African grey parrots are known for their ability to mimic speech. Rocco's owner Marion Wischnewski told the Daily Mail that she was shocked when she realized what Rocco has done. "I have to check the shopping list when I come in from work and cancel all the items he's ordered," she said. Earlier to this, according to BBC, Rocco was re-homed from the National Animal Welfare Trust sanctuary in Berkshire because of his foul language.

It was then, sanctuary worker Marion Wischnewski volunteered to foster him and took him to her home in Oxfordshire. Despite being mischievous, Wischnewski thinks that her parrot is the sweetest. "I've come home before and he has romantic music playing," Ms Wischnewski said. "He loves to dance and has the sweetest personality." [7]

From the above example it is clear that if the system would be properly tested and would have proper algorithm would have recognized to compare between the original and dummy voice then the parrot would have not ordered stuff's and system that there is a change in the frequency so it could be someone else voice rather than the original voice.

IV. Future enhancement

In this research paper voice recording also records the background voices which need to be reduced so that voice will come properly without any disturbance. Voice reduction algorithm which is already implemented in MATLAB which can be modified so that the background voice can be cleared with more accurate data and clearing of unwanted voices from background. Using this algorithm, the voice can be identified in more accurate manner and result will be more appropriate. Objective for future development was to make a voice engine which can also work for recognizing the other local language like Hindi or Marathi. [6]

V. Conclusion

In last few years the accuracy and complexity of voice recognition technology have grown rapidly. From the above study of the different processes it can conclude that voice assistants are always not produce 100% accurate results. It needs more structured framework to solve the problem and to improve so that it can give accurate result. And also security needs to be improved before the Voice Assistant Technology is used for something crucial and for confidential purpose.

References

- [1]. Matthew B. Hoy (2018) Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants, *Medical Reference Services Quarterly*, 37:1, 81-88, DOI:10.1080/02763869.2018.1404391
- [2]. Aarthi Easwara Moorthy and Kim-Phuong L. Vu and S. Yamamoto (Ed.): HIMI 2014, Part II, LNCS 8522, pp. 324–334, 2014. © Springer International Publishing Switzerland 2014
- [3]. Knote, R.; Janson, A.; Eigenbrod, L. & Söllner, M. (2018): The What and How of Smart Personal Assistants: Principles and Application Domains for IS Reserach. In: Multikonferenz Wirtschaftsinformatik 2018, March 06-09, 2018, Lüneburg, Germany.
- [4]. Imrie P, and Bednar P. (2013). 'Virtual Personal Assistant', in: Martinez M. and Pennarola Cecilia F. (editors). 'ItAIS 2013. Proceedings of 10th Conference of the Italian Chapter of AIS, 'Empowering society through digital innovations', Università Commerciale Luigi Bocconi in Milan, Italy, December 14th, 2013. ISBN: 978-88-6685-007-6.
- [5]. Dr. Kshama V. Kulhalli, Dr. Kotrappa Sirbi, Mr. Abhijit J. Patankar : Personal Assistant with Voice Recognition Intelligence, *International Journal of Engineering Research and Technology*. ISSN 0974-3154 Volume 10, Number 1 (2017).
- [6]. Ms. Ayushi Y. Vadwala, Ms. Krina A. Suthar, Ms. Yesha A. Karmakar, Prof. Nirali Pandya: Intelligent Android Voice Assistant - A Future Requisite, *International Journal of Engineering Development and Research*. © 2017 IJEDR | Volume 5, Issue 3 | ISSN: 2321-9939
- [7]. Parrot mimics owner's voice to trick Alexa, orders itself goodies from Amazon (Tuesday, 18 Dec)
<https://m.dailyhunt.in/news/india/english/dna-epaper-dna/parrot+mimics+owner+s+voice+to+trick+alexa+orders+itself+goodies+from+amazon-newsid-104209112?s=a&ss=wsp>