

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/336808968>

Automatic Robot Processing Using Speech Recognition System

Chapter · January 2020

DOI: 10.1007/978-981-32-9949-8_14

CITATIONS

3

READS

933

2 authors, including:



Suseendran G.

VELS INSTITUTE OF SCIENCE, TECHNOLOGY & ADVANCED STUDIES (VISTAS), CHENNAI.

149 PUBLICATIONS 870 CITATIONS

SEE PROFILE

Automatic Robot Processing Using Speech Recognition System



S. Elavarasi and G. Suseendran

Abstract Nowadays, speech recognition is becoming a more useful technology in computer applications. Many interactive speech-aware applications exist in the field. In order to use this kind of easy way of communication technique into the computer field, speech recognition technique has to be evolved. The computer has to be programmed to accept the voice input and then process it to provide the required output, using various speech recognition software. Speech recognition is the process of converting speech signal to a sequence of words using appropriate algorithm. This provides an alternative and efficient way for the people who are not well educated or not having sufficient computer knowledge to access the systems and where typing becomes difficult. This speech recognition technique also reduces the manpower to accept and process the commands. In our research work, we have to implement this speech recognition technique in customer care center, where many queries have to be processed every day. Some of the queries are repeated often and the responses also seem to be the same. In such cases, we have to propose a methodology to automate the query-processing activities using this speech recognition technique. The ways of how to automate the system and how to process the queries automatically are explained in our methodology with suitable algorithm.

Keywords Feature extraction · Speech modeling · Speech automation · Robot processing · Automate query processing · Interactive speech-aware applications

S. Elavarasi (✉) · G. Suseendran
Department of Information and Technology, School of Computing Sciences, Vels Institute of Science Technology & Advanced Studies (VISTAS), Chennai, India
e-mail: elavarasi_msc@yahoo.co.in

G. Suseendran
e-mail: suseendar_1234@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2020
N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_14

1 Introduction

In this decade, people want to make all the actions to be carried out in a much easier and faster way. In order to accomplish this, most of the things have to be modernized and computerized to perform the action efficiently. As the new trends emerge in each and every field, things have to be changed to accommodate the trend. In such a way, many new technologies have emerged in the computer field also. One such technology is *speech recognition* or *voice recognition*.

Speech recognition is a technology that has emerged in order to make easier the process of providing the requirements to the system by the user through speech rather than text. That is, speech recognition accepts the user's spoken words as input and then translates the words to appropriate text or commands to accomplish the task. It is also called as a software to accept and understand dictation to undertake the commands. While accepting the voice input, the system also gathers the information such as gender, expression or emotion and in some cases, the identity of the speaker. Thus, speech recognition converts the speech to text, and so it is also termed as *automatic speech recognition* or *ASR* or *speech to text*. This is also a potential for human-computer interaction.

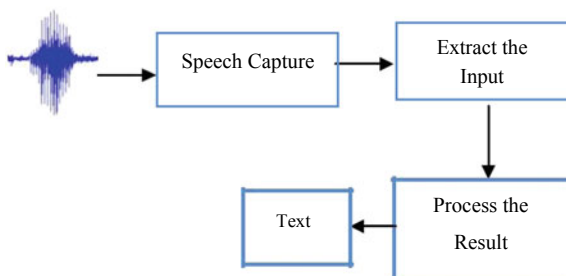
Speech recognition is mostly used in many areas, since it is more useful for those who are not well educated or not having sufficient computer knowledge. The essential requirement for the user is to have sufficient language skills and proper pronunciation efficiency such that the user is able to process the commands. The major research challenges in this area are noise, speaker variation, language variation and vocabulary (Fig. 1).

While designing the speech recognition system, the researcher requires providing special attention toward the challenges:

- Speech module and illustration
- Speech pre-processing
- Feature extraction
- Database maintenance
- Performance evaluation.

In recent years, the enhancement and effective progress of speech recognition technology have been implemented in many areas to simplify the process of user requirements and to improve the performance. Some of the applications are:

Fig. 1 Speech to text conversion



- Aerospace
- Automatic subtitling
- Hands-free computing
- Robotics
- Video games and so on.

In our research work, we have to implement the speech recognition technology in *customer care service* area to improve the customer satisfaction and experience some pre-defined operations with reduced manual efficiency. The major premise is to promote better customer service and to improve customer relationships.

Customer care service is the act of taking care of customer requirements by offering and delivering high-quality service in a short period of time. In order to provide such kind of response, more manual power has to be empowered. To avoid this dilemma, we have to provide suitable methodology to accomplish the process of responding the customer requirement through speech recognition technology automatically. This paper describes the development of an efficient speech recognition system over customer care service using various technologies.

2 Related Works

Gupta et al. discussed that with development in the requirements for installed figuring and the interest for rising implanted stages, it is necessary that the discourse acknowledgment frameworks (SRS) are accessible on them as well. PDAs and other handheld gadgets are ending up increasingly ground-breaking and reasonable also. It has turned out to be conceivable to run interactive media on these gadgets. In that paper, different methods about discourse acknowledgment framework were discussed. Likewise, it displayed the rundown of procedure with their properties of feature extraction and feature coordinating. Through this audit paper, it is discovered that MFCC is generally used to include extraction and VQ is better over DTW [1].

Singh et al. expressed that using fake neural systems (ANNs), numerical models of the low-level circuits in the human mind, to enhance discourse acknowledgment execution, through a model known as the ANN-hidden Markov model (ANNHMM) have indicated guarantee for expansive vocabulary discourse acknowledgment frameworks. Accomplishing higher recognition precision, low word mistake rate, creating discourse corpus relying on the idea of dialect and tending to the issues of wellsprings of fluctuation through methodologies like missing data techniques and convolutive non-negative matrix factorization are the significant contemplations for building up an effective ASR. In this paper, an exertion has been gained to feature the ground made so far for ASRs of various dialects and the mechanical viewpoint of programmed discourse acknowledgment in nations like China, Russian, Portuguese, Spain, Saudi Arab, Vietnam, Japan, UK, Sri Lanka, Philippines, Algeria and India [2].

Arora et al. endeavored to portray a writing audit of automatic speech recognition. It talked about past years' propels gained in order to give ground that has been

refined here of research. One of the vital difficulties for scientists is ASR precision. The speech acknowledgment system centers around troubles with ASR, essential building squares of discourse preparing, highlight extraction, discourse acknowledgment and execution assessment. The primary target of the survey paper is to expose the advancement made in ASRs of various dialects and the mechanical perspective of ASR in various nations and to thoroughly analyze the procedures used in different phases of speech acknowledgment and recognize the examined subject in this testing field. They are not displaying comprehensive depictions of frameworks or numerical definitions, yet rather, they are introducing unmistakable and novel highlights of chosen frameworks and their relative benefits and demerits [3].

Rashmi conveyed a review of various calculations that can be used in discourse acknowledgment in light of the points of interest and disservices. Likewise, it helps in picking the better calculation in light of the examination done [4].

Gamit et al. depicted that speech acknowledgment is the procedure of consequently perceiving the talked expressions of individual in view of the data content in discourse flag. This paper presents a concise overview on automatic speech recognition and examines the different characterization procedures that have been refined in this wide region of discourse handling. The target of this survey paper is to outline a portion of the notable strategies that are broadly used in a few phases of discourse acknowledgment system [5].

Shaikh Naziya et al. examined about speech advances that are limitlessly used and have boundless employments. These advances empower machines to react accurately depending on human voices, and give helpful and profitable administrations. The paper gave a diagram of the discourse acknowledgment process, its fundamental model, its application and approaches, and furthermore, examined near investigation of various methodologies that are used for discourse acknowledgment framework. The paper additionally gives a review of various strategies of discourse acknowledgment framework and furthermore demonstrates the rundown portion of the notable techniques used in different phases of discourse acknowledgment system [6].

Navneet et al. introduced the programmed discourse acknowledgment framework and examined the significant subjects and advances made in the previous 60 long stretches of research, in order to give a mechanical point of view and energy about the principal advance that has been proficient in this vital region of discourse correspondence. Following quite a while of innovative work, the exactness of programmed discourse acknowledgment stays as one of the imperative research challenges. The outline of speech recognition framework requires watchful considerations to the accompanying: definition of different sorts of discourse classes, discourse acknowledgment process, ASR configuration issues and discourse acknowledgment strategies. The target of this audit paper is to outline and look at a portion of the notable techniques used in diverse periods of talk affirmation system and perceive an investigation on subject and applications which are at the front line of this stimulating and testing field [7].

Lawrence et al. looked into some of the significant features in the innovative work of programmed discourse acknowledgment amid the most recent couple of decades to give a mechanical point of view and a valuation for the principal advance that has been made in this imperative region of data and correspondence technology [8].

Prabhakar et al. gave a description of major inventive perspective and valuation for the vital progression of talk affirmation; gave graph technique made in each period of talk affirmation, besides condense; took a gander at changed talk affirmation systems; and perceived an investigation on subjects and applications which are at the front line of this stimulating and testing field [9].

Bhavneet Kaur gave a concise presentation of SRS portraying how the innovation functions, and after that, talks about the general engineering of SRS, strategies, benefits of using this framework, major mechanical point of view and valuation for the principal advancement of discourse acknowledgment. It gives a way to deal with the acknowledgment of discourse flag using recurrence ghashly data with Mel recurrence for the change of discourse that includes portrayal in a HMM-based acknowledgment approach, and furthermore, gives a review of strategies created in each phase of discourse acknowledgment alongside the momentum and explores future options on the same. This paper depicts the real difficulties for SRS framework which have been went over by clients' criticism and different examinations, which must be settled as quickly as time permits for better execution outcome [10].

Swati Atame et al. discussed that one of the technologies used in these fields is automatic singer identification, which is used to recognize from features of the audio signal, the singer of the song or who is the one who is singing, and the genuine singer [5]. This same area of singer identification and recognition can be used in the bioinformatics whereby the voice of the singer is used to gain access to a particular singer. The system would be very much useful to singers of the song who are actual singers and also to the common man who can store their speech and can later use this input signal to access the system. For the professional singers, there are many possibilities that the original singers' voice is get mimicked which may in this situation lead to pirated copies of the voice of the singer which are then sold into the market [11].

Hori et al. exhibited strategies for discourse-to-content and speech-to-discourse programmed synopsis in light of discourse unit extraction and connection. For the previous case, a two organized outline strategies comprising essential sentence extraction and word-based sentence compaction are examined. Sentence and word units which augment the weighted total of etymological probability, measure of data, certainty measure and linguistic probability of linked units are separated from the discourse acknowledgment results and connected for star ducing outlines. For the last case, sentences, words and between-filler units are researched as units to be removed from unique discourse.

Renals et al. portrayed the advancement of a framework to translate and outline voice messages. The consequences of the exploration introduced in this paper are two overlays. Initially, a cross-breed connectionist way to deal with the voicemail interpretation undertaking demonstrates that focused execution can be accomplished using a setting autonomous framework with less parameters than those in view of blends of Gaussian probabilities. Second, a successful and vigorous blend of factual with earlier learning hotspots for term weighting is used to remove data from the decoder's yield keeping in mind the end goal to convey rundowns to the message beneficiaries by means of a GSM short message service (SMS) entryway [12].

3 Proposed Methodology

3.1 Proposed Method

Speech recognition has been the most explored point since mid 1960s and is a standout among the most famous and dynamic territory of research. In speech recognition, there exists lot of fields for research such as:

- Speech recognition
- Speaker recognition
- Speech conversion
- Feature extraction
- Noise reduction.

Speech recognition is used in many applications since it does not need any syntax, procedure or coding to access the commands. It just simply accepts the spoken words as input and then translates it into appropriate text to perform the task. Thus speech recognition technology has unlimited users with exciting range of tasks. The main theme of the speech recognition technology is to “*listen*”, “*identify*”, “*understand*” and “*respond*” to the spoken information. It has the potential to act like an interface between the humans and computer, that is, human–computer interface.

In our research work, we choose this speech recognition technique to be implemented in a customer care organization in order to provide better performance in a short period of time. In customer care organization, they have the task of providing response for the queries submitted by the user. The queries may be of different types such as account details, server problem, server restart, password reset, account lock, account balance, account transaction and last transaction.

In some situations, the customer may submit the queries continuously, for which the executive provides response for it. Some of the queries may be repeated often that the executive must be processed every time. In this case, it requires huge manual power with 24×7 services and high cost.

In order to overcome these circumstances, we have to propose a methodology to automate the process of providing responses for the queries with some pre-defined activities. This has to be carried out by programming the customer care server with some set of instructions with appropriate keywords in the database. The methodology can be executed as follows:

When the customer calls for sending query, the server accepts the query as voice input and records it. Upon receiving the input, the server converts the voice input into text and then starts to search in the database for the keyword found in the text. As a result of this database search, the appropriate instructions for the query have been found from the database, and the response to it is sent to the customer. Thus, the query has been automatically processed by the speech recognition technology like ROBOT processing, and hence our methodology can also be termed as automatic ROBOT processing (ARP) (Fig. 2).

In case the customer needs to change his/her password means, it comes under the category “*Password Reset*”. The customer first calls the customer care service and asks for password reset through voice command as: “*I Need Password Reset*”. This command has been recorded by the server and then processed to convert into text. Then the server starts to search from the database for the keyword “*Password Reset*”. In database, there exist several instructions for the password reset category that are to be programmed by the programmer to process the request in an appropriate manner.

In that database, if there exists as “*Please tell your Username*” and “*Please tell your Date of Birth*” for “password reset” category, then the server gets the instructions and converts it into speech and then transfers to the customer as to submit their username and date of birth through speech. Upon receiving the details from the customer, the server starts converting the details into text from speech and then validates the details. Only if the information provided by the customer matches with the database, the server allows the customer to reset their password. Otherwise, he/she has to be denied (Fig. 3).

Thus, the small pre-determined activities in the customer care service have been automated with the speech recognition technique to simplify the process and to reduce the manual power. It also reduces the cost of processing the pre-determined

Fig. 2 Processing the customer query

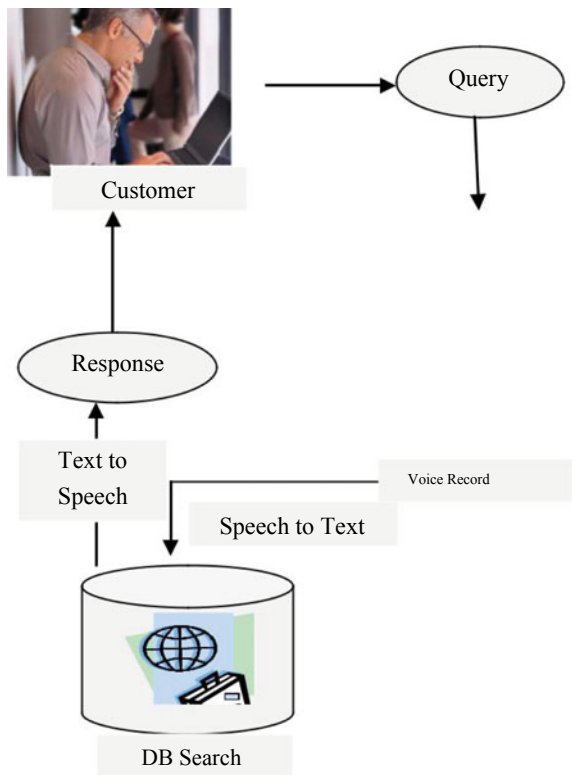
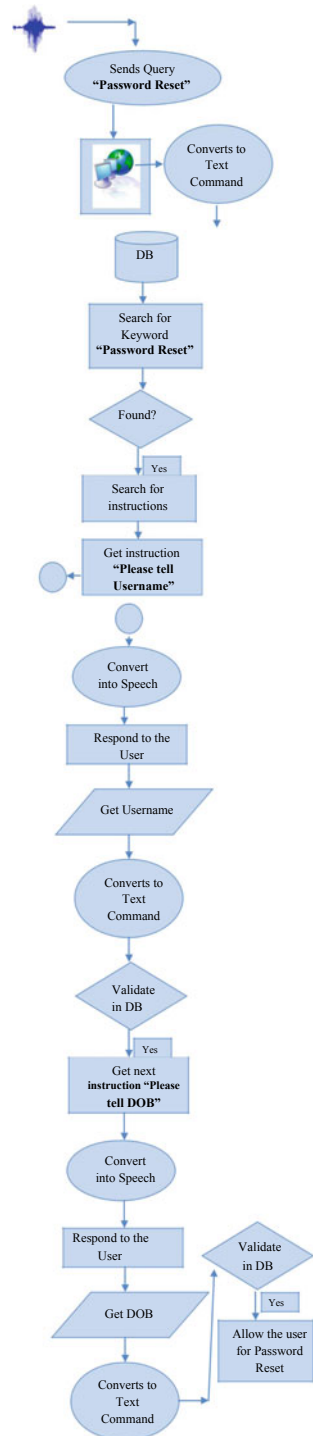


Fig. 3 Processing query for password reset



repetitive query processing. This kind of automatic processing the customer queries resembles ROBOT-type processing and thus we termed our methodology as ***automatic robot processing (ARP)***.

Along these lines, our proposed strategy acknowledges the voice information and follows the procedures to process the client inquiries effectively. The proposed strategy comprises a calculation to depict the stream of the procedure.

3.2 Algorithm-The Flow of Automatic Robot Processing (ARP)

B. Algorithm- The Flow of Automatic Robot Processing (ARP)

Begin

Load the server with Voice Recognition Software

Load Database with instructions for Query Processing

If customer enters then

do

Get the voice Command

Translate the Voice Command to Text Command

Search for relevant keywords

Get appropriate instructions from the database

Translate the instructions into speech commands

Responds the customer with the speech command

until the query processing ends

End If

End

4 Experimental Setup and Performance Metrics

We have to evaluate the performance of our proposed methodology in order to prove that our ARP query processing performs better than the existing techniques. This has to be carried out by taking some set of queries, and those queries have to be processed both manually and through our proposed ARP algorithm. The performance ratio is analyzed and the result has to be tabulated in order to prove that our ARP algorithm performs much better than the existing technique. While the

Fig. 4 Comparison chart

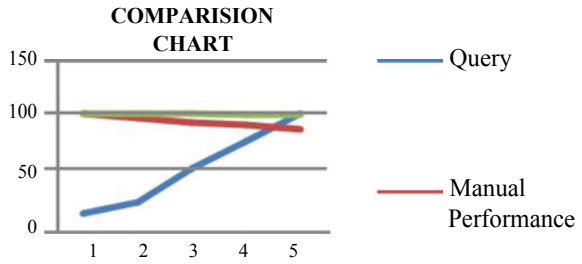


Table 1 Comparison results

| Query | Manual performance | ARP performance |
|-------|--------------------|-----------------|
| 10 | 100 | 100 |
| 20 | 96 | 100 |
| 50 | 92 | 99.5 |
| 75 | 90 | 99 |
| 100 | 86 | 98 |

queries have to be processed manually, as the number of queries becomes increased, the ratio of performance seems low. However, when the queries have to be processed through our proposed algorithm, the performance becomes invariable. The results are tabulated and the comparison chart is shown below. With these results, it becomes clear that our ARP algorithm performs better and the queries are processed successfully (Fig. 4; Table 1).

5 Conclusions

Speech recognition is one of the most assimilating areas of artificial intelligence. This speech recognition technique helps the people who were uneducated or not having sufficient English knowledge or computer knowledge. After undertaking deep research, it has been proved that speech recognition has made a role for handling the speech-oriented activities and has been seen in many walks of life. In this computer world, various techniques are discussed about speech recognition system. This speech recognition is one of the most challenging problems to deal with. We have attempted in this paper to provide a solution to automate some pre-defined activities in the customer care organization. We also provide suitable algorithm to process this robot processing. In this case, it seems easy to access the queries in customer care center without any manual resources.

In future, our approach must be improved with much propelled route in order to handle more operations automatically with reduced manpower.

References

1. Gupta, S., Pathak, A., Saraf, A.: A study on speech recognition system: a literature review. *Int. J. Sci. Eng. Technol. Res. IJSETR* **3**(8), 2193–2194 (2014)
2. Ghai, W., Singh, N.: Literature review on automatic speech recognition. *Int. J. Comput. Appl.* **41**(8), 44–46 (2012)
3. Arora, S.J., Singh, R.P.: Automatic speech recognition: a review. *Int. J. Comput. Appl.* **60**(9), 37–41 (2012)
4. Rashmi, C.R.: Review of algorithms and applications in speech recognition system. *Int. J. Comput. Sci. Inf. Technol.* **5**(5), 5258–5261 (2014)
5. Gamit, M.R., Dhameliya, K., Bhatt, N.S.: Classification techniques for speech recognition: a review. *Int. J. Emerg. Technol. Adv. Eng.* **5**(2), 59–61 (2015)
6. Shaikh Naziya, S., Deshmukh, R.R.: Speech recognition system—a review. *IOSR J. Comput. Eng.* **18**(4), 3–8 (2016)
7. Kaur, I., Kaur, N., Ummat, A., Kaur, J., Kaur, N.: Automatic speech recognition: a review. *Int. J. Comput. Sci. Technol.* **7**(4), 44–46 (2016)
8. Juang, B.H., Rabiner, L.R.: Automatic speech recognition—a brief history of the technology development. *Int. J. Biomed. Eng. Technol.* **3**(2), 4–17 (2005)
9. Prabhakar, O.P., Sahu, N.K.: A survey on: voice command recognition technique. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **3**(5), 578–584 (2013)
10. Kaur, B.: Major challenges of voice command recognition technique. *Int. J. Sci. Eng. Res.* **5**(8), 216–224 (2014)
11. Atame, S., Shanthi Therese, S., Gedam, M.: A survey on: continuous voice recognition techniques. *Int. J. Emerg. Trends Technol. Comput. Sci.* **4**(3), 38–40 (2015)
12. Koumpis, K., Renals, S.: Transcription and summarization of voicemail speech. In: *Proceedings of ICSLP* (2000)