ISSN:2348-2079

Volume-7 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Review of speech recognition in artificial intelligence

Mr.K.Gunasekar[1], M.Deepika[2], M.Usharani[2]

[1]Associate Professor, Department of computer Science and Engineering, Nandha Engineering College

[2]PG Scholars, Department of Computer Science and Engineering, Nandha Engineering College

ABSTRACT

This paper aims to present a retrospective study about speech recognition systems and artificial intelligence. Speech recognition is one of the widely used technologies, as it offers great opportunities to interact and communicate with automated machines. Precisely, it can be affirmed that speech recognition facilitate users and also helps them to perform their daily routine tasks, in a more convenient and effective manner. This research looks to present the illustration of recent technological advancements, which are associated with artificial intelligence. Recent researches has reveals the fact that speech recognition is found to be the utmost issue, which affects the decoding of speech. In order to get over these issues, different statistical models were developed by the researchers. Some of the most spectacular statistical models include acoustic model (AM), language model (LM), lexicon model, and hidden Markov models (HMM). This research helps in understanding all of these statistical models of speech recognition. Researchers also formulated different decoding methods, which are being utilized for realistic decoding tasks and constrained artificial languages. These decoding methods contain pattern recognition, acoustic phonetic, and artificial intelligence. It has been recognized that artificial intelligence is the most efficient and reliable method, which is being used in speech recognition.

Keywords: Speech recognition, Acoustic phonetic, Artificial intelligence, Hidden Markov Models (HMM), Atatistical models of speech recognition.

INTRODUCTION

The aim of this paper is to present the illustration of different advancement in artificial intelligence, in the perspectives of speech recognition. It has been established from the analysis of the research, which was conducted by [5] that speech recognition is one of the most advanced concepts in electrical engineering and computer science. Fundamentally, his approach deals with the conversion of spoken words into text. Speech recognition also referred as ASR (automatic speech recognition), STT (speech to text) or just computer speech recognition. Contrariwise, it had been claimed by [8] that speech recognition can also be interpreted as a field of computer science in artificial intelligence.

It deals with the designing and development of computer systems, in order to recognize the spoken

Words. In the regard, [14] has declared that speech recognition or computer speech recognition or ASR is nothing more than the approaching of converting a speech signal into the sequence of words, by the help of different algorithms and techniques. It has been certified in the research, which was carried out by [6] that these approaches includes artificial intelligence approach, pattern recognition approach, as well as acoustical phonetic approach. In accordance with the views

Author for correspondence:
Department of computer Science and Engineering, Nandha Engineering College

and perceptual experience of [7], artificial intelligence is the most developing and impressive techniques, which supports flawless,perfect and accurate speech recognition. Because artificial intelligence incorporates certain algorithmic approaches, which promote coherent conversion and transformation of speech into readable patterns, and vice versa.

This research will helps in understanding these concepts, which are associated with speech recognition. Artificial intelligence is found to be a most effective and integrated approach, which has strong and improved speech recognition practices [15]. The proceeding manuscript will admirably help in illustrating the core concept of artificial intelligence, as well as a technological advance, which has been occurred in artificial intelligence. In addition to this, this paper will also help in understanding and identifying the statistical models for speech recognition.

SPEECH RECOGNITION

Systems

Speech recognition can also be understood as an approach, which deals with the translation of the spoken words into the text. It has been constituted by [4] that speech recognition can also be referred to as ASR, as a technique offers to recognize the speech automatically. In accordance with the views and perceptual experience of [13], various speech recognition systems utilize speaker independent speech recognition. On the other hand, other speech recognition system utilizes training method, in which an indusial speaker reads section of text into the speech recognition system. It has been established by [9] that the speech recognition system recognizes the specific voice of a person and use it to modify the recognition of person's speech; hence, resulting into more coherent and integrated transcription. It is significant to notice that the systems, which make use training, are referred to as "speaker-dependent systems". However, such a systems, which don't use training, are known as "speaker-independent systems". It has been claimed by [12] that speech recognition can found a application in different areas and sectors. The most prominent applications of speech recognition includes aircrafts (direct voice input), speech-to-text processing (emails or work processors), formulation of structured documents (radiology reports), simple data entry (credit card number entry), smart search (podcast), domestic appliance control, call routing and voice dialing.

According to [11], incorporated and smooth performances of speech recognition mainly depends on utilization of appropriate statistical models. It is due to the concept that these statistical models convert speech into readable form and vice versa [6]. Thereby, the adoption and usages of the inappropriate statistical model may affects the integrity of speech recognition. Proceeding sections incorporate the analysis of different statistical models, which are being used in the speech recognition.

STATISTICAL MODELS OF SPEECH RECOGNITION

Acoustic Model (AM)

One of the most outstanding and widely adopted models of speech recognition is acoustic model (AM). It has been established that the acoustic models of speech recognition capture the characteristics of basic recognition units. According to the [14], the recognition unit can be at phoneme level, syllable level, and at the word level. Several insufficiency and constraints come into consideration with the selection of each of these units. Reference [7] has claimed that for LVCSR (large vocabulary continuous speech recognition) system, speech sound is the most favorable unit. Hidden Markov models and neural networks (NN) are the widely used approaches, which are being utilized by the acoustic modeling of speech recognition systems.

Language Model (LM):

Language model is some other most significant statistical model of speech recognition. One of the main objective of language model is to convey or transmit the behavior of the language. It is due to the fact that it look to forecast the existence of the specific word sequences within the target speech. According to [6], from the view of recognition

engine, this statistical model of speech recognition assists in minimizing the search space for a trustworthy and credible combination of words. It is significant to notice that the language model was developed by the help of CMU statistical LM toolkit.

Lexicon Model

According to [10] that lexicon models provide the pronunciation of the words within the target speech, which has been recognized. In accordance with the perceptual experience of [5], the lexicon model play an inevitable and indispensable role in automatic speech recognitions. Because of the fact that the operations of lexical model are based on two parameters, i.e., whole-word usage, and decomposition of entire speech into small chunk. This process eventually results into appropriate recognition of the speech. For example, if speech recognition models are in domestic language, the lexicon model has to be formulated in the domestic languages, in order to produce valuable and useful result. Regards, artificial neutral network phoneme can also be considered as one of the greatest approached, because it assists in developing the native lexicon from the foreign lexicon. Hence it results in mapping the phone of English to the phones of native languages [12]. It is important to notice that entire processes is conducted, while considering the contextual information.

Hidden Markov Models (HMM):

According to [14] Markov model is the most popular statistical tool, which is being used for the modeling of data. It has been found that the hidden Markov model has played a commendable role in the reduction of issues in speech classification, which is one of the midpoint issue, within the speech recognition approaches. In accordance with [11], hidden Markov model incorporate various issues, which is used to affects the accuracy of speech recognition. To resolve these issues, subspace projection algorithm and weighted hidden Markov models are proposed [8]. Subsequently, this model has become the basic of modern HMM-based continuous speech recognition technology.

DISTINCTIVE DECODING METHODS OF SPEECH RECOGNITION

It has been certified in the research, which had been carried out by [6] that various techniques of decoding can be used, in order to recognize the speech. Some of the most popular and widely used methods include acoustic phonetic method, pattern recognition method and artificial intelligence approach [10]. All of these methods are illustrated in the proceeding section.

Pattern Recognition

Pattern recognition was found to be most common and widely used techniques of speech recognition. This method mostly incorporates two important steps, includes pattern comparison and pattern training. It has been constituted from the studies of [10] that the leading characteristics of this method are to utilize a well-structured and integrated mathematical framework [5]. This statistical framework assists in formulating consistent representations of speech patterns; hence result into the acquisition of more accurate results. Pattern recognition is then further divided into two more approaches, i.e., stochastic approach and template approach.

Acoustic Phonetic

In Accordance with [14], the most primitive approache of speech recognition are mainly based on the process of locating sounds and speeches. One of the main objectives of such activities was to provide adequate labels to the sample sounds, in order to recognise the pattern of the sounds. It is fundamental to notice that such methods are found as the foundation of the acoustic phonetic approaches. As per the opinion of acoustic phonetic approach, there exists phonemes (phonetic units) and finite units within spoken language. These unit of acoustic phonetic approaches are extensively categorized by the collections of acoustic properties that are normally evident in the speech signal.

Artificial Intelligence

In accordance to [11], the approach of artificial intelligence is the almost famous method of speech

recognition, which is being used for decoding. Artificial intelligence was understood as the combinations of pattern recognition approach and acoustic phonetic approach. It is due to the fact that it incorporates the concept and idea of pattern recognition method and acoustic phonetic approaches. It has been accepted that artificial intelligence is also referred to as knowledge based approach and it utilise the information, that is related to spectrogram, phonetic and linguistic. In accordance to [1], the approaches of artificial intelligence play an indispensable role in various activities of speech recognition, consider designing of recognition algorithm, manifestation of speech units, and representation of suitable and appropriate inputs. It is important to bring the notice that, among all these methods of speech recognition, artificial intelligence is the most believable and efficient methods.

In general, artificial intelligence was understood as a rising and continually developing fields of computer science. It has been analyzed that artificial intelligence mostly emphasizes on the development of such machine, which is capable enough to get engage the behaviors of human beings.

It has been found in the research, which was carried out by [14] these artificial machines collect information's from their various environments and respond as an intelligent manner, calculating suitable and adequate steps, speculate answers, and presents desirable results. It has been found from the studies of [2],[7], that artificial intelligence is widely used in various areas, includes pedestrian signals and traffic signals, robotic household equipment, credit card transactions ,maintenance systems and home security, healthcare robotics, video games , and cell phones (smart phones). In addition to all of these applications, artificial intelligence is extensively used in speech recognition.

Application of artificial intelligence in speech recognition

It has been discovered from the evaluation of studies, which given by [3] that artificial intelligence is presently being used in different fields of life, includes transportation, aviation, remote sensing ,law, robot control, stock trading, scientific discovery ,medical diagnosis, and even technologies that have been developed by the researchers, which has made it possible to accomplish reasonable accuracy of words. Precisely, rising approaches and technological paradigms plays a commendable role in steadily enhancing the honesty of speech recognition. On the contrast [7] has asserted the fact that these technologies are not capable to compete with the quality of human listeners. Hence, it is one of the most difficult tasks for the researchers to design and develop perfect and highly efficient speech recognition technique. In such conditions, the approaching of artificial intelligence can be considered as one of the superior opportunities, in terms of recognizes the patterns of speech, perfectly. Because of the fact that artificial intelligence incredibly transform the speeches into well-structured algorithms, by suitably following all stages of [15].Most important stage, that is involved in speech recognition including representation of speech units, formulating and developing of recognition algorithms, as well as demonstration of the correct inputs (speech).

Technological advances in the field of artificial intelligence

Artificial intelligence gained enormous boost in the past two decades, as it has been used in extensive areas and fields of life. It has been discovered from the profound analysis of the research of [5] that the technique of artificial intelligence has been advocated by various developers and researchers. Some of the most popular and well-developed technique of artificial intelligence includes fuzzy logic, neural networks,data mining and knowledge based systems. One of the main objectives behind this process was to enhance and improve the software development process, in order to fight with today's fast paced and volatile environment. In Accordance to [13], some of the main evident advancement in artificial intelligence including the development of object-oriented programming, artificial neural network systems, graphical user interfaces, obstacle avoidance , fuzzy logic, speech recognition ,object recognition, rule based expert systems for aviation sector, text recognition, robot

navigation, intelligent transportation, as well as natural language processing.

CONCLUSION

From above discussion, it could be concluded that the applications of the speech recognition are becoming extensively essential important and useful nowadays. After performing this research, it has been concluded that speech recognition is the process of transforming the input signals (speech) into the well-structured sequences of words. It is important to notice that these sequences are developed in the form of algorithms. Basically, these algorithms convert the speech into the words and vice versa, hence results are more coherent, accrete, and correct recognition of speech. It has been found that speech recognition has become one of the greatest challenges and several techniques and approaches have been developed, to overcome this issue. Among all of those paradigms and models, artificial intelligence was considered as one of the most reliable and sufficient approaches. This research study has in corporate the in-depth evaluation of the core concept of artificial intelligence. In addition, this paper has also illustrated the speech recognition systems. In addition, different statistical models of speech recognition have also been concentrated in the paper, including lexicon model, acoustic model (AM), Hidden Markov models (HMM) and language model (LM).

These models play a major role in developing the algorithms and patterns of speech, which has to be recognized. In addition to this, it also analyzes the different decoding methods used for speech recognition. Some of the most common and extensively used methods including acoustic phonetic, artificial intelligence and pattern recognition. Among all of these approaches or methods, artificial intelligence can be considered as a most integrated and effective approaches, as this technique provides highly accurate and reliable results. This paper also demonstrated the applications of artificial intelligence in speech recognition, while assessing the technological advances in the field of the artificial intelligence.

REFERENCES

[1]. Ammar, Hany H., Walid Abdelmoez, and Mohamed Salah Hamdi. "Software engineering using artificial intelligence techniques: Current state and open problems." Proceedings of the First Taibah University International Conference on Computing and Information Technology (ICCIT 2012), Al-Madinah Al-Munawwarah, Saudi Arabia. 2012, 52.

[2]. Anusuya, A.M. and Katti, K.S. “Speech Recognition by Machine: A Review”, (IJCSIS) International Journal of Computer Science and Information Security, 2009.

[3]. Beigi, Homayoon. "Hidden Markov Modeling (HMM)." Fundamentals of Speaker Recognition. Springer, 2011.

[4]. Besacier, Laurent, et al. "Automatic speech recognition for under-resourced languages: A survey." Speech Communication, 2014.

[5]. Chen, Chi-hau, ed. Pattern recognition and artificial intelligence. Elsevier, 2013.

[6]. Chen, Lijiang, et al. "Speech emotion recognition: Features and classification models." Digital signal processing 22.6 2012.

[7]. Choudhary, A. and Kshirsagar, R. “Process Speech Recognition System using Artificial Intelligence Technique, International Journal of Soft Computing and Engineering (IJSCE), 2012.

[8]. Dolby, Jonathan, and Diane Kelley-Port. "Explicit pronunciation training using automatic speech recognition technology." Calico Journal, 2013.

[9]. Deng, Li, and Xiao Li. "Machine learning paradigms for speech recognition: An overview." IEEE Transactions on Audio, Speech and Language Processing, 21,.5, 2013.

[10]. Hinton, Geoffrey, et al. "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups." Signal Processing Magazine, IEEE 29.6 2012.

[11]. Mikolov, Tomas. "Statistical language models based on neural networks." Presentation at Google, Mountain View, 2012.
[12]. Morgan, Nelson. "Deep and wide: Multiple layers in automatic speech recognition." Audio, Speech, and Language Processing, IEEE Transactions on20.1, 2012.
[13]. Rawat, Seema, Parv Gupta, and Praveen Kumar. "Digital life assistant using automated speech recognition." IEEE, 2014.
[14]. Saini, Preeti, and Parneet Kaur. "Automatic Speech Recognition: A Review." International journal of Engineering Trends &Technology, 2013.
[15]. Saon, George, and Jen-Tzung Chien. "Large-vocabulary continuous speech recognition systems: A look at some recent advances." Signal Processing Magazine, IEEE 29.6, 2012.