# SPEECH RECOGNITION MODULE OF ROBOTIC LANGUAGE AUTOMATION

## ABSTRACT

As a cross-disciplinary, speech recognition is based on the voice as the research object. Speech recognition allows the machine to turn the speech signal into text or commands through the process of identification and understanding, and also makes the function of natural voice communication. Speech recognition involves many fields of physiology, psychology, linguistics, computer science and signal processing, and is even related to the person's body language, and its ultimate goal is to achieve natural language communication between man and machine. The speech recognition technology is gradually becoming the key technology of the IT man-machine interface[1].The paper describes the development of speech recognition technology and its basic principles, methods, reviewed the classification of speech recognition systems and voice recognition technology, analyzed the problems faced by the speech recognition.

Keywords-basic principles;method;speech recognition; application

## INTRODUCTION

Speech recognition is the machine on the statement or command of human speech to identify and understand and react accordingly. It is based on the voice as the research object, it allows the machine to automatically identify and understand human spoken language through speech signal processing and pattern recognition. The speech recognition technology is the high-tech that allows the machine to turn the voice signal into the appropriate text or command through the process of identification and understanding. Speech recognition is a cross-disciplinary and involves a wide range. It has a very close relationship with acoustics, phonetics, linguistics, information theory, pattern recognition theory and neurobiology disciplines. With the rapid development of computer hardware and software and information technology, speech recognition technology is gradually becoming a key technology in the computer information processing technology. Products to develop speech recognition technology is also widely used in voice-activated telephone exchange query information networks, medical services, banking services, industrial control every aspect of society and people's lives. Many experts believe that speech recognition is one of the 2000-2010 IT field ten scientific and technological developments.

## THE DEVELOPMENT PROCESS AND CURRENT SITUATION OF THE SPEECH RECOGNITION TECHNOLOGY

Speech recognition research work began in the 50's, Bell Labs speech recognition system-Audrey system first identifies the ten English digits. But it really made substantial progress, and as an important issue in conducting research in the late 60's the early 1970s. Further speech recognition in the 1980s, the HMM model and artificial neural network (ANN) are successfully used in speech recognition. 1988ˈFULEE Kai and others use the VQ/IüIMM method to achieve speaker-independent continuous speech recognition system-SPHINX, including 997 vocabulary. This is the first of the world speech recognition system, it is a high-performance, non-specific, large vocabulary continuous speech recognition system. People finally breakthrough of the three major obstacles, including a large vocabulary, continuous speech and non-specific. And it identified the mainstream of statistical methods and models in speech recognition and language processing. Speech recognition system has already begun from the laboratory to practical; there have been more mature market products. Many developed countries such as the United States, Japan, South Korea, as well as IBM, Apple, Microsoft, AT&T and other well-known companies to invest heavily in research and development of practical speech recognition system [2].

## BASIC PRINCIPLES AND METHODS OF SPEECH RECOGNITION TECHNOLOGY

The speech recognition system is essentially a pattern recognition system, including feature extraction, pattern matching, the reference model library. Its basic structure is shown in Figure 1:

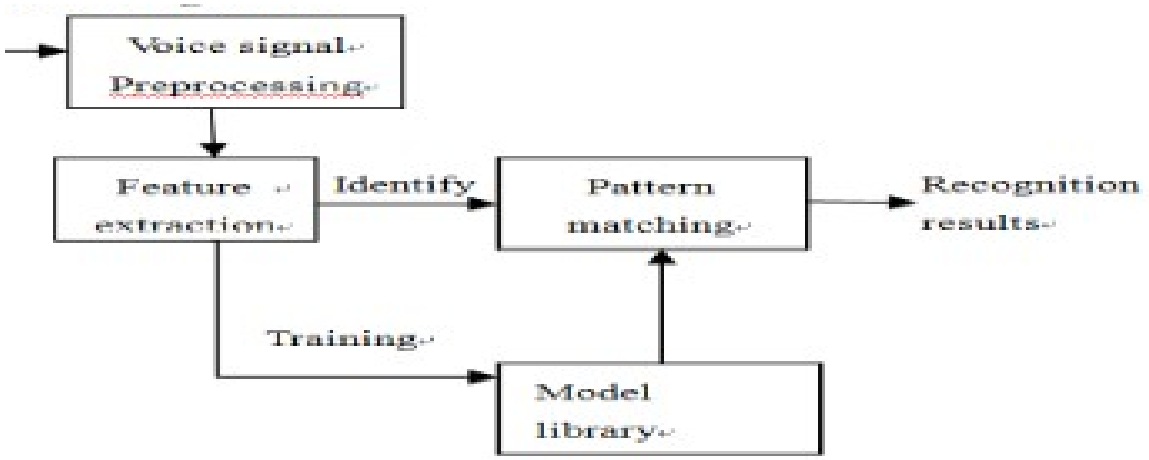


Figure 1 The basic principles of speech recognition system

The unknown voice through the microphone is transformed into an electrical signal on the input of the identification system, the first after the pretreatment. The system establishes a voice model according to the human voice characteristics, analyzes the input voice signal and extracts the required features on this basis, it establishes the required template of the speech recognition. Computer is used in the recognition process according to the model of the speech recognition to compare the voice template stored in the computer and the characteristics of the input voice signal. Search and matching strategies to identify the optimal range of the input voice matches the template. According to the definition of this template through the look-up table can be given the recognition results of the computer. Representative speech recognition methods include dynamic time warping (DTW), hidden Markov model (HMM), vector quantization (VQ), artificial neural network (ANN), support vector machine (SVM) and so on.The article focuses on artificial neural network (ANN).

**Artificial neural network**

ANN (based Artificial Neural Networks), analogous to the way biological nervous systems process information, using a large number of simple processing units connected in parallel to form a complex information processing system. This system has the training, highly parallel, rapid judgment, fault tolerance features applies voice signal processing. Speech recognition neural networks are usually divided into two categories, a class of neural networks or neural networks with the traditional HMM, the DP combination of hybrid network, the other is the establishment of the auditory neural network model based on human auditory physiology, psychology research. Neural network model that more commonly used and has the potentiating of speech recognition mainly include single-layer perception model, multi-layer perception model, Kohonen self-organizing feature map model , radial basis function neural network , predictive neural network etc. In addition, in order to make the neural network reflects the dynamic of the speech signal time-varying characteristics, delay neural network, recurrent neural network and so on. Artificial neural network technology in voice recognition applications mainly the following aspects: a)Reduce the modeling unit, generally in the phoneme modeling to improve the recognition rate of the entire system by improving the recognition rate of phonemes. b)Depth study of the acoustic model, the auditory model, the brain operation mechanism, the introduction of context information, in order to reduce

the impact of changes in voice more than the speech signal. c)Extracted from the speech signal in a variety of features, a hybrid network model (HMM + NN), and apply a variety of knowledge sources (phonemes, vocabulary, syntax and meaning of the word), for voice recognition to understand the research, to improve system properties[5]. Speech recognition using artificial neural network technology, including e-learning process and the speech recognition process, shown in Figure 2. The network learning process is to known speech signal as a learning sample, self-learning neural network, and ultimately a set of connection weights and bias. The speech recognition process is to test the voice signal as network input, the recognition results obtained through the network of associations. The key of tthese two processes is to strike a speech characteristic parameters and neural network learning.

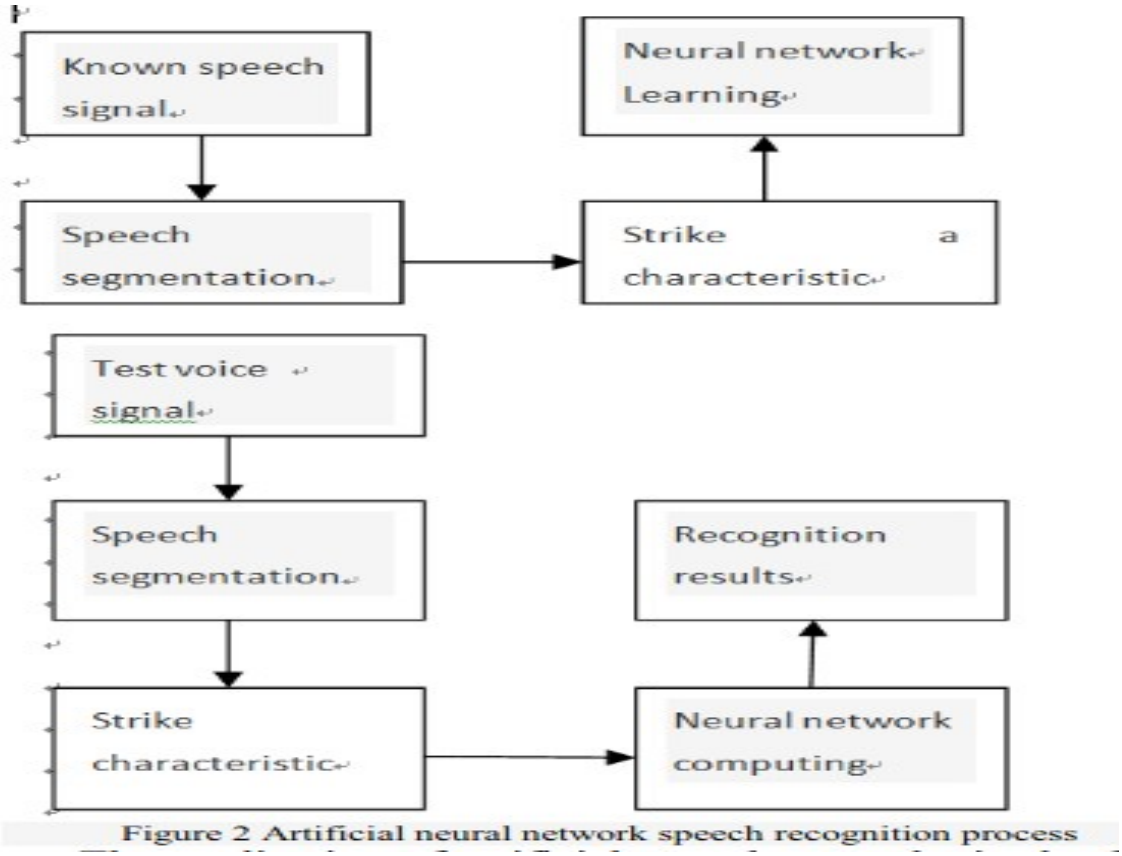

Figure 2 Artificial neural network speech recognition process

The application of artificial neural networks in the field of speech recognition has been greatly developed in recent years, artificial neural networks in speech signal processing can be divided into the following areas: firstly, improve the performance of artificial neural networks. Secondly, artificial neural network has been developed method combines a hybrid system. Thirdly, explore the use of newly emerging or widespread concern mathematical methods constitute the unique nature of the neural network, and applied to the field of speech signal processing [6]. The application of artificial neural networks in speech recognition has become a new hotspot. Artificial neural network technology has been successfully applied to solve pattern classification problems, and was shown to have enormous energy, we can predict that in the last decade, artificial neural network-based speech recognition system products will appear in the market, people will adjust their own way of speaking to accommodate a variety of recognition system.

# APPLICATION OF SPEECH RECOGNITION TECHNOLOGY

Modern speech recognition system has higher precision on identification, not to mention speech of a specific group. Today speech recognition technology is widely used, especially in electronic information industry, such as large-scale integrated circuit technology. There are special speech recognition systems for particular use such as industrial, telephone, mobile phone, communications, medical, children's toys, electronic products and so on. People even can use the speech recognition system of the phone network to inquire about related information.

**A.Application in Automobiles**

Driver in the car must take hold of the steering wheel all along, but if recognition system is installed in the vehicle, it will be convenient to make phone calls without hand, and can also be used to control doors, windows and other equipments of car by voice, improving the drive safety coefficient.

**B.Application in toys**

Now many intelligent toys which can do simple dialogues appear in the market. People can dictate the toys what to do through the voice. Such as a house dog with speech recognition function is very popular for children. Speech recognition system provides a good direction for toy development.

**C.Application in telephone communication**

Voice dialing of telephone uses the voice chip which is based on speech recognition technology. There are more and more technical phone have the voice recognition function. It brings convenience for telephone communication. People can make a phone call which only needs simple words .

**D.Application in information retrieval**

Speech recognition technology combine with retrieval and query a large number of database makes information retrieval more easily. For example, the information from library compared with user speech then transformed into a response of the instruction and get the result. It is also more convenient for company leader to express himself by speech recognition to get materials needed.

**E.Application in communication between human and computer**

Windows system interface is all too familiar. Now the operation of the computer is not new, but some hidden commands are still unclear. The speech recognition technology make the computer gradually become partner of the people to communicate, as long as several simple sentences or words can operate the Windows system. Nowadays IBM Company has products already with very mature technology. Speech recognition will gradually replace the keyboard and mouse, becomes a new way for human and computer communication. PDA (Personal Digital Assistant) is referred to as personal digital assistant. Because of its small volume and carrying easily, PDA is widely used. But it is not convenient to use the keyboard on PDA, the handwriting recognition method is mostly used. If a speech recognition system is applied in PDA interface, it will bring great facility for the communication between man and machine. Now parts of PDA adopt speech recognition technology, we believe voice control will be the mainstream in the development of PDA in the future.

### F.Application in Education

Speech recognition technology help user to practice language better and be more convenient. Formerly people feel difficult to correct pronunciation by contrast, only to learn by simple system. With the development of the speech recognition technology, now learning a language people can speak to a computer to find whether the pronunciation is correct by comparison. The computer will show the differences between standard pronunciation and pronunciation of users after the user talk with a computer, which enable the user to correct pronunciation in time, making it a better learning method. Now the voice recognition technology is also applied in the products of early childhood education, which helps children learn language and stimulates the learning interest of the children.

### G.Application in household appliances

Speech recognition technology has been widely used in modern electronic information industry, such as radio, television, communication, radar, computer, electronic components, electronic instrumentation and so forth. At the same time, information processing, pattern recognition, probability theory and information theory associate with speech recognition technology are widely used in various fields. For example, Microsoft Corporation develops its own speech recognition engine for Office and Vista. Microsoft speech recognition engine is completely free to use, so many speech recognition application software is developed based on it. General electric appliance products are implanted with a speech recognition system, as long as the user issues a command, the machine can understand, and then perform in accordance with the command. For example, when you are cooking at home, a few words may be just make the cooking tool in kitchen automatically start cooking food; when you feel sleepy, music will be played in the room to make you feel relaxed. In shortly, speech recognition system has become a part of our life gradually, and made us feel more relaxed and comfortable.

## THE FACING PROBLEMS

At present, speech recognition research progress has been slow, mainly in theory has been no breakthrough. Although a variety of new amendments continue to emerge, but also the lack of general applicability. Mainly in:Poor adaptability of the speech recognition system is mainly reflected in the dependence on the environment, If you collected speech training system in certain circumstances, the system can only be application in this environment, otherwise the system performance will be a sharp decline, anther problem is that this system does not respond correctly for the error input of users. Additionally, the progress of speech recognition in noisy environments is very difficult, because at this time people's pronounce varies greatly , like voice, slow speech rate, pitch and formant changes, which is the Lombard effect, must find a new signal analysis and processing approach. Understanding of the human auditory comprehension, the accumulation of knowledge and learning mechanism and system of the brain control mechanism is still unclear, and secondly, the existing achievements of this aspect is used in speech recognition also remains a difficult process.

## CONCLUSION

From the problems faced by the speech recognition, speech recognition systems in order to be widely used still have a lot of areas for improvement. However, it is foreseeable in the near future that, with the voice recognition technology continues to progress, the speech recognition system will be more in-depth, the application of speech recognition systems will be more extensive[8]. A

variety of speech recognition systems will appear in the market, people will adjust their speech patterns to adapt to a variety of recognition system Human beings in the short term is also impossible to create a people comparable to the speech recognition system, to build such a system is still a big challenge facing humanity, we can only forward step by step direction to improve the speech recognition system.

## REFERENCES


[1] Yu Tiecheng. The current development of speech recognition [J]. Communication World, 2005.

[2] Ren Tianping. Application of speech recognition technology [J]. Henan Science and Technology, 2005.

[3]L A Liporace.Maximum Likelihood for Multivariate Observation of MarkovSources. IEEE.Trans. IT, 1982, 28(5): 729-734

[4] Zhang Ping, Zhang Qiong. Based on HMM and BP neural network for speech recognition [J]. Cross-century, 2008.

[5] Yin Peng, Li Tao, Wang Haibing.Intelligent neural network system composed of the principle in speech recognition.Mini-Micro Systems,2000,21(8):836-839.

[6] Jiang Ming Hu, in the Yuan Baozong, Lin Biqin. Neural networks for speech recognition research and progress. Telecommunications Science,1997,13(7):1-6.

[7] Huang Shan. Voice recognition systems in the telecom prepaid business applications [J]. Information Science, 2010.

[8] Yangshang Guo, Yang Jinlong. The speech recognition technology overview [J]. Computer, 2006.