Research on English Speech Instruction Recognition Technology

Wu Hao

Jiangxi Police Institute, Nanchang, Jiangxi, China

Keywords: speech recognition, English, ATC instruction recognition

Abstract: It is precisely because of the continuous improvement of high-tech products, the development of speech recognition technology is getting better and better, and the most iconic building in the information field, speech recognition technology is the way of human-computer conversation. Today, with the advancement of science, speech recognition technology is at a critical moment in the march toward commercialization, and its position in the IT industry is becoming more prominent. Up to now, most colleges and universities in China have used a lot of resources to study the English speech instruction recognition technology. People are also beginning to pay attention to this technology. The technology already has a molded product. In life, because people attach great importance to this technology, many large companies including Microsoft and TBM are conducting research on this identification technology, but there is still a long way to go before the real concept. This important technology is linked to many categories, including occurrence and auditory mechanisms, probability theory and information theory, pattern recognition, and advanced mechanisms of people and machines.

1. Introduction

English speech instruction recognition has a very good development prospect and is a discipline with high theoretical value. The technology will use a dynamic time warping template and a hidden Markov model to complete the speech recognition of a single word. It also specifically discusses some of the identification techniques related to hardware implementation and some other complex issues. For the system construction, the theoretical model of English instruction recognition has a very far-reaching level. This article first analyzes the lower half of the speech recognition software and clarifies the differences in the choice of job mode. Then, according to the steps of the scheme, the process of speech recognition is discussed in detail, and the dynamic time warping model is applied to realize the isolated word recognition.

2. Definition and classification of English speech recognition

2.1 Definition of English Speech Recognition

English recognition technology is an emerging edge discipline. The way to improve it is how to use words to actively identify and verify the most fundamental and important meanings provided by phonetic symbols. At a general level, speech recognition also includes speaker recognition. The important content is to comprehend the information about personal characteristics in the speech signal, that is, the unique characteristics of the speech (such as: temperament characteristics, etc.), which refers to meaningful, content-based recognition, and is also part of the speech processing for the subject. . We can clearly see that the important and fundamental information is the fundamental of the phonetic phonological features, that is, the common features of speech, so it can also be called voice information. English speech recognition in physics (acoustics), signal processing, pattern matching, phonetic phonetics, psychology, computer science (the study of software and hardware algorithms to meaningfully present the way in F-recognition systems), psychology it was widely used in areas such as learning.

  DOI: 10.25236/icemeet.2019.220

2.2 Classification of English Speech Recognition

2.2.1 According to the voice recognizer to the user's practical situation

The regular pattern or pattern of speaker-dependent speech recognition is only suiTable for human use. In general, people build (called exercise: training) templates or models based on each word in the input word list or the sound of the phrase. When people use it, they need to create their own formal styles or slabs. Non-specific person speech recognition (independent of the speaker) standard template or speech recognition model is applicable to specific categories of speakers (eg, standard Mandarin) and through standard templates or models generated by multi-person training in categories. It can be used by speakers (internal) who can participate in training, or by speakers of the same category who do not participate in training (outsiders).

2.2.2 Press the type of speech recognizer

(1) Isolated word recognition

For Chinese speech, people recognize that the unit of discourse is the vocabulary, word or phrase that forms the recognition (vocabulary), and each of them trains into a standardized template or model or the words that recognize the speech. One of the words or phrases.

(2) Continuous speech recognition. Connected word recognition

With fewer words, each word is fully recognized. Identify vocabulary and formal templates or slate also words, words or phrases, but they can be continuous between them. Typically, the vocabulary includes ten digits from "0" to "9". When it is recognized, it can be said to be "3", "27", "659", and the like. Continuous speech recognition and understanding (session Speach recognition) is targeted at most vocabulary, and the speech to be recognized is some complete statement. Although each word cannot be completely accurately identified, its important meaning is that it can be understood, and continuous speech instruction discrimination is also called dialogue speech recognition. We can also understand that after speech recognition, the meaning or content of speech is derived from the judgment of language knowledge.

2.2.3 by the size of the phonetic vocabulary

Because recognition is basically based on the readable syllables of all the characters in Chinese culture, it is often referred to as full-spectrum speech recognition. Syllable word list: Leico full syllable speech recognition is the root of the endless phrase or Chinese kanji input.

3. The importance of English speech instruction recognition technology

3.1 The necessity of voice information processing and recognition

Is this the most meaningful fundamental function of all people, spreading people's various emotions and voice messages. People use language to convey information, including the tone and text of the language. In modern society, with the rapid growth of the information society, people and mechanical devices need to exchange a large amount of exchange information at any time, not to mention the exchange of information between people. When people use voice communication, the speaker's brain will have some ideas. After the change of people's language, the pronunciation organ emits a corresponding sound. Sound can be transmitted to people's ears through the air, and then the sound is transmitted to the brain through the auditory organs.

3.2 The importance of English speech instruction recognition technology

The most significant feature of the intelligent computer system is voice command recognition. This technology will become an important part of the advancement of electronic information. The application of this technology is a huge change in the human-machine interface of computers, and has far-reaching significance for the development and use of computers. In the voice control industry, English voice command recognition will inevitably play an important role in promoting office automation, editing and layout, voice control technology for industrial engineering and machine operation.

4. Research on English Speech Instruction Recognition Technology

4.1 The key to the basic language of air calls (English) in recognition

The ATC indication refers to the trouble caused by the English phonetic recognition command in the actual scene in the air traffic control associated with the safe driving of the aircraft. The design of the indicated specifications was first pointed out and provided in accordance with the fundamental characteristics of the ATC instructions. Then the use of this type of language description also gives specific steps in the grammar plan. The analysis is then applied to the special speech recognition system indicated by the ATC and then discerned at the beginning of the interpretation. In the final step, a study was conducted on acoustic model training.

4.2 Air call formal language (English) instruction recognition technology application

The ATC indication is the official language of the call between the controller and the aircraft. It is a language based on English but not ordinary English. Its pronunciation is unique. The fundamental feature is clear vocalization, word meaning and imperative structure. The ATC Directive deals with the formal ATC terminology specified in the standard. The application scenarios of the ATC instructions cover multiple aviation phases, such as rising, landing, air line flight and ground sliding. After studying the formal AC command, we found that the special communication language between air traffic controllers and aviation personnel followed certain rules. Therefore, the complete ATC instruction format is designed as: "flight call sign". + "Control Area Code" + "Command Action" + "Action A. The flight number indicates the flight number of the ATC order received, usually consisting of the airline call sign and flight number. The flight company call number is the IC phone number according to the national civil aviation authority. Such as Air China (CA), Eastern Airlines (CES), etc., generally use string style. The flight number is a random four-digit string, such as the control area code is the sender of the ATC command, it is usually the name of the control unit and The control type (two strings) is composed. The name of the control unit indicates the position of the control area, such as 2838, 4677, etc. The control unit, such as Beijing, Guangzhou, Chengdu, etc. The control type indicates the scope of the control indication, such as the tower, entering the field. , ground, etc. Command actions and actions are used to indicate specific commands. An indication may only include a special command operation and operational adverbial speech recognition. Due to the high accuracy of the ATC instructions in land and air calls, on land and In the air call, you must clearly distinguish between letters or words with similar English mid-range, such as B and D, G and J, especially a string of numbers or The pronunciation of the mother, such as the flight number, the route code and the name of the navigation station in the order. It must be ensured that it does not cause misunderstanding between the controller and the crew. In the "standard", according to ICAO regulations, A series of special pronunciations, including letters, numbers, special nouns and their pronunciation rules, such as A reading Apha, 9 reading niner, etc. In this experiment, the 39 basic phonemes provided by the speech recognition engine Sphinx4 are used to form these strange The vocalization, that is, the mapping between the phrase and the special vocalization can be arranged by the combination of phonemes (ie, the utterance order). The letters, numbers, words and combinations thereof that enable the speech recognition engine to be useful to identify specific utterances are realized. The discriminating process is the path code in the ATC command, the navigation field name, the runway number, etc. are composed of many letters (AZ) or letters/numbers, and multiple letters send or receive commands in the process. It is read bit by bit. (for example, the routing code AWT03 is pronounced Alpha Whiskey Tango zero and thre). In summary, in "regular", Prevent flick utterance, to the 26 letters of the utterance given the other rules, i.e., the pronunciation of letters AZ ALPHA word pronunciation substituted -ZUU example, the route codes BARC pronunciation is Bravo Alpha Romeo Charlie.

5. Conclusion

In general, this paper isolates speech recognition and research begins and studies the basic theory

and speech recognition systems, in particular summarizing and summarizing the understanding of speech recognition models, and decomposing speech recognition from a hierarchical perspective. The application structure has a certain practical significance for both speech recognition and future system construction. Dynamic time-regular style and hidden Markov plated main lines. It is also the two most widely used recognition algorithms for single word and continuous word recognition. The article recognizes the comparison of the two methods for implementation. The front-end signal is pre-solved and refined in speech recognition, and the most important example is the comparison of all conclusions and successes. These basic theories are to further study and learn to do well. Speech recognition technology is geared towards the actual direction. One is to use the digital signal processor as a general-purpose hardware platform, and the other is a general-purpose speech recognition chip design. This paper also discusses the practice of speech recognition algorithms on digital signal processing hardware. The main work is the simulation and demonstration phase. At the same time, the characteristics of digital signal processing hardware are also used. But this task has just begun, and the perfection of theory and practice still requires us to do a lot of homework. From the full text, the author did not give any conclusions on the measurement of speech recognition design results. The purpose of the study was not to require identification of recognition rates to improve the details of speech recognition, but rather to require speech recognition theory. Mastery and learning; Secondly, due to time and laboratory conditions, a large number of voice acquisition and pre-processing work cannot be completed, and the statistical recognition results lack scientific and truthfulness. Third, the existing Chinese speech recognition language database is incomplete. There is no uniform standard for the measurement results of the identification system. In fact, the training and test corpus of most Chinese speech recognition systems differ greatly in terms of data size, sampling format, test environment and task content. This is also the challenge and difficulty faced by Chinese voice recognition.

References

[1] Zhang Jian, Tan Jingxin. Voice command recognition technology and its application in radar simulator [J].Computer Engineering and Design 2010, 31(3):655659

[2] Voice Xu system design of Xu Xiaohao air traffic control simulator [D] Nanjing: Nanjing University of Aeronautics and Astronautics, 2002

[3] Nie Dangmin, Zhang Linchang. Development and application of speech recognition subsystem in tower simulator [D] Beijing: Beijing University of Aeronautics and Astronautics, 2003