International Journal of Database Theory and Application
Vol.7, No.2 (2014), pp.45-58
http://dx.doi.org/10.14257/ijdta.2014.7.2.05

Field Information Acquisition System Research Based on Offline Speech Recognition

Zhengyong Huyan, Likui Xu, Shuai Fang, Zhe Liu, Xiaodong Zhang and Lin Li*

China Agricultural University, People's Republic of China Ministry of Agriculture Key Laboratory of access to agriculture information, College of information and electrical engineering, No.17 Qinhuadong Rd, Haidian District, Beijing, China, 100193
lilincau@126.com

Abstract

In the mobile Internet era, the use of intelligent mobile terminal equipment will become more and more popular, especially in terms of using mobile phones for information collection in outdoor environments. Speech recognition technology is a very friendly method of interpersonal interaction, and using it for data acquisition will greatly improve collection efficiency. In this study, based on the theory of the HMM model, we build a wild information acquisition system which is based on offline continuous speech recognition through open source tools such as HTK and Sphinx. The system is able to efficiently solve some difficult problems, such as bad signals, inability to recognize rarely used terminology, and low speed of recognition.

Keywords: *offline speech recognition; Chinese word segmentation; filed information acquisition*

1. Introduction

With the rapid development of mobile Internet technology, many types of mobile terminal devices have appeared, the prices of which are constantly decreasing, the functions are becoming more and more powerful, and the performance is getting better and better. In particular, smart phones and touchpad computers have become very popular, and new application software for these devices has also appeared frequently, greatly changing people's life and work styles, thus signifying that we have entered the era of mobile Internet. As for field information acquisition, in the past people only used manual records, of which the efficiency is quite low. Now we may consider using our smart phone or touchpad, greatly improving the efficiency of information acquisition.

For smart phones, if using traditional Chinese language input signifies being written by hand, it actually requires two hands; if using the Chinese pinyin input method, a single word will require the input a large number of letters. Neither of these input methods is very efficient. With speech recognition, a very friendly input method, the user only needs to say what he or she wants to enter to the mobile terminal device, then it can automatically quickly translate the voice data into the text data, and realize data entry. In the speech recognition technology file, speech recognition technology, which is based on cloud sharing, is currently used widely. Some companies have been more successful in this area, such as IFLYTEK, Baidu and Tencent. However, these programs can only identify words which are used commonly in our day-to-day life. Moreover, they require continuously connecting to a network and the speed of network also affects its recognition speed. In the field of information acquisition, sometimes there are no network signals in

[*] Corresponding author: Lin Li, Email: *lilincau@126.com*

ISSN: 2005-4270 IJDTA

some places, thus we cannot guarantee that the device is continuously connected to a network, and for information acquisition in some industries we will encounter some uncommon words, and these programs cannot meet these requirements. However, offline speech recognition trains the speech model for a specific application, and the model saved in the mobile terminal does not need to be connected when the speech recognition is used, so that it can effectively meet the requirements of field information acquisition.

The purpose of field information acquisition is saving data within the mobile terminal device after collecting the data, then synchronizing the local data to the server when connected to a network. But the results of speech recognition form a string that cannot be directly deposited in the database, a task which would require use of Chinese word segmentation ion technology. Using this technology we may divide the string into phrases, then show them at the corresponding positions of the location system collection page, and finally save them in the local database.

2. Speech Recognition

The goal of speech recognition is to allow the machine to understand the words used by people, that is, in all cases, to accurately identify the content of speech, and execute orders of people according to their information for various purposes [1]. Speech recognition is an interdisciplinary field including related signal processing, mode recognition, probability theory, information theory, voice detection, artificial intelligence, and so on. Its goal is to convert the content of the speech to the input so that the computer can recognize it and achieve a more natural human-computer interaction. Speech recognition technology is based on the cloud networked processing and offline processing methods. The current mainstream speech recognition software uses the former method, the Internet-based cloud-processing method; specifically, the client inputs speech, the server performs speech recognition, and sends the results back to the client. The advantages of this technology include the following: 1) the powerful voice processing capabilities of the server; 2) saving client storage space of the linguistic model, the acoustic model and the dictionary; and 3) identifying a large amount of common vocabulary. However, it cannot identify uncommon industry-specific terminology, and is unable to guarantee processing speed or recognize uncommon words when not connected to a network. Accordingly, in this paper we mainly introduce offline-based speech recognition technology.

The speech recognition process, as shown in Figure 1, includes the model preparation phase and speech recognition phase. The main task in the model preparation phase includes completing the text data training, voice data training, establishment of language models and dictionaries according to text training data, and establishing acoustic models based on voice training data. Speech recognition mainly includes the input of the speech signal (i.e. words spoken by a person), preceding treatment, and feature extraction of the processed speech signal, then it searches and matches feature data with the dictionary and model base, and finally outputs the recognition result.

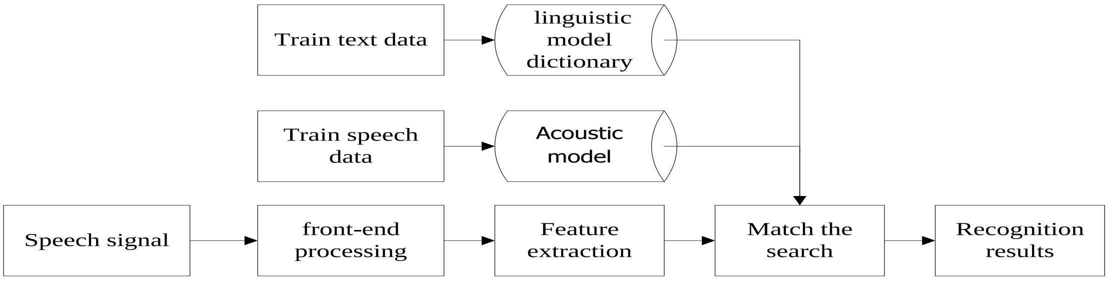


Figure 1. Speech Recognition Process

There are many speech recognition algorithm models, among which the more important are DTW, HMM, SVM (Support Vectors Machine, SVM) [2], ANN (Artificial Neural Network, ANN) [3], and Gaussian Mixture Model (Gaussian Mixture Model, GMM). DTW and HMM are the most widely used and best types of algorithms. This paper mainly applies HMM models and uses HTK (Hidden Markov Model Toolkit) tools to achieve offline speech recognition. HTK is a toolbox for the modular processing hidden Markov model. Developed using C language at the Intelligence Laboratory in the Department of Engineering of the University of Cambridge, after Entropic Corporation and Microsoft Corporation's continuous improvement, as well as providing data preparation, HMM training, speech recognition and data analysis tools in four phases [4], it has become a leading voice recognition toolbox.

Data Preparation: The data preparation phase mainly involves completion of recording (record) and labeling (label), dictionary definitions, tasking of syntax definition (task syntax definition), creating script files (script file) and feature extraction, etc. [5], which are the foundation of voice identification. HTK provides the following tools for data preparation: HSLab for recording and annotation, HDMan to create the dictionary, HParse for generative grammar syntax for the task network, HLed for converting the word-level annotation files into phoneme level annotation files, and HCopy to feature extraction for speech, such as MFCC, LPCC, *etc.*, [6].

Data Training: The main task of this phase is to create the HMM model for voice recognition units. The HMM model parameters are extremely important for speech recognition, and directly determine the quality of the speech recognition results. Before training the HMM model parameters, each model must define an initial structure, including the number of states, mean and variance of each state, and initial state transformation matrix [6]. In the HTK toolkit, Hinit uses the piecewise K-means algorithm to initialize the HMM model parameters, HRset uses the classic Baum-Welch algorithm for the revaluation of HMM parameters, HCompV is mainly used for the calculation of the base macro, and HERest is used for the embedded training of the entire training set.

Speech Recognition: HTK mainly completes speech recognition through HVite tools, which completes tasks based on probability and statistics rules [6]. Using the prepared grammatical structure file, dictionaries and acoustic models which have been trained, the corresponding results are outputted in accordance with the probability of identifying the voice data.

Analysis: Analysis is very important part of speech recognition, and mainly refers to the evaluation of the recognition speed and recognition accuracy. HResults is used for the analysis of results in HTK tools, and will match the identified test results with the completed HMM annotation files, and output the corresponding correct rate and accuracy (accuracy refers to the correct rate based on the inserted error) [6].

3. Chinese Word Segmentation

At present Chinese word segmentation technology is based on the string matching method, semantic understanding of law, and the statistical language modeling method. String based matching method is the simplest type of word segmentation, in which the most representative is the "dictionary method". [7] Its specific method is as follows: from left to right it scans the entire sentence, any words in the dictionary encounter will be marked out, and if it comes across a compound word with a long match, it does not know the string to split the compound word into individual words. Understanding the word signifies having computers learn to think like human beings, through some syntactic and semantic knowledge, to achieve Chinese word segmentation. To make computers learn to think like human beings, in the case of context-free grammar it can be handled easily, but for context-sensitive grammar issues, understanding of the word is still in the experimental stage, and this problem has yet to be solved. Statistical models of the main lexical idea are as follows: the word combination is stable, thus in context, so are the adjacent occurrences of the word, and the more likely they constitute a word. Therefore, between words the adjacent frequency of occurrence can reflect their credibility into words. The training text can appear in the respective adjacent frequency of word combinations statistics, and is calculated based on the mutual information between them. Mutual information reflects the current relationship between characters combined with tightness. When the close is higher than a certain threshold, it may be considered that this word may constitute a word. This paper is based on the achievements of the Chinese word string matching approach. Figure 2 shows the process of specific word segmentation.

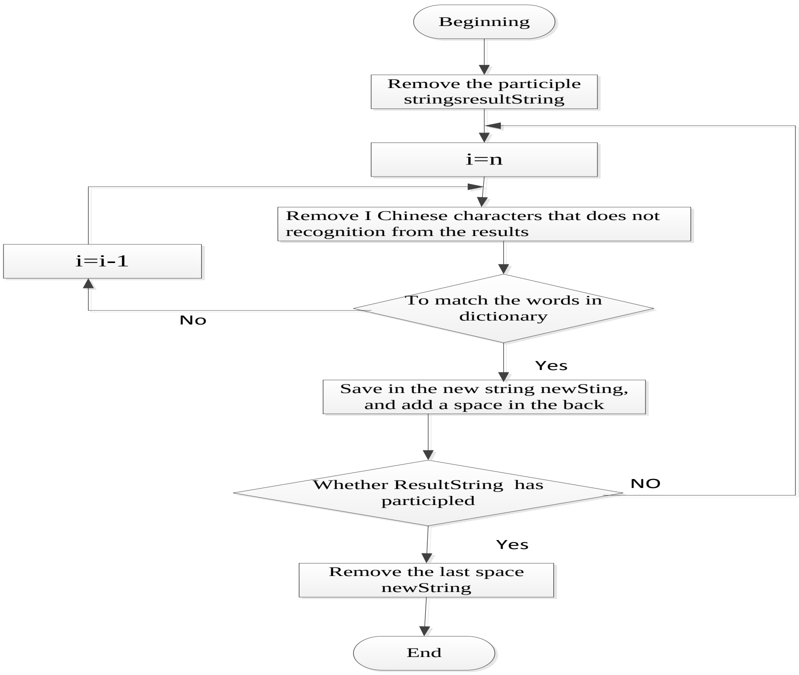

Beginning
Remove the participle stringsresultString
i=n
Remove I Chinese characters that does not recognition from the results
i=i-1
To match the words in dictionary
No
Yes
Save in the new string newSting, and add a space in the back
Whether ResultString has participled
NO
Yes
Remove the last space newString
End

Figure 2. Chinese Word Segmentation Process

4. Field Information Acquisition System Research Based on Offline Speech Recognition

4.1. System Characteristics

What the system mainly realizes is that using offline speech recognition technology and Chinese word segmentation technology implement field information acquisition. Through the integration of offline speech recognition technology and Chinese word segmentation technology, the system realizes a plug-in which is similar to the system input method, and by means of results of Echo word segmentation technology, what the system achieves is that the result of the word segmentation precisely shows the position of the corresponding system interface bivariate table.

4.2. System Architecture

The system architecture is shown in Figure 3, including data acquisition, data management, data synchronization and user management.

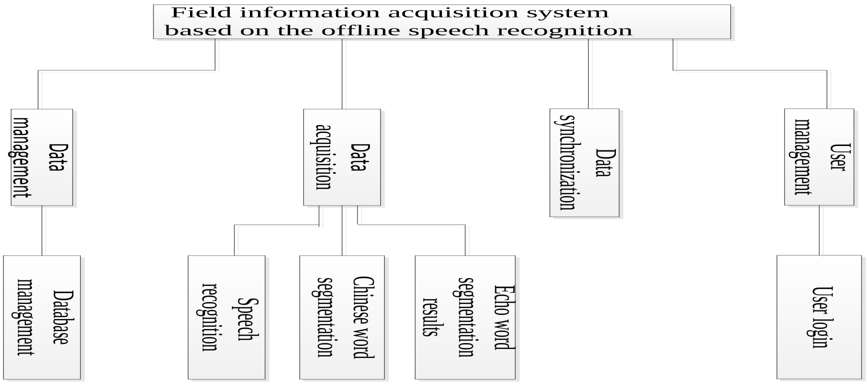

Field information acquisition system based on the offline speech recognition
Data management
Data acquisition
Data synchronization
User management
Database management
Speech recognition
Chinese word segmentation
Echo word segmentation results
User login

Figure 3. System Structure

4.2.1 Data Acquisition: Data acquisition mainly realizes three functions, namely speech recognition, Chinese segmentation, and the results of the segmentation shown in the bivariate table interface.

4.2.1.1 Speech Recognition: What the speech recognition module mainly realizes is the training of the acoustic model data and speech model data, establishment of the acoustic model and speech model by means of HTK tools, and setting up the text dictionary through text editing tools. The module saves model data in the SD card of the mobile terminal device, and then the use of HTK tools fulfills the actual speech recognition and result analysis.

4.2.1.2 Chinese Segmentation: Chinese segmentation is mainly based on the maximum matching method of string matching, compares the string that speech recognition which previously created the word group in the dictionary one by one, according to a certain set of rules. The separator which was selected afterwards has successfully matched connecting phrases after segmentation, then made them into a new string.

4.2.1.3 Segmentation Results Shown Back to Bivariate Table: What the segmentation results show back to the bivariate table is based on the separator which connects the different phrases into a string when Chinese segmentation is performed, and divides the string into different phrases according to partition function, saving them into an array of string. Then it saves the phrases mentioned above into a Hashmap according to the rules which the collection information files match the value of gathering information. It shows the collection data at the correct position of the bivariate table according to the key of the Hashmap match, with the two-dimensional table header field shown last.

4.2.2 Data Management: What data management accomplishes is writing data into the table of the corresponding database, i.e. query data, as well as modifying and deleting the corresponding data in the database through all the objects of the data multiplication layer.

4.2.3 Data Synchronization: What data synchronization accomplishes is that the information which was collected by the mobile terminal is transmitted to the corresponding service and written to database when the mobile terminal connects to a network, and the corresponding data which was increased and changed is quickly downloaded to the mobile terminal device.

4.2.4 User Management: In the user login module of the system, the mobile terminal must be connected to the Internet when the user login is used for the first time. The user's information and related basic information are used to collects information after going through the network verification. The system will download the corresponding information which was newly increased and changed, upon every network login after the first.

4.3. System Implementation

This system based on the score of positioning experiments, for example, in rice. In this case, the system mainly completes the collection of the relative characteristics of wild rice such as plant type, fade, ripe period, lodging resistance, disease resistance, seed setting rate, yield performance and promotion prospect indicators. The information system of the score of positioning experiments in rice, which is based on offline speech recognition, was achieved through offline speech recognition technology. The development tools of the mobile terminal are HTK and Eclipse, the mobile terminal system was developed, compiled and debugged by Android and Sqlite3 as the platform in a Windows XP environment, and the server-side WebService was developed, compiled and debugged with VisualStudio 10 and SQLServer 2005 as tools.

4.3.1. Database Implementation: The system database uses the small embedded mobile database SQLite3, and creates a database and all the tables of database when the program runs. Database implementation of the system is as shown in Figure 4.

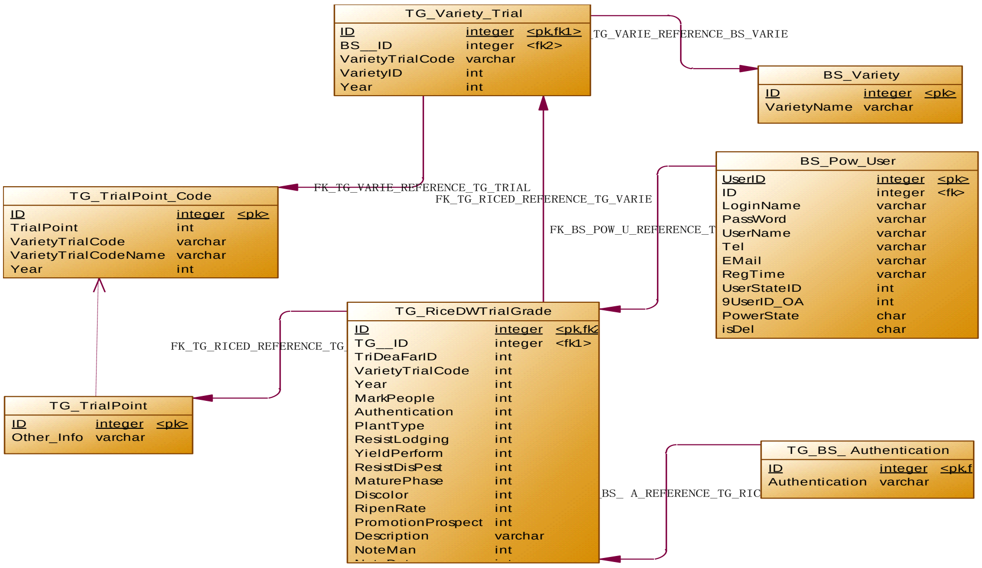
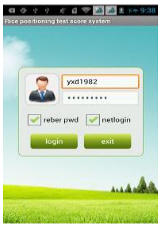


Figure 4. Database Implementation Finger

4.3.2. User Login: The user login includes network login and offline login. The network login requires connection to the Internet, and will download the data which are related to the user from the server after successful login. The offline login does not require connection to the Internet, and only verifies whether or not the user and corresponding password are correct. The software achieved is as shown in Figure 5. The default login type is offline login, and if the network login checkbox is selected, the system logs in using the network login, but we must log in to the system using the network login when system is logged onto for the first time. The system interface after successful login is as shown in Figure 6.

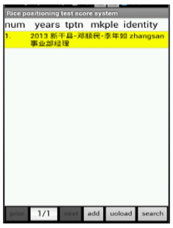


Rice positioning test score system

yxd1982

•••••••••

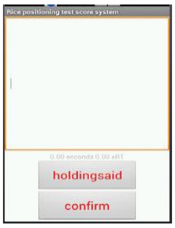
Figure 5. Login Interface



Rice positioning test score system

```
num   years tptn  mkple identity
1.    2013 新干县-邓顺民-李年如 zhangsan
      事业部经理
```

Figure 6. List Interface



Rice positioning test score system

0.00 seconds 0.00 xRT

Figure 7. Speech Input Interface

4.3.3. Speech Recognition: Speech recognition is manly achieved by the HTK tool. Before speech recognition, we must complete the following tasks: First, the training of the speech data is completed by the HTK tool, then the acoustic model, language model and dictionary are created based on the training of the speech data, and finally the models are saved in the system configuration file asset. The data will be written to the specified folder of the SD card when the system is installed to the mobile terminal. The vocabulary content which we have trained includes plant type, fade, ripe period, lodging resistance, disease resistance, seed setting rate, yield performance and promotion prospects indicators, 0, 1, 2, 3, 4, 5 in the system. The speech recognition function module of the system consists of FrontEnd, Decoder and Linguist composition, as shown in Figure 8. The Configuration Manager shows the file configuration manager, FrontEnd shows that the speech data are obtained from the front end, Feature shows that the characteristic data are extracted from the speech data, Linguist shows the model data, Dictionary shows the dictionary, Acoustic Model shows the acoustic model, Language Model shows the speech model, Search Graph shows that the data are received from the model, Scorer shows the scoring, and Pruner shows the pruning. The Scorer and Pruner are the focus of the speech recognition, and the matching portion of speech recognition is composed of the two parts. The software which has been implemented is as shown in Figure 8, speech data are input when the said button is being held, and the system will run speech recognition automatically after the button is released.

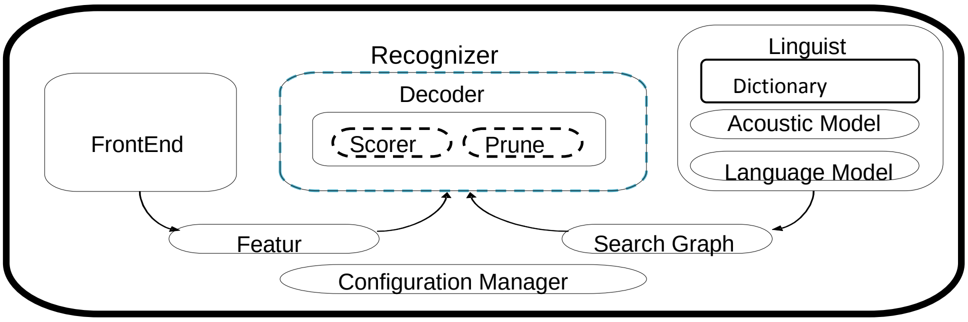
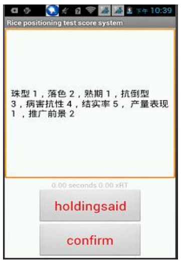


Figure 8. Speech Recognition Model

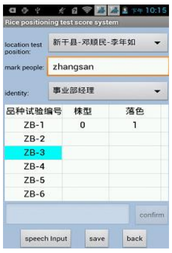


Rice positioning test score system

珠型 1，落色 2，熟期 1，抗倒型 3，病害抗性 4，结实率 5，产量表现 1，推广前景 2

0.00 seconds 0.00 xRT

Figure 9. Output Interface after Segmentation



Rice positioning test score system

location test position: 新干县-邓顺民-李年如

mark people: zhangsan

identity: 事业部经理

品种试验编号	株型	落色
ZB-1	0	1
ZB-2		
ZB-3		
ZB-4		
ZB-5		
ZB-6		

Figure 10. Segmentation Result shows Back Interface

4.3.4. Chinese Word Segmentation and Show Back of Results: Chinese word segmentation is mainly performed on the string after the speech recognition based on the maximum matching method of string matching has been performed. Based on the vocabulary of the data dictionary, the segmentation first removes n words as a new string, then matches them between the new string and the data dictionary words one by one. The recognition results are saved into the new arrays of strings, as a new string, after matching successfully. n-1 words will be removed from the string, also as a new string. They will be matched with the dictionary data until a successful match is reached, then the string will be removed from the original string. n words are then removed from the string as a new string, and the above operation is repeated, until the string is empty. The system uses the style in which Chinese word segmentation will be completed immediately after speech recognition, and the results will be shown immediately after word segmentation. The details of the process of Chinese word segmentation are shown as Figure 3, and the results of the word segmentation realized are shown as Figure 9.

The results shown back consist of the string arrays data which are the same as the collection phrase name, are deposited in the Hashmap key, and those which are not the same as the collection phrase name are deposited in the Hashmap value, through which the information phrases of the system required for collection are compared one be one, such as plant type, fade, promotion prospects indicators, *etc*. Finally, the value of the Hashmap is deposited in the accurate location of the system interface of two-dimensional tables, through the key of the Hashmap one by comparing it with the two-dimensional table head. The detailed process is shown as Figure 11, and the rendering after the soft is realized is shown as Figure 10.

4.3.5. Data Management: Data management mainly includes the management of the mobile terminal database and data synchronization. The management of the mobile terminal database mainly realizes that the collection data is deposited in the local database, as well as the saving data modify function, query function and delete function. The system uses the two-dimensional table as the method of entering collection data. Therefore, one may achieve saving entry data by the bulk style and by saving collection data separately.

Data synchronization refers to the fact that the collected data are locally synchronized to the server database when the equipment connects to the network. The system realizes that the data can be synchronized immediately after collection, and that all the collection data can be centrally synchronized. In the synchronous mode, the system uses the method by which the data were required to synchronize to package the data together, then converted them into an XML file. The XML file is transmitted to the Web server, through the TCP/IP network protocol, using the connection of http technology. Then the data are updated to the server database, after the XML file has been parsed. The data obtained from the server which exits in the form of XML are saved to the local Sqlite database through the local parsing.

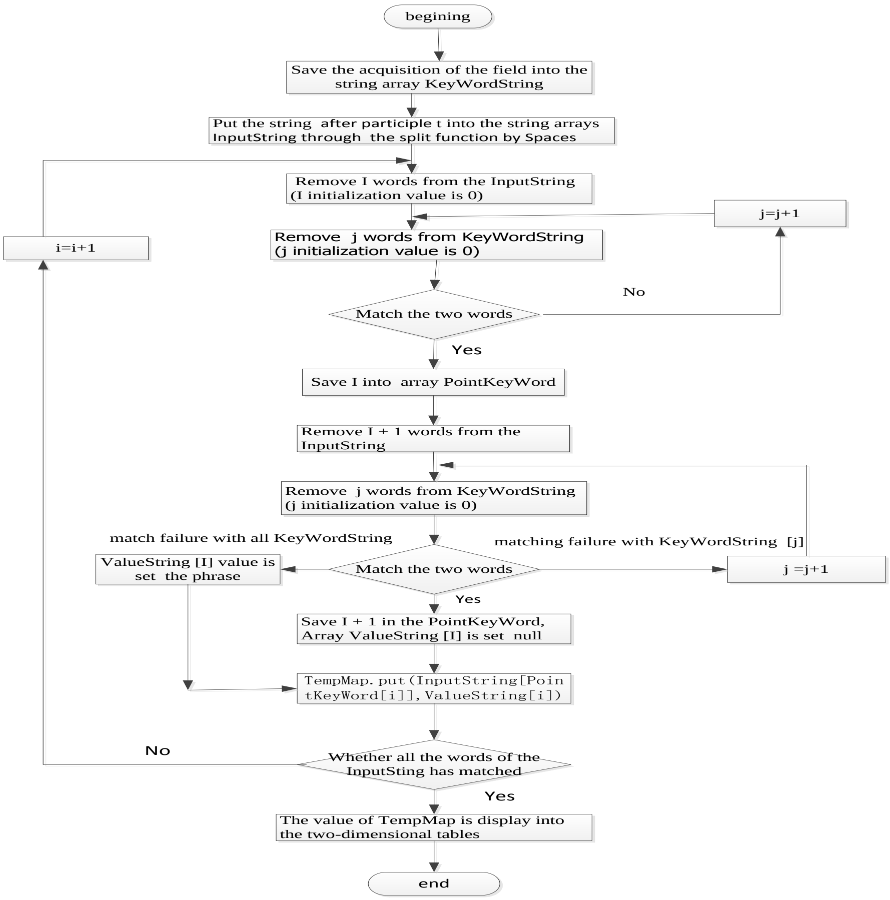

begining
Save the acquisition of the field into the string array KeyWordString
Put the string after participle t into the string arrays InputString through the split function by Spaces
Remove I words from the InputString (I initialization value is 0)
j=j+1
i=i+1
Remove j words from KeyWordString (j initialization value is 0)
No
Match the two words
Yes
Save I into array PointKeyWord
Remove I + 1 words from the InputString
Remove j words from KeyWordString (j initialization value is 0)
match failure with all KeyWordString
matching failure with KeyWordString [j]
ValueString [I] value is set the phrase
Match the two words
j =j+1
Yes
Save I + 1 in the PointKeyWord, Array ValueString [I] is set null
TempMap. put (InputString[PointKeyWord[i]], ValueString[i])
No
Whether all the words of the InputSting has matched
Yes
The value of TempMap is display into the two-dimensional tables
end

Figure 11. Word Segmentation Results is shown Back into Two-dimensional Table

5. Analysis of Experimental Results

5.1. Comparison in Different Environments and with Different Speaking Speeds

The main performance parameters of the speech recognition are error rate, accuracy and speed of recognition [8].

$$\text{error rate} = \frac{D+S}{N}\ , \quad \text{accuracy} = \frac{N-D-S}{N}\ , \quad \text{speed of recognition} = \frac{\textit{recognition time}}{\textit{speak time}}$$

In the formulas above, N represents all vocabulary, D represents deleted words, S represents replaced words; time is identified by Running Time as a unit, *e.g.*, if you speak for two hours, identifying six hours, then the recognition speed is 3xRT.

Different experimental environments can lead to different speech recognition outcomes. The speech recognition results of different experiments are shown in Table 1, and the number of experimental samples is 25 (5 persons, 5 sentences spoken by each person). In an indoor environment, with or without noise, the system may reach the using requirements (practical requirement is error rate ≤ 5%, recognition speed ≤ 2RT); outdoors with no noise or wind, the system can also reach the normal use requirements, but outdoors with a continuous noise or strong wind, the error rate exceeds 5%, then the identification time is more than two seconds, and the system is unable to meet the normal requirements.

Table 1. Results in Different Environment

Experimental environment	Error rate（%）	Recognition（RT）
indoor（no noise）	3.5	1.116
Indoor(with continuous noise around)	4.875	1.118
outdoor（with the breeze）	4.5	1.117
outdoor(with Strong winds）	24.75	1.119
outdoor（with continuous noise）	19	1.200

In speech recognition, different speaking speeds have a great effect on speech recognition, and the speech recognition results of different speech are shown in Table 2, the experimental samples of which is 25 (5 persons, 5 sentences spoken by each person). When the number of the words spoken per second is more than four, then the recognition error rate is quite large, and if the speaking speed increases, the error rate also increases. When the speaking speed is less than four words per second, the recognition accuracy is higher, and as the speaking speed reduces, the recognition accuracy increases. However, if the speed is less two characters per second, the speed has little effect on the recognition accuracy.

Table 2. Results in Different Speeds

Speed（words/second）	Error rate（%）	Recognition time（RT）
more four	23.25	1.160
3-4	10.875	1.154
2-3	4.625	1.118
1-2	3.5	1.113

The different environments play a decisive role in the speech recognition results, and the different speaking speeds have a great impact on the speech recognition accuracy. What can be concluded through the experimental data is that the system may be used normally indoors, or outdoors with no noise or breeze, and when the speak speed does not exceed four words per second, the recognition rate is higher, and the system can be used normally.

5.2. Comparison between Online Voice Recognition and Offline Voice Recognition

Not only are offline voice recognition and online voice recognition different in terms of both identification speed and recognition rate in 2G, 3G, and wifi environment, but they are also different in terms of uncommon industry-specific vocabulary recognition rate and speed. 25 samples were used in this test, with 5 speakers, each saying 5 sentences at one or two words per second, and the test was performed in an indoor noise-free environment. Offline voice recognition uses the speech input method, and the results of test are as follows: online voice recognition cannot recognize Plant type, Lase, Maturity, or lodging resistance specialized vocabulary, and can only recognize Digital “1~5” as the Chinese characters for “one~five”; in addition, if not connected to the Internet, it cannot recognize voice. For recognized vocabulary, such as disease resistance, seed rate, yield performance, and promotion prospects, the results of the online voice recognition and offline voice recognition in the different environments are respectively shown below:

Table 3. Results in Different Networks

	Offline	Wifi	3G	2G
Error Rate (%)	0	28.37	28.37	28.37
Recognition Rate（S）	1.1136	1.274	1.233	6.5

The experimental data show that online voice recognition cannot recognize voice when not connected to the Internet, identification speed is slow with unstable Internet or in a 2G Internet environment, and recognition can only be performed quickly with 3G and strong wifi Internet, but the accuracy of recognizing words in some industry-specific is not good and cannot meet the basic requirements.

6. Conclusion

Based on the basic theory of Continuous Speech Recognition in Hidden Markov Models (HMM), and using some open source tools such as HTK, Sphinx, and so on, this paper implies speech recognition technology based on the offline method, and have completed the Chinese Participle of recognition results by using Maximum matching in Carve segmentation technology. Finally, this technology is used in the Field Data Collection System, via the Chinese Participle to write accurate results to a two-dimensional table of collection system interface after recognition, and implied the Field Data Collection System based on Offline Speech Recognition.

The information vocabulary acquisition implied in this text is based on offline voice recognition technology, Speech model, Acoustic model and stored in the dictionary is smaller, but it is more accurate if the vocabulary in Data Dictionary increase continuously, and the differences among words will become small, thus it is necessary to improve that how to hold a better recognition rate, relatively fast identification speed and relatively high recognition rate when noise is relatively large in outdoor environments.

Acknowledgments

This paper is supported by the “National Key Technology R&D Program under the Twelfth Five-Year plan of P.R. China” (Grant No. 2012BAK19B04-03). And National 863 Program (Grant No. 2011AA10A103-1).

References

[1] X. Wang, “Practical speech recognition basis”, National defence industry press, **(2005)**.

[2] Z. Zhu, B. Zhang and W. Liu, “Speech recognition method based on fuzzy support vector machine”, computer project, vol. 2, **(2006)**, pp. 180-182.

[3] G. Fu and R. Zhao, “Fuzzy self-organizing neural network application in speech recognition”, Journal of Northwestern Polytechnical University, vol. 17, **(1999)**, pp. 600-602.

[4] Y. Steve, E. Gunnar and G. Mark, “The HTK Book”, version 3.4. Cambridge University Engineering Department, **(2005)**.

[5] S. Yang, “The Chinese digital speech input system research based on HTK”, Computer and digital engineering, vol. 4, no. 36, **(2012)**.

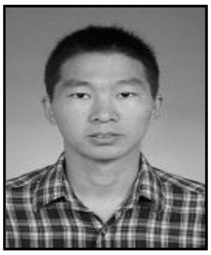
[6] M. Nicolas, HTK(V3.1).Basic tutorial, **(2012)**.

[7] L. Xu and L. Li, “The comparison of the HMM training to improve algorithm of speech recognition”, Computer CD Software and Applications, vol. 23, **(2012)**, pp. 30-32.

[8] N. Moreau, HTK (v.3.1).Basic Tutorial, **(2002)**.

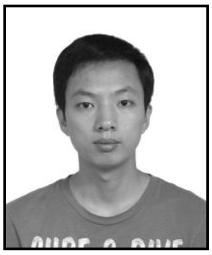
Authors

ZhengYong HuYan, received his Bachelor degree in Information System and Information Managment from Mangement College, at Tianjin University, China in 2010. Currently, He is a candidate for the degree of Master of Agricultural electrification and automation in China agricultural university of College of information and electrical engineering. His research interests include Mobile data synchronization, offline speech recognition etc.



Likui Xu, received his master degree in Computer Science and Technology from College of Information and Electrical Engineering of China Agricultural University,China 2013.He is a software engineer of some software company, His main research field include Speech Recognition, Natural Language Processing and Language Model.

Shuai Fang, received his Bachelor degree in Computer Science and Technology from School of Computer Science and Engineering, at Beihang University, China in 2012. Currently,He is a candidate for the degree of Master of Computer Science and Technology in China agricultural university of College of Information and Electrical engineering. His research interests include Natural Language Processing, Network Reptile etc.



Zhe Liu, Liu Zhe works as Lecturer of Geographic Information Science Department, College of Information and Electrical Enginneering at China Agricultural University(CAU),China. He obtained his Ph.D. from CAU. His current research interest focuses on statistical methods and information technology about plant breeding, phenotyping, seed production and variety extension. Until now, he has published more than 20 peer-reviewed papers.

Xiaodong Zhang, Zhang xiaodong works as prpfessor of Geographic Information Science Department, College of Information and Electrical Enginneering at China Agricultural University(CAU),China. She obtained her Ph.D. from Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, China. Her current research interest includes Application of GIS, Agricultural information system. Until now, she has published more than 50 peer-reviewed papers.

Lin li, Li Lin works as professor of Computer Science and Technology Department,College of Information and Electrical Enginneering at China Agricultural University(CAU),China.She obtained her Ph.D. from China Agricultural University (CAU), China. Her current research interest includes software engineering(SE),software automation(SA), mobile internet technology and so on. Until now, she has published more than 50 peer-reviewed papers.