The research on Uighur speaker-dependent isolated word speech recognition

Wushour • silamu   Caiqin   Nuominghua

College of information science and engineering

Xinjiang University, Urumqi   830046

Abstract: A small vocabulary, isolated word speech recognition system in java has been realized. In this system we have done the extraction of feature parameter, the training of speech model parameter and the recognition of the recorded speech. MFCC is used as feature parameter, HMM model is used for speech model in this article and do some study on Uighur speech recognition based on them.

Keywords：Speech recognition.   Hidden Markov models.   Speech feature parameter.

1.Introduction

Based on achieving English speech recognition system ,we study in Uighur language and Chinese language voice recognition system. We using Jbuilder and Jdk 1.4 development tools to study how to deal with the Uighur text characters.

Isolated word speech recognition system, with broad application prospects, such as computer control, industrial control orders, family services, banking services, personal information Identification, personal items such as mobile communications, the application of this technology will greatly facilitate the daily lives of people. The work of Uighur language recognition system further study instructive and help.

11. Discussion Uighur language speech recognition in the development process to address issues

In system development process, I base on Chinese speech recognition system transform into Uighur language speech recognition system, which is facing the biggest problem how to deal with Uighur text characters in applications processing, Uighur language of a different editorial direction, wide-ranging characters, ranging from long code, even set spacing, and other characteristics of the word, all these specific issues to be resolved. In java programming, each character stored in the code , Internal code is Unicode (all languages are corresponding codes, internal computers character string expressed internal code, in general computer character string encoding relates platform, the Java unrelated platform using Unicode). When Java get a character string form a byte stream, associated with the platform byte changes Unicode byte character string which independent on platform. Exporting, Java will be changes Unicode character string into the byte stream which relates with platform. if a Unicode characters do not exist in a given platform will be exporting a '? '.

1, Java Translation devices translate the source document, source document first will be converted to Unicode coding. So before translation of documents, we must "tell" Translation devices what coding method used.

For example : Our source document is preserved in UTF-8 method, in translation, but translation devices regard as preserving GBK method .the Translation devices will converted accordance with GBK to Unicode coding method of source documents, and then translation, so certainly wrong, Indeed Translation devices should follow the converted accordance with UTF-8 to Unicode coding method.

For control panels procedures, translation device will source documents as system acquiescence coding type to coding (the coding system acquiescence depend on control panels configuration, Chinese Win2K type is usually GBK), can use -encoding parameters to configure, such as : Javac -encoding UTF-8, such translation device will be used as the source document UTF-8 encoding (Telling Translation devices source codes type ,but not change source codes). In a variety of languages platform used -encoding designated with the same source code method in the translation, the internationalization issue would not exist.

2, When dealing with export and import, export and import flow's coding type are according to users import and export of equipment display coding type. As JRE in handling export and import content will be encoded conversion, the import will be converted to Unicode and then send into, and therefore it is necessary to correctly match the actual contents of the coding method and inform JRE coding type; For export, Unicode codes converted to other coding type and send to other procedures. Therefore it is necessary to correctly match the export of equipment that use coding method and inform JRE codes.

For example, in the procedure import flows coding is Uighur text characters. *new InputStreamReader(System.in,"Uighur Basma");* And the establishment of operational procedures, users used the Simplified Chinese input method, inputting the GB2312 codes content .In JRE ,GB2312 coding content as the Uighur Basma been operating.GB2312 convert to Unicode coding. This result is obviously not users want to import content. Acquiescence cases, JRE will import and exports content coding as acquiescence coding type.

At the same time, we found although Java platform support majority additional

characters treatment, but in Java applications procedures, Uighur text characters ﺔ

Can not be very good support. If correct display must be in a special handling procedures. Figure 1, figure 2 is the small vocabulary isolated word Uighur speech recognition system training interface and identify interface, Figure 3, Figure 4 is the corresponding Chinese language training and identification interface. During the training process parameters will be set and each word can be select the number of training options, the rate of identification of specific laboratory conditions could reach

80%.

Figure 1 Uighur speech recognition system identify interface

figure 2　Uighur speech recognition system training interface

Figure 3 Chinese speech recognition system identify interface

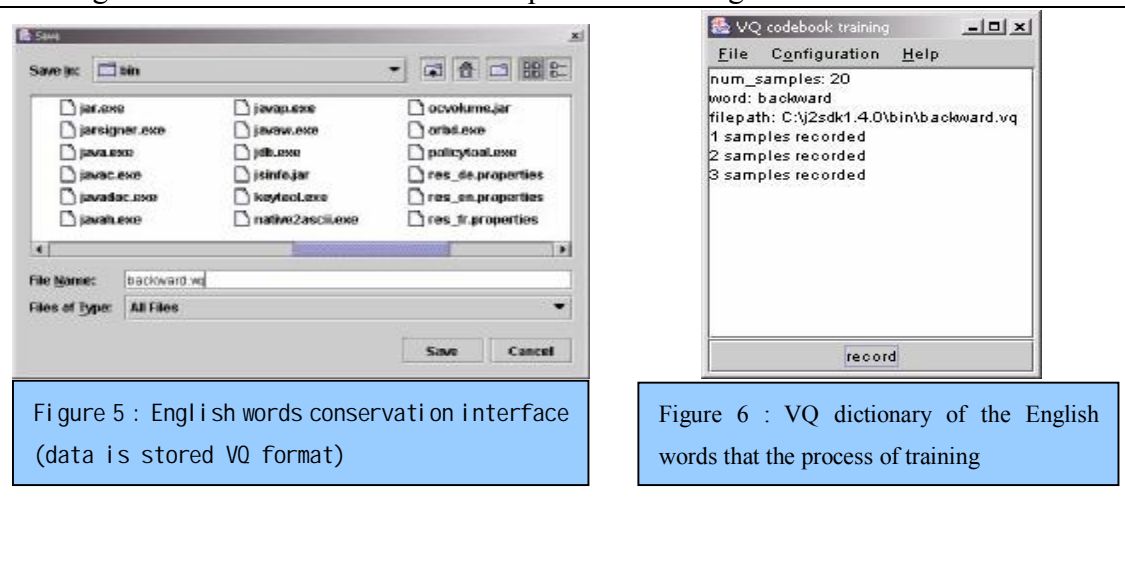figure 4 Chinese speech recognition system training interface

Ⅲthis work and innovation as follows

1. For front-end voice signal processing : from two aspects of the domain and frequency domain to voice signal parameters LPCC and MFCC extraction feature of the analysis ,analysis of the end testing of the impact of speech recognition system. Traditional band signal detection methods will be the true voice of energy combined with a zero rate to be judged, and the system made small improvements ,in process we added variables and controlled window wide, used a dynamic window long end voice detection methods to improve choose the precise starting point.

2. From HMM models the three basic issues (assessment, decode, training) to

proceed as a detailed analysis our system models of the basic principles. the largest clustering algorithm applicant templates training process, to speed up the pace of Hmm parameters optimized, through reference extensive documentation that the largest clustering only have a greater impact on the first iterative process, but follow that the iterative accelerate effects are not obvious, on the contrary, an increase of the complexity and fewer training data may increase errors. Based on the above reasons, the system also made small improvements, the first iterative process applied the largest clustering algorithm.

3. Java language achieve to the basic problems of Hmm training, decode, assessment algorithms. Achieving cross-platform functionality, and integration of the practical application of voice, increased front-end processing, such as end detection, filtering and other processes, forming a complete voice processing include voice input ,training, identification and thus may deal with a large number of voice data in time. Procedures interface friendly, easy to operate. Run on Windows2000 operating system and Red hat Linux 9 system, JDK 1.4 environment translation passed. Software is divided into two modules : voice training and voice recognition. As with similar with Uighur and English (both choose the factors of recognition based), the first realization system identify English words ,and then processing the Chinese language and Uighur language operations, the highest rate of identification English. As figure 5 and Figure 6 shows the data interface is preserved in English :

Figure 5 : English words conservation interface (data is stored VQ format)

Figure 6 : VQ dictionary of the English words that the process of training

Ⅳ summing up

The small vocabulary and isolated word speech recognition technology has done preliminary exploration, the establishment of a foundation for the actual engineering applications. But far from practical level, due to experimental conditions and time constraints, the system still has many deficiencies, and there are still many areas which need further research :

(1) This issue is only for small glossary and isolated word speech recognition research, did not consider language model, which is necessary condition of large vocabulary continuous speech recognition. Currently Uighur language speech recognition technology research still in initial stage, should be conjunction with the Uighur language text different characteristics and special issues focusing on the study large vocabulary, non-specific recognizer, continuous speech recognition , Uighur key technologies, new ideas, new methods.

(2) improve the experimental system language corpus. The use of voice is invited several friends in the laboratory environment collected, fewer in number and geographical distribution is uneven, and HMM is based on the identification of statistical models, sufficient data for training speech recognition results with greater impact. Collection needs more rich voice samples, voice recognition will eventually have to go out of the laboratory, the real application to the lives of people, the voice laboratory environment better, almost no unexpected noise in the practical application of these is inevitable, so the collection also needs some simulation of the actual circumstances voice samples for further study.

References :

1.Yangshixing,Chihuisheng, voice digital signal processing, Beijing : electronics Press, 1995

2.Yikechu, Tianbin,Voice signal processing, defense industry Press, 2001

3. Xuxiaopen,Wuji, Isolated word speech recognition algorithms to improve the performance of research and computer engineering and applied 2001,21

4. Wangrangding, Chaipeiqi, Voice opposing spectral characteristics of the study, computer engineering, 2003,29 (13), 31~33

5. Lawrence Rabiner，Biing—Hwang Juang A photocopy version of the basic principles of voice recognition, Tsinghua University Press, 1999.

The author profiles :

Wushour.Silamu:1946-8-20;Male; Uygurs; Urumqi City, Xinjiang Uygur Autonomous; Xinjiang University College computer information science and engineering professor; Graduate, and doctor tutor; Major research multilingual information processing, intelligent applied research, analysis and development direction ;E-mail : wushour@xju.edu.cn; fax : 0991-8582762 detailed address : Xinjiang University Information Engineering College; zip code : 830046; Tel : 0991-8582762 (office)

Caiqin:1977-11-23; Female, Han nationality; Tianjin; Xinjiang University computer information science and engineering department at Reading College Graduate; Research direction for Uighur language information processing ;E-mail: cq_cqin@163.com

Nominghua:1981-11-21；Mongolian; Wu Ming Siziwangqi Inner Mongolia Autonomous Region; Xinjiang University computer information science and

engineering department at Reading College Graduate; Research direction for Uighur language information processing ;E-mail: nuominghua@163.com