Design of Intelligent Pronunciation Training System based on Android Pslatform

Mingming-Ding
Henan University of Animal Husbandry and Economy, Zhengzhou, Henan,450000， China
1377960836@qq.com

ABSTRACT
In recent years, with more and more Chinese learning English, there are more and more English learning software, but most of the software are lack of good pronunciation evaluation and feedback correction. However, in English pronunciation learning, especially for non-native English learners, effective feedback motivation is very important. This has become a bottleneck restricting intelligent English learning software. Based on Android smart phone platform, this project develops a set of English pronunciation training system with multiple functions such as pronunciation following, pronunciation evaluation and pronunciation correction by using related speech recognition technology, aiming to realize the intelligence, popularization and portability of English pronunciation learning.
Keywords: *Speech recognition; Pronunciation evaluation; English learning; Android*

1 INTRODUCTION

Dialogue database management: training data and corpus of model a (structured and unstructured). The business problems consulted by users are generally managed in the form of question and answer library, and some are stored and trained in the form of table \ dB, Atlas, document and log. Table: if users ask for product information, because there are many products and many parameters, such as 10000 products, 10 parameters need to be disassembled into 100000 Q & A. once the parameter value is modified, the workload will increase exponentially. Similar problems can be stored in the form of a two-dimensional table, just set semantic information for various parameters or fields. Knowledge map: the knowledge map Q & a publicized in the market is mainly the Q & A of attribute of in the form of the above table, or the gossip in the user's general field (such as who starred in x x film) documents (reading and understanding): training, learning documents, policy documents, etc. through the ability of reading and understanding, the information is extracted to form the knowledge map[1]. At present, most of them only realize the extraction of entity classes and right and wrong classes. Log (data interface): it is generally used as a source to extract hot Q & A and business scenarios, and learn the Q & A ability of excellent customer service and related recommendations from the logr.

The related research of auxiliary pronunciation learning has great theoretical research value and application prospect. Nowadays, the development of mobile Internet technology is rapid. Mobile communication extends the network to everyone's pocket, and it is truly ubiquitous, which provides a stage for the application of smart phones. Smart phones not only have the traditional functions of voice call, but also have more powerful data processing ability, more gorgeous graphical user interface and stronger software loading ability. The most important point is that the price of smart phones is not different from that of ordinary mobile phones, so the popularization of smart phones is an inevitable trend[2]. Compared with traditional computer software, the software based on smartphone platform has great convenience and practicability. Android system, as an excellent smartphone operating system, has developed rapidly in recent years, and has reached more than 80% of the smartphone. Therefore, the development of an intelligent English spoken English learning software with sound feedback function based on Android smartphone platform can greatly improve the efficiency of learners' oral learning, and has great social and economic value[3].

2 RELATED WORK

Qianyu cao proposed optimization of intelligent english pronunciation training system based on android platform[3]. Yaohua bu, tianyi ma, weijun li, hang zhou, jia jia, shengqi chen , kaiyuan xu, dachuan shi, haozhe wu, zhihan yang, kun li, zhiyong wu, yuanchun shi, xiaobo lu, ziwei liu[4] proposed a computer-aided personalized pronunciation training sys-tem with exaggerated audio-visual corrective feedback. Ayush vishnoi, preeti mishra, charu negi, sateesh kumar peddoju [6] proposedandroid

mal-ware detection techniques in traditional and cloud computing platforms: a state-of-the-art survey.
Android is the original meaning of "robot", which is the name of the open source mobile phone operating system opened by Google on November 5th, 2007. It is composed of Linux system kernel, middleware, user interface and application software. It is the first fully open and customized operating system for mobile platform. At present, the latest version of Android is called jelly bean (android4.1Jelly Bean).
Android has a very important significance. Its emergence not only makes Google enter the ranks of mobile Internet, but also has set off a revolution in the mobile Internet world, and further promotes the concept of "providing information for everyone anytime and anywhere" as a mobile interconnection concept. Android allows developers to freely access and modify the source code, which is a set of mobile terminal solutions with open source code nature. Since Android is completely open source and free of charge, businesses can improve or customize the platform, and can also apply their own proprietary technologies to the platform, and create better things based on the platform[7].
This paper consists of the following parts. The first part introduces the related background and significance of this paper, the second part is the related work of this paper, and the third part is data analysis. The fourth part is example analysis. The fifth part is conclusion.

2.1 *architecture of Android system*

Android is a fully open and free open source mobile phone operating system based on Linux platform. Its system framework adopts a hierarchical structure. As shown in Fig 1, Android system is divided into four layers, from low to high, which are Linux core layer, system runtime layer, application framework layer and application layer.

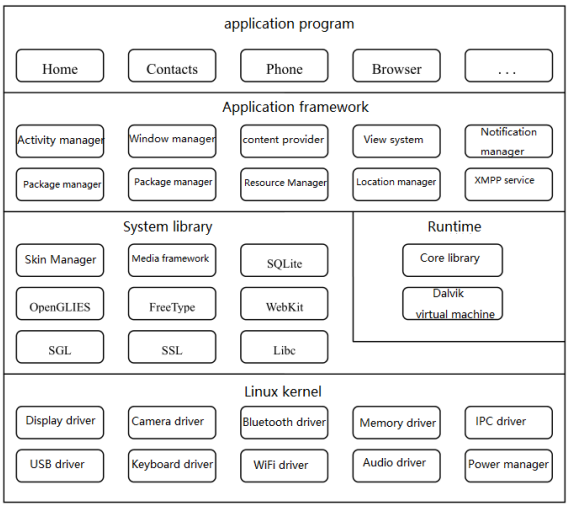

Fig. 1. Android system architecture

2.2 *Android application construction principle*

Android applications are written in Java language. Like ordinary Java se, Android programs need to use sun JDK to compile source java files into bytecode files (. Class files). After that, all bytecode files can be further transformed into DEX files by using special tool software DX. In order to form the final application package, Android comes with a package tool AAPT. Through AAPT software, all the DEX files, layout files (XML files) and various resource files of the program can be packaged together to form a file with APK suffix, that is, application installation package (APK). Apk can be decompressed and installed on the mobile phone to form the final executable program[8]. Dalvik virtual machine reads instructions and data from it to make the application run.

2.3 *Android application development environment*

The development of Android program usually uses Java as the programming language. This topic uses eclipse as the integrated development environment for system development[9-10].
later began to support the integration of a variety of plug-ins. Because of its good plug-in extensibility, eclipse is more flexible than other traditional ides.
and other plug-ins. Eclipse platform is the foundation and core of the whole system. Java development tool (JDT) can provide functions such as view, edit, compile, debug and run. It is a special plug-in using java code. The plug-in development environment (PDE) is based on Eclipse platform and JDT, and provides tools for plug-in development.

3 DATA ANALYSIS

3.1 speech signal feature extraction

for subsequent processing. The feature extraction of signal not only highlights the data features of pattern matching and improves the recognition rate of the system, but also compresses the information and reduces the storage and calculation of the system[11].

Research shows that the level of sound perceived by human ear is not consistent with the actual frequency of sound. In fact.But in the high frequency part, the human ear's feeling is more and more rough. The relationship between sensitivity and frequency is linear below 1000Hz, and logarithmic above 1000Hz.

The calculation process of MFCC is shown in Fig 2. The specific steps of extracting MFCC feature parameters are as follows:

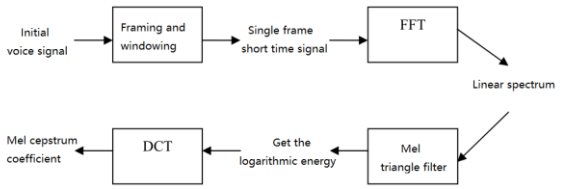

Fig. 2. Extraction process of MFCC parameters

(1) The original speech signal is transformed into a single frame short-time signal x (n) after pre-processing such as framing and windowing.

(2) The short-time signal x (n) is analyzed in frequency domain. Using discrete Fourier transform (DFT), the formula is as follows:

$$X(k) = \sum_{n=0}^{N-1} x(n) \bullet e^{(-j\frac{2\pi kn}{N})} \quad (1)$$

(3) After the linear spectrum X (k) passes through the Mel frequency triangular filter banks, the logarithmic energy s (I) of each filter output is obtained. Mel filter bank is a number of triangular filters distributed averagely on Mel frequency in the frequency range of speech signal. The center frequency of each filter is f (i), 0 ⩽ i ⩽ m, where m is the number of filters.

The transfer function of each triangular bandpass filter is as follows:

$$H_i(k) = \begin{cases} 0, k \le f(i) or k \ge f(i+2) \\ \dfrac{k - f(i)}{f(i+1) - f(i)}, f(i) < k \le f(i+1) \\ \dfrac{f(i+2) - k}{f(i+2) - f(i+1)}, f(i+1) < k < f(i+2) \end{cases} \quad (2)$$

(4) The logarithmic energy s (I) is transformed into the cepstrum domain by discrete cosine transform (DFT), and the MFCC characteristic parameters of a single frame speech signal are obtained[8].

3.2 speech signal pattern matching

In the pronunciation evaluation system, in order to compare the speech to be evaluated with the reference standard speech, the similarity between them can be reflected by calculating the difference of their characteristic parameters. However, in general, the pronunciation length and speed of the tested speech and the reference standard speech are not the same, so it is not possible to compare them directly. At this time, we need a certain matching method to match the feature parameters. At present, the common matching algorithms for speech recognition mainly include dynamic time warping (DTW) and hidden Markov model (HMM).

HMM is a kind of probability model, which is used to describe the statistical characteristic parameters in the random process. It uses a lot of speech data in the training stage, and only after repeated operation can the parameter model be obtained.

$$\Delta u_{k+1}(t) = u_d(t) - u_{k+1}(t) = u_d(t) - \Big(u_k(t) + L(t)\big(\dot{e}_{k+1}(t) + e_{k+1}(t)\big)\Big) = u_d(t) - u_k(t) - L(t)\big(\dot{e}_{k+1}(t) + e_{k+1}(t)\big) = \Delta u_k(t) - L(t)(\dot{C}(t)\Delta x_{k+1}(t) + C(t)\Delta\dot{x}_{k+1}(t) + C(t)\Delta x_{k+1}(t)) \quad (3)$$

$$E(t)\dot{x}_{d+1}(t) - E(t)\dot{x}_{k+1}(t) = E(t)\Delta\dot{x}_{k+1}(t) = f\big(t, x_d(t)\big) + B(t)u_d(t) - f\big(t, x_k(t)\big) - B(t)u_k(t) = f\big(t, x_d(t)\big) - f\big(t, x_{k+1}(t)\big) + B(t)\Delta u_{k+1}(t) \quad (4)$$

4 EXAMPLE ANALYSIS

4.1 system analysis

According to the requirement analysis, the functions of the system mainly include:

(1) Demonstration of pronunciation. Pronunciation demonstration refers to that when learning pronunciation, first play the standard pronunciation animation video or standard pronunciation sound, at the same time, cooperate with the pronunciation structure diagram and introduction text, so that learners have a correct understanding of the pronunciation, and can fully understand the key points of pronunciation, the movement characteristics of mouth shape and tongue position, etc.

(2) Follow the pronunciation. The system first plays the correct pronunciation animation or sound, and then prompts the learners to follow it; The system records the pronunciation to the memory of the mobile phone for subsequent processing.
(3) Pronunciation contrast. The system first plays the standard pronunciation animation video or sound, and then plays the recorded learner's pronunciation. Pronunciation contrast function is mainly through the comparison of standard reference pronunciation and learners' pronunciation, so that users can have a direct understanding of the gap between pronunciation and standard pronunciation.
(4) Pronunciation score. Pronunciation score is one of the core functions of the system. It mainly uses speech recognition technology and related pronunciation score algorithm to have a quantitative evaluation of learners' pronunciation results. Accurate and reliable pronunciation score can make learners have an accurate understanding of their pronunciation performance, and then continuously improve their pronunciation and improve their pronunciation level.
Through the analysis of the functional requirements of the system, the core of the system should include the following modules: voice recording module, voice and video playback module, voice scoring module based on AP, and image display module of voice formant, as shown in Fig 3.
Application service integration mainly includes six smart campus application services: user identity management, library management, campus public facilities management, campus consumption management, campus security and campus Mobile SMS notification management. Data fusion layer data fusion refers to the establishment of a smart campus information resource sharing platform through the integration of smart campus information resources. The establishment of smart campus information resource sharing platform includes not only the integration of identity information, application data, messages, teaching content and perception information, but also data storage, middleware and supporting software

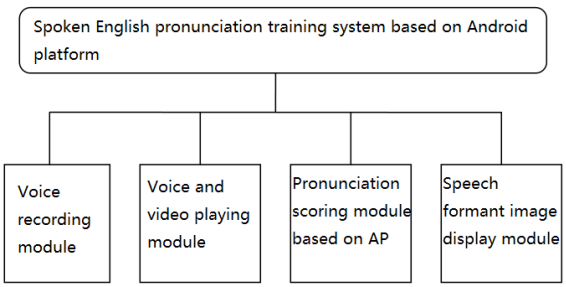

Fig. 3. Module structure of the system

4.2 IO module design

However, it is difficult for developers to set the sampling and coding of the underlying audio. Audiorecord class is a solution proposed by Google to solve the limitations of mediarecorder class. Through audiorecord class, audio can be recorded in the buffer for later processing. Developers can flexibly set the basic parameters of voice signal, such as the size of audio buffer, sampling rate and sampling bits. The audiotrack class corresponds to the audiorecord class and is used to play voice signals. IO module design is shown in Fig 4. Two recording methods are implemented in the Android SDK, as shown in Fig 5.
The classroom monitoring subsystem uses single chip microcomputer as the controller of the monitoring module to realize the functions of perception, control, display and service. The monitoring module is arranged in the classroom, uses the RF card reader to identify the personnel information, collects the teaching temperature by the temperature sensor, and collects the brightness inside and outside the classroom through the photosensitive device. On the one hand, after being processed by the controller, the data taken from the sensor is submitted to the environmental monitoring server for use by the whole system; On the other hand, it can intelligently adjust the switching state of classroom lights, fans and other equipment to minimize energy waste. At the same time, the indoor temperature, brightness, number of people and other relevant information collected by the sensor will also be displayed on the monitoring display screen installed in the teaching space. By viewing the relevant data sent by the information acquisition module to the remote control server, the classroom administrator releases some or all of the intelligent control mode, enters the manual control mode, and opens the control right of the data desk lamp and audio-visual education equipment. The control terminal in the teaching slaughterhouse is also allowed to release part or all of the intelligent control mode through the remote control and enter the manual control mode. In this way, after viewing the environmental information and equipment usage of each classroom, the administrator can remotely control the classroom according to the actual situation.

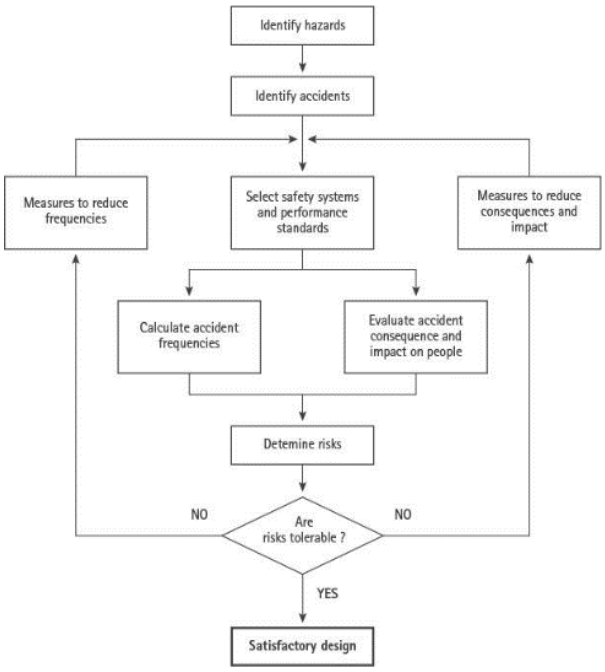

Fig. 4. IO module design

This design unifies the portal of web and WAP, realizes the integration of application services and data, adopts the idea of three networks in one, and integrates various network access means. The smart campus system based on Internet of things is a large integration of heterogeneous subsystems with multiple functions, including classroom monitoring subsystem, security inspection system, canteen management, student dormitory water and electricity management, daily teaching management, intelligent library, laboratory management, etc. The school should consider the actual situation of its campus from the aspects of human, material, financial and technical support, so as to realize it step by step and steadily, and improve the pace and effectiveness of the construction of intelligent campus.

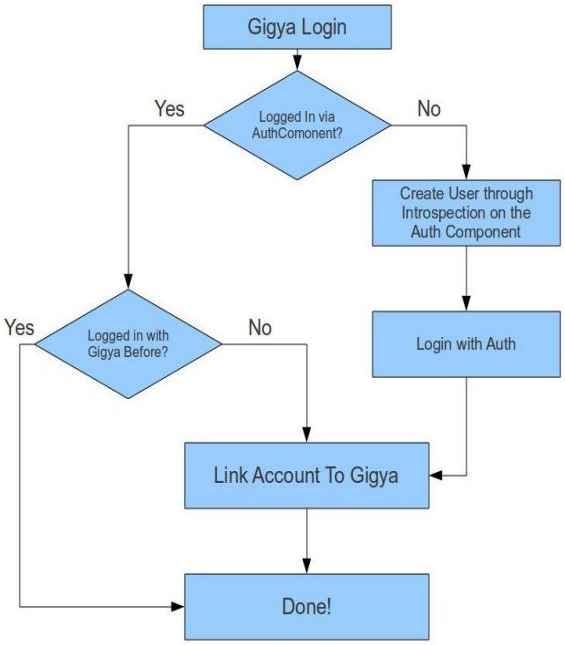

Fig. 5. IO module design

5 CONCLUSION

Intelligent language learning system based on speech recognition is the development trend and frontier hot spot of computer-aided language learning. Aiming at the current research status, this paper deeply studies the related theories of speech recognition, discusses the speech scoring and pronunciation correction algorithm, and designs and develops a set of speech recognition system based on Android mobile platform The multi-functional English pronunciation training system, such as pronunciation correction, aims to realize the intelligence, popularization and portability of English pronunciation learning.

REFERENCES

[1] Wang Bingxi, Qu Dan. Fundamentals of practical speech recognition [M]. National Defense Industry Press, 2005.
[2] Zhao Li. Speech signal processing [M] China Machine Press, 2011.
[3] Qianyu Cao , Hanmei Hao:Optimization of Intelligent English Pronunciation Training System Based on Android Platform. Complex. 2021: 5537101:1-5537101:11 (2021).
[4] Yaohua Bu, Tianyi Ma, Weijun Li, Hang Zhou, Jia Jia, Shengqi Chen , Kaiyuan Xu, Dachuan Shi, Haozhe Wu, Zhihan Yang, Kun Li,

Zhiyong Wu, Yuanchun Shi, Xiaobo Lu, Ziwei Liu:PTeacher: a Computer-Aided Personalized Pronunciation Training System with Exaggerated Audio-Visual Corrective Feedback. CHI 2021: 676:1-676:14.

[5] Aayush Vishnoi, Preeti Mishra, Charu Negi, Sateesh Kumar Peddoju:Android Malware Detection Techniques in Traditional and Cloud Computing Platforms: A State-of-the-Art Survey. Int. J. Cloud Appl. Comput. 11(4): 113-135 (2021).

[6] Sorinel Capusneanu, Melinda Timea Fülöp, Dan Ioan Topor, Nicoleta Magdalena Iacob, Alina Georgiana Solomon, Ana Maria Ifrim:Designing a Mobile Chat Application Based on Android Platform. Int. J. Innov. Digit. Econ. 12(4): 43-63 (2021).

[7] Guanlin Chen, Kunlong Zhou, Yubo Peng, Liang Zhou, Yong Zhang:A novel network intrusion prevention system based on Android platform. Int. J. Internet Protoc. Technol. 14(2): 67-76 (2021).

[8] Sharma A , Kumar M . Feature Selection and Classification Based on Swarm Intelligence Approach to Detect Malware in Android Platform[M]. 2021.

[9] Mahajan G , Prajapati V . Fertilizer calculator – NEW!: A generic app for the android platform for soil test-based fertilizer recommendations. 2021.

[10] Sengupta A , Sivasankari S . A Novel Approach for Analysing and Detection of Obfuscated Malware Payloads in Android Platform Using DexMonitor[M]. 2021.

[11] Novaliendry D , Subagia A , Sriwahyuni T , et al. Designing Augmented Reality-Based Computer Devices Learning Media on Android Platform[C]// Proceedings of the 4th Forum in Research, Science, and Technology (FIRST-T1-T2-2020). 2021.