Design and Implementation of Adaptive Speech Control Intelligent Conference Room Based on Intelligent Speech Technology

Qiuyong Yang¹,*, Jianyong Wang², Yanghua Pan³
¹Guangdong Power Grid Co., Ltd.; Guangzhou Guangdong 510062 China
²Guangdong Electric Power Information Technology Co., Ltd.; Guangzhou Guangdong 510062 China
³Guangzhou Yueneng Information Technology Co., Ltd.; Guangzhou Guangdong 510062 China

*Corresponding author: liminlm26@126.com

Abstract. With the development of multimedia and network technology, more and more enterprises use network conference room system. This paper implements a multi-functional intelligent conference room system based on streaming media server. This system is developed on the basis of the prototype conference system. In addition to the functions of the traditional conference system, it can also access the video monitoring system and allow intelligent mobile terminals to join the conference room. The intelligent conference room sends management commands to the user through the interface to realize the management function of the voice conference room. At the same time, the video information can be exchanged through the video processing module. Finally, the mobile terminal video access and conference room video extrapolation function are realized. The experimental data show that the intelligent conference room can improve the recognition of conference information. This provides a reference for the design and implementation of adaptive speech control intelligent conference room.

Keywords: Streaming Media, Intelligent Conference Room, Voice Management, Mobile Terminal.

1. Introduction
With the advent of globalization, companies and enterprises generally set up offices in multiple locations [1-2]. Because the offices are scattered, cross domain communication problems are brought about. With the development of multimedia and network technology, cross domain communication can be well solved by network consultation system [3]. The network consultation system can transmit voice, video, pictures and other information through the network, which can not be limited by the region, improve office efficiency and save administrative costs for companies and enterprises [4-5]. But the traditional network consultation system can provide limited functions, has been unable to meet the growing needs of users.

With the development of video compression and codec technology, cameras and other monitoring
equipment are becoming cheaper and cheaper, people's security awareness is getting higher and higher, video monitoring system is more and more widely used in all walks of life [6]. Users with monitoring system hope to watch the monitoring image in real time at the conference site, which leads to the need to integrate the monitoring system into the conference system.

Now smart phones are very popular, users do not need to access the network through the cable, they can surf the Internet at high speed. For some outdoor workers or business people who often travel, if they can participate in video conference through intelligent mobile terminal, it can not only improve their work efficiency, but also save time and administrative expenditure.

This paper implements a multi-functional network conference system based on streaming media server. This system is developed on the basis of the prototype conference system [7-10]. In addition to the functions of the traditional conference system, it can also access the video monitoring system and allow intelligent mobile terminals to join the conference room.

2. System analysis and design

2.1. System architecture design
Bigbluebutton open source framework is mainly composed of four parts: client interface, BBB-Web, red5 streaming media server and FI-EE switch voice conference server [11-12]. Its overall architecture is shown in Figure 1:

![Figure 1 Overall architecture of bigbluebutton conference room](image)

Client interface refers to the conference room interface that users see through the browser. This part mainly displays the video, sound and data information in the conference room to users in the form of graphics. At the same time, the interface program also has a certain data structure, realizes the complex application logic, and can achieve certain functions.

Bbb-web is a conference room management platform of bigbluebutton consultation system. When users enter the conference room, they can verify their rights through this platform. In addition, this platform receives the user's file upload request, and converts the format of the received file.

Red5 streaming media server is the core of bigbluebutton system, which is responsible for creating conference rooms, maintaining user data and forwarding streaming media data.

Freeswitch voice conference room is the core of bigbluebutton voice conference room, which is responsible for forwarding users' voices. Users must join the voice conference room before they can communicate with each other.

2.2. Design of video playing function module
In the original consultation system, pictures and documents can only be displayed on the whiteboard, but video files cannot be uploaded and played. In order to make the system function more complete,
video playback function is added on the basis of the original system. Firstly, the storage module of video file is designed. Video files need to be uploaded to the back-end server before they can be shared, but the original system did not design the function of playing video files, so there is no directory to store video files on the server side, so it is necessary to design a file storage scheme. This storage scheme can not only make the conference room accurately and conveniently locate the files, but also save the files permanently, so that users can view them when they enter the conference room next time. Based on the above two requirements, a new directory named "" is created under the directory where the conference room files are stored to store the conference room audio-visual files. In order to accurately locate the audio-visual files in the conference room, a secondary directory is established under the name of the conference room number. This directory is the final directory where the audio-visual files are stored. The conference room is uniquely identified by the conference room number, so users can easily view the history and upload audio-visual file records when entering the same conference room through the same conference room number.

After the video files are uploaded and stored in the server, the client can play them through the streaming media server. In the design of video file playing architecture, it is divided into video playing control module and video information processing module. The video playback control module can control the video file "play" and "stop". The architecture design of this function module is the same as that of the information release function module in the conference room management, which calls the client method through the streaming media server to achieve the effect of information broadcast. In addition, the module is also responsible for the video file storage directory to obtain the basic information of the video file for the client to query.

3. System implementation

3.1. Video file playing architecture

Video and audio file play refers to the video and audio files shared by users in conference room. This function module is realized by adding video player module to the client and server on the basis of clarifying the architecture and data flow of bigbluebutton conference room.

The function of playing video and audio files is realized through the cooperation of three modules: client, bigbluebutton App and playerApp. The organizational structure of the three modules is shown in Figure 2:

![Figure 2 Video file playing architecture](image)

In the client application, a new player function module is created, which is divided into four parts: View, Map, Event and SOservice. The View part is the display layer of the audio-visual playback module, which mainly includes two display interfaces: one interface is used to display the audio-visual file list, which contains the information of uploading audio-visual files in the conference room for users to choose; Another interface is the player interface, which reuses the video playing area to play audio and video files. Mate in the Map part is the event map in the framework, which connects the view part with the application logic in series by means of configuration file. The Event part contains
some user-defined events, which are unique in the system and represent a certain meaning, such as video list request event, which requests a list of audio and video files on behalf of the client. Events defined in the event section generally need to be matched with the Map configuration map. The SOservice part is the logic implementation part, the map part receives the events of the View part, searches in the configuration file, and finally the SOservice calls some methods, most of which interact with the server.

3.2. Video file encoding and decoding operation

Playerapp application is the core of video file playing function. It not only creates and forwards video stream, but also encodes and decodes different formats of video by using codec library.

Video is the dynamic effect of playing the static image continuously in the form of more than 24 frames per second, so the video is realized by the dynamic display of the picture in a certain period of time. So as long as the pictures contained in the video are stored in a certain order, the video can be saved. However, the video contains dozens of pictures per second, and a video may last several hours. In this way, the saved video will occupy a lot of storage space. Under the above background, a variety of video compression technologies, namely video coding technology, have emerged. Due to the different information structure of each video coding technology, there are different formats of video files. When playing video files, we need to choose the appropriate decompression method according to the file format, namely video decoding. Generally, the player has its own codec library. After reading the suffix and header information of the video file, the player can determine the organization format of the video file, find the appropriate decoding program, and restore the compressed video file to a static image, which is displayed to the user. In this system, there is no video codec library. In order to read different formats of video information, an open source and powerful video codec Xuggle is applied. The background server transfers video and audio images to the client through rtmp data, which is one more step than the player encoding and decoding process, that is, not only the video and audio files need to be decoded into pictures, but also the decoded images need to be compressed and encoded into the data format specified by rtmp protocol. The coding and decoding operation flow of the system is shown in Figure 3.

![Figure 3 Server code operation flow chart](image)
Xuggle codec is the core of this process. Xuggle codec is written with C language and Java language. In the bottom layer, C language is used to realize a variety of codec libraries, and in the upper layer, Java language is used to package, which provides a good interface for java development applications. In this system, we can play video and audio files in mkvb, avi, MP4, flv and other common formats.

4. System test
Software system contains a large number of data structures and complex logic implementation, which will inevitably lead to errors and defects. Software testing is to find these potential errors and defects. The purpose of testing is not to prove that the designed software system is perfect, but to find the most errors and defects with the least manpower and time. The test generally follows the principle from part to whole. In this system, the unit test of each functional module is carried out first to solve the logical and functional errors and defects of each functional module. Then, the multiple functional modules are integrated into a subsystem for integration testing, and the structural errors of the subsystem are found and eliminated. Finally, the subsystems are assembled as a whole to check whether the system meets the functional requirements.

4.1. System testing process and principle
System testing determines the quality of the system, so it is very important, generally accounting for about half of the workload. System testing should be carried out in all stages of software development, and it needs the participation of developers, testers and end users. System testing is a reciprocating process. If there is a problem in the testing, the programmer may need to modify it again. The specific process is shown in Figure 4:

![Figure 4 System test structure diagram](image)

The test principle of the system is as follows:
1. Whether the legal input gets the expected result;
2. As many test cases as possible;
3. Write illegal and unexpected test cases, whether the system processing is reasonable;
4. Legitimate input cases only test the realization of system functions, while illegal test cases are the key to discover potential errors and defects of the system;
5. When designing test cases, don't think that there are no errors or defects in the system.

4.2. System test method
1) Unit testing
Unit test is the lowest level test in the software development process. As the whole system is accumulated by units, unit test is very important in the whole software development process, which determines the basis of system implementation. Generally, a function module is divided into a unit. Unit testing is to verify the function realization and potential errors and defects of unit modules. Unit testing has the following advantages:
1. It is the lowest level of testing behavior, after finding problems or defects, programmers can
easily modify it;

(2) Unit testing is usually designed from the perspective of callers to find out whether the function implementation meets the needs of users;

(3) Unit test needs to write documents to save data for operation and maintenance;

(4) It is regressive;

The system is divided into five functional modules: shared data, video playback, monitoring access, voice conference room and conference room locking. Each functional module is tested independently to find the internal errors and potential defects of each functional module, and then integrated testing is carried out.

2) Integration testing

After the unit test is completed, each functional module can be assembled together for integration test according to the design. When each functional module carries out unit test, the data flows in a closed environment, and there is no need to interact with other modules, so some problems cannot be detected. Integrated test provides an environment for multiple functional modules to work together, which can detect the problems when each functional module cooperates. The problems that need to be paid attention to in integration testing are as follows:

(1) Whether the data flow between each functional module is normal;

(2) Whether each functional module will affect each other;

(3) Whether the combined function modules achieve the expected functions;

(4) Whether there are problems in the global data structure;

(5) Whether the integration of each function module will amplify the accumulated errors.

4.3. System function test

Video file play is divided into two parts: get video file list test and video file play test.

When the user clicks the audio-visual file list access button at the bottom of the playback window, a drop-down list will appear on the left side of the playback window, which displays the information of the audio-visual files used.

Figure 5 Get audio and video file list test

Figure 6 Video file playing test
The video file will be demonstrated in the play window area. The play window has three modes: single screen, quarter screen and nine screen. When the user needs to play the video file, double-click the file name in the video list to display the video in the selected play window. When playing video files, you can play, stop and pause. Video and audio files can be played in multiple ways at the same time.

5. Conclusion
This system is developed on the prototype conference bigbluebutton. Because bigbluebutton can not meet the requirements of enterprise application in function, it adds conference room management function module, video and audio playing function module and video monitoring access function module. And the original voice conference room is improved by adding voice conference room management system, terminal image access function and video extrapolation function.

In the video file playing function module, the user uploads the files to the back-end server, and the server establishes a directory according to the conference room number to store the video files. When playing the files, Xuugle Codec is used to encode the video files of different formats into RTMP packets, which are forwarded by red5 streaming media server. Finally, multiple users can share the video files.

References
[1] Zhang Ya, Zhou Yue, Tang Jimin. Design of Intelligent Illumination Controller Based on Lonworks Technology. Journal of Shenyang Jianzhu University (natural Science Edition), 2008
[2] Wang Xiaocheng. Discussion on System Integration Design of Multimedia Intelligent Conference Room. Neijiang Science and Technology, 2013 (01): 164-165
[3] Hong Yuan, Yin Dandan. Application and Value of Intelligent Voice Technology in Adaptive Speech Control Intelligent Conference Room. Intelligent Building, 2016, 12: 35-36
[4] Wen Xianli. Application of Mobile Platform Technology in the Design of Intelligent Conference Room Reservation System. Information Technology and Informatization, 2019, 2016, 42 (019): 30-31
[5] Niu Xin. Research on the Design of Intelligent Conference Room. Henan Science and Technology, 2018, 1: 40-41
[6] Chen Han. Environment Intelligent Control Technology of Conference Room and Conference Hall. Intelligent Building, 2008 (11): 49-52
[7] Cai Qingong. Improvement of Lighting System for Multi Function Conference Room in Hotel. Electronic World, 2020, No. 583 (01): 159-160
[8] Fan Gongrui, Huang Siqi, Liu Wensheng. Audio and Video System Design of Multi-function Hall. Intelligent Building, 2015, No. 182 (10): 71-76
[9] Zhong Guojian. Application and Development of AI in Smart Park -- Exploration and Construction of Intelligent Conference Room. China Public Security, 2019, 9: 39-41
[10] Huang Yuanyuan, Chen Chaofan, Cheng Zhigang. Design and Implementation of Intelligent Conference Room Reservation System Based on Internet of Things. Computer Knowledge and Technology, 2018, V.14 (20): 68-70
[11] Lu Xiuqiong. Small Intelligent Multimedia Conference Room Solutions. Digital Communication World, 2016, 3: 106-107
[12] Dong Shan, Zhao Haojie, Luo Rongliang. Design of Conference Room Management System Based on Wechat Applet. Electronic World, 2019, No. 577 (19): 161-162