Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

Investigation on hands-free VoIP Based on Hardware AEC

Yi Liu [1, a], Hanghang Du [2, b]

[1]Engineering Teaching Training Center, Tianjin Polytechnic University, Tianjin, China

[2] Engineering and Technology Training Center, Civil Aviation University of China, Tianjin, China

[a]344044758@qq.com, [b]hhdu@cauc.edu.cn

Abstract: Voice acquisition, compression, Acoustic Echo Cancellation (AEC), noise suppression, coding and decoding, network logic control and other tasks of the hands-free Voice over Internet Protocol (VoIP) system require high performance of the processor. Aiming at the low performance embedded system, a design scheme of VoIP based on hardware AEC was proposed. FM1288 was utilized to complete voice data acquisition, compression, AEC, noise suppression, encoding and decoding tasks. The controller was only responsible for voice data transmission and logical control tasks, which greatly reduced the requirements of controller performance and RAM. Experiments show that the scheme proposed in this paper has high integration degree, the quality of the voice communication meets the demand, and improves the integration degree of the hands-free VoIP.
Keywords: Hands-free VoIP; hardware AEC; FM1288; ARM.

---

1. INTRODUCTION

Generally voice data is transmitted in TCP/UDP mode in a hands-free Voice over Internet Protocol (VoIP) system. In order to reduce network load, voice data is sent in frames, the length of which will determine the basic delay of the VoIP system, and the network delay can contribute partially, which consequently cause the acoustic echo phenomenon [1,6,7].

In order to solve the echo problem effectively, scholars have carried out in-depth research. Basing on the traditional NLMS (Normalized Least Mean Square) algorithm, Yao et al. [1] implemented a fast algorithm FDNLMS (Normalized Least Mean Square) in frequency domain by using Fast Fourier Transformation, which greatly reduced the computational complexity and improved the real-time performance of the algorithm. Aiming at cancelling echoes effectively and suppressing howling rapidly, Liang et al. [2] proposed the Variable Step Normalized least mean square-Notch Filter and Wang et al. [3] proposed a normalized least mean square algorithm that combined with howling suppression. Zhang et al. [4] and Wang et al. [5] have designed VoIP from the functional point of view, without considering the quality of intercom and other issues.

With the gradual penetration of industrial Ethernet into various industries, the demand of VoIP based on Ethernet bus will be increasing, and its cost will be lower and lower. With respect to Hands-free VoIP, it involves voice data acquisition and compression, data transmission, acoustic echo cancellation, noise suppression, voice data decompression and playback, etc. For embedded chips with limited resources, excessive computation will lead to the reduction of real-time performance, affecting the quality of the intercom seriously.

To solve the problems above, we propose a design of hands-free VoIP based on hardware AEC. The hardware completes the tasks of voice data acquisition, compression, echo cancellation, noise suppression, data decompression and playback, which improves the real-time performance of VoIP.

2. SYSTEM DESIGN

2.1 Hardware design

Traditional VoIP hardware includes MIC gain control module, acoustic echo cancellation and noise suppression module, audio encoding and decoding module, audio power amplifier circuit, network communication module and core control circuit. This structure can accomplish the basic functions of network telephone, which is low integrated though.

The hardware block diagram of the hands-free VoIP presented in this paper is shown in Fig. 1. The design replaces the first three parts of the traditional network telephone hardware structure with a single chip of FM1288. It completes the functions of automatic gain control and band-pass filtering of MIC input signal, acoustic echo cancellation and noise suppression, and audio encoding and decoding. FM1288 collects the voice signal input by MIC and sends the near end voice data to the ARM through the I2S interface after digital signal processing. The ARM compresses the voice data and sends it to the far end through the Ethernet. Meanwhile, it receives the far end voice data. After decompression, the data is sent to FM1288 through the I2S interface. After FM1288 decoding, the data is converted into analog signal and sent to the power amplifier circuit to play. ARM sends configuration parameters to FM1288 through Universal Asynchronous Receiver/ Transmitter (UART) interface.

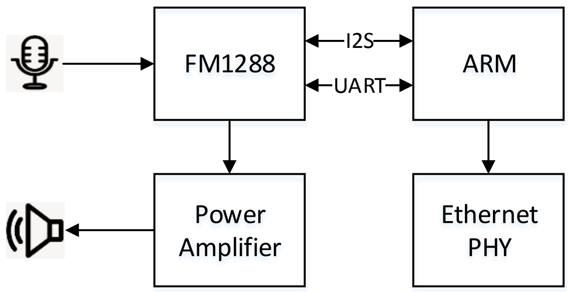

Fig. 1 Hardware structure diagram of hands-free VoIP

FM1288 is a new generation of Fortemedia products, which is utilized to eliminate environmental noise and acoustic echoes and facilitate product integration. The internal audio channel is divided into two parts, MIC to LINE-OUT and LINE-IN to SPK. The former consists of MIC power gain amplifier, mute control, low pass filter, pre-emphasis filter, gain amplifier, noise suppression unit, non-linear AEC, noise paste back unit, volume control unit, AGC, high pass filter and LINE-OUT power amplifier. The latter consists of LINE-IN power amplifier, far end noise suppression, bright voice, SPK high pass filter, AGC, dynamic range control, Volume control unit, mute control, power amplifier.

The ARM uses STM32F407 with CORTEX-M4 core. Its maximum CPU frequency is up to 168MHz, with three SPI interfaces and two extended I2S modules, which can be used for full-duplex communication of voice data. The core of the Ethernet circuit is a 10/100 Mbit/s adaptive Ethernet chip LAN8720, and the connector is RJ45 with internal Ethernet transformer. LM386N audio power amplifier is used in power amplifier circuit. The voltage gain is as high as 46dB. When driving 8Ω loudspeaker, the maximum power can be 700 mW, which meets the system requirements. The connection diagram of FM1288, ARM and Ethernet circuit is shown in Fig. 2.

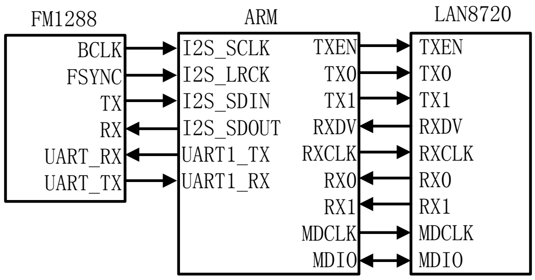

Fig. 2 Hardware connection diagram

2.2 Software Design

The flow chart of the system software is shown in Fig. 3. The system takes the ARM processor as the core, initializes the hardware resources of the system after power-on, including clock system, I2S interface, UART, Ethernet MAC interface, etc., and then initializes LAN8720 to set network parameters. By sending configuration information to FM1288 through UART, FM1288 works normally. Called and calling tasks are processed in the main loop. Near end voice is sent to the target through Ethernet, and the target voice is played synchronously.

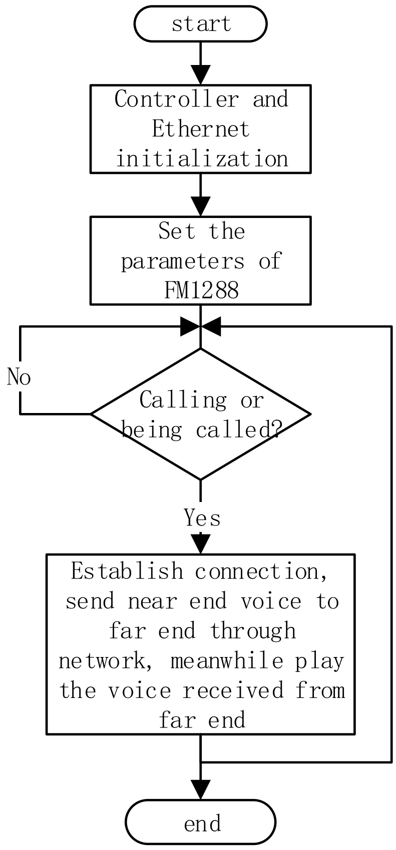

Fig. 3 Software flow chart

There are many registers in FM1288. The main parameters are shown in Table 1, which means to enable MIC0, MIC1, MIC bias voltage output, speaker of FM1288. FM1288 is set to work in I2S master mode, using PCM A-Law encoding. MIC dynamic range control, idle noise suppression, pre-emphasis, 180Hz high pass filter, linear AEC are all enabled. For the stability of audio power amplifier, input volume and speaker volume gain are 0. The gain of AEC reference channel is - 3dB, and the convergence rate of echo cancellation is set to 0x1FFE to accelerate the convergence rate. The ambient noise rejection remains the default, the estimated length of the echo tail coverage is 96 ms, and the idle noise rejection is - 4 dB.

Table 1 Main parameters of FM1288

Address	Value
0x22F9	0x619F
0x22FA	0x010F
0x2301	0x0012
0x2303	0x4DF1
0x2304	0x03CF
0x2305	0x0231
0x2307	0x0000
0x230C	0x0100
0x230D	0x0100
0x232F	0x0080
0x2337	0x1FFE
0x23B6	0x0009
0x23B7	0x0006
0x23BA	0x3000
0x23EE	0x32F5

FM1288 can automatically complete the configuration at startup by reading the configuration information in EEPROM, which is omitted in order to simplify the system hardware. ARM initializes FM1288 by sending configuration commands through UART interface after power-on. The communication format is shown in Table 2, among which the synchronization word is fixed as "0xFCF3" and the command word "0x3B" is the write register.

Table 2 UART Configuration Communication Protocol for FM1288

Synchronous word	Command word	Registers' address	Data Domain
2 bytes	1 byte	2 bytes	2 bytes

When the VoIP is used in the same LAN, in view of the low packet loss rate, in order to ensure the real-time performance of voice data and reduce the network bandwidth occupancy rate, UDP mode is adopted. When the VoIP is working in the point-to-point mode, voice data is transmitted through the point-to-point mode; when the VoIP needs emergency voice broadcasting, voice data is broadcasted. The feature is suitable for applications with security alarm requirements.

3. RESULTS AND DISCUSSION

In the laboratory environment, the performance of the hands-free VoIP is tested by using a 100 MHz bandwidth oscilloscope and a laptop with Intel i5 processor. The experiments include: network bandwidth test, packet loss rate test, voice data transmission accuracy test and acoustic echo cancellation evaluation.

The hands-free VoIP is connected with laptop by HSYV. The speed of network connection is 100.0 Mbps, which proves that the network state of the hands-free VoIP is normal and works in 100 Mbps mode. Using the ping command of the command line in WINDOWS to test the packet loss rate. The test results show that the round trip time is less than 1 ms and the packet loss rate is 0%, which indicates that the communication status of the Ethernet module of the hands-free VoIP is normal.

Using MATLAB 2016b to produce a sinusoidal sound with a frequency of 840Hz, and the MIC of hands-free VoIP A to collect the audio signal, then the audio data is sent to hands-free VoIP B through Ethernet, and is played through its loudspeaker. The oscilloscope is used to observe the MIC input signal of hands-free VoIP A and the speaker output signal of hands-free VoIP B respectively, as is shown in Fig. 4(waveforms of Label 1 and Label 2). It can be seen that the frequency of the two signals is the same and the voice is clear and distinguishable. The waveform is partially distorted due to the high frequency interference of switching power supply and the absence of shielding layer on MIC conductor.

The voice is sent to the hands-free VoIP B through the hands-free VoIP A, which are all in full duplex mode. The distance between the loudspeaker and MIC is 30 centimeters, and the angle between them is 90 degrees. No echo is heard in the loudspeaker of A, which proves that the hands-free VoIP system is with the function of acoustic echo cancellation in line with practical application.

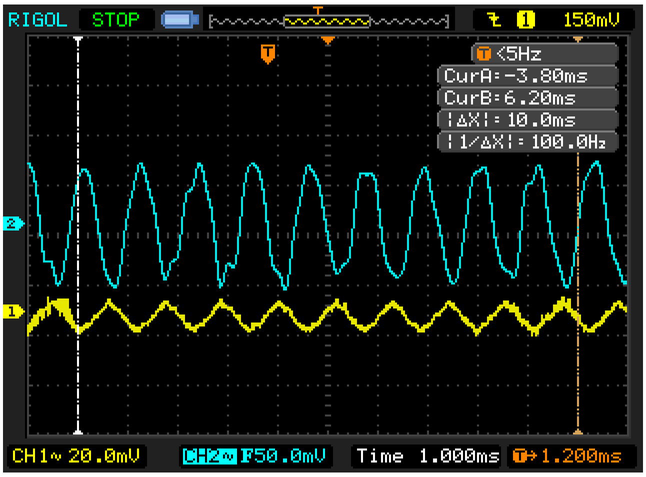

Fig. 4 Hands-free VoIP audio transmission test

4. CONCLUSION

Aiming at the problem that VoIP based on pure software algorithm requires high performance of embedded controller, a design of hands-free VoIP based on hardware AEC was proposed, which transfers the task of improving voice performance to FM1288, which greatly reduced the requirement of speed and RAM of the controller. Experiments show that the hands-free VoIP can complete the basic functions of voice communication and guarantee the quality of the call. On the premise of quality, it improves the integration of the system and provides a new hardware design reference for the telephone system based on LAN.

ACKNOWLEDGEMENTS

This paper was supported by the Fundamental Research Business Fee Project of Central University (3122018S001).

REFERENCES

[1] Yao Yueqin, Yang Yan, Chen Lin and others. Echo cancellation algorithm in VoIP phones [J]. Chinese Journal of Electron Devices, 2018, 41(06):1618-1621.

[2] Liang Ruiyu, Wang Xia, Wang Qingyun, et al. An echo cancellation algorithm for whistle rapid suppression in hearing aids [J]. Acta Acustica, 2016, 41(02):249-259.

[3] Wang Guozheng, Liu Quanli. Echo control algorithm with howling suppression on metro intercom system and its application [J].Control Theory & Applications, 2018, 35(04):539-546.

[4] Zhang Haichun, Gao Wei. Network Telephone communicate system base on VxWorks [J]. Microcomputer Information, 2008, 24(35):62-64.

[5] Wang Shiyan, Miao Changyun, Liu Xiaoxue, etc. Research on Network Telephone of Belt Conveyor Monitoring System [J]. Instrument Technique and Sensor, 2018(10):55-58.

[6] Huang F, Zhang J, Zhang S. Affine Projection Versoria Algorithm for Robust Adaptive Echo Cancellation in Hands-Free Voice Communications [J]. IEEE Transactions on Vehicular Technology, 2018:1-1.

[7] Sunitha T, Suja M M R. Nonlinear Acoustic Echo Cancellation Based on Multichannel Adaptive Filters: A Novel Approach [J] Wireless Personal Communications, 2018.