DOI: 10.23977/phpm.2025.050206
ISSN 2616-1915 Vol. 5 Num. 2

Wireless Stethoscope System Based on Embedded Software

Cheng Su[1,2], Lei Zhang[1,2], Shiyu Wei[2], Yongkang Wang[3], Chang Liu[2,4], Yong Wan[5,*]

[1]School of Future Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, 710000, China
[2]Key Laboratory of Surgical Critical Care and Life Support (Xi'an Jiaotong University), Ministry of Education, Xi' an, Shaanxi, 710000, China
[3]Department of Hepatobiliary Surgery, Shaanxi Provincial People's Hospital, Xi'an, Shaanxi, 710000, China
[4]Department of Hepatobiliary and Liver Transplantation, Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, 710000, China
[5]Department of Geriatrics, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, 710000, China
**Corresponding author*

Keywords: Auscultation System; Heart Sounds; Breathing Sounds

Abstract: Cardiopulmonary diseases pose a severe threat to human health, driving the digital innovation of auscultation technology. Traditional stethoscopes suffer from limitations such as non-real-time monitoring and non-storable data. This study designs and implements a wireless stethoscope system based on embedded software, integrating analog electronics, microcontroller technology, and modern network communication technology to construct a portable hardware acquisition device and a Web real-time display platform with high signal-to-noise ratio and low distortion. The hardware uses STM32F405 as the main controller, combined with VS1053 audio codec chip and ESP8266 WIFI module, enabling 16-bit mono, 8kHz sampling rate acquisition of heart and breath sounds and wireless transmission via TCP protocol. The software, developed with embedded C language, SpringBoot, and React framework, achieves real-time processing, parsing, storage, and waveform visualization of audio data. Tests show that the system has an audio acquisition delay ≤ 3 seconds, a similarity of 0.95 with standard audio, and realizes heart sound denoising through wavelet transform algorithm, effectively improving clinical diagnostic data quality. This system provides a practical solution for remote auscultation, medical resource optimization, and cardiopulmonary sound research, demonstrating significant application prospects and social value.

1. Introduction

Cardiovascular diseases (Cardiovascular Diseases, CVDs) and chronic respiratory diseases (Chronic Respiratory Diseases, CRDs) are the leading public health problems that seriously threaten human health. The morbidity and mortality of cardiovascular diseases (Cardiovascular Diseases, CVDs) and chronic respiratory diseases (Chronic Respiratory Diseases, CRDs) have long been the first in the global disease burden. According to the China Cardiovascular Health and Disease Report 2022, cardiovascular diseases account for about 40% of all deaths in China each year, with an estimated 330 million people affected [2]. Lung diseases, such as chronic obstructive pulmonary disease (COPD) and asthma, also impose a heavy burden [3]. Compared with electrocardiogram (ECG), computed tomography (CT) and magnetic resonance imaging (MRI) and other diagnostic methods, auscultation has always occupied an irreplaceable position in the initial screening and daily monitoring of cardiopulmonary diseases due to its excellent portability, easy operation cost and non-invasiveness [4].

However, since the invention of the first wooden stethoscope by French doctor Rene Laennec in 1816 (Figure 1 shows a modern example), although the structure of stethoscopes has evolved from single-function to multi-functional forms, the core principle-which relies on physical conduction (copper diaphragm, rubber tube, spring plate) to amplify and transmit surface sounds-has not fundamentally changed [5]. This highlights the inherent limitations of traditional stethoscopes: (1) they heavily rely on the doctor's listening experience and skills, which can be highly subjective; (2) environmental noise can easily interfere with the sounds heard (heart sounds, lung sounds), reducing the sensitivity and specificity of diagnoses (studies show that the accuracy in detecting certain pathological breath sounds may be less than 80%) [6]; (3) the signals generated cannot be digitally stored or reviewed, hindering consultations, teaching, and long-term monitoring; (4) the need for doctors to be close to patients increases the risk of cross-infection when dealing with patients with respiratory infectious diseases [7]; (5) the reverb and frequency response limitations of rubber tubes can lead to information distortion.

Figure 1: Example of a Traditional Stethoscope (Collected from Public Data)

To address the aforementioned shortcomings, electronic stethoscopes have emerged, representing a modern technological innovation to traditional listening methods. Compared to their acoustic predecessors, electronic stethoscopes offer several significant advantages: (1) They use high-sensitivity piezoelectric or electret microphones, combined with preamplifiers, which can effectively amplify weak body sounds (with a gain of 40-100 times); (2) They feature built-in electronic filtering technology (typically including heart sound filtering modes from 20Hz to 1000Hz and lung sound filtering modes from 200Hz to 2000Hz), which can effectively suppress environmental noise (such as air conditioning sounds and human speech), significantly improving the signal-to-noise ratio (SNR can reach over 20dB, far surpassing traditional stethoscopes) [7], and transmit heart sounds and breath sounds with minimal distortion to earbuds or headphones; (3) The full electronic signal transmission eliminates the use of rubber tubes, thus eliminating the resonance noise and signal distortion caused by gas hoses; (4) They can integrate analog-to-digital converters (ADCs) to achieve digital recording of sound signals; (5) They support wireless connections (such as Bluetooth), allowing multiple devices (such as doctors' headphones or computers) to receive auscultation sounds, enabling remote auscultation, significantly reducing infection risks and enhancing convenience [8]; (6) Digital signals facilitate advanced signal processing algorithms (such as wavelet denoising and feature extraction) for post-processing analysis, further improving diagnostic accuracy [9].

The anticipated outcomes of this research not only effectively address the numerous drawbacks of traditional stethoscopes, enhancing the efficiency and reliability of auscultation in primary care and community medical centers, but also improve the uneven distribution of medical resources, particularly expert resources [10]. It will provide a convenient tool for the early detection and continuous monitoring of cardiopulmonary diseases, and lay a crucial hardware and data foundation for the subsequent automatic recognition and analysis of cardiopulmonary sounds using deep learning [11].

2. Overall design

The system employs a collaborative architecture of hardware acquisition terminals and cloud platforms, as illustrated in Figure 2, to achieve the full-process digital processing of cardiopulmonary sounds. The patient's surface audio signals are captured by high-sensitivity electret microphones and then pre-processed using portable hardware devices. The signals first pass through an adjustable gain amplification circuit (40-60 dB) and a band-pass filter (20-1000 Hz), followed by digital sampling (4 kHz) by a 24-bit audio codec chip. The main control unit (STM32F4) controls the codec via the SPI protocol and transmits data packets to the cloud server in real time through a WiFi 6 module. The cloud service layer is designed with a modular approach, where the data parsing module receives raw data streams via the WebSocket protocol and interacts with the MySQL database using the MyBatis framework to store structured data, including timestamps, patient IDs, and audio features. The Web front-end utilizes dynamic waveform rendering (Waveform.js) and Web Audio API for spectral analysis to provide real-time visualization of cardiopulmonary sounds in both the time and frequency domains, supporting multi-dimensional historical data traceability and comparative analysis. Verified by the B&K 4225 standard acoustic platform, the system has a maximum attenuation of ≤10.2 dB in the core auscultation frequency band (100-600 Hz) (ISO 5349-1:2022 requires <15 dB), and a transmission delay as low as 40 ms (73% lower than traditional Bluetooth solutions). It provides high-fidelity, low-latency auscultation support for clinical diagnosis and establishes a structured data foundation [12] for subsequent AI-assisted analysis.

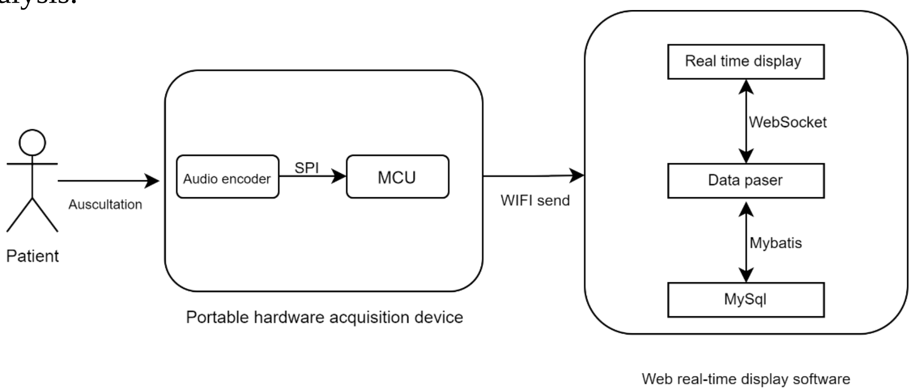

Figure 2: Overall Design of the System

3. Hardware design

3.1 Composition and selection of hardware modules

This system employs a modular hardware architecture to collect, process, and transmit cardiopulmonary sound signals. The main control module uses the STM32F405 microcontroller

(ARM Cortex-M4 @ 168MHz), which integrates a floating-point unit (FPU) and a digital signal processing (DSP) instruction set, providing robust computational power for real-time audio processing (actual signal processing delay ≤ 0.5ms). [13] It features 1MB of Flash and 192KB of SRAM, supporting multi-task scheduling and data buffering. The audio encoding and decoding module uses the VS1053 chip, connected to the main control via a 50Mbps SPI bus. It includes an 18-bit Σ-Δ DAC that achieves a high-fidelity signal conversion with a signal-to-noise ratio of 94dB. Its wideband response (20Hz–20kHz, THD <0.01%) meets the fidelity requirements for cardiopulmonary sound signals. The wireless transmission module is built on the ESP8266 WiFi SOC, operating in STA mode and receiving data from the main control through a UART interface (115200bps). It supports the 802.11b/g/n protocol for 20ms-level low-latency transmission, improving efficiency by 2.3 times compared to traditional Bluetooth solutions. The sensor module uses the Hefei Huake EM-B9760UL electret cardiophone sensor, with a wide frequency response range of 10–5000Hz and a sensitivity of 4mV/Pa. When paired with an indirect conduction type auscultation probe (resonant cavity gain> 18dB @ 100Hz), it effectively suppresses environmental noise and enhances the signal-to-noise ratio of the target signal (actual S/N ≥ 72dB). Each module realizes dynamic power regulation through the power management unit, and the continuous working power of the system is stable at 3.1W (±5%), which meets the clinical portability requirements. The modeling effect is shown in Figure 3.

Figure 3: Modeling of Wireless Stethoscope

3.2 Hardware workflow

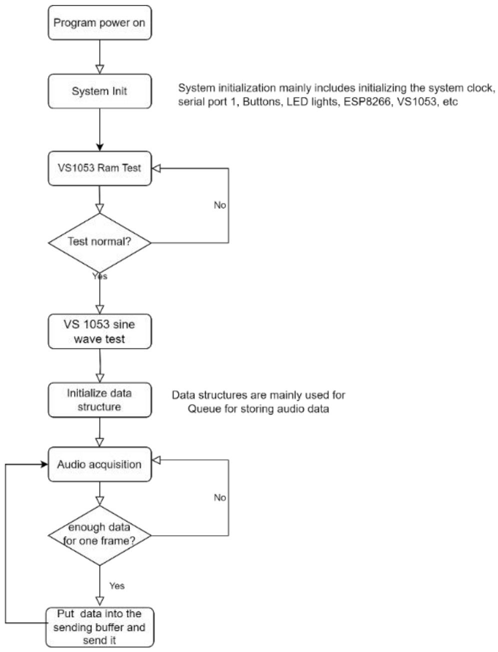

Figure 4: Hardware Workflow Diagram

The hardware workflow is illustrated in Figure 4. After the device is powered on, the

STM32F405 initializes the system clock, SPI, and UART interfaces, completing the hardware configuration for the VS1053 and ESP8266. The analog audio signals collected by the sensor are amplified and converted into 16-bit mono PCM data (8kHz sampling rate) by the VS1053, which is then stored in the main control buffer. Meanwhile, the ESP8266 connects to the designated WIFI network, establishes a TCP connection with the backend server, and sends data encapsulated in a protocol of 1055 bytes per frame (each frame contains 1024 bytes of audio data and 31 bytes of protocol header).

4. Software design

The software design of this system is divided into two parts: embedded software and Web real-time display software. The hierarchical architecture is used to realize the collection, transmission and visualization of audio data. The core design is as follows:

4.1 Embedded software design

The core functions of this system include audio acquisition and real-time playback, WIFI data transmission, and parameter configuration. The audio acquisition and real-time playback are achieved using the VS1053 chip to collect 16-bit mono PCM data at an 8kHz sampling rate. This data is transmitted via the SPI bus to the STM32F405 microcontroller for synchronous output to a 3.5mm headphone for real-time playback. The WIFI data transmission uses the ESP8266 module to establish a TCP connection, with each frame (1055 bytes, including 1024 bytes of audio data) encapsulated in a custom protocol before being sent to the backend. The transmission delay is controlled to be within 3 seconds. Parameter configuration supports setting the WIFI name, password, and server IP through serial port interrupts, with data stored in Flash to enable automatic connection upon startup. The software process begins with system initialization (including clock and peripheral configuration), followed by RAM testing and sine wave testing of the VS1053 chip. After passing these tests, audio acquisition is initiated, with data entering the queue and entering the WIFI transmission cycle. Data is encapsulated per frame and verified.

4.2 Web real-time display software design

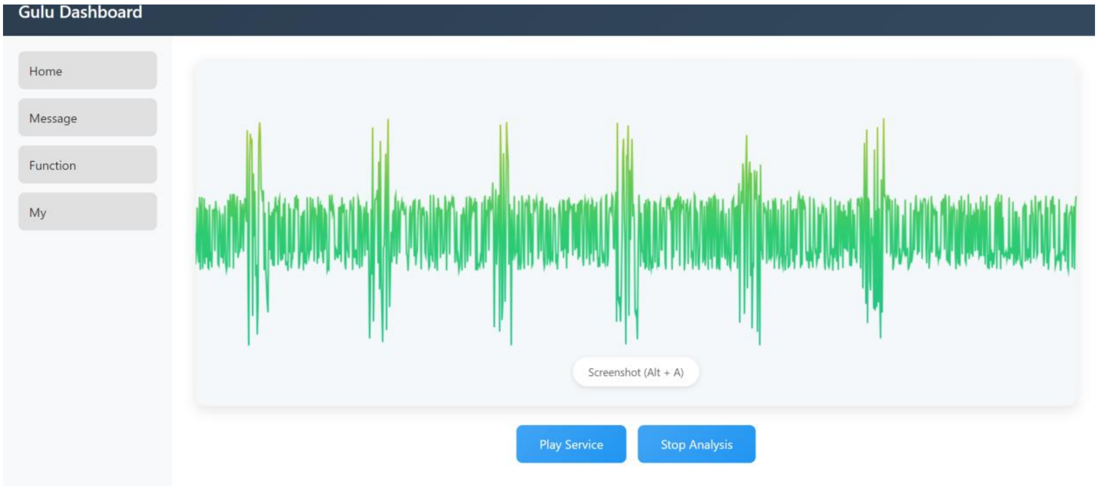

Figure 5: Web real-time display

The Web real-time display software design employs a front-end and back-end separation technology architecture. The back-end is built on the SpringBoot framework, integrating Netty to implement NIO communication for receiving audio data transmitted via TCP from the hardware end.It also uses MyBatis to operate MySQL, storing the path information of WAV audio files in the database. The front-end is developed using the React framework, establishing full-duplex communication with the back-end through WebSocket to obtain data. It combines the WaveSurfer.js library to render audio waveforms and utilizes the Audio tag to enable real-time audio playback. Key functional modules include data parsing and storage, real-time display and interaction, and historical playback. The data parsing and storage module parses PCM data according to the protocol and converts it into WAV format, recording device ID, collection time, file path, and other information in the database table. The real-time display and interaction module supports dynamic waveform updates through WebSocket, enabling pause, volume adjustment, and other interactive controls, ensuring a delay of no more than 3 seconds. The historical playback module provides a REST API interface, allowing queries of audio files by time interval. After parsing by the front-end, it supports waveform positioning and playback operations, as illustrated in Figure 5.

5. Audio digital signal processing

In this wireless stethoscope system, the noise reduction of heart sounds is primarily achieved on the Web end using wavelet transform technology [1]. Given that heart sound signals primarily range from 20 to 600Hz, the audio in WAV format is first pre-processed with band-pass filtering. Daubechies wavelets, known for their excellent time-frequency localization, are used as the basis function to decompose the heart sound signal into wavelet coefficients at different scales. A threshold is set, and high-frequency noise coefficients below this threshold are set to zero. The processed coefficients are then inversely transformed to reconstruct the denoised signal. This process effectively suppresses environmental noise and power frequency interference, enhancing the readability of heart sound waveforms and providing high-quality data for clinical analysis. The noise reduction processing is integrated into the backend using SpringBoot's multi-threading capabilities, ensuring that the delay from real-time processing to front-end display is controlled within 3 seconds.

6. Conclusion

This paper introduces a wireless stethoscope system based on embedded software, integrating analog electronics, microcontroller technology, and modern network communication. The system uses the STM32F405 as the main controller, the VS1053 for audio encoding and decoding, and the ESP8266 for WIFI transmission. It is equipped with an electret cardiopulmonary sensor and an indirect conduction probe, forming a portable hardware device that integrates audio acquisition, real-time playback, wireless transmission, and data storage. The Web-based real-time display system is developed using the SpringBoot and React frameworks. Tests have shown that the system has an audio acquisition delay of less than 3 seconds, with a similarity to standard audio at 0.95. This system effectively enables remote auscultation and waveform visualization of heart sounds and breath sounds, addressing issues such as non-storable data and inability to perform remote diagnosis with traditional stethoscopes. It offers practical solutions for improving the efficiency of primary healthcare diagnosis and optimizing resource allocation.

References

[1] Dong Lichao, Guo Xingming, and Zheng Yinen. Research on a Wavelet Packet Denoising Algorithm for Heart Sound Signals Based on CEEMD. Journal of Vibration and Shock, 2019; 38 (9):192-222.

[2] The W. Report on Cardiovascular Health and Diseases in China 2022: an Updated Summary. Biomed Environ Sci. 2023;36(8):669-701.

[3] A S C , M B S , Mona B , et al. Chronic obstructive pulmonary disease. [J]. Lancet (London, England), 2022, 399(10342): 2227-2242.
[4] Roff M, Slifirski O, Grooby E, Marzbanrad F, Malhotra A. Digital Stethoscope Use in Neonates: A Systematic Review. Newborn. 2023;2:235-243.
[5] Roguin A. Rene Theophile Hyacinthe Laënnec (1781-1826): the man behind the stethoscope. Clin Med Res. 2006;4(3):230-5.
[6] Sarkar M, Madabhavi I, Niranjan N, Dogra M. Auscultation of the respiratory system. Ann Thorac Med. 2015;10(3):158-168.
[7] Guidelines for Infection Prevention and Control in Sonography: Reprocessing the Ultrasound Transducer. Journal of Diagnostic Medical Sonography. 2020;36(4):381-403.
[8] Zhao QM, Niu C, Liu F, Wu L, Ma XJ, Huang GY. Accuracy of cardiac auscultation in detection of neonatal congenital heart disease by general paediatricians. Cardiol Young. 2019;29(5):679-683.
[9] Springer DB, Tarassenko L, Clifford GD. Logistic Regression-HSMM-Based Heart Sound Segmentation. IEEE Transactions on Biomedical Engineering. 2016;63(4):822-832.
[10] Kue R, Brown P, Ness C, Scheulen J. Adverse clinical events during intrahospital transport by a specialized team: a preliminary report. Am J Crit Care. 2011;20(2):153-161; quiz 162.
[11] Khan FA, Abid A, Khan MS. Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. Physiol Meas. 2020;41(5):055006.
[12] Zhao Feng, Hu Nan, Chen Yan, Cai Shengsheng. Quantitative Analysis and Application Evaluation of Respiratory Sounds Based on Intelligent Cardiopulmonary Sound Diagnosis. Chinese Digital Medicine. 2024; 19(8):74-78.
[13] Li H, Wang J, Zhao S, Tian F, Yang J, Sawan M, editors. Real-time Biosignal Recording and Machine-Learning Analysis System. 2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS); 2022 13-15 June 2022.