Scholar Publishing Group
International Journal of World Medicine
https://doi.org/10.38007/IJWM. 2026.070103.
ISSN 2789-7133 Vol. 7, Issue 1: 18-27

AI traditional Chinese medicine Assistant Based on MCU

Baizhen Zhu

School of Electronics and Information Engineering, Ningbo University of Technology, Ningbo, China

Keywords: Traditional Chinese Medicine, Artificial Intelligence, Microcontroller

Abstract: To address TCM decoction challenges including error-prone manual dosage/frequency recording and medication psychological burden caused by complex operational procedures, this study designs an AI medical assistant based on STM32F407 MCU and ESP32 module with integrated voice recognition technology. The system uses STM32F407's high-performance processing core, integrates WiFi for real-time cloud data upload, and Bluetooth for mobile terminal synchronization. It features voice recognition for hands-free dosage input, voice prompts for medication guidance, buzzer alarm, button input, and LCD display. The built-in AI assistant employs a golden spiral microphone array to optimize audio recognition, reducing Word Error Rate (WER) by3.75% compared to traditional uniform circular arrays at0.50-4.00m, thereby enhancing medication management intelligence and user-friendliness through intuitive voice interaction.

1. Introduction

AI, IoT, and TCM preparation technology convergence advances intelligent medical devices reshaping traditional health management[1-2]. In home medication scenarios, context-aware medical assistants are critical[3-5]. Current AI health assistants have two forms:

(1) E-commerce-embedded general health Q&A robots with commercial integration but limited clinical process depth[6-8];

(2) Platform-based assistants (e.g., Alexa) offering basic health queries via cloud NLP, yet consumer-grade hardware struggles with far-field pickup in noisy homes. This study adopts a distinct technical and functional pathway (Table 1).

Home TCM decoction requires precise liquid proportion control for efficacy, yet traditional experience-based methods risk errors for elderly patients[9]. This study designs an intelligent decoction assistant integrating Kalman filter algorithm and golden spiral microphone array. Given AI voice interaction's healthcare potential—affecting acceptance[10], reshaping doctor-patient interactions[11], and improving experience via familiar sounds [12]—the system emphasizes natural voice interaction to enhance accuracy, safety, and user experience.

Table 1. Comparison between Alexa and self-designed AI medical assistant

Contrast dimension	Alexa Medical Assistant	Self-designed AI medical assistant
Technical form	Software-service led, relying on consumer hardware and cloud computing	Hardware-software integrated system (STM32F407/ESP32 + DeepSeek LLM)
Core functions	General health Q&A, medication reminders, appointment management (broad but shallow)	TCM decoction automation with pressure sensing, heating control, voice interaction, and closed-loop medication reminders
Interactive innovation	Standard arrays with limited far-field pickup in noisy home environments	Golden spiral non-uniform array optimizing physical-layer acoustics for long-distance recognition/anti-interference
Scene pertinence	Broad scenario coverage with insufficient single-scenario optimization	Focused on TCM decoction, solving water-drug ratio monitoring and process control

2. Overall Design of System

The system integrates pressure sensing, heating control, voice interaction, and medication reminders into a closed-loop TCM decoction process, addressing key limitations through two innovations: a golden spiral-based non-uniform microphone array improves acoustic interaction for reliable far-field voice recognition in home environments, and focused decoction control enables real-time water-herb ratio monitoring to overcome traditional experience-based errors. A cantilever pressure sensor(uses Kalman filtering to reduce thermal interference and track decoction mass changes, with data transmitted via Wi-Fi to a local server. The golden-spiral array ensures clear voice capture via golden angle spacing and non-uniform layout, enhancing sound localization and noise suppression. This integrated solution reduces herbal waste, ineffective decoction, and burning risks, improving home safety, efficacy, and user experience.

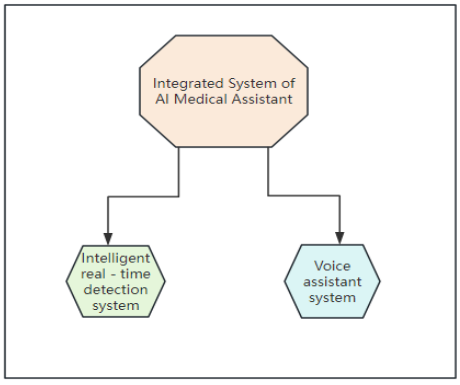

Figure 1. Overall design drawing of the system

To address accessibility limitations of professional services in home medical scenarios, this study proposes an intelligent voice medical assistant system based on a golden spiral microphone array. The solution integrates advanced acoustic design with AI technology to develop a cost-effective home medical terminal: a 10-microphone golden spiral array enables 360° omnidirectional sound coverage and voice enhancement at the physical layer, with its golden ratio structure significantly improving

effective pickup radius within 4 meters for reliable voice interaction in complex home acoustic environments. At the algorithmic layer, the "Xiaozhi" terminal, built on the Deepseek deep learning framework, integrates voiceprint recognition, emotion computing, and multi-turn dialogue technologies to achieve precise semantic analysis for medical scenarios. The overall framework is shown in Figure 1.

3. Pressure Detection System Design

Hardware Design as shown in Figure 2.

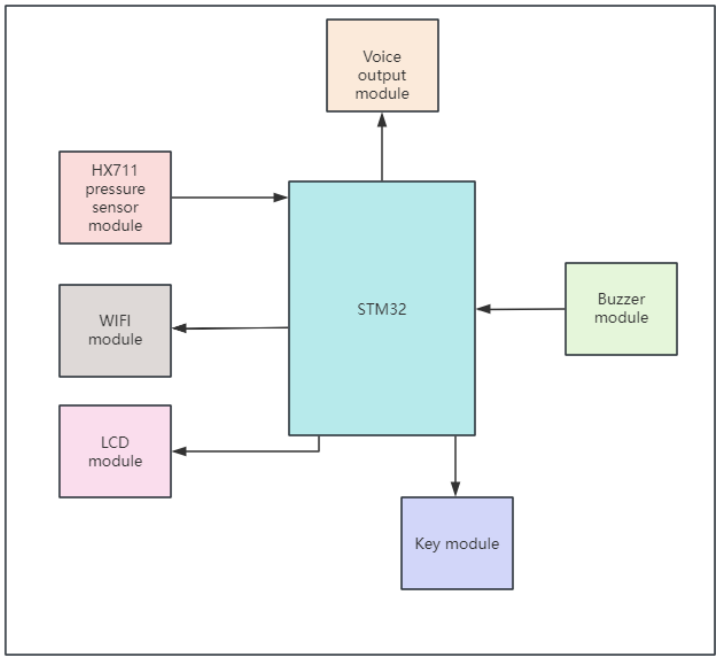

Figure 2. Hardware design system

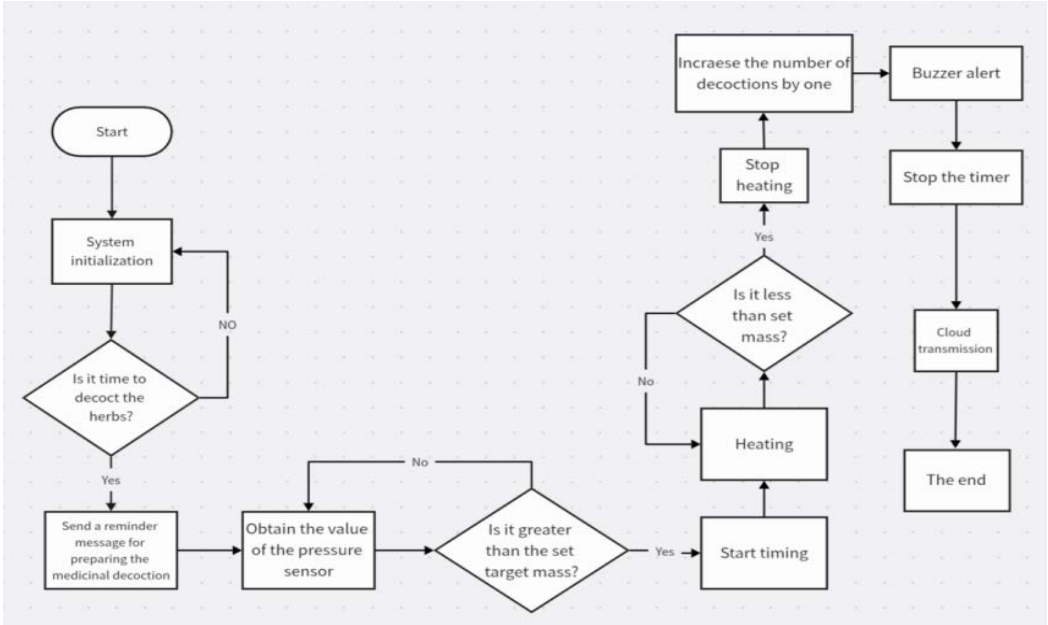

Figure 3. Software flow chart

The system uses an HX711 high-precision pressure sensor for intelligent decoction monitoring (Fig. 3). The HX711 bridge sensor captures decoction mass signals via low-noise programmable amplifier and 24-bit ADC. MCU filters/compensates data, transmits to cloud via WiFi, and displays key parameters on LCD. Software flow includes: system initialization, decoction-time judgment with reminders, pressure-based heating control/timing, cycle counting with buzzer alarms, and cloud upload. Core logic automates heating control, timing, multi-cycle monitoring, and data recording via pressure feedback.

3. Optimization of Pressure Detection System

3.1 Modeling of Problem

In high-temperature/pressure TCM decoction, boiling-induced vibration transmits to the pressure sensor and introduces high-frequency noise into the mass signal. This structural vibration often exceeds the compensation range of conventional filters, complicating the separation of mechanical disturbances from the actual mass signal, which can result in deviations that may trigger false "boiling-overflow" or "dry-burning" alarms.

3.2 Kalman Filter Design

The main idea is to reduce the sensor acceptance error caused by noise and shaking in the process of decoction by Kalman filtering:

System modeling Let the true mass be xThe measured value is z_k, the system equation is:

Equation of State (considering Mass Change Dynamics):

$$x_k(\text{True value at this time}) = x_{k-1}(\text{Last True}) + u_k + w_k \tag{1}$$

u_k:Theoretical estimates of evaporation/loss mass (based on energy conservation model)

$w_k \sim \mathcal{N}(0, Q)$:Process noise (evaporation model uncertainty)

Observation equation (including measurement noise):

$$z_k = x_k + v_k \quad \Rightarrow \quad v_k \sim \mathcal{N}(0, R) \tag{2}$$

v_k: Measurement noise (including gas jitter interference)

Among z_k For the first k Secondary measurements,R Is the measurement noise covariance.

Kalman filter recursive formula

The prediction step (prior estimation) has constant quality, and the predicted value and covariance are unchanged:

$$\hat{x}_k^- = \hat{x}_{k-1} + u_k \tag{3}$$

$$P_k^-(\text{Initial uncertainty at current time}) =$$

$$P_{k-1}(\text{Last moment uncertainty}) + \text{Update step (posterior estimation)} \tag{4}$$

Kalman gain:$K_k = \frac{P_k^-}{P_k^- + R}$

Filtered quality:$\hat{x}_k = \hat{x}_k^- + K_k(z_k - \hat{x}_k^-)$

Core Correlation

$$\hat{x}_k = (1 - K_k)\hat{x}_k^- + K_k z_k \tag{5}$$

Filtered value $= K_k \times$ Current measured value $+(1 - K_k) \times$ Predicted value. $K_k \in [0,1]$ Dynamically adjust the weights of measured and predicted values.

3.3 Experimental Results

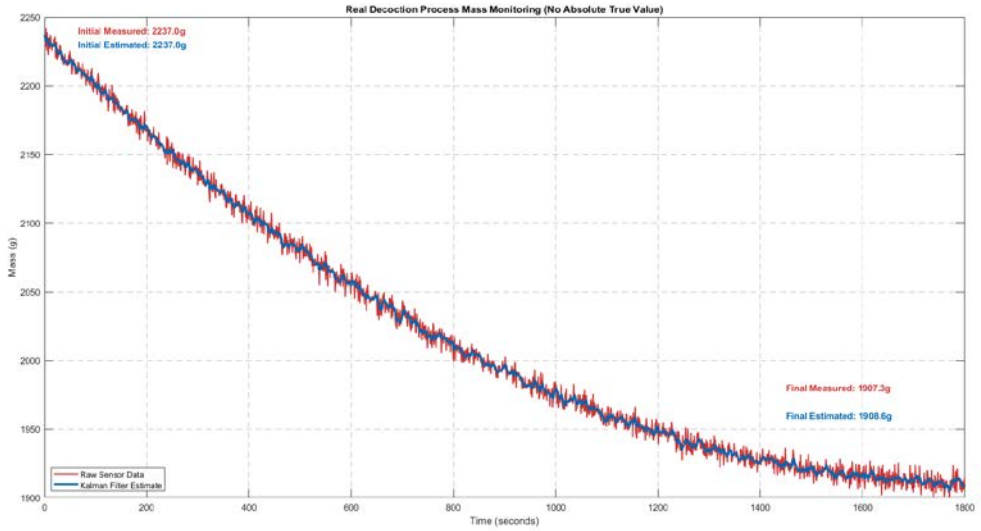

Figure 4. Test Results of Experiment 1

Figure 4 shows original measurement data (red) and Kalman-filtered data (blue).

4. AI Voice Assistant Design

4.1 Hardware Design

The core control module uses Lexin ESP32-S3 (240MHz, 8MB PSRAM, 16MB Flash) supporting Wi-Fi/Bluetooth dual-mode for real-time voice processing (Fig. 5).

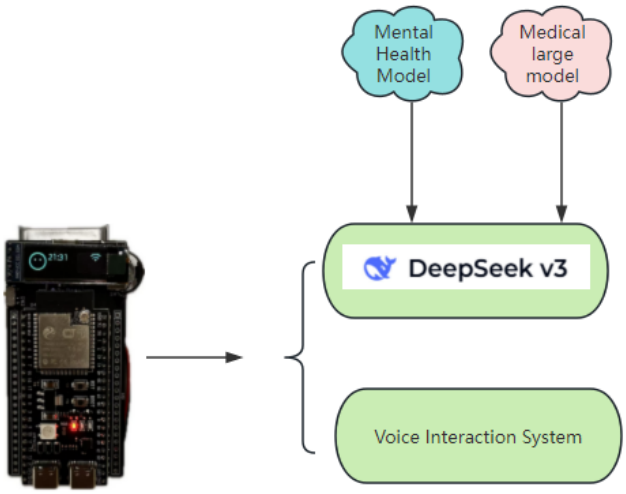

Figure 5. Schematic diagram of AI voice assistant

Expansion modules include microphone array, speaker, OLED/LCD display (dialogue/expression), and optional temperature/humidity sensors or cameras for environmental awareness (Figs. 5-6).

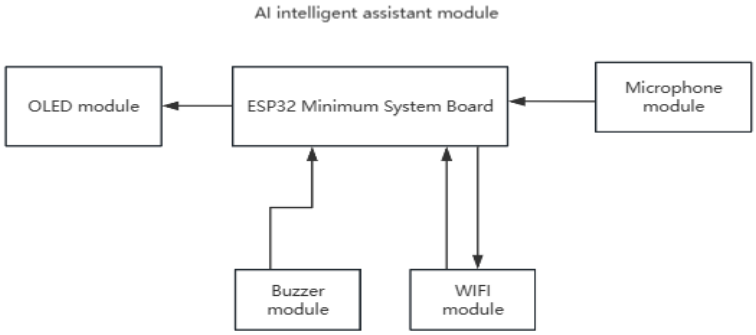

Figure 6. Hardware design system diagram

4.2 Software Design

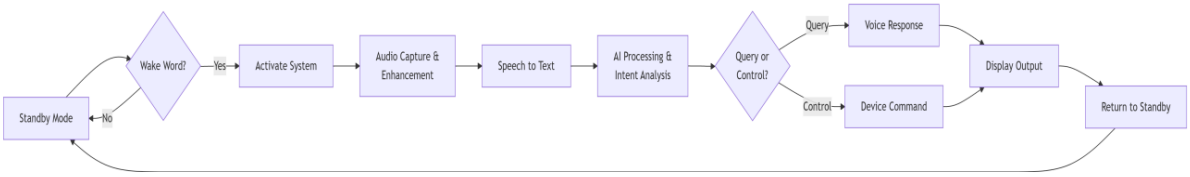

Figure 7. Software flow design diagram

Fig. 7 shows the voice assistant workflow: low-power standby monitors wake-word; upon detection, it activates, captures/enhances audio, converts speech to text, and performs AI intent analysis. Queries trigger voice/display responses; control commands send device instructions. Post-task, it returns to standby for continuous interaction.

4.3 Improvements and Optimizations in Speech Recognition

Traditional uniform circular arrays degrade in long-distance (>4m) scenarios (Fig. 8) due to uniform symmetry causing:

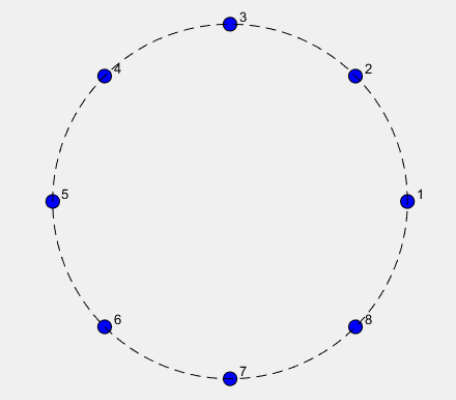

Figure 8. Schematic Diagram of Uniform Circular Microphone Array Arrangement

The core disadvantage of one-dimensional microphone arrays stems from their single-dimensional geometry, resulting in inherent limitations in direction estimation, resolution, anti-jamming capability, and so on. 2D/3D array or multi-sensor fusion schemes are more

appropriate in scenes requiring 3D spatial perception or complex acoustic environments; One-dimensional arrays are suitable for simple scenarios with high cost, complexity and space requirements.

1. Innovation Goal: Design Non-Uniform Circular Array and Realize It Through Geometric Optimization: 1. Increased speech recognition distance 2. Improve the accuracy of speech recognition.

Improve the accuracy of speech recognition.

The microphone unit [13-16] is arranged in a golden spiral structure.

2. Core Innovation and Technical Verification:

Uniform Circular Array (UCA)

Position of the nth microphone (polar coordinates):

$$r_n = R \quad (\text{Fixed radius}), \quad \theta_n = \frac{2\pi n}{N} \quad (n = 0,1, \dots, N-1) \tag{6}$$

For the frequency f And direction (ϕ, θ), the response vector is:

$$a_{UCA}(f, \phi) = \left[e^{-j\frac{2\pi f}{c} R\cos(\phi - \theta_n)} \right]_{n=0}^{N-1} \tag{7}$$

Among c Is the speed of sound.

The golden spiral array

Position of the nth microphone (polar coordinates):

$$r_n = \alpha\sqrt{n}, \quad \theta_n = n \cdot \varphi \quad (\varphi = 137.5°1.\text{golden corner}) \tag{8}$$

Array response vector:

$$a_{spiral}(f, \phi) = \left[e^{-j\frac{2\pi f}{c} r_n \cos(\phi - \theta_n)} \right]_{n=0}^{N-1} \tag{9}$$

3. Analysis of Valvular Inhibition:

Evenly aligned periodic structures result in grating lobes (repeating main lobe directions) at a particular frequency, which occur when:

$$\frac{2\pi f}{c} d\cos\phi = 2\pi k \quad (k \in Z) \tag{10}$$

Among d Is the microphone spacing.

4. The Non-Cyclical Advantage of the Gold Spiral:

Radius of the golden spiral by $r_n = a\sqrt{n}$ Increase, the angle increases by the golden angle, and its spacing $\Delta r_n = a\left(\sqrt{n+1} - \sqrt{n}\right)$ And angular separation $\Delta\theta_n = \varphi$ All of them are non-uniform. This aperiodicity breaks the symmetry of spatial sampling and makes the frequency response of the array more sparse in the spatial frequency domain, thus suppressing the Grating Lobe.

Mathematical verification:

By calculating the beam patterns of the two arrays, $B(\phi) = |w^H a(f, \phi)|^2$:

Uniform circular arrangement: in $f > \frac{f_c}{2\pi}$ The beam pattern exhibits periodic Grating Lobes.

Golden spiral array: Non-uniform spacing results in aperiodicity of the wavenumber response and a significant reduction in the Grating Lobe amplitude.

5. Directional Index (DI) Comparison:

The directivity index measures the degree of energy concentration of the array in the direction of the principal axis and is defined as:

$$DI=10\log_{10}\left(\frac{4\pi|B(\phi_0)|^2}{\int_0^{2\pi}\int_0^{\pi}|B(\phi)|^2\sin\theta\,d\theta d\phi}\right) \quad (11)$$

Among ϕ_0 Is the direction of the principal axis. The DI of a uniform circular array for N A uniform circular array of elements whose theoretical maximum DI is:

$$DI_{UCA} \approx 10\log_{10}N \quad (12)$$

6. Broadband Performance Analysis:

Spatial aliasing frequency The spatial aliasing frequency of a uniform array is:

$$f_{alias} = \frac{c}{2d} \quad (13)$$

If the microphone spacing $d = \frac{2\pi R}{N}$, then the high frequency signal ($f > f_{alias}$Aliasing can occur. Broadband advantage of the golden spiral The spacing of the golden spiral increases with increasing radius, and its minimum spacing $d_{min} = a(\sqrt{1} - \sqrt{0}) = a$, Maximum Spacing $d_{max} = a\sqrt{N}$。 Through reasonable choice a? Allows the array to avoid aliasing over a wide frequency band. The validation formula is designed to target the highest frequency $f_{max} = 8\,\text{kHz}$ No aliasing, to meet:

$$a < \frac{c}{2f_{max}} - \frac{343}{2 \times 8000} \approx 0.0214\,\text{m}$$

Choose $a = 0.02\,\text{m}$, then $d_{min} = 0.02\,\text{m}$ Satisfied f_{max} Request.

Through formula derivation and numerical verification, the golden spiral microphone array and the traditional uniform microphone array are arranged.Its core advantages are as follows:

1. Word Error Rate (WER);
2. Directional improvement: the main lobe is sharper;
3. Wide band adaptability: support wider frequency range through parameter optimization;
4. Anti-aliasing capability: The minimum spacing design avoids high frequency aliasing.

Arrange the microphone array as shown in the following Figure 9:

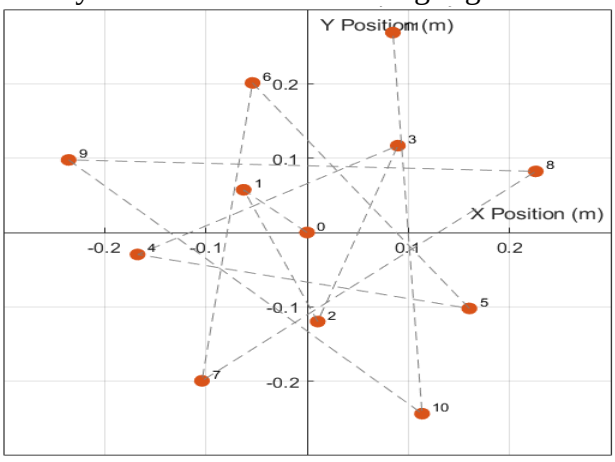

Figure 9. Schematic Diagram of Non-linear Golden Spiral Microphone Array Arrangement

Audio recognition distance was tested in a controlled room. At 6.5m total distance, detection was performed every 0.5m from the array center (multiple positions), with accuracy judged by voice-to-text error rate.

1. Word Error Rate (WER)

Primary metric calculating word-level differences between recognized and original text:

$$WER=(S+D+I)/N\times100\% \quad (14)$$

Sound sources were recorded via mobile phone, played back at 100% volume with consistent SNR.

Test data are shown in Figure 10and Table 2.

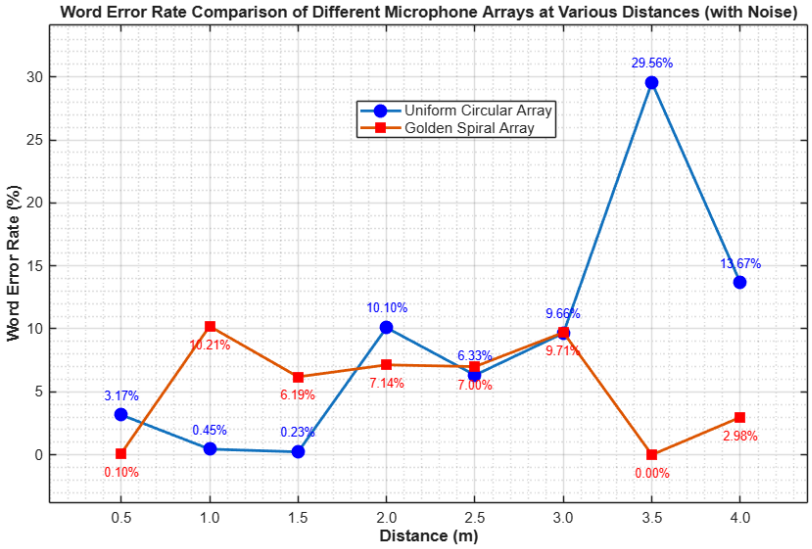

Figure 10. Experimental test data

Experimental results are presented in Table 2:

Table 2 Analysis of experimental data and conclusion

	Relative advantage section (m)	WER comparison (0.50-4.00m, %)	Conclusion
Uniform circular	0.50-1.50	9.16	Golden spiral array has larger effective aperture and lower WER than uniform circular array within 0.50–4.00m.
Golden spiral	2.00-4.00	5.41	

In 0.50–4.00m range, golden spiral array outperforms uniform circular array with 3.75% lower WER.

5. Conclusion

An AI-powered TCM decoction management assistant is proposed, with key contributions:

1. Integrated hardware-software system with pressure sensing and AI voice interaction.

2. Kalman filter application to enhance mass detection accuracy by suppressing decoction vibration interference.

3. Innovative golden-spiral microphone array with 3.75% lower WER than uniform arrays (0.50–4.00 m), improving speech interaction reliability.

Acknowledgements

This work was supported by National Undergraduate Innovation and Entrepreneurship Training Program (No. 202411058005).

Reference:

[1] *O. Ibrahim, Y. Nejat, A Novel Spaceborne Antenna Repositioning Method for Reliable Communication in Inclined Satellites, International Journal of Satellite Communications & Networking, 43 (2025) 318-332.*

[2] *S.J. Peale, The Origin of the Natural Satellites - ScienceDirect, Treatise on Geophysics, 10 (2007) 465-508.*

[3] *X. Zhou, H.P. Lee, H. Li, Y. Fan, X. Ma, Innovative Wrinkle-Suppressing Design for Satellite Phased Array Antennas, International Conference on Computational & Experimental Engineering and Sciences, 3 (2025) 899-909.*

[4] *M.A. Ingram, R. Romanofsky, R.Q. Lee, F. Mir, Z. Popovic, J. Langley, W.C. Barott, M.U. Ahmed, M. Dan, Optimizing Satellite Communications With Adaptive and Phased Array Antennas, Earth Science and Technology Conference, USA, (2004) 1-7.*

[5] *K. Ohata, K. Kobayashi, K. Nakahira, M. Ueba, Broadband and scalable mobile satellite communication system for future access networks, Acta Astronautica, 57 (2005) 239-249.*

[6] *A. Destounis, A.D. Panagopoulos, Dynamic Power Allocation for Broadband Multi-Beam Satellite Communication Networks, IEEE Communications Letters, 15 (2011) 380-382.*

[7] *S. Verma, Advanced Satellite Communications Systems & Services, in: E.D. Re, M. Ruggieri (Eds.) Satellite Communications and Navigation Systems, Springer US, Boston, MA, (2008) 513-516.*

[8] *M. Sarkar, A. Nayak, S. Nanda, S. Patnaik, Non-orthogonal Multiple Access (NOMA) Channel Estimation for Mobile & PLC-VLC Based Broadband Communication System, International Journal of Sensors, Wireless Communications and Control, 15 (2025) 30-39.*

[9] *J. Peng, H. Zhu, C. Gao, H. Liang, D. Du, S. Fan, R. Wang, Z. Qian, A Low-Profile 2-D Multibeam Multifeed Antenna Array With Millimeter-Wave Applications, Antennas and Propagation, IEEE Transactions on, 73 (2025) 588-593.*

[10] *Y.F. Hestrio, M. Soleh, A. Hidayat, H. Afida, H. Gunawan, A. Maryanto, Satellite data receiving antenna system for pleiades neo observation satellite, Journal of Physics: Conference Series, 1763 (2021) 012019.*

[11] *W. Sun, Z. Chen, L. Li, C. Shen, K. Yu, S. Li, J. Long, H. Zheng, L. Wang, T. Long, A chip-integrated comb-based microwave oscillator, Light: Science & Applications, 14 (2025) 276811.*

[12] *Xu, J.; Fan, S.; Zhao, Z.; Li, F.; Zhang, Y. Deep reinforcement learning-based dynamic multi-beam power allocation for GEO-LEO co-existing satellites. Acta Astronautica, 223 (2024) 197-209.*

[13] *Hu, X.; Liu, S.; Wang, Y.; Xu, L.; Zhang, Y.; Wang, C.; Wang, W. Deep reinforcement learning-based beam Hopping algorithm in multibeam satellite systems. IET Communications, 13 (2019), 2485-2491.*

[14] *S. Zhai, T. Hui, X. Gong, Z. Zhang, X. Gao, K. Yang, High performance receiving and processing technology in satellite beam hopping communication, Systems Engineering and Electronics Technology, 35 (2024) 815-828.*

[15] *W. Ren, B. Wang, J. Sun, Y. Hong, Inter-satellite link allocation in low-earth-orbit mega-constellation networks, The Journal of Supercomputing, 81 (2025) 1163.*

[16] *Q. Yang, D.I. Laurenson, J.A. Barria, On the Use of LEO Satellite Constellation for Active Network Management in Power Distribution Networks, IEEE Transactions on Smart Grid, 3 (2012) 1371-1381.*