ANR-2 基於ESP32-S3的迷你AI降噪模組 — 產品規格書

> 版本: v1.0 | 日期: 2026-07-17 | 狀態: DRAFT
> 作者: Honey | 用途: 法庭錄音系統(RecordStudio)前端降噪

---

1. 項目概述

1.1 目標

設計一個基於ESP32-S3的超小型AI降噪模組，每個麥克風掛一個，在類比音頻通路中做實時RNNoise降噪，為RecordStudio法庭錄音系統提供前端信號淨化。

1.2 核心價值

- **錄音質量優先**：48kHz全頻帶保留，非語音辨識降采样
- **非穩態噪聲強**：GRU擅長處理翻紙聲/椅子聲/關門聲
- **迷你形態**：~25×20mm級別，類似Seeed XIAO/DIO
- **極簡BOM**：兩顆主芯片(ESP32-S3+ES8311)+阻容

1.3 應用場景

- 法庭錄音：法官/原告/被告/證人各一個麥克風，每個掛一個ANR-2
- 會議室：多麥克風場景前端降噪
- 任意需要Line In→降噪→Line Out的場景

---

2. 音頻通路SPEC

2.1 信號流

```
電容麥 → 前置放大器 → [ANR-2 Line In] → ES8311 ADC → I2S → RNNoise(ESP32-S3) → I2S → ES8311 DAC → [ANR-2 Line Out] → 錄音設備
```

2.2 關鍵約束

- **兩端都是類比線路信號**：Line In來自電容麥+放大器，Line Out給錄音設備
- **不加數字麥(INMP441)**：輸入已是放大後的Line In
- **不加功放(MAX98357A)**：輸出是Line Out不是驅動喇叭
- **採樣率**：48kHz（錄音質量，非16kHz語音辨識）
- **位深**：24bit（ES8311原生支持）

- **端到端延遲**：≤20ms（RNNoise 1幀10ms + 處理3-5ms + I2S緩衝）

2.3 增益架構

級節	實現方式	範圍	步進	特點
前增益	ES8311 ADC PGA（硬體）	0~+36dB	3dB	類比域放大，提高信噪比，零CPU零延遲
後增益	軟體數字乘法	+3~+12dB	0.1dB	補償降噪後音量損失，1行代碼實現

增益控制：2個4位DIP開關，分別控前/後增益，斷電不丟，現場免工具調。

---

3. 硬件SPEC

3.1 主芯片

ESP32-S3-MINI-1 模組

參數	值
尺寸	15.4×20.5mm
芯片	ESP32-S3FN8
Flash	8MB
PSRAM	8MB
核心	雙核LX7 240MHz
GPIO	39
天線	PCB天線
價格	¥22.55/片（優信電子 深圳）

ES8311 Codec

參數	值
封裝	QFN-20 3×3mm
ADC SNR	100dB
DAC SNR	110dB
ADC PGA	0~+36dB 3dB步進
DAC Volume	-95dB~0dB 0.5dB步進
採樣率	8~96kHz（目標48kHz）

參數	值
位深	24bit
接口	I2S(音頻) + I2C(控制)
價格	¥1.63/片（ICGO 深圳）

3.2 I2S引腳分配（迷你模組）

信號	GPIO	方向
MCLK	GPIO0	輸出
BCLK	GPIO1	輸出
WS	GPIO2	輸出
DOUT(ADC入)	GPIO3	輸入
DIN(DAC出)	GPIO4	輸出
I2C_SDA	GPIO5	雙向
I2C_SCL	GPIO6	輸出

3.3 接口SPEC

鳳凰端子（3.5mm間距螺絲壓接型）

- 型號：2EDGK-3.5系列
- 間距：3.5mm
- 方案：3×2位端子
- 佈局：
- 端子1：VCC + GND（電源輸入）
- 端子2：Line In + GND（音頻輸入）
- 端子3：Line Out + GND（音頻輸出）
- 板寬約25mm（3個端子並排）
- 價格：¥0.30/個（mt數碼配件企業店）

3.4 BOM（迷你模組終版）

序號	器件	型號	數量	單價	小計
1	ESP32-S3模組	ESP32-S3-MINI-1	1	¥22.55	¥22.55
2	音頻Codec	ES8311 QFN-20	1	¥1.63	¥1.63
3	鳳凰端子	2EDGK-3.5 2P	3	¥0.30	¥0.90
4	DIP開關	4位DIP	2	¥0.40	¥0.80

序號	器件	型號	數量	單價	小計
5	LDO	ME6211 3.3V SOT-89	1	¥0.50	¥0.50
6	晶振	32.768kHz	1	¥0.30	¥0.30
7	電容	10μF/100nF各2	4	¥0.05	¥0.20
8	電阻	4.7k×2(I2C上拉)+其他	5	¥0.03	¥0.15
9	PCB	嘉立創4層 25×20mm	1	¥10.00	¥10.00
-	**合計**	-	-	-	**¥37.03**

3.5 兩步走採購方案

步驟	目的	採購	費用
Step 1	快速驗證通路	ESP32-S3 DevKitC-1 N16R8（¥23）+ ES8311+NS4150B音頻模塊（¥13.8）+ 杜邦線（¥3）	¥39.8
Step 2	迷你模組打板	上表BOM	¥37.03
總計	-	-	**¥76.83**

> Step 1 採購說明：DevKitC-1為樂鑫官方核心板，ES8311音頻模塊為小智DIY出品（綠深LSSZ，河南鄭州），I2S五線+I2C兩線全引出，NS4150B功放部分不接即不影響codec功能

3.6 電源SPEC

- 輸入：5V DC（經鳳凰端子VCC）
- LDO降壓：5V→3.3V（ME6211，最大500mA）
- ESP32-S3功耗：~240mA@240MHz
- ES8311功耗：~10mA
- 預估總功耗：~260mA@3.3V ≈ 0.86W

---

4. 軟件SPEC

4.1 RNNoise模型

參數	值
模型大小	88KB
幀長	10ms（480樣本@48kHz）
延遲	15-20ms
CPU佔用	~8-15%@240MHz

參數	值
推理時間	3-5ms/幀
許可證	BSD-3
語依賴	純C，無外部庫
VAD	自帶VAD輸出

4.2 雙核調度模型

核心	任務	優先級
Core 0	I2S驅動 + ES8311 I2C控制 + FreeRTOS調度	Normal
Core 1	RNNoise推理（綁核）	Priority 7（最高）

4.3 FreeRTOS任務劃分

任務	核心	優先級	棧大小	說明
i2s_rx_task	Core 0	5	2KB	I2S DMA接收，填輸入環形緩衝
i2s_tx_task	Core 0	5	2KB	I2S DMA發送，從輸出環形緩衝讀
rnnoise_task	Core 1	7	4KB	讀480樣本→推理→寫輸出
es8311_task	Core 0	3	2KB	I2C初始化+增益控制+狀態監控

4.4 記憶體布局

區域	內容	大小	存放
模型權重	RNNoise 88KB	88KB	Flash(XIP)
GRU狀態	3層GRU隱藏態	~20KB	SRAM
特徵buffer	FFT+band特徵	~15KB	SRAM
I2S DMA	雙通道×2緩衝	2×4KB	PSRAM
任務棧	4個任務	~14KB	SRAM
SRAM佔用	-	**~50KB**	/512KB可用
PSRAM佔用	-	**~8KB**	/8MB可用

4.5 I2S音頻通路實現

```
ES8311 ADC (48kHz 24bit)
  → I2S DMA RX (雙緩衝, 480樣本/幀)
    → 環形緩衝 input_buf
```

```
    → RNNoise處理 (Core 1, 每10ms一幀)
      → 前增益PGA已由硬體完成
      → rnnoise_process_frame() 降噪
      → 後增益: out[i] *= post_gain (數字乘法)
      → Bypass模式: 直接旁路RNNoise
    → 環形緩衝 output_buf
  → I2S DMA TX (雙緩衝, 480樣本/幀)
→ ES8311 DAC (48kHz 24bit)
```

4.6 ES8311驅動接口

```
// 初始化
esp_err_t es8311_init(i2c_port_t port, uint8_t addr);

// 增益控制
esp_err_t es8311_set_mic_gain(int gain_db);    // 前增益 PGA 0~36dB 3dB步進
esp_err_t es8311_set_voice_volume(int vol_db);  // DAC音量 -95~0dB 0.5dB步進

// 採樣率
esp_err_t es8311_set_sample_rate(int rate);    // 設48kHz

// Bypass
esp_err_t es8311_bypass_adc(bool enable);      // ADC直通DAC
```

4.7 後增益控制

```
// 後增益：RNNoise輸出後的數字乘法
// 範圍: +3~+12dB (1.41~3.98倍)
// 精度: 0.1dB
// 由DIP開關4位編碼: 0000=+3dB, 1111=+12dB
float post_gain = powf(10.0f, (dip_value * 0.6f + 3.0f) / 20.0f);
for (int i = 0; i < FRAME_SIZE; i++) {
    out[i] = rnnoise_out[i] * post_gain;
}
```

4.8 Bypass模式

- 觸發：GPIO中斷（按鍵）或I2C命令
- 行為：ADC直通DAC，跳過RNNoise處理
- 用途：對比降噪前後效果、故障安全旁路

4.9 VAD指示

- RNNoise自帶VAD輸出（rnnoise_process_frame返回值）
- VAD > 閾值 → GPIO輸出高電平（可接LED）
- 用途：指示有無人說話

5. PC虛擬驗證SPEC（沙箱可執行）

5.1 驗證分層

Layer	目的	環境	狀態
Layer 1	RNNoise離線WAV降噪驗證	PC沙箱(Python)	🔄 進行中
Layer 2	即時音頻pipeline驗證	PC聲卡	⏳ 待硬體
Layer 3	ESP-IDF交叉編譯+benchmark	ESP-IDF v5.4	⏳ 待環境

5.2 Layer 1 驗證項目

#	驗證項	方法	通過標準
1.1	RNNoise源碼可編譯	gcc直接編譯	無error
1.2	48kHz全頻帶保留	降噪前後頻譜對比	語音頻段衰減<3dB
1.3	非穩態噪聲消除	翻紙/椅子/關門聲降噪	SNR提升>6dB
1.4	穩態噪聲消除	HVAC/白噪聲降噪	SNR提升>10dB
1.5	延遲測量	幀處理計時	<10ms/幀
1.6	語音素失真	降噪後語音PESQ/STOI	PESQ>3.0
1.7	增益補償	後增益數字乘法	輸出RMS與預期一致
1.8	Bypass模式	跳過降噪直接輸出	輸出=輸入

5.3 Layer 1 驗證輸入

- **測試音頻**：48kHz 5秒模擬法庭錄音
- **語音**：多諧波合成+語音包絡（模擬說話停頓）
- **穩態噪聲**：60/120/180Hz HVAC嗡嗡聲
- **非穩態噪聲**：翻紙聲(50ms脈衝)、椅子聲(300ms低頻)、關門聲(衝擊衰減)
- **環境白噪聲**

5.4 Layer 1 驗證工具

- **noisereduce** (Python)：頻譜門降噪（對比基線）
- **numpy/scipy**：頻譜分析、SNR計算、PESQ近似
- **soundfile**：WAV讀寫
- **ffmpeg**：格式轉換
- **gcc/make**：RNNoise C源碼編譯（如可解決autotools依賴）

6. RNNoise vs 樂鑫NS 對比（決策依據）

維度	RNNoise	樂鑫AFE NS
採樣率	48kHz ✅	16kHz ❌
非穩態噪聲	GRU擅長 ✅	AEC+NS適配慢 ❌
模型大小	88KB	~740KB起(PSRAM)
CPU佔用	~8-15%	~9-12%
延遲	15-20ms	~30ms
依賴	純C零依賴 ✅	需esp-skainet框架
VAD	自帶 ✅	需額外配置
適配場景	錄音降噪 ✅	語音辨識/通話

結論：RNNoise更適合法庭錄音場景。

---

7. PCB設計約束（預留）

約束	值	說明
尺寸	25×20mm	類XIAO迷你形態
層數	4層	信號完整性+電源完整性
板厚	1.0mm	迷你模組標準
阻抗	50Ω	I2S高速信號
音頻走線	類比/數字分區	AGND/DGND單點連接
電源	LDO 3.3V	5V輸入→3.3V

---

8. 開發里程碑

階段	內容	交付物	依賴
M1	PC Layer 1 離線驗證	驗證報告+降噪樣本	沙箱環境
M2	PC Layer 2 即時驗證	pipeline demo	PC聲卡
M3	Korvo-2 V3 真機驗證	固件+測試結果	硬體到貨
M4	迷你模組PCB設計	Gerber+BOM	M3通過
M5	迷你模組打板+調試	樣品	M4+嘉立創

階段	內容	交付物	依賴
M6	RecordStudio整合	系統測試報告	M5+RecordStudio

9. 風險與踩坑預判

#	風險	影響	緩解
R1	RNNoise移植ESP32-S3無先例	中	純C零依賴，移植難度低；沙箱先驗證算法
R2	I2S 48kHz 24bit雙工穩定性	高	Korvo-2 V3先驗證；雙緩衝DMA
R3	ES8311 I2C控制可靠性	中	樂鑫有現成driver；加超時重試
R4	迷你板音頻信號完整性	中	4層板；類比/數字地分區；短走線
R5	240MHz跑RNNoise效能不足	低	預估8-15%，實測確認
R6	PSRAM訪問延遲影響即時性	中	I2S DMA放PSRAM；模型權重Flash XIP
R7	熱設計（迷你板散熱）	低	ESP32-S3功耗~0.5W；無需主動散熱

10. 可借鑑項目

項目	倉庫	借鑑價值
xiaozhi-esp32	78/xiaozhi-esp32	I2S+雙核調度+Opus編解碼+OTA
Seeed XIAO+reSpeaker	Seeed-Projects/Xiaozhi_Esp32S3_reSpeaker	XMOS方案參考
ESP-SR/esp-skainet	Espressif官方	AFE框架+I2S配置
RNNoise	xiph/rnnoise	核心算法(BSD-3, 88KB, 48kHz, 純C)
立創ESP32-S3開發板	oshwhub.com/cfengls/esp32s3-wroom-dev	硬體參考

附錄A：Korvo-2 V3 參考引腳

信號	GPIO
MCLK	GPIO16
SCLK(BCLK)	GPIO9
LRCK(WS)	GPIO45
DSDIN	GPIO8
I2C_SDA	GPIO17

信號	GPIO
I2C_CLK	GPIO18

附錄B：DIP開關增益編碼

前增益（DIP-SW1, 4位）

SW1[3:0]	增益dB	說明
0000	0dB	最小增益
0001	3dB	
0010	6dB	
...	...	3dB步進
1100	36dB	最大增益

後增益（DIP-SW2, 4位）

SW2[3:0]	增益dB	說明
0000	+3dB	最小補償
0001	+3.6dB	
...	...	0.6dB步進
1111	+12dB	最大補償