

Shenzhen Waytronic Electronic Co., Ltd.

WT3000TX Speech Synthesis Chip and module introduction

V1.09



Disclaimer:

Shenzhen Waytronic Electronic Co., Ltd. hereby declares: The product manual shall be subject to the official website version. No individual updates will be notified if the content is modified. We shall not be liable for any infringement of third-party patents or rights caused by the use of our ICs. In the event of significant property damage, personal injury, or fatalities resulting from the use of our ICs in aviation, satellite, military equipment, or personal safety applications, we shall not be held responsible.

catalogue

VERSION UPDATE 2	3
1. PRODUCT INTRODUCTION 3	4
2. PRODUCT FEATURES 3	4
3. INTRODUCTION TO PROCEDURE AND MODULE SELECTION 4	5
3.1. MODULE INTRODUCTION 4	5
3.2. PROGRAM INTRODUCTION 4	5
4. PIN DESCRIPTION 4	5
4.1. CHIP PACKAGE PIN DIAGRAM 4	5
4.2. WT3000T M01 MODULE 6	7
4.2.1. Pipe Foot Description 6	7
4.2.2. Circuit Design Reference 7	8
5. HOW TO USE THE SPEECH SYNTHESIS FUNCTION 8	9
5.1. SIMPLE USAGE 8	9
5.2. STANDARD CALL METHOD 8	9
5.3. METHODS FOR QUERYING THE WORKING STATUS OF CHIPS 8	9
6. FIXED VOICE FUNCTION 8	9
7. FORMAT OF AGREEMENT ORDER 8	9
INSTRUCTION FORMAT 9	10
INSTRUCTION RETURN CODE FORMAT 9	10
8. ELECTRICAL PARAMETERS 10	10
8.1. ABSOLUTE MAXIMUM RATED PARAMETERS 10	11
8.2. PMU CHARACTERISTICS 10	11
8.3. ELECTRICAL LOGIC CHARACTERISTICS OF I/O INPUT/OUTPUT 10	11
8.4. SIMULATION OF DAC CHARACTERISTICS 11	11
8.5. ADC CHARACTERISTICS 11	12
9. PACKAGING INFORMATION 11	12
9.1. WT3000T8-32N PACKAGE SIZE 11	12

Version update

version number	Revision Notes	date changed
V1.00	original version	2023-12-26
V1.01	Transition Version	2024-01-30
V1.02	Add new features, modify code verification and calculation methods, and include electrical parameters and chip packaging	2024-02-29
V1.03	Add module description and correct partial descriptions	2024-03-15
V1.04	Add a module description section to define unused parts	2024-05-31
V1.05	Increase the deep sleep power consumption and remove the appendix, which is now incorporated into the instruction set.	2024-06-14
V1.06	Add Chinese and English version descriptions	2024-06-26
V1.07	VCM Capacitor Parameter Correction	2024-12-09
V1.08	modified deep sleep power	2025-08-14
V1.09	Update module image	2026-01-29

1. Product Introduction

The WT3000TX series is a high-performance voice chip series featuring a 32-bit processor with a maximum frequency of 240MHz. Its highly integrated speech synthesis chip supports Chinese, English letters, and bilingual speech synthesis, while also incorporating voice encoding and decoding functions for both synthesis and playback. The chips are characterized by low cost, low power consumption, high reliability, and strong versatility. Available in two package sizes: WT3000T8-32N QFN32 (compact 4*4mm) and WT3000T3-32N QFN32 (4*4mm). Additional features include address-based playback, insert playback, single track loop, all tracks loop, and random playback. The system offers 31-level volume adjustment and supports up to 128MB of external Flash memory.

2. Product Features

- Control mode: UART, with a default baud rate of 9600.
- Power-up defaults to no playback; features a BUSY status indicator where BUSY playback defaults to low level and no playback defaults to high level (default configuration can be modified via code transmission).
- Audio output mode: the sample defaults to DAC output.
- Supports high-quality voice audio formats (8kbps~320kbps) with clear sound quality, including MP3 and WAV formats.
- Supports random playback of instructions and seamless loop playback.
- It supports up to 128Mbit Flash memory.
- Volume adjustable with 31 levels;
- High-power IO driving capability, capable of directly driving up to 32mA.
- The WT3000T8 A version supports synthesis of any Chinese text and English letters, and allows mixed reading of Chinese and English letters. It does not currently support variable speed or pitch modulation for English letters using markers. In contrast, the WT3000T3 D version supports mixed reading and synthesis of any Chinese and English text, and allows variable speed or pitch modulation for English letters using markers.

The chip supports synthesizing any Chinese or English letters using the GB 2312 encoding. Each synthesis can produce up to 2K bytes of text.

The chip analyzes text and accurately identifies and processes common formats such as numbers, dates, times, and measurement symbols based on its built-in text matching rules.

- Supports multiple control commands

The controller supports functions including text synthesis, stop/suspend/restart synthesis, status queries, sleep mode activation, and wake-up. Through communication interfaces, it issues control commands to manage the chip. The chip's commands are user-friendly: For example, it can play audio prompts and bilingual text synthesis by referencing corresponding

instruction manuals, while enabling parameter adjustments for synthesized speech through text tagging. The controller supports functions including text synthesis, stop/suspend/restart synthesis, status queries, sleep mode activation, and wake-up. Through communication interfaces, it issues control commands to manage the chip. The chip's commands are user-friendly: For example, it can play audio prompts and bilingual text synthesis by referencing corresponding instruction manuals, while enabling parameter adjustments for synthesized speech through text tagging.

- Supports multiple methods to check the chip's operational status

It includes: query the status pin level, read the working status word automatically returned by the chip, and send the query command to get the return data of the working status of the chip.

- For single-chip operation (using built-in capacity), the built-in voice must be pre-programmed before factory release.
- The power consumption is less than 20uA in deep sleep mode.

3. Introduction to Program and Module Selection

3.1. Module Introduction

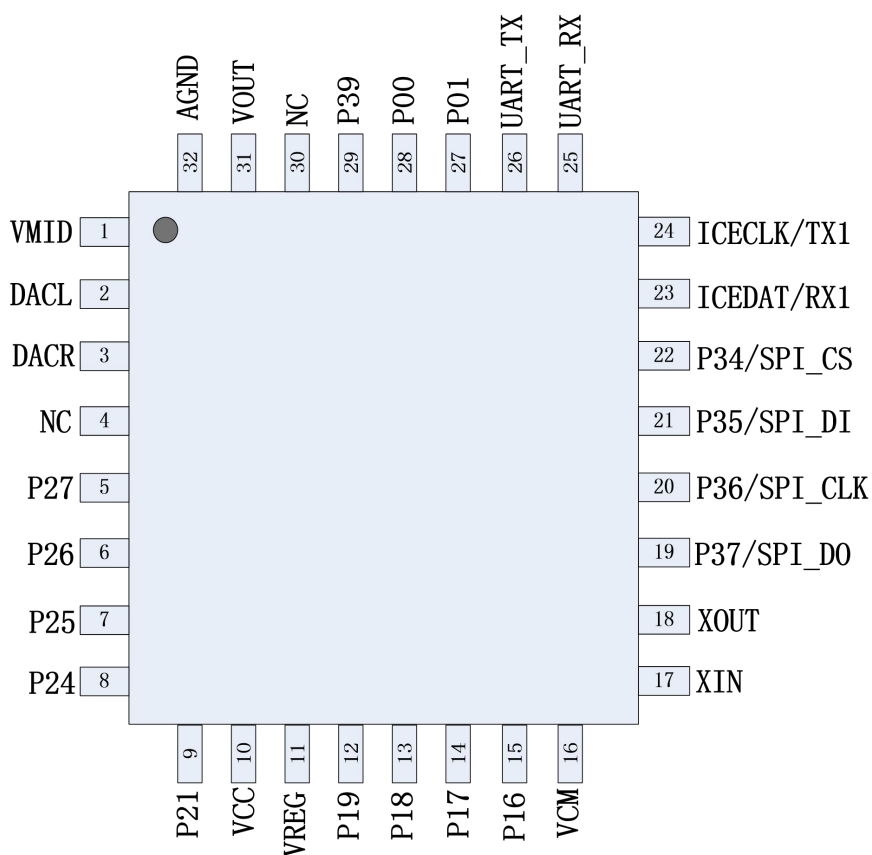
series	Function code	communication mode	module selection	Audio output	function declaration
WT3000T series	T001	UART (9600)	M01	DAC	This module uses the WT3000T8-32N chip by default, supporting Chinese/English text synthesis playback. It only supports built-in Flash playback and can store 30 seconds of fixed voice recordings.

3.2. Program Introduction

slug	Function code	communication mode	edition	Audio output	function declaration
WT3000T8-32N	T001	UART (9600)	A	DAC	Chinese/English text-to-speech, supports built-in Flash playback only, can store 30 seconds of fixed voice
WT3000T3-32N	T001	UART (9600)	D	DAC	Chinese and English audio synthesis playback. Only built-in Flash is supported. Supports storing up to 500 seconds of fixed audio.

4. Pin Description

4.1. chip package pin diagram



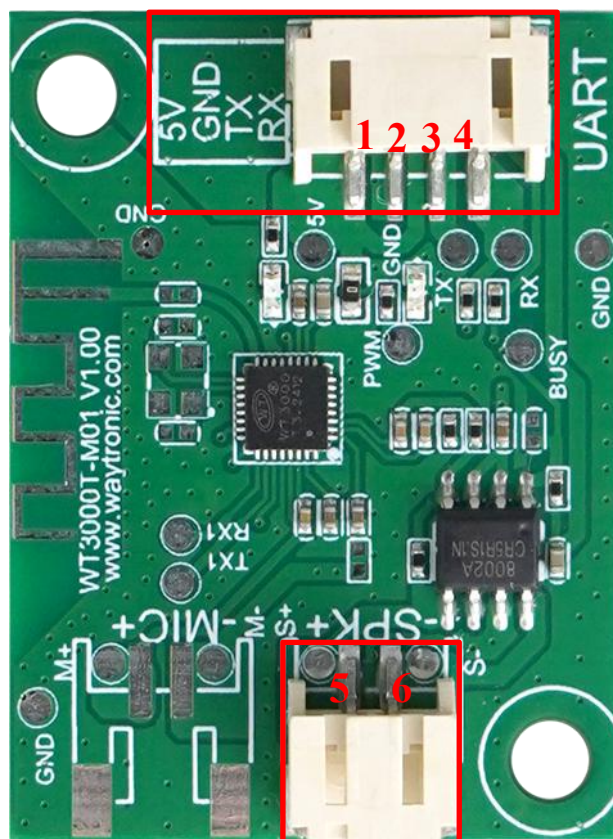
WT3000TX-32N

PIN-NO.	name	function declaration
0	GND	Chip ground (must be grounded)
1	VMID	Connect a 1uF external capacitor to ground.
2	DACL	DAC left audio output
3	DACR	DAC right channel audio output
4	NC	Keep this pin floating
5	P27	IO mouth
6	P26	IO mouth
7	P25	IO mouth
8	P24	IO mouth
9	P21	IO mouth
10	VCC	Power input: connect 10 μ F and 0.1 μ F capacitors to ground
11	VREG	Connect a 1uF capacitor to ground
12	P19	IO mouth

13	P18	IO mouth
14	P17	IO mouth
15	P16	IO mouth
16	VCM	Connect a 1uF external capacitor to ground at the reference voltage output
17	XIN	not have
18	XOUT	not have
19	P37/SPI_DO	SPI Flash Data Output
20	P36/SPI_CLK	SPI Flash clock
21	P35/SPI_DI	SPI Flash Data Input
22	P34/SPI_CS	SPI Flash chip selection
23	ICEDAT/D-/RX1	UART Rx, voice chip data reception
24	ICECLK/D+/TX1	UART TX, voice chip data transmission
25	UART_RX	IO mouth
26	UART_TX	IO mouth
27	P01	Busy show
28	P00	IO mouth
29	P39	IO mouth
30	NC	Keep this pin floating
31	VOUT	The chip's internal IO power supply outputs 3.3V and must be connected to ground via a 10μF capacitor.
32	AGND	simulated ground

4.2. WT3000T M01 module

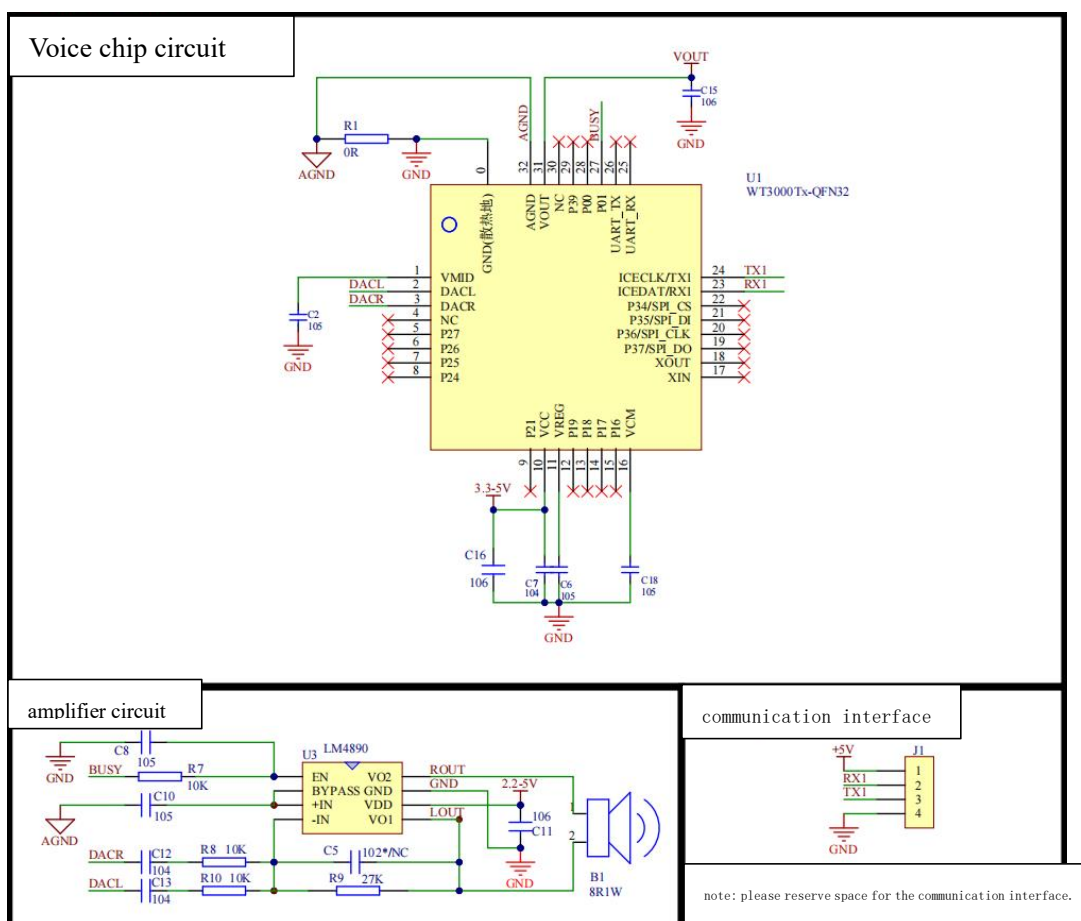
4.2.1. Pin Description



joggle	name	function declaration
1	5V	power input
2	GND	power supply ground
3	TX	UART TX, communication port
4	RX	UART Rx, communication port
5	SPK-	Internal power amplifier PWMN output
6	SPK+	Internal power amplifier PWMP output

Note: The MIC interface, antenna, and crystal oscillator are reserved for compatibility with other models and future product upgrades. These components are not required when using the TTS function. Serial port and speaker interface are sufficient for debugging.

4.2.2. Circuit Design Reference



Notes:**Voice IC:**

- 1.1 The voice IC supports UART serial control, with a 3.3V communication interface. If the MCU I/O level is 5V, a 220 Ω series resistor or a level-shifting circuit is required.
- 1.2 The programming pins ICECLK and ICEDAT should be reserved to facilitate firmware updates.
- 1.3 Power supply: the recommended operating voltage range is 3.3–5V. Both VCC and VOUT of the voice IC must be connected to decoupling capacitors to ground, and the trace length should be less than 1 cm.
- 1.4 During PCB layout, all power pins of the voice IC must be connected to capacitors to ground. Power traces must pass through the capacitor before connecting to the IC power pins.
 - The distance between the capacitor GND network and the IC GND pin, and between the 3.3–5V capacitor network and the IC VCC pin, should each be less than 5 mm.
 - The distance between the capacitor GND network and the IC GND pin, and between the VOUT capacitor network and the IC VOUT pin, should also be less than 5 mm.
- 1.5 For DAC audio output signals, traces and related components should be surrounded by ground as much as possible. Avoid routing near RF or digital signal traces, and avoid parallel routing to reduce interference.
- 1.6 This IC includes an analog ground. Analog ground and digital ground should be routed separately. Split GND copper pours may be used, and AGND and GND should preferably be connected at a single point near the power supply when using an external amplifier to reduce noise.

Power Amplifier IC:

- 2.1 During PCB layout, the power supply of the amplifier IC must be connected to a capacitor to ground, and the power trace must pass through the capacitor before connecting to the IC power pin.
 - The distance between the capacitor GND network and the amplifier GND pin, and between the 2.2–5V capacitor network and the amplifier VDD pin, should each be less than 5 mm.
- 2.2 In PCB design, to avoid affecting power system stability during amplifier operation, it is recommended to use a branched (Y-shaped) power routing scheme, branching from the positive terminal of the input power capacitor.
- 2.3 Ground separation: when the system includes digital or RF signals, it is recommended to keep the ground of the voice IC and the amplifier IC separate and connect them at a single point.
- 2.4 To eliminate “pop” noise, the amplifier enable pin can be controlled via the customer’s MCU I/O or the BUSY pin of the voice IC.

5. How to use voice synthesis

5.1. Simple invocation method

Simple call means that the user does not need to care about the current working state of the chip,

only need to send text to the chip to synthesize the voice output.

In simple call, the host computer only needs to establish UART communication with the chip, and can send the synthesis command to realize the text synthesis, the host computer does not need to judge the chip's return data or the output of the status pin.

Note: If the previous frame text is not fully synthesized, sending text to the chip again will interrupt the previous synthesis and start a new one.

5.2. standard call method

When the host computer must verify the integrity of the previous text synthesis before sending the next command frame, it requires confirming the chip's operational status through feedback. For example, if the application requires synthesizing 5k bytes of text—exceeding the maximum

capacity (2016 bytes) of a single command frame in the WT3000TX chip—it must transmit the text in multiple segments. The procedure is as follows:

1. The host computer first transmits a text synthesis command frame to the WT3000TX chip, containing up to 2016 bytes of text.
2. The host computer waits for the WT3000TX chip to automatically return the transmitted data, until it receives the data, indicating that the previous Text synthesis is complete; or use the status pin of the query chip and the method of sending query commands to confirm that the previous frame of text synthesis is complete.
3. The host computer sends another text synthesis command frame to the chip, transmitting the remaining text. This process repeats until all data is sent.

5.3. Method of Querying Chip's Working Status

The working status of the WT3000TX chip is queried by both hardware and software.

Hardware method: The chip's operational status is determined by monitoring the BUSY pin level. A low BUSY level indicates text synthesis, while a high level signifies idle operation.

Software implementation: The chip's operational status is queried by transmitting a status query command frame. Upon receiving this command from the host computer, the chip automatically transmits its current status back. The host computer then determines whether the chip is idle or in text synthesis mode based on the status data.

6. Fixed voice function

The WT3000T8-32N features a 30-second preset voice storage, offering bulk-purchasing customers customized voice prompt services.

The WT3000T3-32N features a 500-second preset voice storage, offering batch-purchasing customers customized voice prompt services.

Note: If you need to add a prompt sound, submit it to the business staff and provide the audio along with the required information.

7. protocol command format

The chip features a built-in standard UART asynchronous serial interface with a default baud rate of 9600 and operates at 3.3V TTL voltage. The communication data structure includes: 1 start bit, 8 data bits, no parity bit, and 1 stop bit. To use the computer serial port debugging tool, configure the serial parameters as shown in the diagram.

串口:	COM1
波特率:	9600
校验位:	无校验
数据位:	8
停止位:	1

☒ 十六进制发送
☐ 字符格式发送

instruction format

start code	Frame length	serial number	response mark	data frame source	N command messages	cumulative sum check	end code
7E	underneath	underneath	underneath	underneath	underneath	underneath	EF

instruction return code format

start code	Frame length	serial number	response mark	data frame source	N command messages	cumulative sum check	end code
7E	underneath	underneath	underneath	underneath	underneath	underneath	EF

Note: After executing each write command, return the corresponding result code according to the communication protocol format.

Frame length: 2 bytes, indicating the sum of frame length, sequence number, acknowledgment flag, data frame source, N command information, and checksum, with the high bits preceding the low bits.

The 流水号 (serial number) is a 1-byte field that increments by 1 per frame to prevent data duplication. Duplicate 流水号 data must be discarded (MCU does not perform this check when transmitting codes to chips/modules).

Response flag: 1 byte, fixed to 00;

Data frame source: 1 byte, where 02 indicates the TTS chip interface and 03 the MCU chip interface.

N command messages: Composed of N command messages, each message contains a 2-byte command, 1-byte data length, and N bytes of data. A single message supports up to 255 bytes of data, but multiple identical commands with different information can be transmitted simultaneously (the information in the return code differs slightly from the sending code; see the 'WT3000Tx Speech Synthesis Command Table' for details).

The sum and check means the sum of frame length, flow number, response flag, data frame source and N command information.

8. electrical parameters

8.1. absolute maximum rated parameter

Symbol	Parameter	Min	Max	Unit
Tamb	Ambient Temperature	-40	+85	°C
Tstg	Storage temperature	-65	+150	°C
VCC	Supply Voltage	-0.3	5.5	V
VOOUT	3.3V IO Input Voltage	-0.3	3.6	V

8.2. PMU characteristic

Symbol	Parameter	Min	Typ	Max	Unit	Test Conditions
VCC	Voltage Input	2.6	3.7	5.5	V	2.8V
VVcc	Voltage output	2.6	3.0	3.4	V	VBAT = 3.7V, 100mA loading
IVOUT	Loading current	—	—	100	mA	VBAT=3.7V

8.3. IO input/output electrical logic characteristics

IO input characteristics						
Symbol	Parameter	Min	Typ	Max	Unit	Test Conditions
VIL	Low-Level Input Voltage	-0.3	—	0.3* Vout	V	VOUT= 3.3V
VIH	High-Level Input Voltage	0.7* VOUT	—	VOUT+0.3	V	VOUT = 3.3V
IO output characteristics						
VOL	Low-Level Output Voltage	—	—	0.33	V	VOUT = 3.3V
VOH	High-Level Output Voltage	2.7	—	—	V	VOUT = 3.3V

8.4. analog DAC characteristics

Parameter	Min	Typ	Max	Unit	Test Conditions
Frequency Response	20	—	16K	Hz	1KHz/0dB 100kohm loading A-Weighted Filter
THD+N	—	-65	—	dB	
S/N	—	95	—	dB	
Output Swing	—	0.54	—	Vrms	
Dynamic Range	—	92	—	dB	1KHz/-60dB 100kohm loading With A-Weighted Filter
Output Resistance		8.3		K	

8.5. ADC characteristic

Parameter	Min	Typ	Max	Unit	Test Conditions
Dynamic Range	—	75	—	dB	1KHz/210mVrms
S/N	—	79	—	dB	line mode :6dB with cap
THD+N	—	-70	—	dB	PGAIS=2

9. Encapsulation information

9.1. WT3000T8-32N package size

unit : mm

