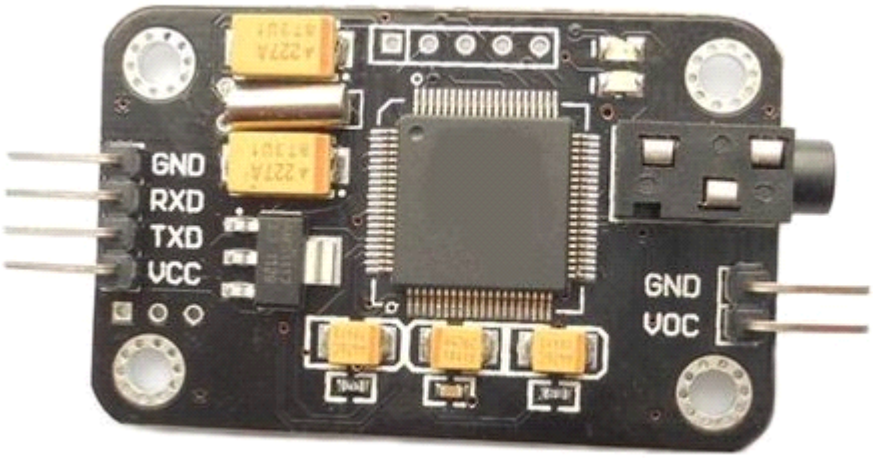
Voice Recognition Module

Speak to control (Arduino Compatible)



Introduction

The module could recognize your voice. It receives configuration commands or responds through serial port interface. With this module, we can control the car or other electrical devices by voice.

This module can store 15 pieces of voice instruction. Those 15 pieces are divided into 3 groups, with 5 in one group. First we should record the voice instructions group by group. After that, we should import one group by serial command before it could recognize the 5 voice instructions within that group. If we need to implement instructions in other groups, we should import the group first. This module is speaker independent. If your friend speaks the voice instruction instead of you, it may not identify the instruction. Please note that speaker independence requires strictly good MIC. The MIC we supply is not good enough for it to be speaker-independent.

Technical

Parameters

- *Voltage: 4.5-5.5V*
- *Current: <40mA*
- *Digital Interface: 5V TTL level UART interface*
- *Analog Interface: 3.5mm mono-channel microphone connector + microphone pin interface*
- *Size: 30mm x 47.5mm*
- *Recognition accuracy: 99% (under ideal environment)*

Serial Command

This module can be configured by sending commands via serial port. Configuration will be not erased after powered off.

Its interface is 5V TTL. The serial data format: 8 data bits, no parity, 1 stop bit. The default baud rate is 9600 and baud rate can be changed.

Command format is "Head + Key". "Head" is a 0xaa, and "Key" is as follows:

Key (HEX format)	Description	Respond in Common Mode	Respond in Compact Mode
0x00	Enter into "Waiting" state	"Waiting! | n" : successful "ERROR! | n" : Instruction error	0xcc : successful 0xe0 : Instruction error
0x01	Delete the instructions of group 1	"Group1 Deleted ! | n" : successful "ERROR! | n" : Instruction error	0xcc : successful 0xe0 : Instruction error
0x02	Delete the instructions of group 2	"Group2 Deleted ! | n" : successful "ERROR! | n" : Instruction error	0xcc : successful 0xe0 : Instruction error
0x03	Delete the instructions of group 3	"Group3 Deleted ! | n" : successful "ERROR! | n" : Instruction error	0xcc : successful 0xe0 : Instruction error
0x04	Delete the instructions of all the 3 groups	" All Groups Deleted !|n " : successful "ERROR! | n" : Instruction error	0xcc : successful 0xe0 : Instruction error
0x11	Begin to record instructions of group 1	"ERROR! | n" : Instruction error "START | n" : Ready for recording, you can speak now "No voice | n" : no voice detected "Again | n" : Speak the voice instruction again. Do not speak until getting the START message "Too loud | n" : Too loud to record "Different | n" : voice instruction confirming failed. Voice for the second chance is different with the first one. "Finish one | n" : recording one voice instruction successfully "Group1 finished! | n" : finish recording group 1	0xe0 : Instruction error 0x40 : Ready for recording, you can speak now 0x41 : no voice detected 0x42 : Speak the voice instruction again. Do not speak until getting the START message 0x43 : Too loud to record 0x44 : voice instruction confirming failed. Voice for the second chance is different with the first one. 0x45 : recording one voice instruction successfully 0x46 : finish recording group 1
0x12	Begin to record instructions of group 2	"ERROR! | n" : Instruction error "START | n" : Ready for recording, you can speak now "No voice | n" : no voice detected "Again | n" : Speak the voice instruction again. Do not speak until getting the START message "Too loud | n" : Too loud to record "Different | n" : voice instruction confirming failed. Voice for the second chance is different with the first one. "Finish one | n" : recording one voice instruction successfully "Group2 finished! | n" : finish recording group 2	0xe0 : Instruction error 0x40 : Ready for recording, you can speak now 0x41 : no voice detected 0x42 : Speak the voice instruction again. Do not speak until getting the START message 0x43 : Too loud to record 0x44 : voice instruction confirming failed. Voice for the second chance is different with the first one. 0x45 : recording one voice instruction successfully 0x47 : finish recording group 2
0x13	Begin to record instructions of group 3	"ERROR! | n" : Instruction error "START | n" : Ready for recording, you can speak now "No voice | n" : no voice detected "Again | n" : Speak the voice instruction again. Do not speak until getting the START message "Too loud | n" : Too loud to record	0xe0 : Instruction error 0x40 : Ready for recording, you can speak now 0x41 : no voice detected 0x42 : Speak the voice instruction again. Do not speak until getting the START message

		"Different | n" : voice instruction confirming failed. Voice for the second chance is different with the first one. *"Finish one | n" : recording one voice instruction successfully* *"Group3 finished! | n" : finish recording group 3*	*0x43 : Too loud to record* *0x44 : voice instruction confirming failed. Voice for the second chance is different with the first one.* *0x45 : recording one voice instruction successfully* *0x48 : finish recording group 3*
0x21	*Import group 1 and be ready for voice instruction*	*"Group1 Imported !\n" : Successful* *"ERROR! | n" : Instruction error* *"Import failed !\n" : Importing voice group failed*	*0xcc : Successful* *0xe0 : Instruction error* *0xe1 : Importing voice group failed*
0x22	*Import group 2 and be ready for voice instruction*	*"Group2 Imported !\n" : Successful* *"ERROR! | n" : Instruction error* *"Import failed !\n" : Importing voice group failed*	*0xcc : Successful* *0xe0 : Instruction error* *0xe1 : Importing voice group failed*
0x23	*Import group 3 and be ready for voice instruction*	*"Group3 Imported !\n" : Successful* *"ERROR! | n" : Instruction error* *"Import failed !\n" : Importing voice group failed*	*0xcc : Successful* *0xe0 : Instruction error* *0xe1 : Importing voice group failed*
0x24	*Query the recorded group*	*"Used group:0\n" : No group is recorded* *"Used group:1\n" : Group 1 is recorded* *"Used group:2\n" : Group 2 is recorded* *"Used group:3\n" : Group 3 is recorded* *"Used group:12\n" : Group 1 and Group 2 are recorded* *"Used group:13\n" : Group 1 and Group 3 are recorded* *"Used group:23\n" : Group 2 and Group 3 are recorded* *"Used group:123\n" : All the 3 groups are recorded* *"ERROR! | n" : Instruction error*	*0x00 : No group is recorded* *0x01 : Group 1 is recorded* *0x02 : Group 2 is recorded* *0x04 : Group 3 is recorded* *0x03 : Group 1 and Group 2 are recorded* *0x05 : Group 1 and Group 3 are recorded* *0x06 : Group 2 and Group 3 are recorded* *0x07 : All the 3 groups are recorded* *0xe0 : Instruction error*
0x31	*Change the baud rate to 2400bps*	*"Baud: 2400\n" : Successful* *"ERROR! | n" : Instruction error*	*0xcc : successful* *0xe0 : Instruction error*
0x32	*Change the baud rate to 4800bps*	*"Baud: 4800\n" : Successful* *"ERROR! | n" : Instruction error*	
0x33	*Change the baud rate to 9600bps*	*"Baud: 9600\n" : Successful* *"ERROR! | n" : Instruction error*	
0x34	*Change the baud rate to 19200bps*	*"Baud: 19200\n" : Successful* *"ERROR! | n" : Instruction error*	
0x35	*Change the baud rate to 38400bps*	*"Baud: 38400\n" : Successful* *"ERROR! | n" : Instruction error*	
0x36	*Switch to Common Mode*	*"Common Mode\n" : Successful* *"ERROR! | n" : Instruction error*	
0x37	*Switch to Compact Mode*	*"Compact Mode\n" : Successful* *"ERROR! | n" : Instruction error*	
0xbb	*Query version information*	*Version information*	*No respond*

*If you want to modify the serial baud rate to 38400, you need to send command: **0xaa35**. If successful, it will return "Baud: 38400 \| n"(in Common Mode) or **0xcc** (in Compact Mode). The baud rate is set to 38400.*

*The main difference between Compact Mode and Common Mode is the returning message. Common Mode response is long string but Compact Mode response is a byte. For example, after sending **0xaa04** to delete all the contents of the 3 groups, in Common*

Mode it will return ***"All Groups Deleted! | n"****, but in Compact Mode it will return a concise bytes such as* ***0xcc*** *which means a successful operation.*

For the first-time use, we need to do some configuration:

1. *Select the serial baud rate (default 9600)*
2. *Select the communication mode: Common Mode or Compact Mode*
3. *Recording five instructions of the first group(or 2nd or 3rd as required)*
4. *Import the group you need to use (only recognize 5 instructions within one group at the same time)*

After all the setting above, you can speak or send voice instruction to it. If identified successfully, result will be returned via serial port in the format: group number + command number. For example, return ***Result: 11*** *(Compact mode returns 0x11) means identified the first command of group 1.*

If voice instruction is recorded, each time after you power it on, you need to import the group before letting it identify voice instructions.

LED

Recording stage:

1. *Record indication: D1 (RED) flashes 3 times within the 600ms, then off for400ms, and then flashes quickly for 4 times within 600ms. Now the recording indication is over.*
2. *Begin to speak: D1 (RED) is off for 400ms, and then is on. Voice during the time while D1 (RED) is on will be recorded by this module.*
3. *Recording a voice instruction successfully for the first time: D1 (RED) off, D2 (ORANGE) on for 300ms.*
4. *Recording a voice instruction successfully for the first time: D1 (RED) off, D2 (ORANGE) on for 700ms.*
5. *Recording failure: D2 (ORANGE) flashes 4 times within the 600ms. In cases that voice instructions detected twice don't match, or the sound is too large, or there is no sound, recording will fail. You need to start over the recording process for that instruction.*

Waiting mode:

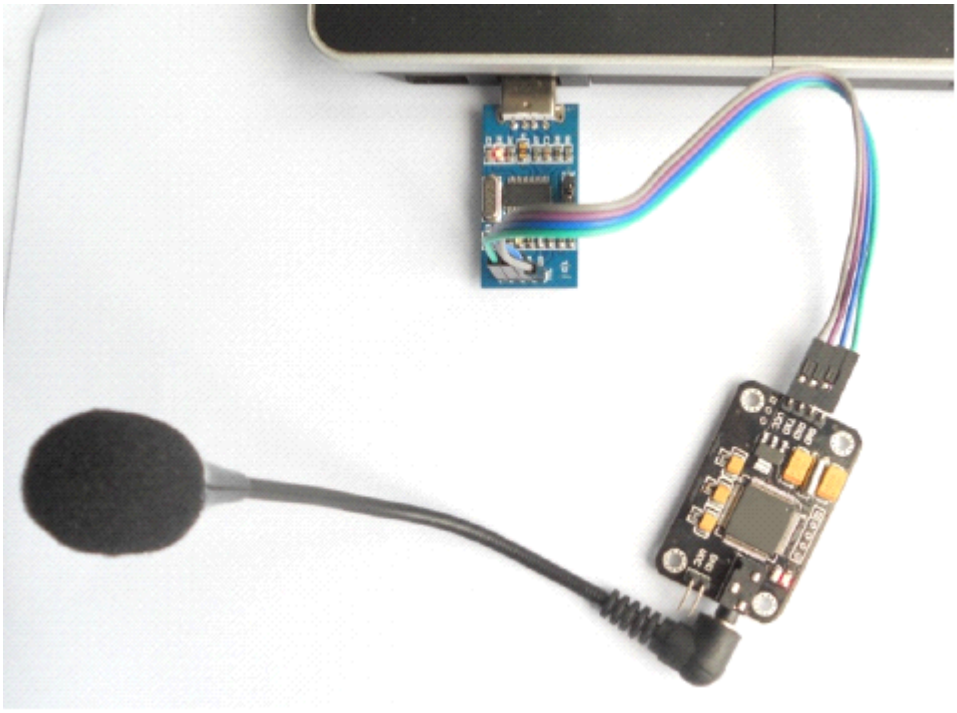
In waiting mode, D2 (ORANGE) is off, and D1 (RED) is on for 80ms every other 200ms, fast flashing. In this mode, it doesn't recognize voice command, only waiting for serial commands.

Recognition stage:

In identification stage, D2 (ORANGE) is off, and D1 (RED) is on for 100ms every other 1500ms, slow flashing. In this stage, this module is processing received voice signal, and if matching, it will send the result immediately via serial port.

Recording

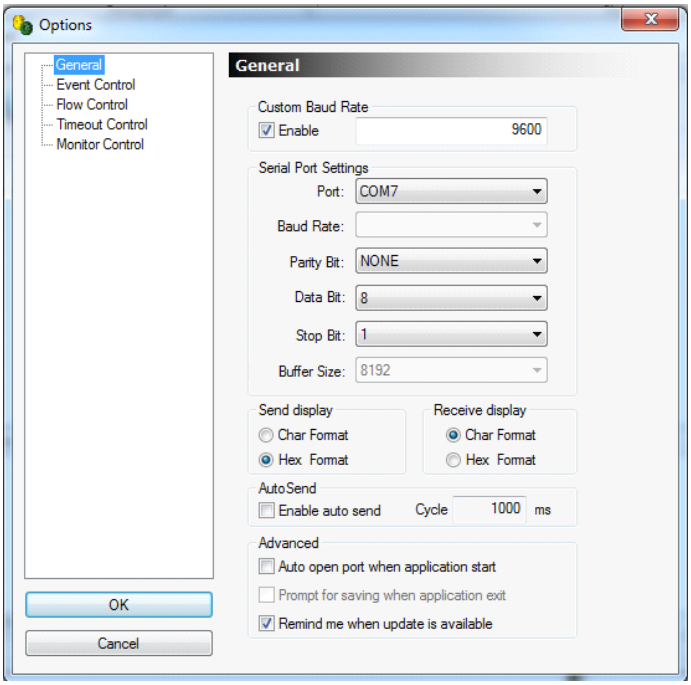
Before using it, we have to record voice instructions. Each voice instruction has the maximum length of 1300ms, which ensures that most words can be recorded. Once you start recording, you can't stop the recording process until you finish all the 5 voice instructions recording of one group. Also, once you start recording, the previous content of that group will be erased. In recording stage, this module doesn't reply to any other serial commands.



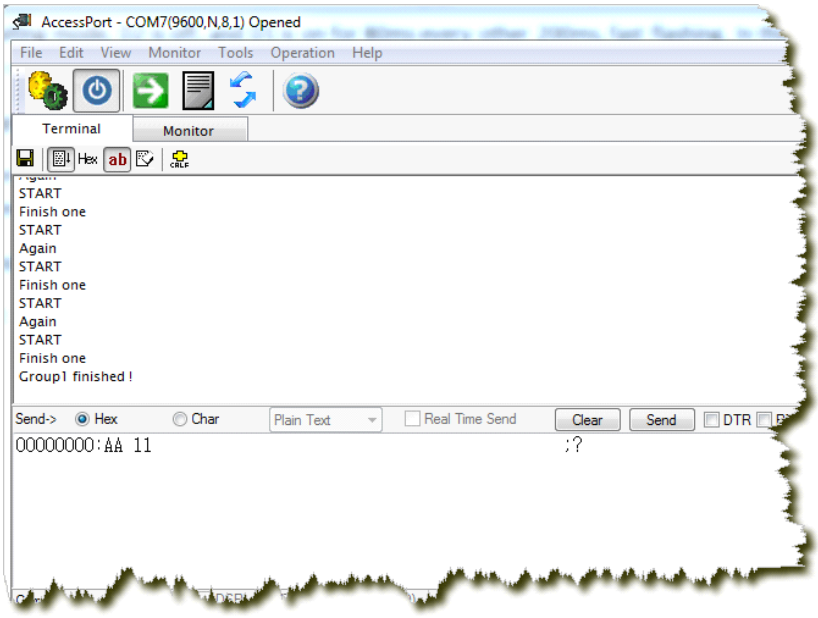
LED will flash to indicate state. Please refer to the LED part.

First, you need a serial tool. Here we use AccessPort (Download page).

1. *Serial port setting:*
    - *Baud rate: 9600*
    - *Parity bit: None*
    - *Data bit: 8*
    - *Stop bit: 1*
    - *Send format: Hex*
    - *Receive format: Char*



2. *Send commands*



Send: ***0xaa11***

Receive (in Common Mode):

START

No voice *// I did not make any sound. So it replied such message*

START

Speak now

Again

START

Speak again now

Different *// I spoke another words for the second time. So it replied such message*

START

Speak now

Again

START

Speak again now

Finish one *// recording one instruction successfully*

START

Again

START

Finish one

START

Again

START

Finish one

START

Again

START

Finish one

START

Again

START

Finish one

Group1 finished ! *// recording group 1 successfully*

By now, you've finished instruction recording for group 1.

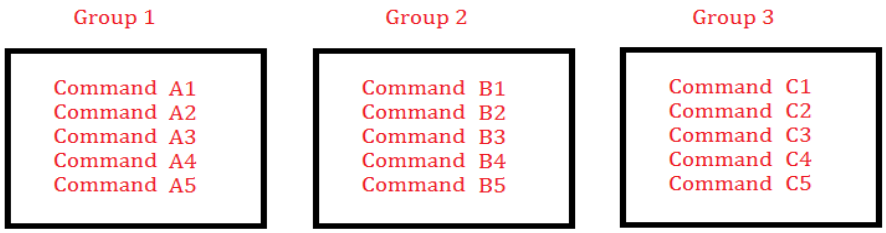
Recognition

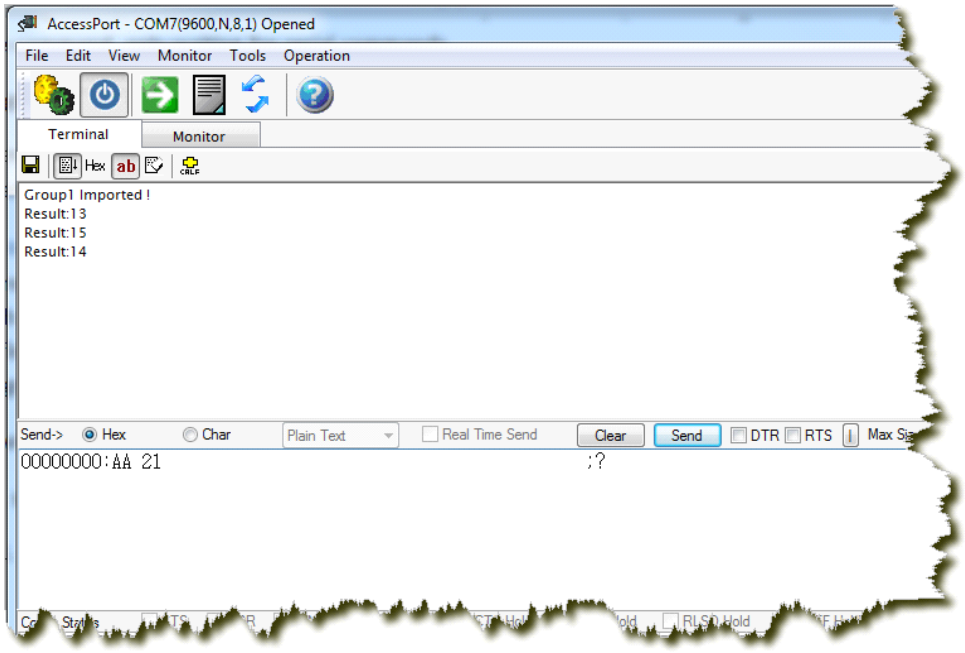
Send command ***0xAA21*** *to import group 1.*

Send command ***0xAA22*** *to import group 2.*

Send command ***0xAA23*** *to import group 3.*

*This command can recognize 5 voice instruction of a group at the same time. It could have 15 voice instructions in 3 groups. Each time you need to import the group before it could recognize instructions in that group. In recognition stage, this module could receive other serial commands. It will not exit the recognition stage until you send **0xAA00**, or delete that group, or begin recording instructions.*





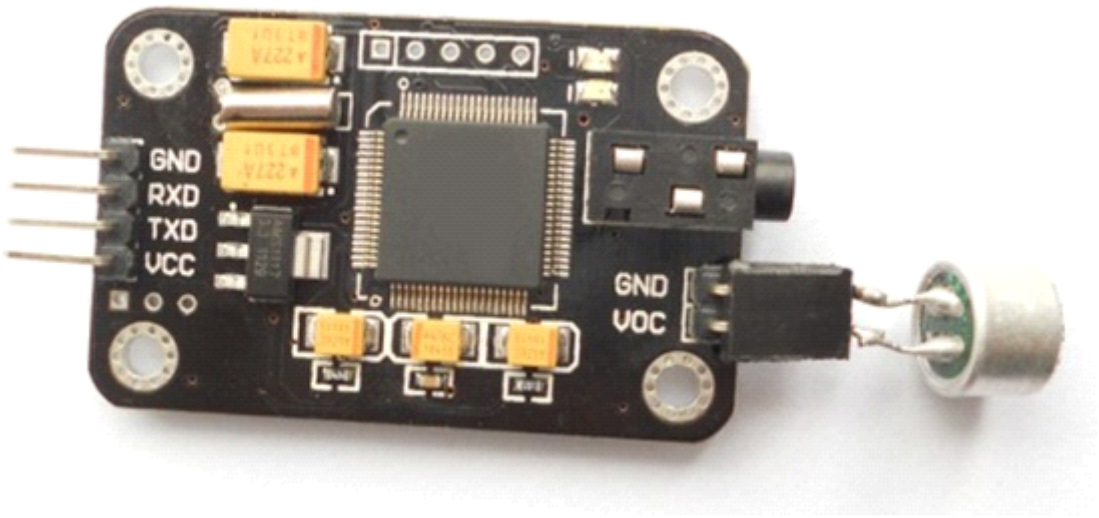
Video: http://www.youtube.com/watch?v=0Zjdu_HSr00

There are some tricks to improve the recognition rate.

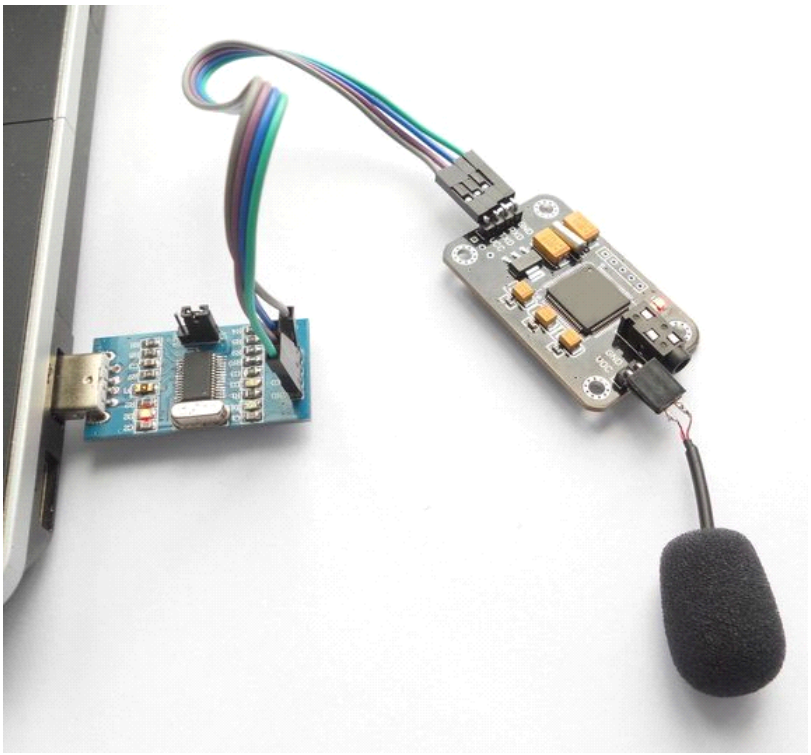
Try to record and use it in the same environment. And noise could greatly reduce the recognition rate. MIC plays very important role. As for the MIC wire connection to this module, the shorter, the better.

In this test, I recorded 5 voice instructions: white, red, green, blue, off.

First time I use a simple MIC as follows:



I found occasionally this module could not distinguish "red" and "green". Then I changed the MIC as follows:



Then the recognition rate is greatly improved.

Also I find that it is a little hard for this module to distinguish "of" and "off". To distinguish those two similar pronunciations, high-sensitive MIC is required and speaking must be very clear, as well as in quiet environment. All those limitation makes low practicality. But there are simple ways to solve it. Make your voice instructions distinctive. Do not use voice instructions of similar pronunciation.