



2.8 Second LOG-PCM Speech

Features

- Operating voltage: 2.4V~5.0V
- Directly drives an external transistor
- Low standby current (1 μ A Typ. for VDD=3V)
- Minimal external components
- 124 words table ROM for key functions
- Programmable silence length and end-pulse width (minimal end-pulse width is 330 μ s at a 6kHz sampling rate)
- 2.8-second voice capacity
- Controllable volume
- FLAG1 options:
 - End-pulse output
 - 3Hz flash
 - 6Hz flash
 - Voice output indication
 - Busy output
- FLAG2 options:
 - 3Hz flash
 - 6Hz flash
 - Busy output
- Four keys
- Key options:
 - Stop key: KEY4
 - Random (only for KEY1)
 - Sequential (only for KEY1)
 - Repeat (for all KEYs, under retriggerable condition)
 - Key debounce time (for all KEYs): 700 μ s, 22ms, 45ms, 180ms (based on a sampling rate of approximately 6kHz)
 - One shot (for all KEYs)
 - Level-trigger
 - Pull-high resistance (for all KEYs)
- Section options:
 - Retriggerable
 - Non-retriggerable

Applications

- Toys
- Alarm clocks
- Public address system
- Alert & warning system
- Sound effect generators

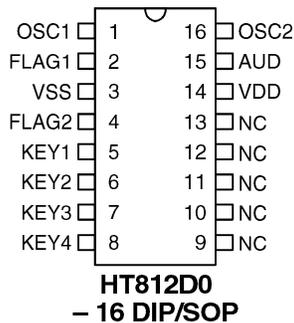
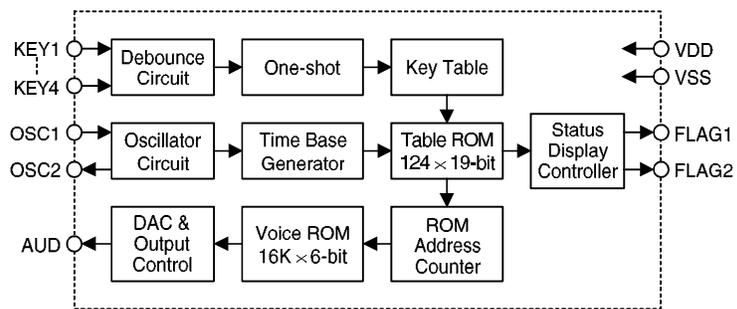
General Description

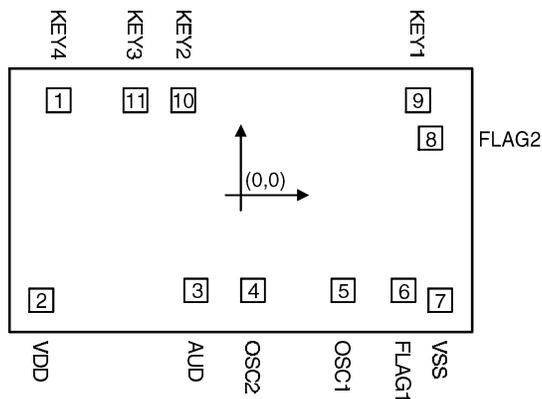
The HT812D0 is a single chip LOG-PCM voice synthesizer LSI with 2.8-second voice capacity at 6kHz sampling rate. The chip when triggered drives a speaker through an external transistor with a current switch D/A converter output. Negligible current will be consumed in the standby state.

The HT812D0 provides four key inputs and two programmable FLAG outputs. With a 2.4V~5.0V power supply, a complete synthesized voice playback system can be easily built with very few external components.

The customer's voice sources are recorded section by section into an internal mask ROM. The sectional playback arrangement instructions of each key are stored in the table ROM. The key features are also programmable. With such a flexible structure, the HT812D0 is excellent for versatile voice applications.



Pin Assignment

Block Diagram

Pad Assignment

 Unit: μm


Pad No.	X	Y
1	-919.00	480.30
2	-1007.00	-530.30
3	-226.99	-480.30
4	61.70	-480.30
5	515.20	-480.30
6	821.90	-480.30
7	1006.90	-530.30
8	957.00	289.10
9	894.20	480.30
10	-291.90	480.30
11	-532.50	480.30

 Chip size: $2270 \times 1380 (\mu\text{m})^2$

* The IC substrate should be connected to VSS in the PCB layout artwork.

Pin Description

Pin No.	Pin Name	I/O	Internal Connection	Description
1	OSC1	I	—	Oscillator input pin
2	FLAG1	O	NMOS Open Drain	3Hz/6Hz flash output, busy output, end-pulse or voice output indication (by mask option). Open drain, active low output
3	VSS	I	—	Negative power supply (GND)
4	FLAG2	O	NMOS Open Drain	3Hz/6Hz flash output or busy output (by mask option), open drain, active low output

Pin No.	Pin Name	I/O	Internal Connection	Description
5~8	KEY1~ KEY4	I	Pull-High	Trigger key, active low. Key features like debounce time, pull-high resistance and repeat can be selected by mask option.
9~13	NC	—	—	No connection
14	VDD	I	—	Positive power supply
15	AUD	O	PMOS Open Drain	Voice output for driving an external transistor
16	OSC2	O	—	Oscillator output pin

Absolute Maximum Ratings*

Supply Voltage -0.3V to 6V Storage Temperature -50°C to 125°C
 Input Voltage $V_{SS}-0.3V$ to $V_{DD}+0.3V$ Operating Temperature -20°C to 70°C

*Note: These are stress ratings only. Stresses exceeding the range specified under “Absolute Maximum Ratings” may cause substantial damage to the device. Functional operation of this device at other conditions beyond those listed in the specification is not implied and prolonged exposure to extreme conditions may affect device reliability.

Electrical Characteristics

(Ta=25°C)

Symbol	Parameter	Test Conditions		Min.	Typ.	Max.	Unit
		V _{DD}	Conditions				
V _{DD}	Operating Voltage	—	—	2.4	—	5.0	V
I _{DD}	Operating Current	3V	No load, f _{OSC} =96kHz	—	200	400	μA
I _{STB}	Standby Current	3V	—	—	1	3	μA
I _O	Max. AUD Output Current	3V	V _{OH} =0.6V	-1.5	-2	—	mA
I _{OL}	FLAG Sink Current	3V	V _{OL} =0.3V	1.5	3.0	—	mA
V _{IH}	“H” Input Voltage	—	—	0.8V _{DD}	—	V _{DD}	V
V _{IL}	“L” Input Voltage	—	—	0	—	0.2V _{DD}	V
f _{OSC}	Oscillating Frequency	3V	R _{OSC} =570kΩ	60	76	92	kHz

Functional Description

The HT812D0 is a mask ROM type voice synthesizer with 2.8-second voice capacity. A group of pre-recorded voice sections is played upon receipt of key trigger input signals. Two output FLAGS send signals while playing voices.

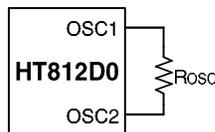
The 2.8-second voice capacity can be divided into sections of arbitrary length. Notice that the silence length and end-pulse width are not included in the memory.

By using HOLTEK's programming tools, the contents and arrangement of sections, key features and FLAG output are all programmable before device fabrication.

The IC provides four key inputs (KEY1~KEY4). Of the four keys, KEY1 can be optioned as a direct, sequential or random trigger key. KEY4 can be selected as a stop or direct key. The remaining two keys (KEY2, KEY3), are used as direct keys exclusively.

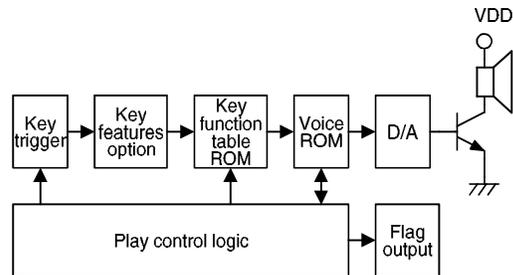
System oscillator

The HT812D0 has a built-in RC oscillator which requires only one external resistor for normal applications. The oscillator frequency is typically 96kHz for an external resistor of 530kΩ. The required oscillator frequency may vary with different sampling rates in the process of voice programming. As a result the value of the oscillator resistor may be different for different items.



The oscillator is turned on when triggered by a key input. After playing, the oscillator is immediately turned off and the chip goes into the standby state.

Play function block diagram



Voice ROM

The voice ROM is originally designed to continuously record the 2.8-second voice data at about 6kHz sampling rate. A higher sampling rate can generate voices of good playback quality, but will shorten the total recording time. On the other hand, a lower sampling rate will result in longer recording time but sacrifice voice quality.

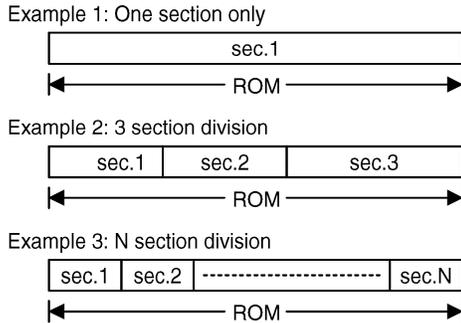
The playback time can be significantly extended by making use of coding efficiency, silence playing, section repeating, section cascade, etc.

Section

Section is the basic element constituting the contents of the voice ROM. During programming, the customer's voice sources can be divided into as many sections as required. A section can be composed of a voice or an interval of silence. However, the silent length is not counted in the voice ROM. The total number of sections included should be less than 124 due to the space limitation of the function table ROM. The total length of the included sections is limited by the voice ROM.

A section, when triggered by a key input, can be played once, repeatedly or cascaded with other sections, depending on the key function table instructions.

The following are examples of section division:



In addition, a section can be set as retriggerable or non-retriggerable depending on the code option.

- **Retriggerable**
When the currently playing section is set as retriggerable, it will stop immediately upon receipt of other key inputs.
- **Non-retriggerable**
When the currently playing section is selected as non-retriggerable, it will keep on playing till the whole section is completed, whether or not there is a key input during voice playing.
For a key group, some sections can be set as retriggerable and some as non-retriggerable. When a retriggerable section of a key group is playing, any key can be triggered to interrupt its playing. On the other hand, if a non-retriggerable section is playing, any key interrupt is invalid.

Group

The HT812D0 plays groups according to the key input. A group can be made up of one or more sections. When a key is triggered, the corresponding group is played immediately. For example, triggering KEY2 plays group 2, and so on. The same section is allowed to appear in different groups. However, KEY1 can be made up of multiple groups when it is optioned as a sequential or random key. Otherwise, each key is composed of one group only.

Key function table

Sections in the voice ROM are played according to the instructions of the key function table. The function table contains group information and the playing order of sections in the groups. Notice that the total amount of sections included in the groups should be less than 124—the space limitation of the function table ROM.

- **KEY1 as a direct key**
Each key is mapped to a group in the function table. If a key is not used, the group mapped to that key is a piece of silence. The following is an example of the function table:

Group 1	sec.1 + sec.2 + sec.3 + sec.5
Group 2	sec.3
Group 3	sec.2 + sec.2 + sec.3 + sec.4
Group 4	sec.5 + sec.3

As illustrated on the above table, the voice ROM is composed of 5 sections, and as a whole comprise 11 sections for the function table. If KEY1 is momentarily triggered, section 1, section 2, section 3 and section 5 are played in sequence and then stopped. Triggering KEY2 plays section 3, and so on.

- **KEY1 as a sequential or random key**
When KEY1 is optioned as a sequential or random key, KEY1 can include multiple groups (sub-groups) in the function table. However, the remaining 3 keys (KEY2~KEY4) are used as direct keys exclusively and comprise only one group in the function table. An example is shown below:

Group 1-1	sec.4 + sec.2
Group 1-2	sec.1 + sec.3
⋮	⋮
Group 1-N	sec.2 + sec.3
Group 2	sec.2 + sec.3
Group 3	sec.3 + sec.5
Group 4	sec.1 + sec.5 + sec.2

As stated on the table, KEY1 can be made up of sub-groups. The corresponding sub-groups is played in sequence each time KEY1 is triggered.

- The playing sequence of sequential KEY1 is:
Group 1-1 → Group 1-2 → Group 1-3 → Group 1-N (the last group) → Group 1-1
- The playing sequence of random KEY1 is:
Group 1-3 → Group 1-5 → Group 1-N → Group 1-3 → Group 1-5

That KEY1 functions as a random key is a special case of sequential key, which combines a particular arrangement of sub-group playing sequence.

- Reset of KEY1 playing sequence

If a sub-group of KEY1 group is being played and one of the keys from KEY2~KEY4 is triggered, the playing sub-group will be terminated, and the newly triggered key group is played instead. The first sub-group will start playing by retriggering KEY1. In other words, the KEY1 playing sequence is reset whenever a key other than KEY1 is triggered (see Figure 1).

- Sub-group selection

When KEY1 is triggered with pulses, the desired sub-group can be selected by controlling its corresponding pulse number.

The features of KEY1 have to be set in the

following ways:

- sequential or random
- retriggerable
- minimum key debounce time
($\approx 700\mu s$, $f_{OSC}=96kHz$)

For instance, if sub-group 1-3 is the previous playing group, sub-group 1-5 will start playing after 2 pulses are applied to KEY1, and so on.

To make easy selection of KEY1 sub-groups, one of KEY2~KEY4 should be programmed as silence. Then this silence key has to be triggered to reset KEY1. By so doing, the playing sub-group of KEY1 is directly specified by the pulse number applied to KEY1 (see Figure 2).

- KEY4 as a stop key (by mask option)

When KEY4 functions as a stop key, any voice output can be stopped by pressing KEY4.

- KEY1~KEY3 as a repeat key

KEY1~KEY3 all function as repeat key if one of the three keys is set as a repeat key. In other words, once the mode of one of KEY1~KEY3 is determined, the remaining two keys are set accordingly.

As a repeat key, the sections included can be played sequentially and repeatedly till other trigger is applied. KEY4 functions as a stop key when KEY1~KEY3 are set as repeat keys.

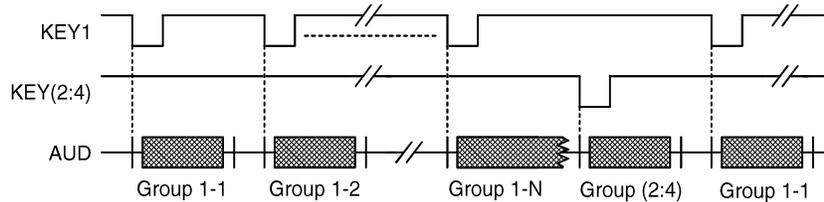


Figure 1 Reset of KEY1 playing sequence

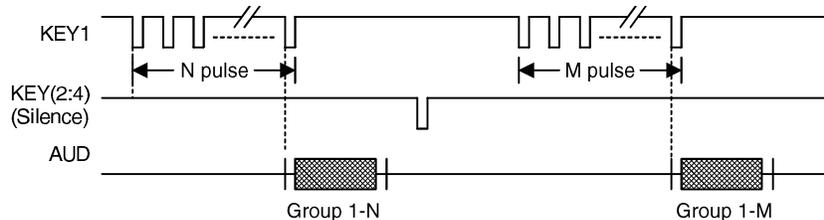


Figure 2 KEY1 sub-group selection

When KEY1~KEY3 are set as repeat keys, the key function should be optioned as retriggerable keys.

Key features

• Key priority

When two or more keys are triggered simultaneously, the output voice is decided by the key priority as shown below:

KEY1>KEY2>KEY3>KEY4

• Key debounce time

There are four kinds of key-in debounce time to be selected by mask option, namely, 700µs, 22ms, 45ms and 180ms. The key debounce time varies with the value of the system frequency.

• Pull-high resistance

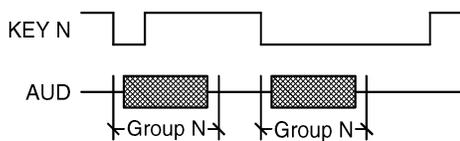
Four kinds of key input pin pull-high resistance can be selected by mask option, namely, 20kΩ, 50kΩ, 100kΩ and 200kΩ. The resistance may vary with VDD, temperatures and the chip itself due to process variations.

• Trigger mode

All of the four keys can be optioned as one-shot trigger mode or level-trigger mode.

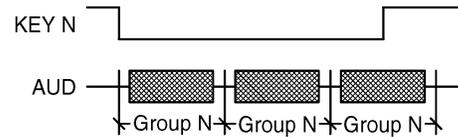
• One shot

When one of the four keys (KEY1~KEY4) is pressed and held down, the group corresponding to that held key will play once. The corresponding group will also play once when one of the four keys is pressed momentarily.



• Level trigger

When one of the four keys is pressed and held down, the corresponding group will keep playing. Once the pressed key is released, the group will not stop till the included sections are all completed.



FLAG

When voices are playing, both FLAG1 and FLAG2 pins are activated to output one of the following signals through code option.

FLAG1 can be optioned as one of the following signal outputs:

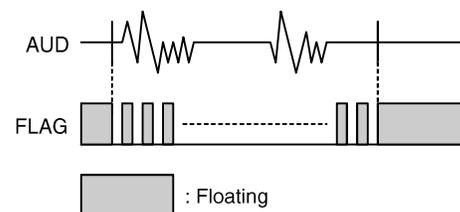
None, 3Hz flash, Busy, 6Hz flash, Voice indicator, or End-pulse output

FLAG2, on the other hand, can be set as one of the following signal outputs:

None, 3Hz flash, 6Hz flash, or Busy output

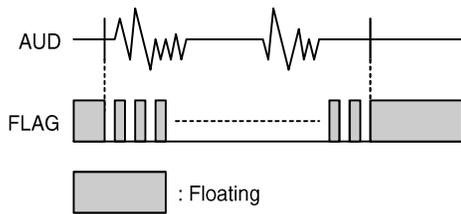
• 3Hz flash

When voices are playing, FLAG1 as well as FLAG2 pin outputs a 3Hz signal to drive an LED. The signal is active low, 25% duty. Once the voice output is terminated, FLAG1 and FLAG2 pins become floating outputs. When FLAG1 and FLAG2 pins are optioned as 3Hz output, they will be alternately signaled at a 3Hz rate output.



• 6Hz flash

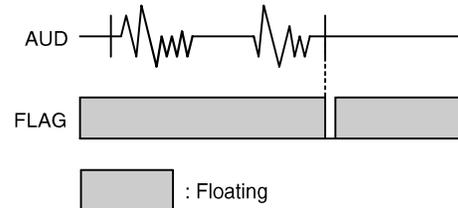
When voices are playing, the FLAG1 pin outputs a 6Hz signal to drive an LED. The signal is active low, 25% duty. Once the voice output is terminated, the FLAG1 pin becomes a floating output. When the FLAG1 and the FLAG2 pins are optioned as 6Hz output, they will be alternately signaled at a 6Hz rate output.



• End-pulse output

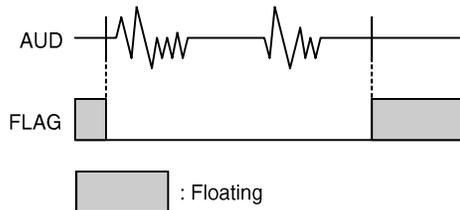
When the voice output is completed, the FLAG1 pin outputs an active low pulse. The pulse width can be programmed depending on the customer's requirements.

FLAG1 and FLAG2 pins are both floating outputs when the chip is in the standby state.



• Busy output

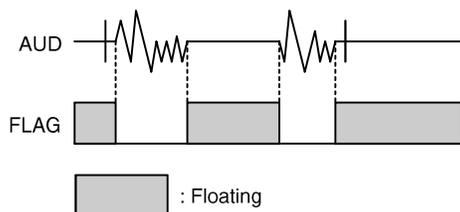
When a voice group is playing, the outputs of both FLAG1 and FLAG2 are turned low, indicating that the chip is busy.



In addition to the above-stated output signals, FLAG1 can also generate one of the following signals by code option:

• Voice indicator output

FLAG1 is active low when voices are playing. It is also turned low during a voice section output. FLAG1 becomes a floating output after the silence section output is completed or the voice output is terminated.



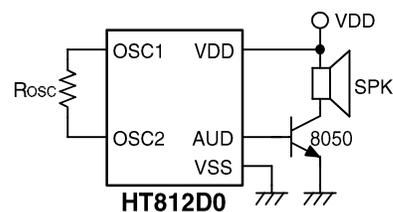
Volume control

The function of the volume control can be set by mask option. A code is written in the function table for the purpose of controlling the volume of each section output after the volume control function is chosen. There are two volume options, namely, full range and half range.

AUD

The AUD pin is a PMOS open drain structure. It outputs voice signals to drive a speaker through an external NPN transistor when the chip is active. However, the AUD pin becomes a floating output when the chip is in the standby state.

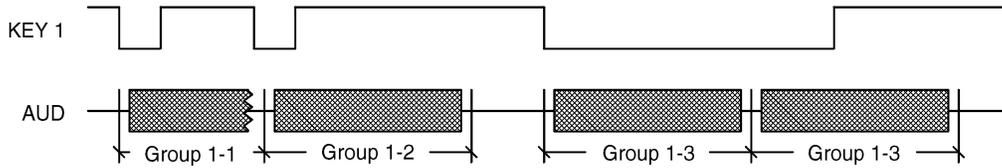
The 8050 type transistor with $h_{FE} \approx 150$ is recommended for an output driver.



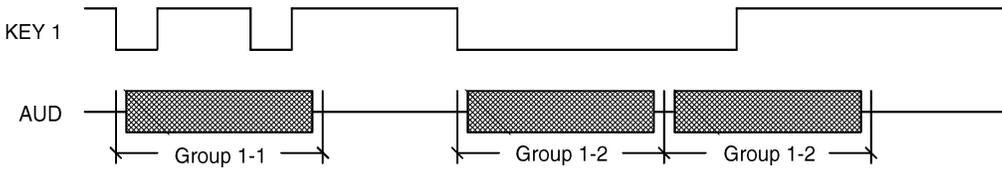
Timing Diagram

One-key operation

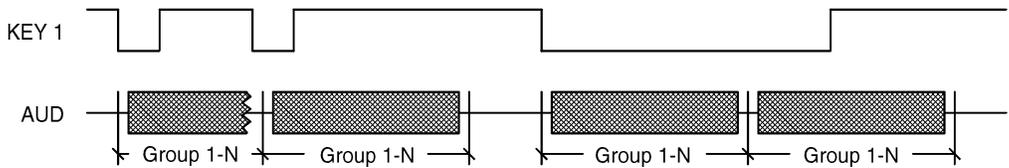
- Sequential-retriggerable



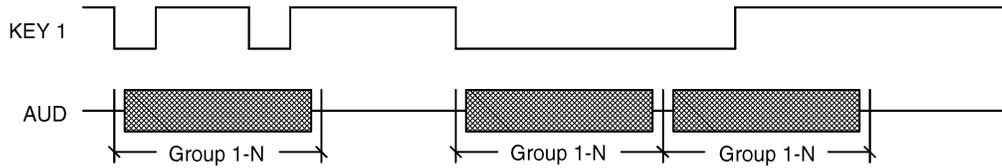
- Sequential-nonretriggerable



- Random-retriggerable



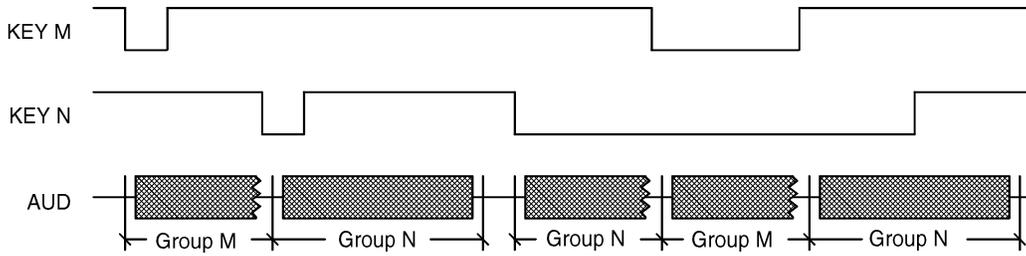
- Random-nonretriggerable



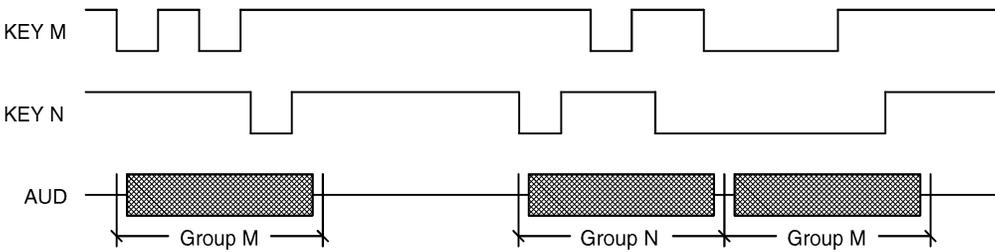
Note: Group 1-N may be any of the KEY1 groups

Multi-key operation

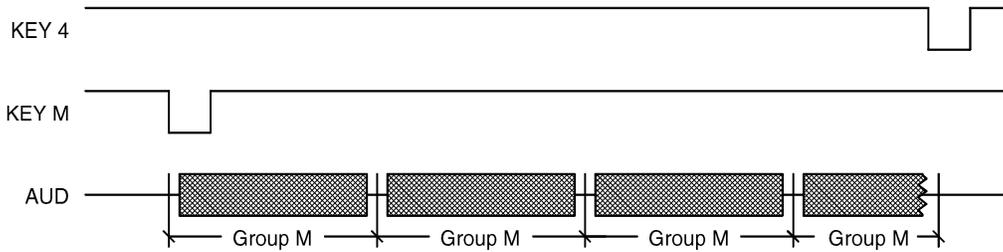
• Retriggerable



• Non-retriggerable



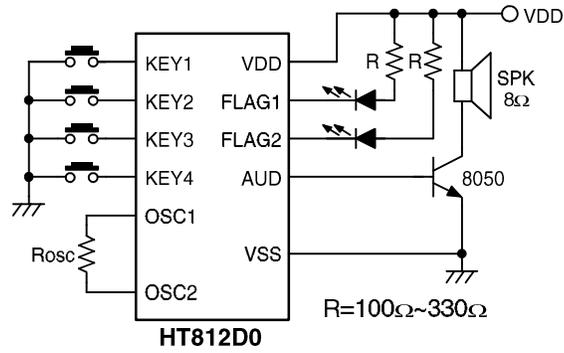
• Repeat



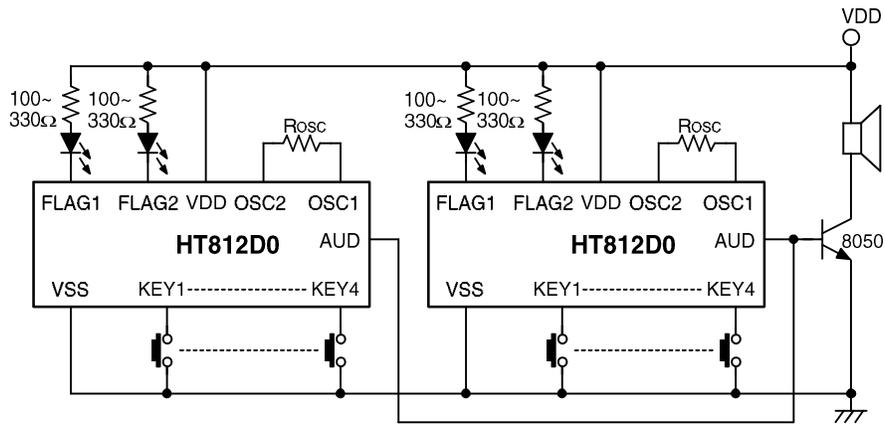
Note: M, N=1:4

Application Circuits

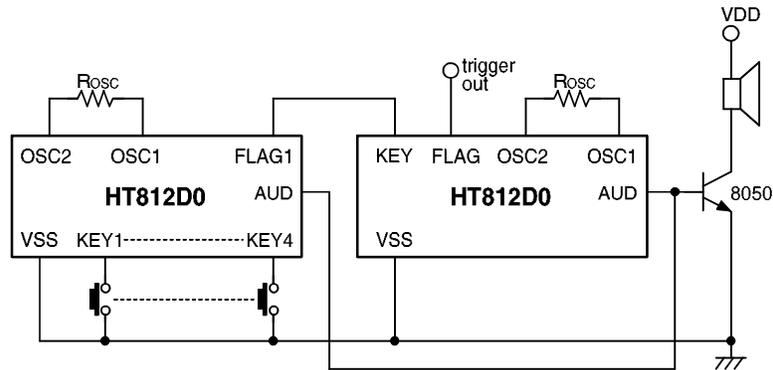
General application



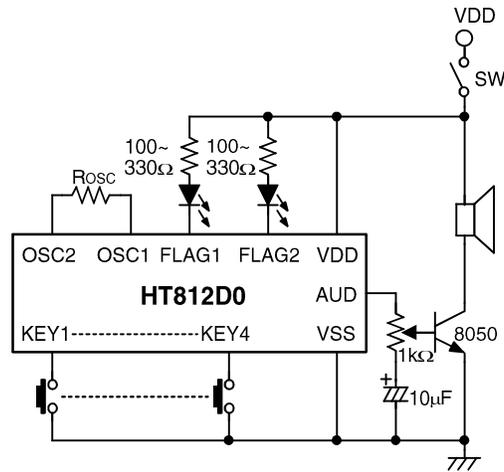
Parallel application



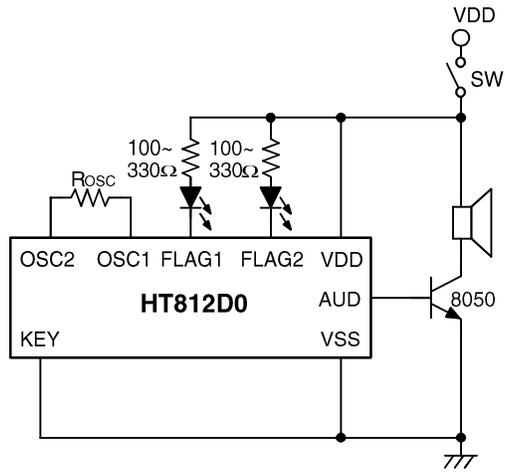
Cascade or external driving



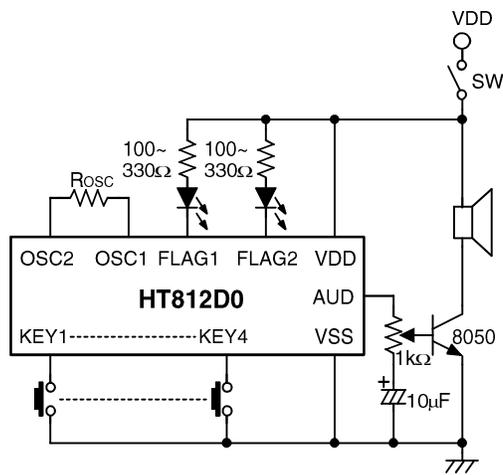
Power-on play (one shot)



Power-on play (continuous)

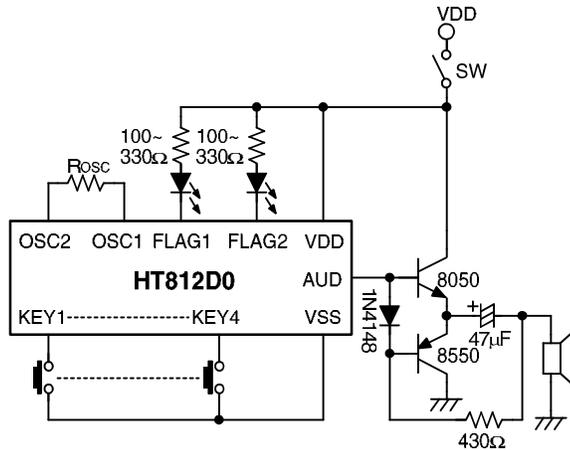


With volume adjustment

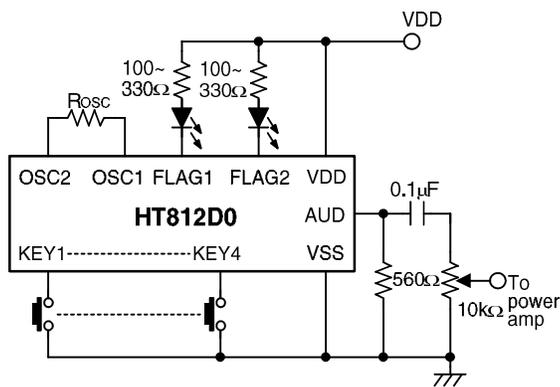


Push-pull output

To prevent the speaker and driver transistor from damage due to excess power dissipation which results from a high voltage power supply (4.5V~5.5V), the following push-pull output stage is recommended.



Coupling to power amplifier



Standard Item List

Item	Name	Rosc	Remarks
HT812D1	Baby Cry	532kΩ	4 keys
HT812D2	Firecracker	370kΩ	4 keys
HT812D3	Train Sound (I)	570kΩ	4 keys
HT812D4	Crazy Laugh	787kΩ	4 keys
HT812D5	Witch Laugh	469kΩ	4 keys
HT812D6*	Glass Broken	—	4 keys
HT812D7	Scream Sound	419kΩ	4 keys
HT812D8	Racing Car Sound	572kΩ	4 keys
HT812D9	Truck Sound	450kΩ	4 keys
HT812DD	Duck	397kΩ	4 keys
HT812DF*	Monkey	—	4 keys
HT812DG	Horse	570kΩ	4 keys
HT812DH	Cat	427kΩ	4 keys
HT812DJ	Lion	472kΩ	4 keys
HT812DK	Elephant	363kΩ	4 keys
HT812DL	Sheep	264kΩ	4 keys
HT812DN*	Train Sound (II)	—	4 keys

* Under development