NEC

USER'S MANUAL
μPD7755
μPD7756

NEC ELECTRONICS (EUROPE) GMBH

USER'S MANUAL

µPD7755

µPD7756

3/85 V1.0

Table of Contents

CHAPTER 1 OUTLINE

The μPD7755 /7756 is a speech synthesis LSI that employs the waveform coding method. By combining the ADPCM coding method and phoneme method, natural synthesized speech has been realized. Also, the wide range of the operating voltages, compact package, standby function, etc., make these LSIs ideal for various application systems such as battery-driven systems.

1.1 Features

Synthesizing method	:Combined ADPCM and phoneme methods
Sampling frequency	:4, 5, 6 or 8kHz
Bit rate (speech)	:8 to 32kbps
Built-in speech data ROM	:96K bits (μPD7755) 256K bits (μPD7756)
Synthesizing time	:12 sec. max. (μPD7755) 32 sec. max. (μPD7756)
Built-in D/A converter	:9-bit resolution current output
Built-in standby function	:Popcorn noise suppressor incorporated
Supply voltage	:2.7 to 5.5V
Process	:CMOS
Package	:18-pin plastic DIP

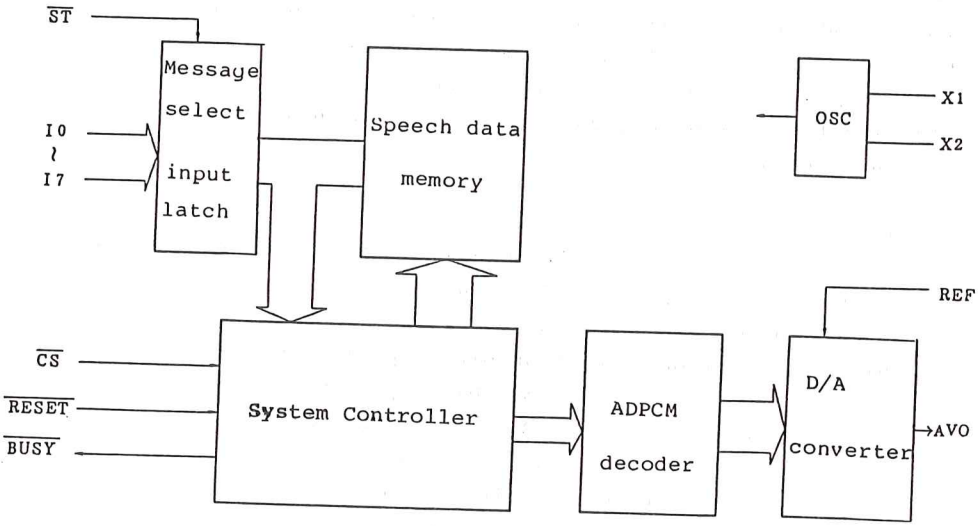
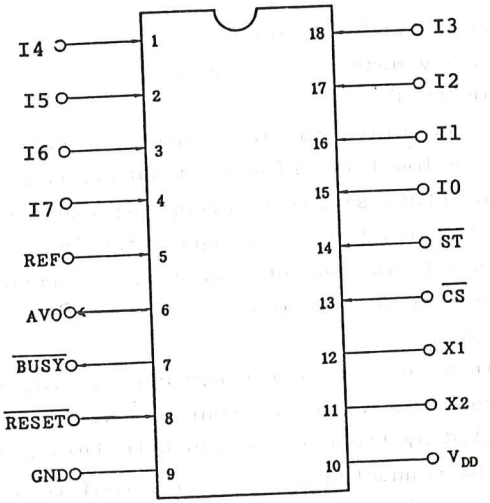


Fig. 1-1 System Configuration

CHAPTER 2 PIN FUNCTIONS

2.1 Pin Configuration (Top View)



2.2 Pin Functions

(1) I0-I7 (Input)

These eight pins together function as the message select input. Positive logic is used for these pins (high level=1). Each input is provided with an input latch that latches at the rising edge of the $\overline{ST}$ input. During standby mode, these pins should be set to either high or low level; If they are biased at or near the typical CMOS swith input,causing excess current drain.

(2) $\overline{CS}$ (Input)

This is the chip select input pin. When a low level signal is input to this pin, the $\overline{ST}$ input is enabled.

(3) $\overline{ST}$ (Input)

This is the start signal input pin. When a low level signal is input to this pin when the $\overline{CS}$ signal is also low level, speech synthesis of the message stored in the speech ROM location addressed by the contents of IO-I7 begins.

When the $\overline{CS}$ and $\overline{ST}$ signals are both set to low level, the standby mode is released.

(4) $\overline{BUSY}$ (Output)

This is an active low level pin that outputs the $\overline{BUSY}$ signal. A low level signal is output from this pin when the start signal is accepted; and once this output becomes low level, the start input will not be accepted. The output from this pin becomes high impedance in standby mode.

(5) REF (Input)

This pin inputs the reference current for the D/A converter. The output current of the D/A current is controlled by the current input to this pin. This pin should be connected to V_{DD} via a resistor. In standby mode, this pin becomes high impedance.

(6) AVO (Output)

This pin outputs the analog speech signal, which is a unipolar, sink current.

(7) $\overline{RESET}$ (Input)

This is the reset signal input pin and is used to initialize the LSI at power up, abort the speech synthesis, and release the standby mode. To reset the chip, this signal must be held low for at least 12 oscillator clocks. When the standby mode is released, at least 12 clocks must be input after clock oscillation completes.

(8) X1, X2

These are the clock pins and are connected to a ceramic resonator (640kHz). During standby mode, the X1 becomes low level and the X2 becomes high level.

(9) V_{DD}

This pin should be connected to the power supply.

(10) GND

This pin should be connected to ground.

CHAPTER 3 OPERATION

3.1 Sample Frequency

The relation between synthesized speech message length and bit rate when the ADPCM method is used is shown in Table 3-1.

Table 3-1 Sampling Frequency and Maximum Message Length

Sampling frequency (kHz)	Bit rate (kbps)	Message length (sec.)	
		μPD7755	μPD7756
4	16	6	16
5	20	5	12
6	24	4	10
8	32	3	8

For normal speech with no background music,about 10% to 20% of the message is made up of silent frames. These frames are compressed in the case of the μPD7755 /7756 so the actual bit rate is about 80% to 90% of that shown in Table 3-1.

Even further reductions in the bit rate have been achieved by combining the ADPCM and phoneme methods. The bit rates for speech synthesis when this combined method is used are shown in Table 3-2.

Table 3-2 Sampling Frequency and Maximum Message Length (with data compression)

Sampling frequency (kHz)	Bit rate (kbps)	Message length (sec.)	
		μPD7755	μPD7756
4	8	12	32
5	10	9	25
6	12	8	21
8	16	6	16

3.2 D/A Converter

The built-in D/A converter is a unipolar, current output type with a 9-bit resolution. The digital input to the converter is converted to a 9-bit offset binary code. This code has a range of 0 to 1FFH with a midrange value of 100H.

The schematic drawing of the D/A converter is shown in Fig. 3-1. As can be seen, a constant-current source is connected in parallel for each bit of the digital input.

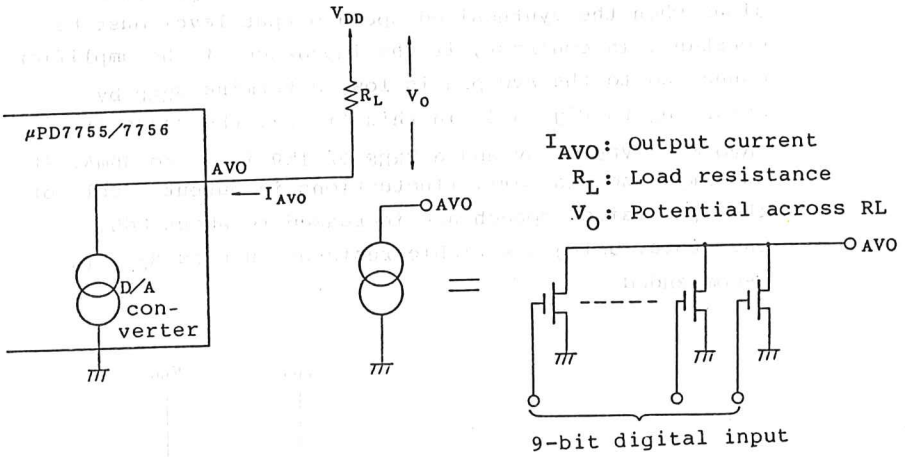


Fig. 3-1 Schematic Diagram of D/A Converter

3.2.1 Output Current and Reference Resistance

The output current of the D/A converter can be controlled by the voltage applied to the REF pin. This is realized by the resistor connected across the V_{DD} and REF pins as shown in Fig. 3-2.

The relations between the voltage to the REF pin, V_{REF}, and input reference current I_{REF} or D/A converter output current I_{VAO} are shown in Figs. 3-3 and 3-4.

If the impedance of an amplifier to be connected to the AVO pin is as high as several kΩ or more, refer to Fig. 3-3 to determine the value of I_{REF}. In Fig. 3-3, the

range of I_{REF} is 60 (at the intersection of the load line, R_{REF}, and the MAX. curve) to 80μA (at the intersection of R_{REF} and MIN.) when R_{REF} is 50kΩ and V_{DD} is 5V. Under this condition, the range of I_{AVO} is 1.9 to 2.9mA because I_{VAO} is 32 to 36 times that of I_{REF}. Also, if the impedance is high, fluctuations in the output level of the synthesized speech are about 4dB, which is not so high. Therefore, a highly precise fixed resistor can be used as R_{REF} instead of the variable resistor except for occasions when the synthesized speed output level must be constant. In contrast, if the impedance of the amplifier connected to the AVO pin is low, determine R_{REF} by referring to Fig. 3-4. In this figure, the range of I_{AVO} at a V_{DD} of 5V and a R_{REF} of 1kΩ is 12 to 30mA. If the impedance is low, fluctuations in output levels of the synthesized speech are increased to about 8dB. Therefore, using a variable resistor such as R_{REF} is recommended.

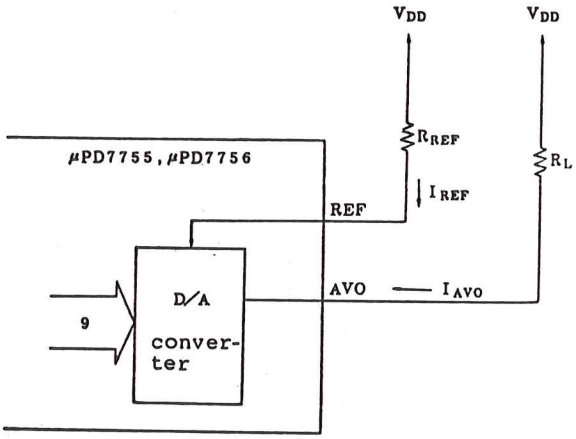


Fig. 3-2 D/A Converter Reference Current

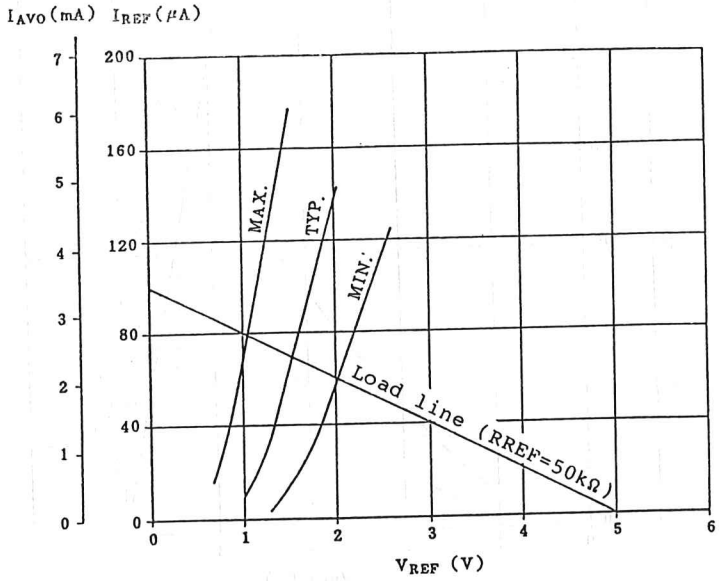


Fig. 3-3 Relation Between V_{REF} and I_{REF}/I_{AVO} (in Low-Current Area)

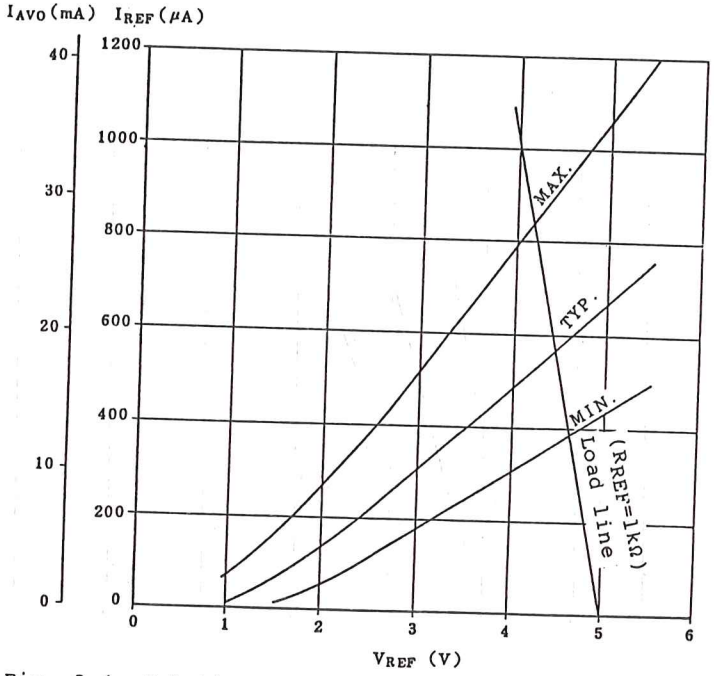


Fig. 3-4 Relation Between V_{REF} and I_{REF}/I_{AVO} (in High-Current Area)

3.2.2 Setting Output Current

Because the output current from the AVO pin (I_{AVO}) is a sink-load current, the load resistance (RL) is connected to V_{DD}. Therefore, the potential at AVO must be within the range at which the output transistor of the D/A converter will operate as a constant-current supply.

The operating voltage range is:

When V_{DD}=5V: 1.5V to V_{DD}

When V_{DD}=3V: 1V to V_{DD}

The load resistance and output current should be determined so that the operating voltage applied to the output transistor falls within this range.

To adjust output current I_{AVO}, observe the following procedure.

(1) Select the message to be synthesized (to allow the D/A converter to output the middle(bias) current).

(2) Input the $\overline{\text{RESET}}$ signal within 3 seconds after the synthesized speech has been output (to prevent the LSI from entering standby mode).

(3) Adjust R_{REF} so that the voltage applied to the I_{AVO} falls within the rated value.

3.3 Standby Mode

When the μPD7755 /7756 is not performing speech synthesis, it is set in standby mode. In this mode, current consumption is reduced to less than 1μA (TYP). The condition of the μPD7755 /7756 in standby mode is as follows:

- The clock stops
- The $\overline{\text{BUSY}}$ pin becomes high impedance
- The REF input pin becomes high impedance
- The X1 becomes low level and the X2 becomes high level.
- I0-I7, $\overline{\text{ST}}$, $\overline{\text{CS}}$, and $\overline{\text{RESET}}$ inputs are all enabled
- The I_{AVO} becomes 0.

3.3.1 Entering standby mode

The μPD7755 /7756 enters the standby mode if the following condition exists for more than 3 seconds after completion of speech synthesis:

(1) $\overline{\text{CS}}$ or $\overline{\text{ST}}$ is high level

(2) $\overline{\text{RESET}}$ is high level

3.3.2 Releasing standby mode

Standby mode is released by the following procedure:

(1) Set $\overline{\text{CS}}$ to low level

(2) Set $\overline{\text{ST}}$ to low level

At this time, the message select code will be input to I0-I7.

3.3.3 Eliminating popcorn noise in standby mode

Because the μPD7755 /7756 uses a unipolar 9-bit D/A converter in operation mode, there is a bias current output even when there is no signal input (AC signal). In standby mode, the input value to the D/A converter becomes 0 as does the output current.

Therefore, when the LSI moves from operating mode to standby mode, popcorn noise may be generated by the sudden changes in the output of the D/A converter. To prevent this, the output of the D/A converter is gradually reduced prior to entering the standby mode (see Fig. 3-5).

Also, when changing from standby to operating mode, the D/A output is gradually raised to AC 0 from the time when clock oscillation starts and completes in response to the standby release signal generated by input of the $\overline{CS}$ and $\overline{ST}$ signals (see Fig. 3-6).

The transition time for the D/A output is approximately 46.5ms and which enough to suppress this type of noise.

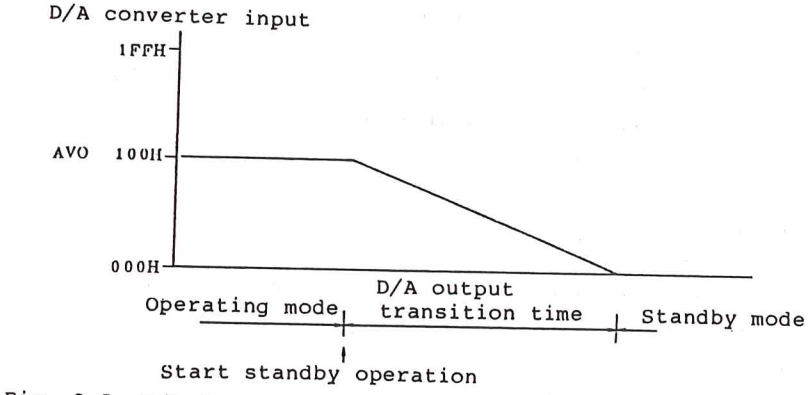


Fig. 3-5 D/A Converter Output when Entering Standby Mode

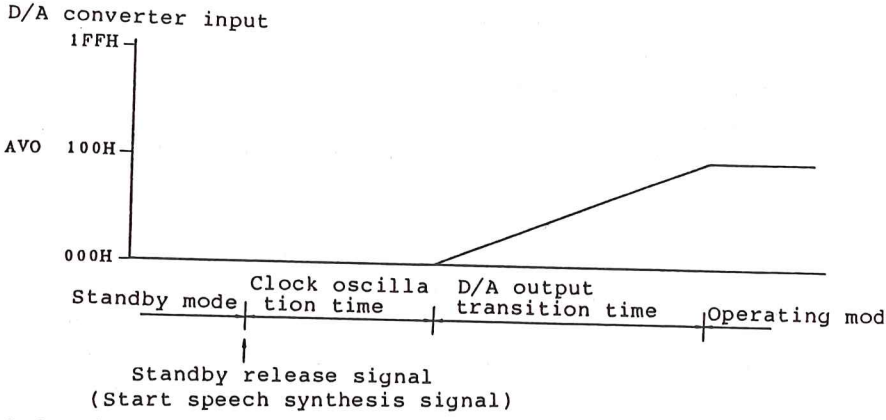


Fig. 3-6 D/A Converter Output when Releasing Standby Mode

3.3.4 Pin settings in standby mode

The following cautions should be observed in regard to the $\overline{\text{BUSY}}$ signal output and the message select code input (I0-I7) in standby mode.

(1) $\overline{\text{BUSY}}$ pin

Because the output from this pin becomes high impedance in standby mode in implementations where this signal is checked the external controller must be configured to insure that this signal is pulled up.

(2) I0-I7 pins

If these pins are not set to high or low level in standby mode, there is a chance that excess current will drain. These pins should therefore be set either to high level or low level during standby mode. If these pins are connected to a circuit (such as a bus) with a floating state, they must be pulled up or down.

3.4 Start Speech Synthesis

Speech synthesis begins when the signal input to the $\overline{\text{ST}}$ pin goes low when the signal at the $\overline{\text{CS}}$ is already low. Note that the operation performed when a pulse is applied to the $\overline{\text{ST}}$ pin is somewhat different from that when it is fixed at low level.

3.4.1 $\overline{\text{ST}}$ pulse input

The timing when the signal input to the $\overline{\text{ST}}$ pin is a pulse is shown in Fig. 3-8. The data on I0-I7 (message select data) is latched at the rising edge of the signal input to the $\overline{\text{ST}}$ pin. After this data is latched, there will be no effect on the speech synthesis operation even if both the $\overline{\text{CS}}$ and the $\overline{\text{ST}}$ signals are set to high level.
Since the speech synthesis start control circuit continues to operate during standby mode, the start procedure is the same in both standby and operating modes.

3.4.2 $\overline{ST}$ fixed input

When the $\overline{ST}$ input is fixed at low level, the speech synthesis operation is performed repeatedly (see Fig. 3-9). In this state, the message select code (the contents of I0-I7) are not latched so this data should not be changed. Changing this data may result in misoperation.

3.4.3 Standby mode

Since the speech synthesis start control circuit continues to operate during standby mode, the same procedure as that described in 3.4.1 can be used to start speech synthesis. In this case, the clock oscillation begins when both the $\overline{ST}$ and $\overline{CS}$ pins become low level and the $\overline{BUSY}$ signal is output after the clock oscillation has become stable. Then, as described in section 3.3, the speech output begins approximately 46.5ms later (See Fig. 3-10).

Table 3-2 Speech Synthesis Start Timing
(V_{DD}= +2.7 ~ 5.5V, f_{OSC}= 640KHz)

Parameter	Symbol	Condition	MIN.	TYP.	MAX.	Unit
ST pulse width	t_{CC}		2			μs
BUSY output time	t_{SBO}	Operation mode		6.25	10	μs
	t_{SBS}	Standby mode * (includes oscillation start time)		4	80	ms
Speech output start time	t_{SSO}	Operation mode		2.1	2.2	ms
	t_{SSS}	Standby mode		2.1	2.2	ms
D/A converter transition time	t_{DA}	Standby mode		46.5	47	ms

Note: Ceramic resonator Kyocera Corp. KBR-640B, C1=C2=150 pF

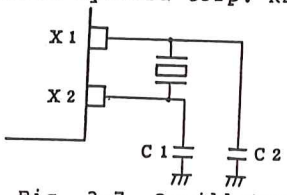


Fig. 3-7 Oscillator

3.4.4 At reset or power up

If the $\overline{\text{RESET}}$ signal is input while a speech is being synthesized, or on power application, the data input to the D/A converter is undefined but within the limit of 0 to 100H.

If speech synthesis is started under this condition with both the $\overline{\text{ST}}$ and $\overline{\text{CS}}$ pins being low level, the input data of the D/A converter is shifted to 100H and then the speech is output.

The input to the D/A converter is shifted gradually to 100H in the same manner as described in section 3.3, the synthesized speech will be output a maximum of 47ms after the $\overline{\text{RESET}}$ signal has been input or the power has been applied.

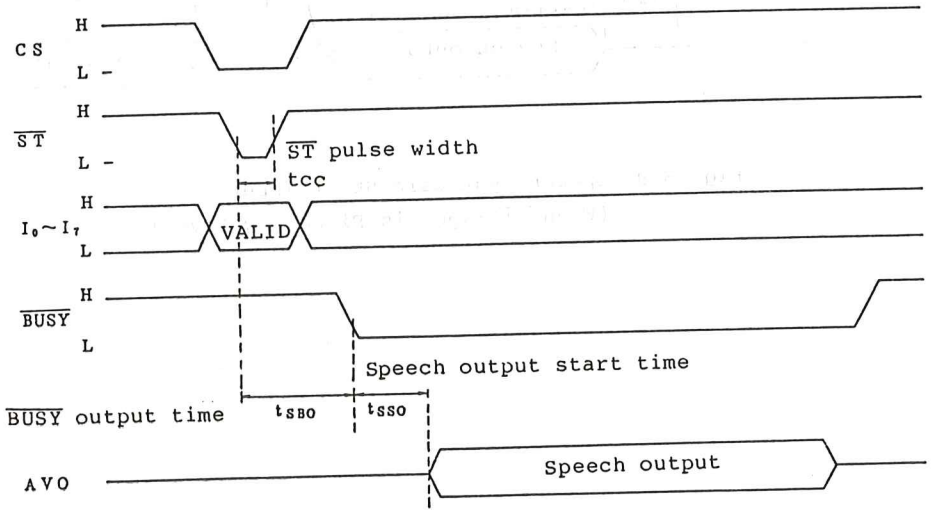


Fig. 3-8 Speech Synthesis Start Input
(When $\overline{\text{ST}}$ Input is Pulse)

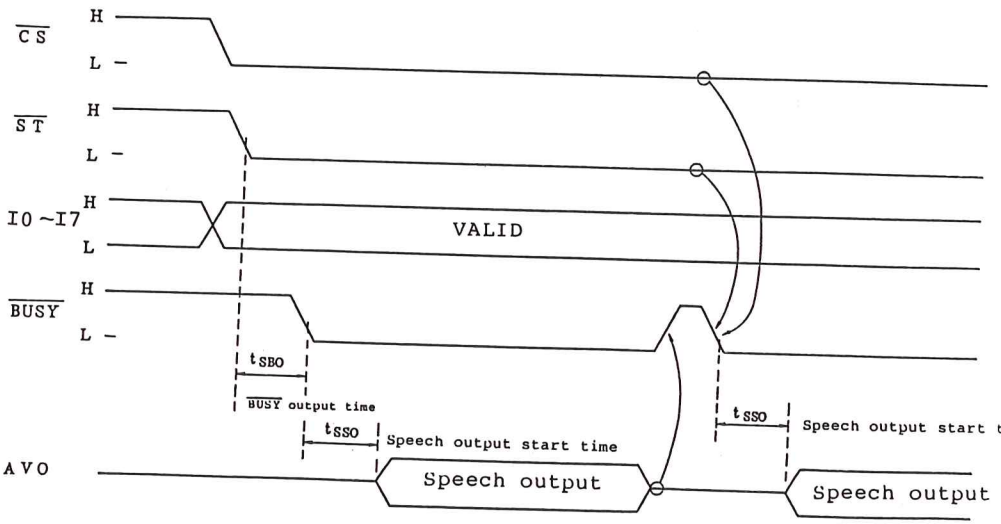


Fig. 3-9 Speech Synthesis Start Input
(When $\overline{ST}$ Input is Fixed Low Level)

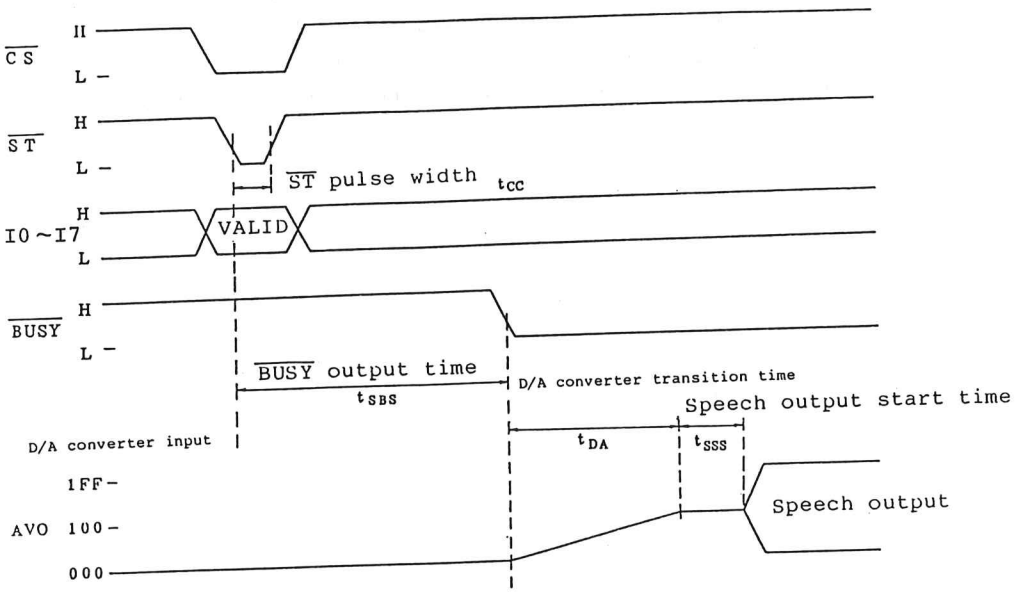


Fig. 3-10 Speech Synthesis Start Operation
(Standby Mode)

3.5 Reset

The $\overline{\text{RESET}}$ signal is used to initialize the LSI on power application, terminate the synthesized speech output, or release the standby mode.

The μPD7755 /7756 is reset by holding the $\overline{\text{RESET}}$ pin to low level for the following period of time.

° During operation	12 oscillator clocks (min.) (18.75μs at 640kHz)
° On power application in standby mode	The same as above after clock oscillation is completed

CHAPTER 4 INTERFACES

4.1 Message Select Input

4.1.1 Host control mode

Fig. 4-1 shows an example when a μPD80C48 is used as the host CPU. The message select code is output to the data bus and this data is written to the μPD7755 /7756 by setting the $\overline{CS}$ and $\overline{ST}$ signals to low.

Because the output of the $\overline{BUSY}$ pin becomes high impedance in standby mode, this signal must be pulled up. (In the example in Fig. 4-1, because the ports of the μPD80C48 are provided with built-in pull-up resistors, the pull-up resistor shown can be omitted.) Also, in standby mode, if the bus to which I0-I7 is connected becomes high impedance, there is a chance that excess current will drain. In this case, these signals should always be connected either to pull-up or pull-down resistors. Fig. 4-2 shows a flowchart when control is performed by the host CPU. The circuit that checks the $\overline{BUSY}$ output in standby mode must be provided with a wait equivalent to the clock oscillation start time.

4.1.2 Key input mode

Fig. 4-3 shows an example of the μPD7755 /7756 used alone. If the $\overline{ST}$ input switch is fixed to ground, the speech synthesis output will repeat indefinitely. Unnecessary I0-I7 inputs should be connected to ground.

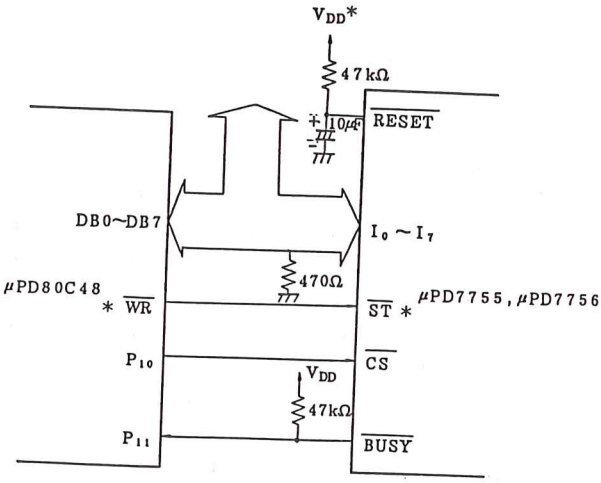
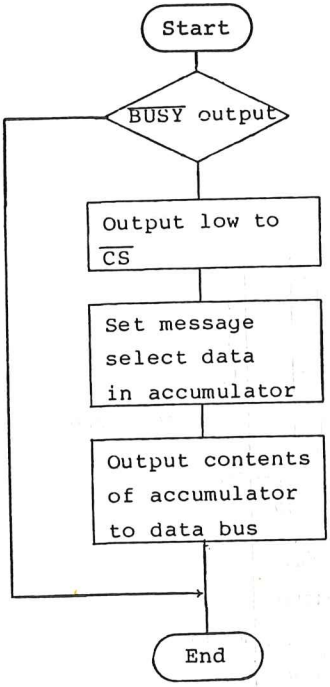


Fig. 4-1 Circuit Example when Control is by Host

* Use 5V±10% for V_{DD}. When a 3V line is used, the minimum pulse width for the μPD7755 /7756 $\overline{ST}$ input becomes 2μs and it no longer possible to use the $\overline{WR}$ output of the μPD80C48.



; Check μPD7755 /7756 BUSY output

; WR is output with output to bus; WR functions as start signal

Fig. 4-2 Flowchart (Control by Host)

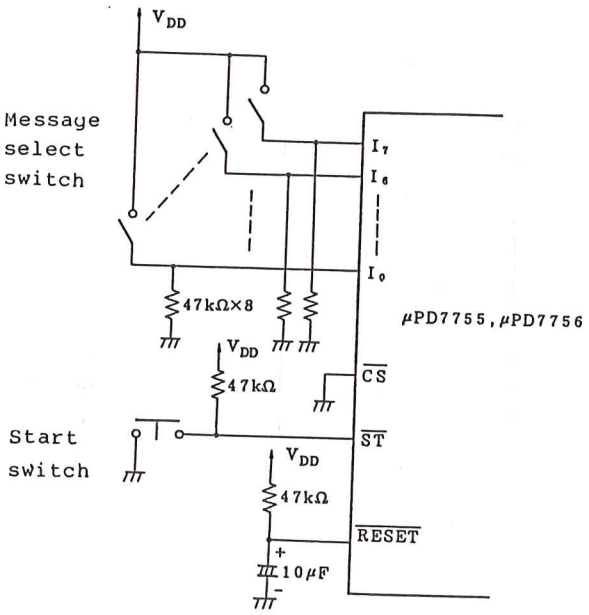


Fig. 4-3 Key Input Mode Circuit Example

4.2 Lowpass Filter

Signals that have been digitally analyzed and synthesized by a sampling method can only reproduce frequency ranges up to one half the sampling frequency. Frequencies above this range appear as unwanted noise and if the D/A output is directly amplified, the result would be a poor quality sound with an extremely unfavorable S/N ratio. It is therefore necessary to use a lowpass filter that will only allow signal frequencies of less than half the sampling frequency to pass.

A filter with sharp bandpass characteristics is desirable and ideally a filter with damping characterics of 48dB/oct or better should be used. The complexity of such a filter, however, is prohibitive and we have opted, in the interest of simplicity and operability, for a filter with characterisics of 24dB/oct for use in the μPD7755 / 7756 . Fig. 4-4 shows a μPC358C analog filter with Butterworth damping characteristics (24dB/oct) that can operate on a +5V single power supply.

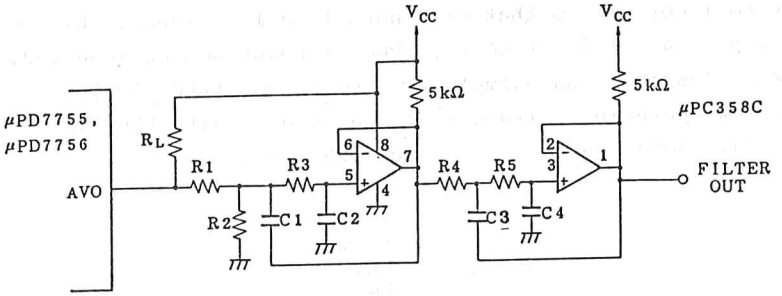


Fig. 4-4 Lowpass Filter Circuit Example

When the μPC358C operates on +5V, input signals in a range of 0 to 3.5V may be used. Resistors R1 and R2 should therefore be used to set the operating point to 2.5V.

Table 4-1 shows the appropriate constants when sampling frequencies (f_{SMPL}) of 4, 5, 6, and 8kHz are used.

Table 4-1 Lowpass Filter Constants
(24dB/oct Butterworth characteristics)

f_{SAMP} (kHz)	R1 (KΩ)	R2 (KΩ)	R3 (KΩ)	R4 (KΩ)	R5 (KΩ)
4	12	13	24	2.7	12
5	10	11	20	2.2	10
6	8.2	9.1	16	1.8	8.2
7	5.6	6.8	12	1.3	6.2

C1	0.022	μF
C2	0.0022	μF
C3	0.022	μF
C4	0.01	μF
RL	1	KΩ

4.3 Power Amp

Because the signal obtained as the output of the lowpass filter described in the preceding paragraph has a waveform identical to that of a normal analog signal, this output can be input to any type of power amp as required. Fig. 4-5 shows an example of a power amp (μPC1212C) with a low operating voltage that can be used with the +5V single power supply of the μPD7755 /7756 .

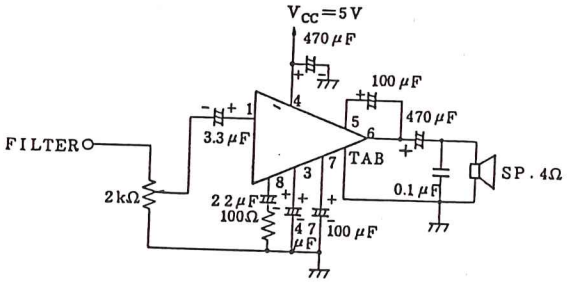


Fig. 4-5 Power Amp (0.7W, 5V)

Because the μPC1212C can operate on voltages in the range of 4.5V to 7.0V, if a 4Ω speaker is used with 5.0V power supply, a 0.7W (THD=10%) output is obtained.

CHAPTER 5 SPEECH ANALYSIS

The procedure for developing a ROM code for the μPD7755 / 7756 is shown in the flowchart in Fig. 5-1. The main points of the procedure are described below.

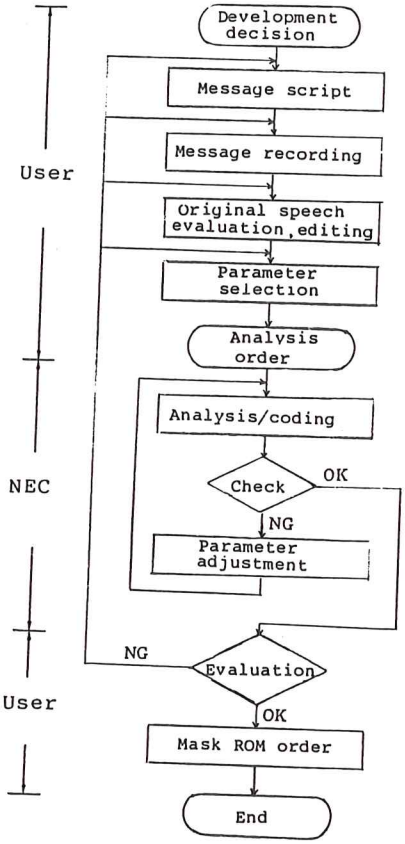


Fig. 5-1 ROM Code Development Flowchart

5.1 Writing Message Script

Because the fequency bandwidth of the reproduced speech is narrower than that of natural speech, care should be taken to write a script that avoids the use of ambiguous expressions. Also, if messages that share common phrases are used, ROM area can be used effectively.

5.2 Original Speech Recording

Because the effective dynamic range for synthesized speech is narrow, the speaker (announcer) should take care to suppress inflections and to speak in an level, even tone of voice.

Because the synthesizing process sometimes emphasizes noise in the original voice recording, care should be taken to keep all noise (particularly hum) to the absolute minimum. For this reason, the use of a professional recording studio using open reel tapes is recommended. Also, to permit selection of the sample with the best speech quality and the least noise, you are asked to provide several recordings of each message.

5.3 Original Speech Evaluation and Editing

Selection should be made from among the recorded messages taking into consideration lack of noise, even tone of voice, sound quality, etc. The messages thus selected should be edited onto an open reel or cassette tape. When an open reel tape is used, either the full tape or two tracks may be used. For a good S/N ratio when using a cassette, the use of metal tape is recommended. The recording should be done at a somewhat high level with peak values are in the range of +5 to +8dB. To ease the analysis process, the format shown in Fig. 5-2 should be used with the messages appropriately divided by blanks.

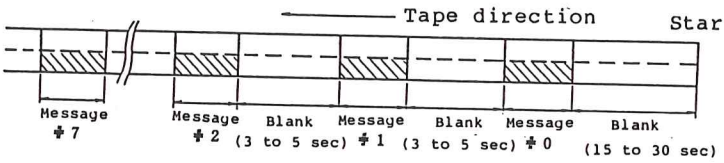


Fig. 5-2 Original Speech Tape Format for Analysis

When a cassette is used, only side A should be used.

5.4 Parameter Selection

The parameters to be used must be determined before the order for analysis is placed. For the μPD7755 /7756 , the only parameter that must be selected by the user is the sampling frequency.

Sampling frequency selection

Please specify 4, 5, 6, or 8kHz as the sampling frequency. When 6kHz is specified, the frequency bandwidth is almost identical to that of a telephone. When 4kHz is selected, the bandwidth drops to only those frequencies under 2kHz and there is some loss of clarity. Although no definitive statement can be made about selection of the sampling frequency, the guidelines laid out in Table 5-1 can be used.

Sampling frequency	Synthesized bandwidth	Application	Quality	Bit rate (bps)
4kHz	up to 1.8kHz	Male voice/ sentences	Distorted	8 to 16K
5kHz	up to 2.2kHz	Male voice/ single words; female voice/ sentences	Slightly distorted	10 to 20K
6kHz	up to 2.7kHz	Female voice/ single words	Equivalent to telephone conversation	12 to 24K
8kHz	up to 3.5kHz			16 to 32K

Note that a great deal of noise in the message will cause the bit rate to be high.

5.5 Ordering Speech Analysis

Place your order for analyis after making the preparations described above. Be sure to include all of the following when placing your order:

Original speech tape (with company name and date clearly indicated)

Message list (message selection codes in the same order as messages are recorded on the tape)

Parameter specification (sampling frequency)

All tapes should include the company name, section, and name of person responsible.

5.6 Evaluation

A cassette tape with a recording of the synthesized speech (result of analysis and coding) will be returned to the client along with the ROM code on an 8-inch floppy disk.

If the result of evaluation is satisfactory, order the mask ROM specifications.

If there are any problems, contact NEC.

Europe) GmbH, Oberrather Str. 4, D-4000 Düsseldorf 30, W.-Germany, Tel. (0211) 650301, Telex 8 58 996-0
Germany) GmbH, Oberrather Str. 4, D-4000 Düsseldorf 30, Tel. (0211) 650302, Telex 858996-0
28/29, D-3000 Hannover 1, Tel. (0511) 881013-16, Telex 9230109
D-8000 München 81, Tel. (089) 416002O, Telex 522971
314, D-7000 Stuttgart 30, Tel. (0711) 890910, Telex 7252220
Benelux), Boschdijk 187a, NL-5612 HB Eindhoven, Tel. (040) 445845, Telex 51923
(Scandinavia) – Box 4039, S-18304 Täby, Tel. (08) 756 7245, Telex 13839
France) S.A., Tour Chenonceaux, 204, Rond Point du Pont de Sèvres, F-92516 Boulogne Billancourt, Tel. (01) 6 09 90 04, Telex 203 544
Italiana s.r.l., Via Cardano 3, I-20124 Milano, Tel. (02) 67 09 108, Telex 315355
(UK) Ltd., Block 3 Carfin Industrial Estate, Motherwell ML1 4UL, Scotland, Tel. (0698) 732221, Telex 777 565