AN ABSTRACT OF THE THESIS OF

PHOJANA SIMASATHIEN for the MASTER OF SCIENCE
(Name) (Degree)

in ELECTRICAL AND ELECTRONICS ENGINEERING presented on ____________
(Major) (Date)

Title: RECOGNITION OF SELECTED SPOKEN DIGITS

Abstract approved: Redacted for privacy
Donald L. Amort

This thesis is concerned with the design of a speech recognition system to recognize digits 1, 2, 3 and 4. The system was designed by using the characteristics of the spectral patterns of amplitude vs. time at discrete frequencies. Data obtained for digits 0 to 9 are presented. The outputs of the recognition system are presented in a Binary Coded Decimal.

A minimum system was evaluated in the laboratory to show feasibility of the technique. The cost of the major components of the system, not including labor work was estimated. The test shows that a 90-95% correct performance was obtained when individual digits were spoken repeatedly. Also there was an 80-85% correct performance when there were series of mixed digits spoken.

The system was also tested by using different speakers, five American, three Thai and two Chinese students. None of them have

been trained. The result obtained was a 50-60% correct performance.

This paper indicates how improved performance can be obtained by using more frequency channels.

Recognition of Selected Spoken Digits

by

Phojana Simasathien

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

June 1969

APPROVED:

Redacted for privacy

Associate Professor of Electrical and Electronics Engineering
in charge of major

Redacted for privacy

Head of Department of Electrical and Electronics Engineering

Redacted for privacy

Dean of Graduate School

Date thesis is presented ______________________

Typed by Clover Redfern for Phojana Simasathien

ACKNOWLEDGMENT

The author wishes to express his most sincere appreciation to Professor D. L. Amort for his valuable advice, from the subject selection of the thesis to its final proofreading.

The author also wishes to thank the friends who helped during the research period and the writing of this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

RECOGNITION OF SELECTED SPOKEN DIGITS

I. INTRODUCTION

Most of the existing man-machine communication means are exclusively oriented to man's hands and eyes and such devices as printers, push-buttons, displays, etc. are well in common usage. With computers, the man-machine communication was developed but it is still essentially written language which is inconvenient and time-consuming.

The man-machine communication is still being developed for convenience, greater speed and economics.

Since speech is the basic means of communication between men, it would be very interesting to use speech to communicate between men and machines.

The recognition system is used as a media to communicate between men and machines. Its function is to recognize the human speech in a certain way and presents to the input of the machine causing the machine to operate according to the command speech.

The recognition systems which have been previously developed for recognizing vowel sounds and spoken words are complicated and expensive (2,3,7,9). Therefore the development of this system is based on low cost as well as simplicity.

II. SYSTEM ORGANIZATION

The designed speech-recognition system consists of six major components as shown in Figure 1.

1. input unit
2. frequency separators
3. detectors
4. quantizers and samplers
5. encoder
6. output indicator

The input unit consists of microphone and audio amplifier. The tape-player was also used for the convenience of the research.

The output signal from the input device is then fed to the frequency separator circuit, which consists of a bank of bandpass filters with center frequencies ranging from 300 to 4000 hz (see Appendix A). Each frequency was chosen so that the various sounds of speech will exhibit different sound spectrum displays (1, 3, 7, 9). Speech signal will be separated into frequency channels according to the center frequency of the filters. Each filter is followed by a detector circuit. The signal is rectified to a d. c. waveform and then smoothed to get the envelope waveform as shown in Figure 2.

The speech waveform from each channel is then quantized in the quantizing circuit into levels, the output of which is a binary coded signal depending on the amplitude of the speech waveform. The quantized waveform is then sampled at regular intervals and stored in the

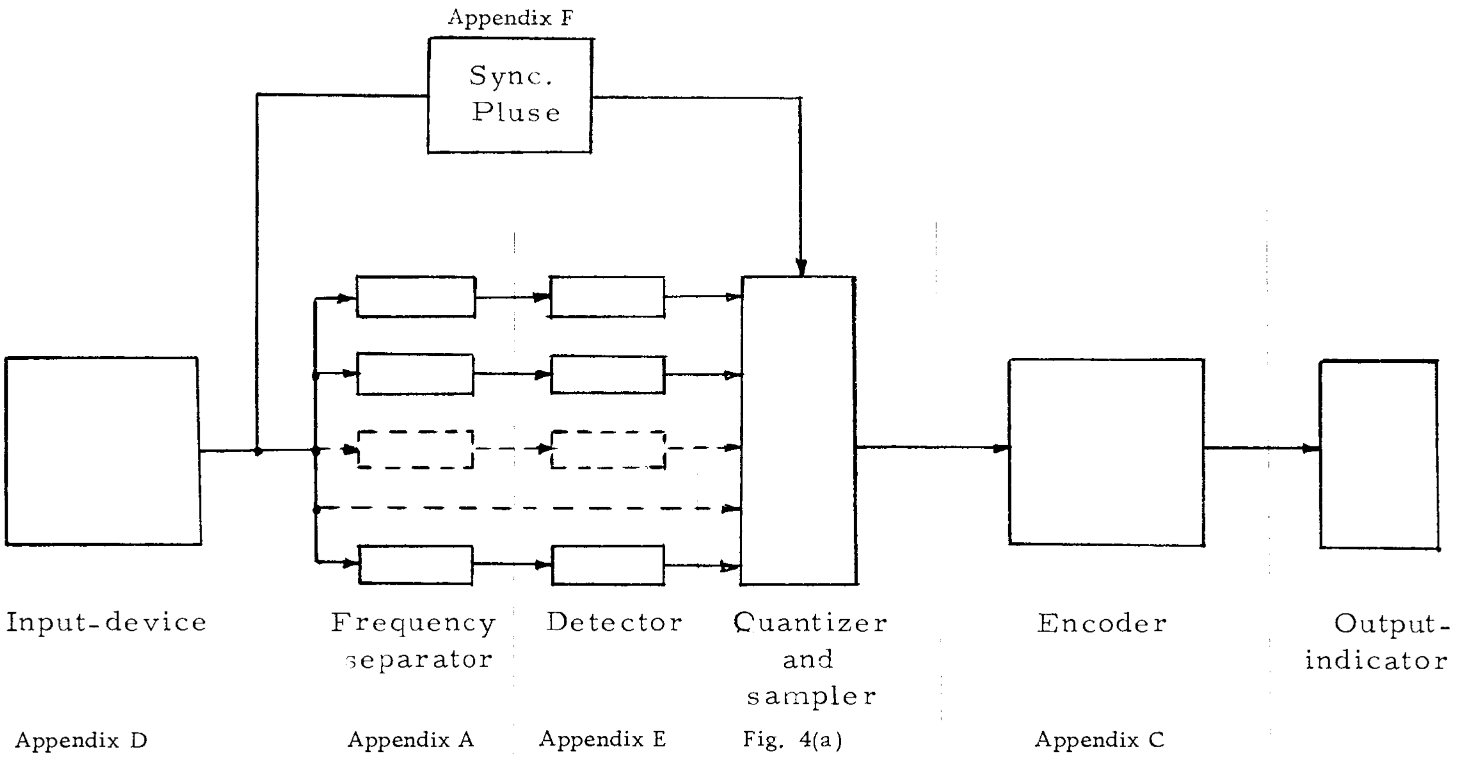


Figure 1. System organization of the speech-recognizer.

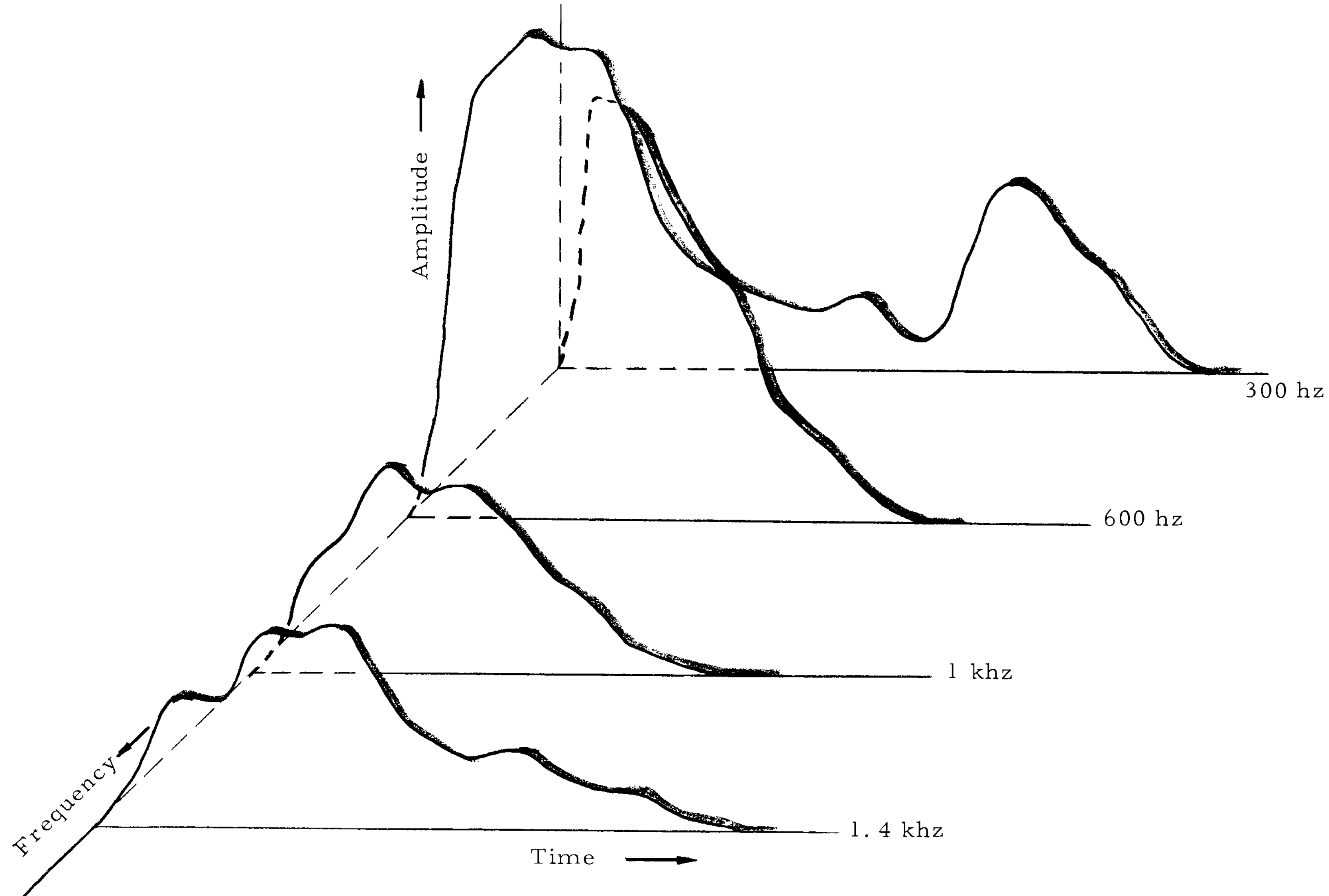


Figure 2. Spectral waveform of the spoken digit "ONE".

shift-register.

Figure 3 shows in more detail how the quantizing and sampling process operates. For simplicity, the output of a single channel is shown. In the figure, the waveform for the 300 hz channel of Figure 2 has been redrawn. One threshold line has been drawn across the waveform, dividing the amplitude scale into two levels. More levels can be used if more detail is required.

The encoder consists of logic circuits. The stored signal is fed to the logic circuits to translate into Binary Coded Decimal indicating by the output indicator.

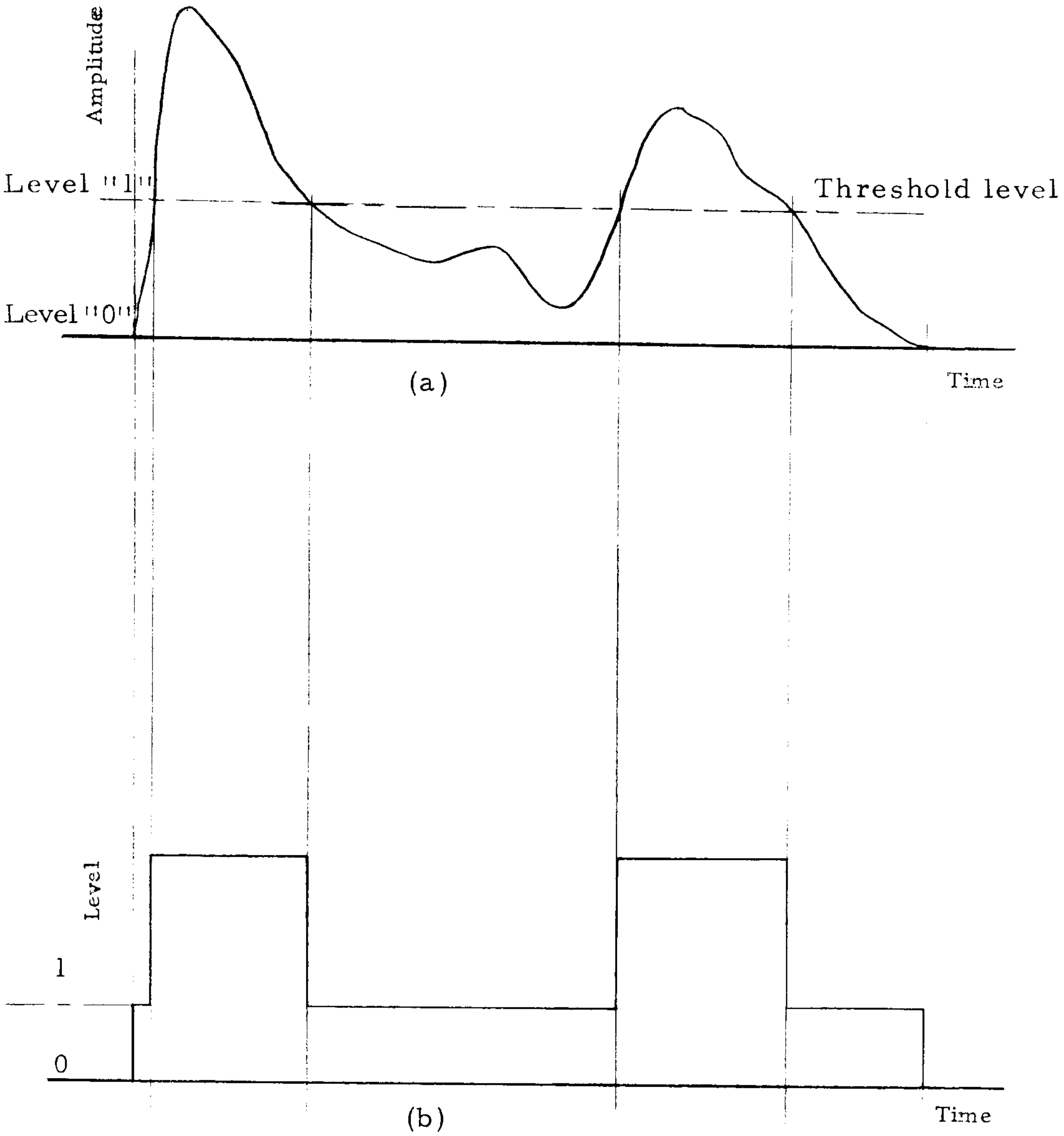


Figure 3a. Detector output waveform of digit "ONE" at 300 hz channel.

3b. Quantized detector output waveform.

III. TECHNIQUES OF REPRESENTATION OF SPEECH WAVEFORM

A sound wave of digits 1, 2, 3 and 4 can be adequately described in terms of amplitude vs. time at discrete frequency intervals as shown in Figure 2 (1, 9). The first step in the recognition procedure is to produce a reference pattern. This pattern is obtained by quantizing the speech waveform.

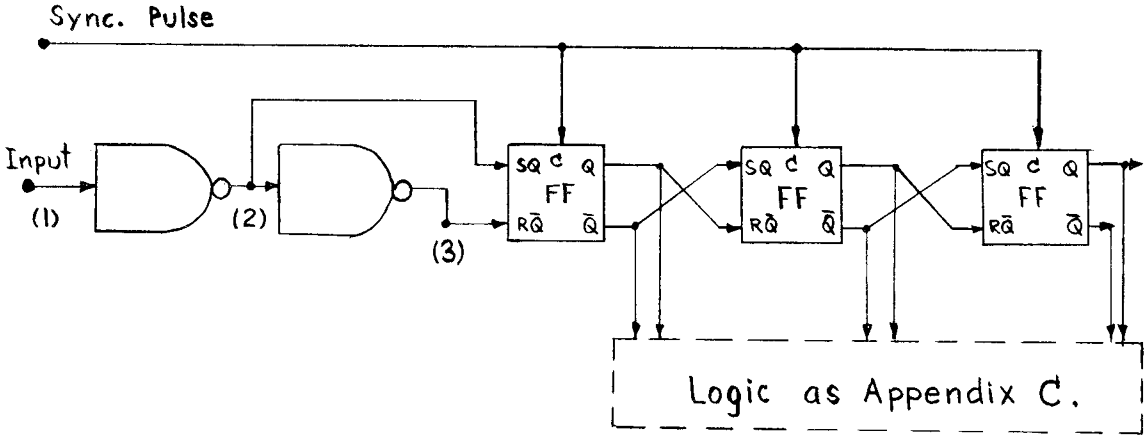
Figure 4 shows the quantizing and sampling circuit and waveforms at the input and gates output.

The quantized waveform can be made to show more detail of the speech waveform by setting up more threshold levels. The speech waveform in Figure 3 is redrawn in Figure 5(a). By setting three threshold levels, the quantized result is shown in Figure 5(b) which shows more detail than using one threshold level (see Figure 3(b)).

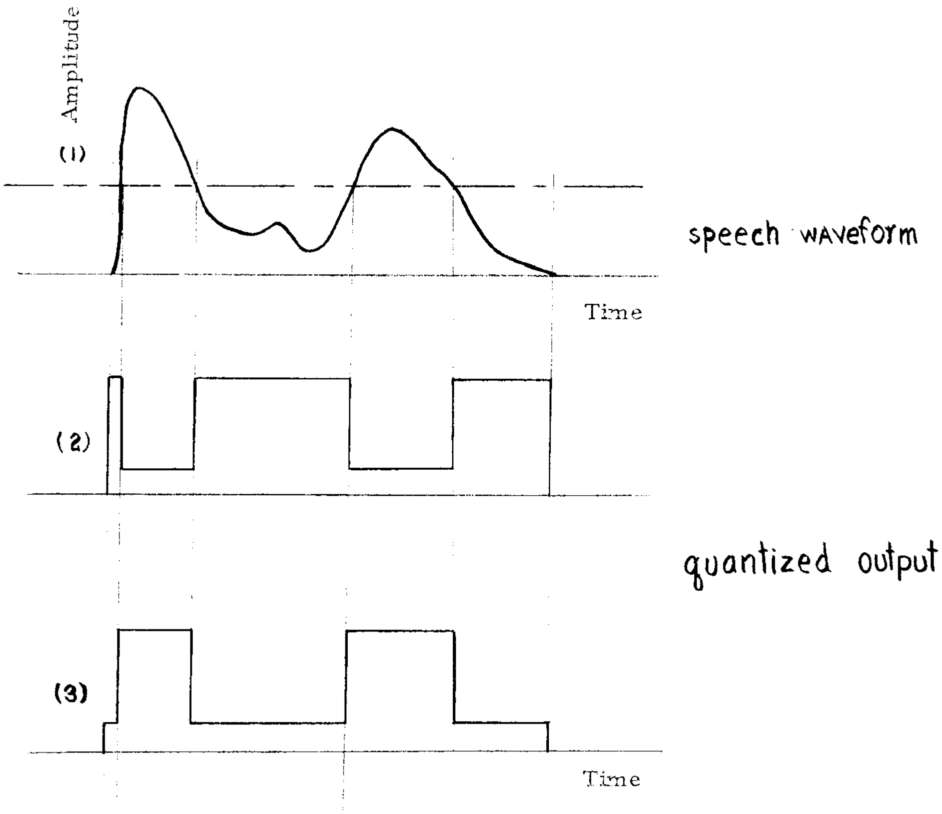
The trouble with using many threshold levels is that, when each digit is spoken at the different loudness, the quantized waveform will give more difference than when using fewer threshold levels.

Figure 5(a) shows the speech waveform of digit "ONE" at 300hz channel in two different loudnesses. The corresponding quantized waveforms are shown in Figure 5(b). The shaded area is the difference from this result. Figure 6(a) and (b) show this difference when a single threshold level is used which is comparatively less than Figure 5(b).

To avoid this difference the single threshold level is used in the



(a)



(b)

Figure 4a. Quantizing and sampling circuit.

4b. Waveform at (1) input
(2) 1st gate output
(3) 2nd gate output.

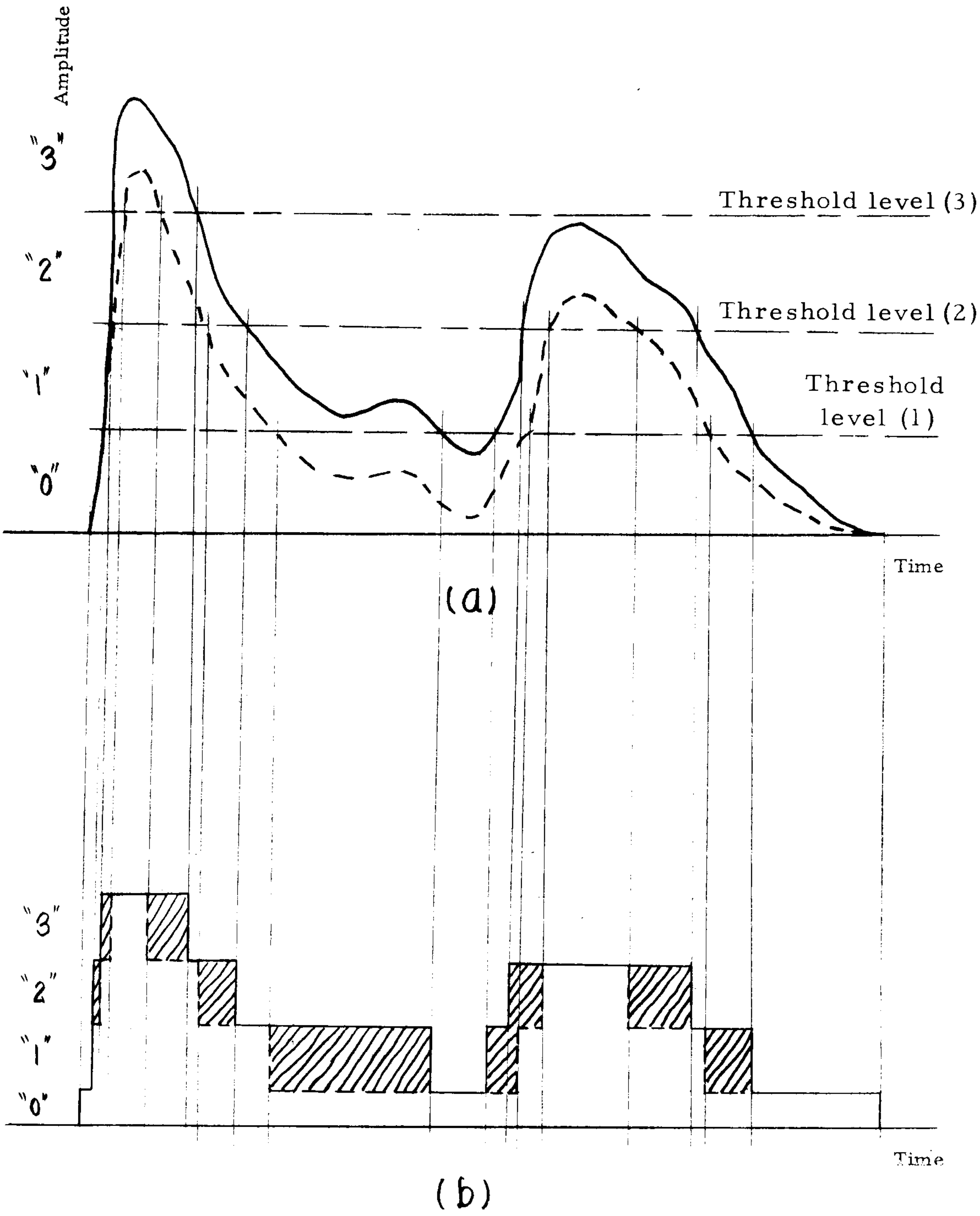


Figure 5a. Speech waveform of digit "ONE" at 300 hz channel.
5b. Three levels quantized waveform.

system. By reducing the threshold level, less detail of each speech waveform is obtained. So, to compensate this loss more frequency channels are used.

To obtain the third parameter (time), the quantized waveform is sampled at a regular rate and stored in the register. The sampled data in the registers are in binary form. These data present the binary signal of each digit. "1" represents the quantized amplitude in 1-level and "0" represents 0-level.

The detail of the speech waveform of digits 0 to 9 at each frequency channel are shown in Appendix B.

Encoding to the Binary Coded Decimal output is done by setting up the logic equations from the binary signals at each frequency channel. These equations are shown in Appendix C. The output indicators X, Y, Z which correspond to 4, 2, and 1 in BCD code will show the "1" or "0" state according to the truth table in Table 1 shown below.

Table 1. BCD code.

8	4	2	1	
	X	Y	Z	Digit
0	0	0	1	1
0	0	1	0	2
0	0	1	1	3
0	1	0	0	4

$$X = 4 \tag{1}$$

$$Y = 2 + 3 \tag{2}$$

$$Z = 1 + 3 \tag{3}$$

From (1) indicates that X has "1" state if and only if digit 4 is spoken. Likewise (2) and (3) indicate that Y has the "1" state when digit 2 OR 3 is spoken, Z has the "1" state when digit 1 OR 3 is spoken.

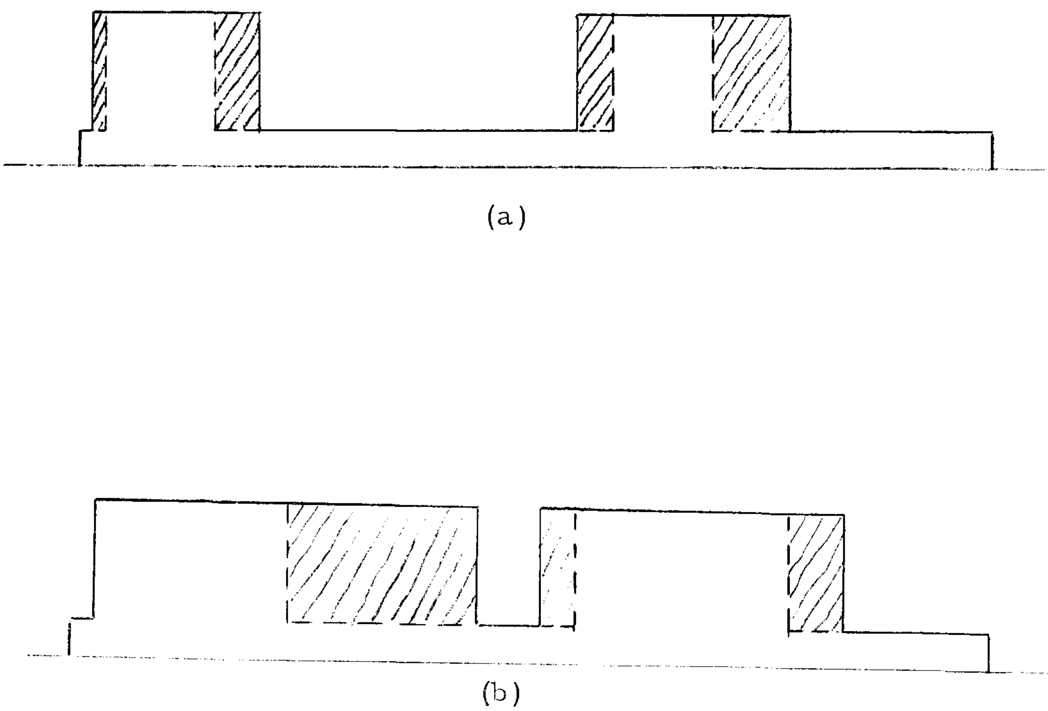


Figure 6. Quantized waveform of Figure 5(a).
(a) Using threshold level (2) only.
(b) Using threshold level (1) only.

IV. EXPERIMENT AND EVALUATION OF THE SYSTEM

In the experiment five American, three Thai and two Chinese students were used to provide the test-material. The speakers were asked to pronounce digits 1, 2, 3 and 4 naturally as telling a telephone number. The utterances of the digits were recorded on a tape recorder. The recordings and all experiments were made in the laboratory. The speech waveforms of five American students were analyzed and used as a standard waveform. The recordings of these speakers were used as the untrained speakers. The result was obtained 50-60% correctly.

Since the recognition required that speech waveforms of each speaker must match the standard waveforms, so the system requires the speaker to learn to pronounce each digit to get the proper response from the system. After learning to speak to the system two types of tests were conducted. Individual digits were spoken repeatedly in normal rate. The result was obtained 90-95% correctly. The other test was done by speaking a series of mixed digits. The result was obtained 80-85% correctly (Appendix G).

For recognizing four digits only three frequency channels (300, 600, and 1000 hz) and 4-bit shift-register are used for each channel. The capability of recognizing more digits can be done by using more frequency channels (Appendix B).

Cost Estimation

The cost of the system is considerably low. For the capability of recognizing four digits, the cost of the major components are:

Quantity	Component	Price $
3	Filter: $9. 00 @	27. 00
3	Operational[1] Amplifier: $4. 00 @	12. 00
3	4-bit shift register: $8. 00 @	24. 00
16	3, 3-input NAND GATE package $1. 50 @	24. 00
1	Sync. circuit $100. 00 @	100. 00
	Total $	187. 00

[1] The operational amplifier can be used in each frequency channel to replace the audio-amplifier in the input circuit.

V. CONCLUSION

In this paper, the method of designing a speech recognition system was described. The system can perform the recognition process in real-time with direct microphone input as well as from a tape-player. The method described here is not the only possible method, but it is one of the simplest and most economical methods.

The spectral patterns from each output of the detector circuit were examined visually by the oscilloscope. The examination showed that each digit formed a distinctive pattern. It was shown that by using only three frequency channels each of two quantized levels, the system can recognize all digits 1 to 4.

For better performance of the system, speakers are required to be trained to adapt to the system.

APPENDICES

BIBLIOGRAPHY

1. Bell, C. G., H. Fugisaki, J. M. Heinz, K. N. Stevens and A. S. House. Reduction of speech spectra by analysis-by-synthesis techniques. Journal of the Acoustical Society of America 33: 1725-1736. 1961.

2. Bezdel, W. and H. J. Chandler. Results of an analysis and recognition of vowels by computer using zero-crossing data. Proceedings of the Institution of Electrical Engineers 112: 2060-2066. 1965.

3. Denes, P. B. and M. W. Mathews. Spoken digit recognition using time-frequency pattern matching. Journal of the Acoustical Society of America 32:1450-1455. 1960.

4. Dersch, W. C. SHOEBOX--a voice responsive machine. Datamation 8(6):47-50. 1962.

5. Forgie, T. W. and C. D. Forgie. Results obtained from a vowel recognition computer program. Journal of the Acoustical Society of America 31:1480-1489. 1959.

6. Marcus, Mitchell P. Switching circuits for engineers. Englewood Cliffs, New Jersey, Prentice Hall, 1967. 338 p.

7. Olson, Harry F. Speech processing systems. IEEE Spectrum 1(2):90-102. 1964.

8. Stevens, K. N. Toward a model for speech recognition. Journal of the Acoustical Society of America 32:47-55. 1960.

9. Talbert, L. R., G. F. Groner, J. S. Koford, R. J. Brown, P. R. Low and C. H. Mays. A real-time adaptive speech recognition system. Stanford, 1963. 18 p. (Stanford University. Stanford Electronics Laboratory. Technical Documentary Report No. ASD-TDR 63-660)

APPENDIX A

Filter Data

Channel	F_0	F_1	F_2	ΔF
1	300	200	400	200
2	600	500	700	200
3	1000	900	1100	200
4	1400	1200	1600	400
5	1950	1750	2150	400
6	2400	2200	2600	400
7	2900	2700	3100	400
8	4000	3500	4500	1000

F_0 = center frequency (hz)

F_1 = lower 3 db frequency (hz)

F_2 = upper 3 db frequency (hz)

ΔF = $(F_2 - F_1)$ bandwidth (hz)

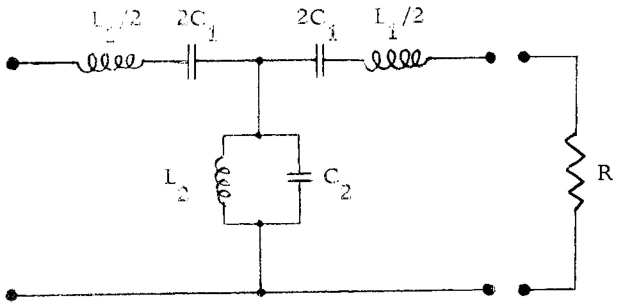


Figure 7. Filter circuit.

$$C_1 = \frac{f_2 - f_1}{4\pi R f_1 f_2}$$

$$L_1 = \frac{R}{\pi(f_2 - f_1)}$$

$$C_2 = \frac{L_1}{R^2}$$

$$L_2 = C_1 R^2$$

APPENDIX B

Data of Digits 0-9

The average data of each digit spoken twice by five male speakers:

Digit	Frequency (hz)	Time (ms.) 50	100	150	200	250	300	350	400
1	300	1. 9	0. 8	0. 4	0. 4	0. 3	1. 6	1. 1	0. 1
	600	2. 0	3. 2	1. 9	1. 7	0. 8	0. 0	0. 0	0. 0
	1000	1. 0	1. 3	1. 3	0. 7	0. 2	0. 0	0. 0	0. 0
	1400	0. 8	1. 4	1. 5	0. 7	0. 5	0. 4	0. 3	0. 1
	1950	0. 1	0. 3	0. 4	0. 1	0. 0	0. 0	0. 0	0. 0
	2400	0. 6	1. 0	1. 0	0. 6	0. 2	0. 0	0. 0	0. 0
	2900	0. 2	0. 4	0. 3	0. 1	0. 0	0. 0	0. 0	0. 0
	4000	0. 2	0. 4	0. 3	0. 0	0. 0	0. 0	0. 0	0. 0
2	300	1. 4	2. 3	2. 4	1. 8	1. 3	0. 9	0. 1	0. 0
	600	0. 7	1. 0	1. 4	1. 0	0. 3	0. 1	0. 0	0. 0
	1000	0. 8	1. 1	1. 0	0. 9	0. 6	0. 4	0. 0	0. 0
	1400	0. 3	0. 6	0. 4	0. 4	0. 2	0. 0	0. 0	0. 0
	1950	0. 2	0. 2	0. 1	0. 1	0. 0	0. 0	0. 0	0. 0
	2400	0. 6	0. 6	0. 5	0. 3	0. 1	0. 0	0. 0	0. 0
	2900	0. 4	0. 2	0. 3	0. 5	0. 2	0. 1	0. 0	0. 0
	4000	0. 2	0. 1	0. 2	0. 0	0. 0	0. 0	0. 0	0. 0
3	300	1. 0	2. 1	2. 8	2. 5	1. 6	1. 0	0. 8	0. 1
	600	0. 4	1. 4	1. 0	1. 3	0. 8	0. 4	0. 0	0. 0
	1000	0. 6	1. 4	1. 2	1. 6	2. 0	0. 6	0. 2	0. 0
	1400	0. 6	0. 9	0. 8	1. 3	0. 7	0. 4	0. 0	0. 0
	1950	0. 2	0. 2	0. 2	0. 1	0. 0	0. 0	0. 0	0. 0
	2400	0. 8	0. 9	0. 7	0. 5	0. 1	0. 0	0. 0	0. 0
	2900	0. 6	0. 4	0. 3	0. 7	0. 4	0. 1	0. 0	0. 0
	4000	0. 3	0. 5	0. 2	0. 1	0. 0	0. 0	0. 0	0. 0

Amplitude (volt)

Digit	Frequency (hz)	Time (ms.)							
		50	100	150	200	250	300	250	400
4	300	2. 0	2. 4	2. 2	1. 5	1. 1	0. 4	0. 0	0. 0
	600	1. 8	1. 5	1. 5	1. 3	1. 2	0. 6	0. 0	0. 0
	1000	1. 6	1. 6	1. 4	1. 6	1. 0	0. 6	0. 0	0. 0
	1400	0. 7	0. 9	0. 6	1. 0	0. 6	0. 0	0. 0	0. 0
	1950	0. 3	0. 3	0. 3	0. 1	0. 0	0. 0	0. 0	0. 0
	2400	0. 7	0. 6	0. 6	0. 5	0. 2	0. 0	0. 0	0. 0
	2900	0. 5	0. 3	0. 4	0. 2	0. 1	0. 0	0. 0	0. 0
	4000	0. 3	0. 2	0. 3	0. 1	0. 0	0. 0	0. 0	0. 0
5	300	0. 6	1. 7	1. 8	1. 9	1. 8	1. 4	0. 6	0. 4
	600	0. 2	1. 9	1. 5	1. 5	1. 7	0. 7	0. 2	0. 1
	1000	0. 3	1. 6	1. 5	1. 7	1. 2	0. 6	0. 3	0. 2
	1400	0. 3	1. 0	0. 7	0. 8	0. 5	0. 3	0. 1	0. 0
	1950	0. 1	0. 2	0. 2	0. 2	0. 3	0. 2	0. 1	0. 0
	2400	0. 2	0. 7	0. 9	0. 9	1. 0	0. 6	0. 3	0. 0
	2900	0. 2	0. 1	0. 4	0. 3	0. 6	0. 2	0. 0	0. 0
	4000	0. 0	0. 3	0. 3	0. 4	0. 3	0. 2	0. 0	0. 0
6	300	2. 2	1. 8	0. 5	0. 2	0. 5	0. 2	0. 4	0. 1
	600	1. 6	1. 0	0. 3	0. 2	0. 1	0. 1	0. 0	0. 0
	1000	1. 2	1. 2	0. 6	0. 4	0. 4	0. 2	0. 2	0. 1
	1400	0. 7	0. 9	0. 4	0. 1	0. 0	0. 0	0. 0	0. 0
	1950	0. 2	0. 0	0. 0	0. 0	0. 0	0. 0	0. 0	0. 0
	2400	0. 5	0. 2	0. 1	0. 2	0. 1	0. 0	0. 0	0. 0
	2900	0. 3	0. 6	0. 1	0. 0	0. 0	0. 0	0. 0	0. 0
	4000	0. 2	0. 0	0. 0	0. 0	0. 0	0. 0	0. 0	0. 0

Amplitude (volt)

Digit	Frequency (hz)	Time (ms.) 50	100	150	200	250	300	350	400
7	300	0. 2	1. 2	2. 0	1. 1	2. 1	2. 1	2. 1	1. 4
	600	0. 1	1. 0	2. 0	2. 1	1. 4	1. 8	0. 8	0. 3
	1000	0. 1	0. 6	1. 5	1. 5	1. 1	1. 2	1. 0	0. 6
	1400	0. 3	0. 6	0. 7	0. 8	0. 5	0. 2	0. 1	0. 0
	1950	0. 0	0. 3	0. 2	0. 3	0. 1	0. 1	0. 0	0. 0
	2400	0. 0	0. 7	0. 7	0. 9	0. 4	0. 4	0. 0	0. 0
	2900	0. 2	0. 2	0. 4	0. 4	0. 3	0. 2	0. 0	0. 0
	4000	0. 1	0. 3	0. 4	0. 4	0. 1	0. 1	0. 0	0. 0
8	300	1. 9	2. 0	1. 9	1. 0	0. 4	0. 0	0. 0	0. 0
	600	1. 9	2. 0	1. 8	0. 6	0. 1	0. 0	0. 0	0. 0
	1000	1. 3	1. 5	2. 4	0. 7	0. 2	0. 0	0. 0	0. 0
	1400	0. 7	0. 3	0. 4	0. 1	0. 0	0. 0	0. 0	0. 0
	1950	0. 3	0. 4	0. 5	0. 2	0. 0	0. 0	0. 0	0. 0
	2400	0. 7	1. 1	1. 2	0. 7	0. 2	0. 0	0. 0	0. 0
	2900	0. 3	0. 6	0. 1	0. 0	0. 0	0. 0	0. 0	0. 0
	4000	0. 5	0. 4	0. 3	0. 0	0. 0	0. 0	0. 0	0. 0
9	300	1. 6	1. 4	1. 8	1. 8	2. 2	1. 5	1. 0	0. 9
	600	0. 6	1. 6	1. 3	1. 6	0. 8	0. 4	0. 1	0. 0
	1000	0. 6	1. 4	1. 7	1. 7	1. 1	0. 6	0. 4	0. 2
	1400	0. 3	0. 8	0. 7	0. 5	0. 2	0. 0	0. 0	0. 0
	1950	0. 4	0. 3	0. 3	0. 5	0. 2	0. 0	0. 0	0. 0
	2400	0. 6	0. 8	1. 0	1. 0	0. 6	0. 2	0. 0	0. 0
	2900	0. 7	0. 8	0. 4	0. 3	0. 2	0. 0	0. 0	0. 0
	4000	0. 4	0. 4	0. 5	0. 3	0. 0	0. 0	0. 0	0. 0

Amplitude (volt)

Digit	Frequency (hz)	Time (ms.)							
		50	100	150	200	250	300	350	400
0	300	2. 2	2. 5	2. 1	1. 6	1. 4	1. 4	1. 0	0. 6
	600	1. 3	1. 4	1. 5	1. 4	1. 6	1. 4	0. 6	0. 2
	1000	1. 1	1. 2	1. 2	1. 4	1. 5	1. 2	0. 7	0. 3
	1400	0. 3	0. 6	0. 4	0. 2	0. 1	0. 0	0. 0	0. 0
	1950	0. 2	0. 1	0. 2	0. 2	0. 1	0. 1	0. 0	0. 0
	2400	0. 5	0. 7	0. 5	0. 5	0. 6	0. 4	0. 0	0. 0
	2900	0. 3	0. 1	0. 1	0. 2	0. 2	0. 0	0. 0	0. 0
	4000	0. 2	0. 1	0. 3	0..2	0. 1	0. 0	0. 0	0. 0

Amplitude (volt)

APPENDIX C

Encoder Logic Equations

The quantized waveform in 4-bit rigister of digits 1, 2, 3 and 4 at frequency channel 300 (A), 600 (B), and 1000 hz (C) are:

Digit	Channel	Bit 1	2	3	4
1	A	1	0	0	0
	B	1	1	1	1
	C	0	0	0	0
2	A	0	1	1	1
	B	0	0	0	0
	C	0	0	0	0
3	A	0	1	1	1
	B	0	0	0	0
	C	0	0	0	1
4	A	1	1	1	1
	B	1	1	1	1
	C	1	1	0	0

(Sampling interval is 50 ms.)

Logic equations of digits 1, 2, 3 and 4 are:

$$1 = (A_1\bar{A}_2\bar{A}_3\bar{A}_4) \cdot (B_1B_2B_3B_4) \cdot (\bar{C}_1\bar{C}_2\bar{C}_3\bar{C}_4)$$

$$2 = (\bar{A}_1A_2A_3A_4) \cdot (\bar{B}_1\bar{B}_2\bar{B}_3\bar{B}_4) \cdot (\bar{C}_1\bar{C}_2\bar{C}_3\bar{C}_4)$$

$$3 = (\bar{A}_1A_2A_3A_4) \cdot (\bar{B}_1\bar{B}_2\bar{B}_3\bar{B}_4) \cdot (\bar{C}_1\bar{C}_2\bar{C}_3C_4)$$

$$4 = (A_1A_2A_3A_4) \cdot (B_1B_2B_3B_4) \cdot (C_1C_2\bar{C}_3\bar{C}_4)$$

since: $X = 4$, $Y = 2 + 3$, $Z = 1 + 3$ (see page 11).

$$\therefore \quad X = (A_1 A_2 A_3 A_4) \cdot (B_1 B_2 B_3 B_4) \cdot (C_1 C_2 \overline{C}_3 \overline{C}_4) \qquad (1)$$

$$\therefore \quad Y = (\overline{A}_1 A_2 A_3 A_4) \cdot (\overline{B}_1 \overline{B}_2 \overline{B}_3 \overline{B}_4) \cdot (\overline{C}_1 \overline{C}_2 \overline{C}_3 \overline{C}_4)$$

$$+ (\overline{A}_1 A_2 A_3 A_4) \cdot (\overline{B}_1 \overline{B}_2 \overline{B}_3 \overline{B}_4) \cdot (\overline{C}_1 \overline{C}_2 \overline{C}_3 C_4)$$

$$Y = (\overline{A}_1 A_2 A_3 A_4) \cdot (\overline{B}_1 \overline{B}_2 \overline{B}_3 \overline{B}_4) \cdot (\overline{C}_1 \overline{C}_2 \overline{C}_3) \qquad (2)$$

$$\therefore \quad Z = (A_1 \overline{A}_2 \overline{A}_3 \overline{A}_4) \cdot (B_1 B_2 B_3 B_4) \cdot (\overline{C}_1 \overline{C}_2 \overline{C}_3 \overline{C}_4)$$

$$+ (\overline{A}_1 A_2 A_3 A_4) \cdot (\overline{B}_1 \overline{B}_2 \overline{B}_3 \overline{B}_4) \cdot (\overline{C}_1 \overline{C}_2 \overline{C}_3 C_4)$$

$$Z = (\overline{C}_1 \overline{C}_2 \overline{C}_3) \cdot [\overline{C}_4 (A_1 \overline{A}_2 \overline{A}_3 \overline{A}_4) \cdot (B_1 B_2 B_3 B_4) \qquad (3)$$

$$+ C_4 (\overline{A}_1 A_2 A_3 A_4) \cdot (\overline{B}_1 \overline{B}_2 \overline{B}_3 \overline{B}_4)]$$

APPENDIX D

Frequency Response of Audio System
(Tape Recorder Input to Filter Input)

Microphone not included was Shure CR-81 (100-7,000 hz response).

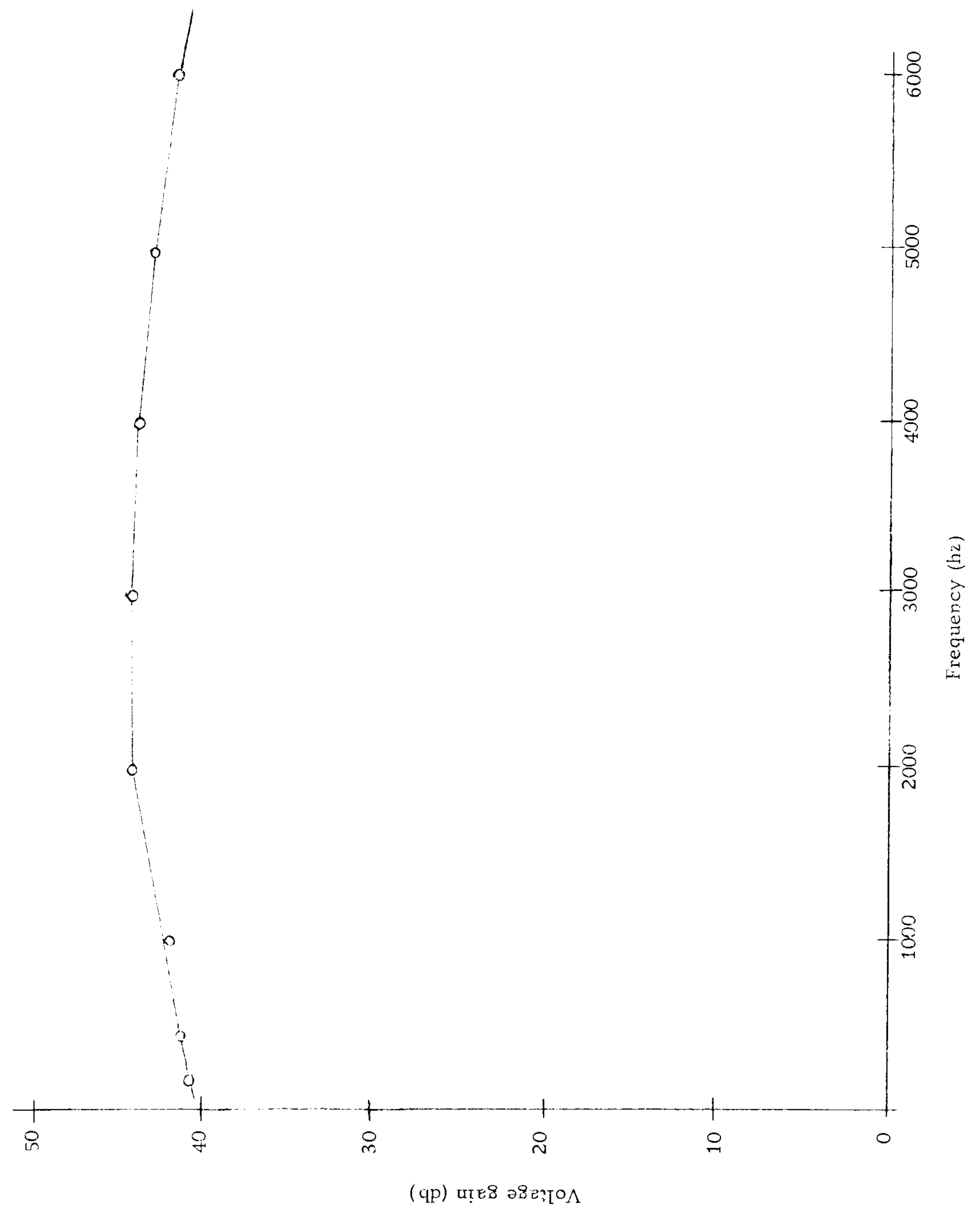


Figure 8. Frequency response of audio system.

APPENDIX E

Detector Circuit

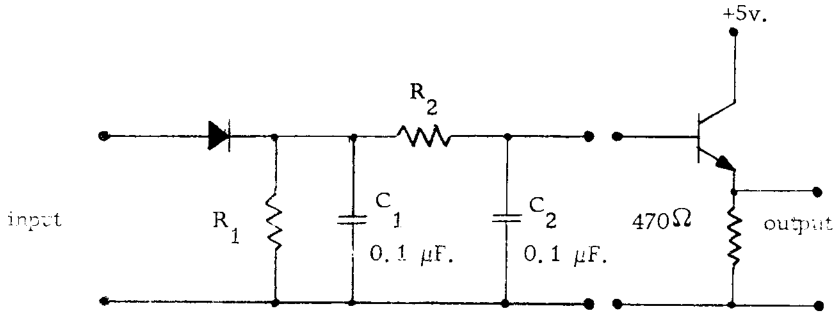


Figure 9. Detector circuit.

Channel	R_1 Ω	R_2 Ω
1	27×10^4	12×10^4
2	27×10^4	58×10^3
3	58×10^4	39×10^4
4	12×10^5	33×10^4
5	27×10^5	10×10^4
6	22×10^5	33×10^4
7	27×10^5	12×10^4
8	82×10^4	12×10^5

APPENDIX F

Synchronize Pulse Generator Circuit

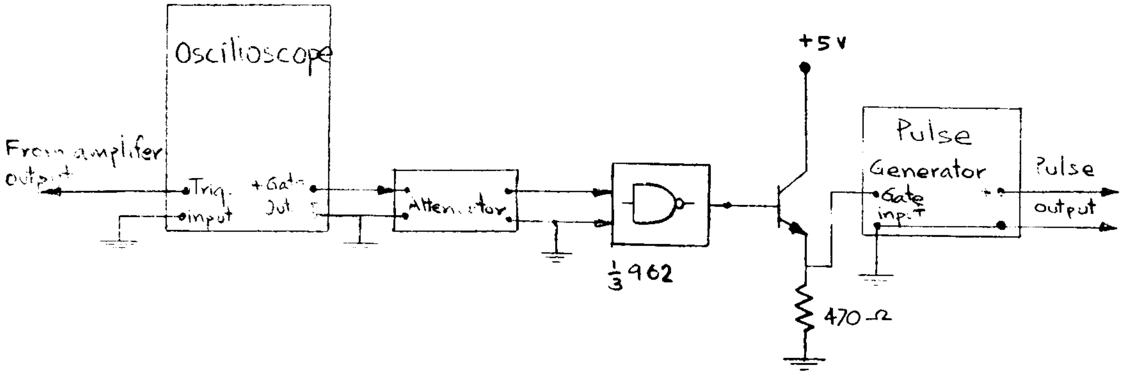


Figure 10. Synchronize pulse generator circuit.

Oscilloscope: Type 549 storage oscilloscope

Triggering setting:

Mode Triggered

Source External

Single Sweep

Pulse generator:

Data pulse 110A Pulse generator

APPENDIX G

Experimental Results

a) Experimental results of digits individually spoken by untrained speakers.

		Recognized as					%
		1	2	3	4	None	Correct
Spoken Digits	1	29	--	--	9	12	58
	2	--	28	11	--	11	56
	3	--	12	25	--	13	50
	4	5	--	1	30	14	60

b) Experimental results of digits individually spoken by trained speakers.

		Recognized as					%
		1	2	3	4	None	Correct
Spoken Digits	1	75	--	--	1	4	93.8
	2	--	75	--	--	5	93.8
	3	--	1	72	--	7	90.0
	4	--	--	--	76	4	95.0

c) Experimental results of series of mixed digits spoken by trained speakers.

		Recognized as					%
		1	2	3	4	None	Correct
Spoken Digit	1	67	--	--	4	9	83.8
	2	--	64	7	--	9	80.0
	3	--	6	67	--	7	83.8
	4	2	--	--	68	10	85.0