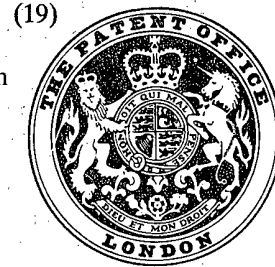


PATENT SPECIFICATION (11) 1 594 412

1 594 412

- (21) Application No. 6753/78 (22) Filed 21 Feb. 1978
- (31) Convention Application No's. 52/017143 (32) Filed 21 Feb. 1977  
53/003702 19 Jan. 1978 in
- (33) Japan (JP)
- (44) Complete Specification Published 30 Jul. 1981
- (51) INT. CL.<sup>3</sup> G10L 1/00
- (52) Index at Acceptance  
G4R 10F 10X 11A 11C 11D 11E 11X 12F  
12X 1F 3G 8G 9D



(54) VOICE INSTRUCTION IDENTIFICATION SYSTEM

(71) I, HIROYASU FUNAKUBO, a Japanese Citizen of 2-17-4, Nishiogi-Kita, Suginami-Ku, Tokyo, Japan 167 do hereby declare the invention for which I pray that a Patent may be granted to me and the method by which it is to be performed, to be particularly described in and by the following statement:-

5 The present invention relates to a voiced instruction identification system for controlling a mechanical or electric device such as powered artificial hands or manipulators by encoding voice signals. 5

10 In the voice identification systems proposed in the prior art, various voice signals from a person constituting a voice source (hereinafter referred to as a source person) have been subjected to frequency analysis to be stored as separate patterns in a memory, and a voice signal subsequently generated subjected to the same frequency analysis to be compared with said stored patterns for determining what kind of voice signal the latter is. However, in these prior art systems, there have been many problems associated with the pattern comparison because voice signals from the same source person differ from day to day or from time to time in respect of the interval between syllables and, besides, these prior art systems have been large in size and expensive to produce. 10 15

20 According to the present invention there is provided a voiced instruction identification system comprising means for converting a series of human voice sounds uttered separately into a series of electrical signals; means for extracting fundamental waves from the respective electrical signals; means for sampling at least one feature of each fundamental wave to obtain a symbolic value; means for forming an information pattern using relative changes of the symbolic values of each fundamental wave from that of other fundamental waves; and means for identifying a voiced instruction from said information pattern. 20

25 The present invention will now be described in more detail by way of example, with reference to the accompanying drawings, in which:- 25

Figure 1 is a block diagram showing the entire arrangement of a voiced instruction identification system embodying the present invention;

Figure 2 is a circuit diagram of the counter circuit of Figure 1;

Figure 3 is a circuit diagram of the symbolic value evaluating circuit of Figure 1;

30 Figure 4 is a circuit diagram of the tonal change detecting circuit of Figure 1; 30

Figure 5 is a circuit diagram of the syllable pause detecting circuit of Figure 1;

Figure 6 is a circuit diagram of the system control unit of Figure 1;

35 Figures 7 to 11 are timing charts illustrating the timed relationship of the operations of the system of Figure 1; 35

Figures 12A to D are diagrams illustrating the changes occurring between syllables;

Figure 13 is a circuit diagram of another preferred embodiment of the total change detecting circuit shown in Figure 4;

Figure 14 is a block diagram of another voiced instruction identification system embodying the present invention;

40 Figure 15 is a wave-form diagram illustrating the components of a voice signal supplied to the system of Figure 14; 40

Figures 16 to 18 are flow charts for use in the description of the operation of the system of Figure 14;

Figure 19 is a diagram showing a pattern formed from instruction words.

45 Referring now to the drawings, and particularly to Figure 1 which is a block diagram of 45

the entire system, *A* is a voiced instruction detecting circuit which converts voice sound wave signals from a source person into electric square wave signals. The voice detecting circuit *A* comprises a microphone 1 provided on the outer wall of the trachea of the source person. The output signal from the microphone is fed to a filter 2 having a pass-band of 50 to 250 Hz, which is the fundamental frequency band of vibration of the vocal chords of the source person. Thus the filter 2 removes frequency components of the voice signal other than the pitch component, the frequency of the pitch component being a parameter representing the tone of the voice sounds, emitted from the source person. The filter 2 includes an amplifier. The output signal of the filter 2 is fed to a Schmidt trigger circuit 3 which converts the signal to a square wave signal. The Schmidt trigger circuit also further removes components other than the pitch component as a result of its hysteresis characteristic.

*B* designates a counter circuit for measuring the period of the square wave output signal from the Schmidt trigger circuit 3. The counter circuit *B* comprises a gating circuit 4 for gating the output signal from the Schmidt trigger circuit 3 and also for selectively gating the output signals from a clock pulse generator circuit 6, which generates two clock pulse signals at frequencies of 10KHz and 300Hz. The output signals from the gating circuit 4 are fed to a counter circuit 5 which counts the number of clock pulses which it receives during each period of the square wave signal for the duration of the signal output from the voice detecting circuit *A*.

*C* designates a symbolic value evaluating circuit which stores a group of numbers counted by the counter 5 and then determines a symbolic value for the group of numbers. In the present embodiment of the invention, the symbolic value is the central value of the group of numbers. The symbolic value evaluating circuit comprises an array 7 of latch circuits for storing the numbers of each group of numbers and a comparator 8 which compares the numbers fed from the respective latch circuit of the array 7 to determine the symbolic value or central value of the group.

*D* designates a tonal change detecting circuit which stores the symbolic value for each syllable, as will be described in more detail below, to detect the tonal change of each syllable. The tonal change detecting circuit *D* comprises an array 9 of latch circuits for storing the symbolic value of each syllable and a comparator 10 for comparing the symbolic values from the respective latch circuits of the array 9 so as to detect the change of symbolic value which occurs between one syllable and the next syllable. This change of symbolic value represents the tonal change occurring between the two syllables.

*E* designates a syllable pause detecting circuit which detects syllable pauses by using the output of one of the latch circuits of the array 7 of the symbolic value evaluating circuit *C*. The syllable pause detecting circuit comprises a comparator 11 which compares the output of the latch circuit with a predetermined value.

*F* designates a system control unit which supplies pulse signals to the circuits *A* to *E* for providing a sequential controlling function over these circuits.

The operation of the voice identification system will now be described further in detail with reference to Figures 2 to 12 on the basis of timing charts included therein.

Assuming that three syllables (for example [tsu], [ka] and [me] which means "grasp" in the Japanese language when pronounced as a single word) are spoken by the source person as shown by voice input envelopes in Figure 7, the microphone 1 provided on the outer wall of the trachea (preferably, at a position directly below the Adam's apple where higher frequency components and vocal tract characteristics have relatively small effect) will receive the sound wave signal thus emitted by the source person. Then, the vocal signal generated by the microphone 1 is fed to the filter 2, having a pass-band of 50 to 250Hz, which passes and amplifies the pitch components of the vocal signal. The output signal from the filter 2 is converted by the Schmidt trigger circuit 3 into a square-wave signal. For a female voice source, the pass-band may be shifted to a higher frequency. Due to the hysteresis characteristic of the Schmidt trigger circuit 3, signal components other than the pitch component are further eliminated.

The square-wave output signal of the Schmidt trigger circuit is shown in Figure 7 and illustrated as an enlarged wave-form (a) in Figures 9 to 11.

The operation sequence of the system is illustrated in Figure 7 and as may be seen, consists of fourteen phases denoted by numerals [0] to [13]. These phases will now be described.

[0] denotes an input-waiting phase, in which all circuits are reset to their initial status. [1], [5] and [9] denote initial period eliminating phases. Each of these phases correspond to the building up period for an associated syllable and in each of these phases the corresponding unstable portion of the square wave output signal for the associated syllable of the Schmidt trigger circuit 3 is eliminated. In the present embodiment the unstable portion is the first square wave output pulse for the associated syllable. [2], [6], and [10] denote phases in

which the period of the square wave signal is measured. In the present embodiment the periods of the 2nd to 4th square wave pulses of each syllable are measured by counting the number of clock pulses generated during each period.

5 In each of the phases denoted by [3], [7] and [11] one of the counted numbers of clock pulses is selected out of the three counted numbers as the symbolic value for the associated syllable. In the present embodiment the central value is selected. However, in alternative 5 embodiments it may be arranged that the symbolic value is the value giving the least squares, the maximum value, or the n-th largest value of a group of counted numbers of clock pulses is selected. In other words, any arrangement may be adopted so long as one 10 value may be selected for each group of numbers of clock pulses under the same conditions. [4], [8] and [12] denote phases in which syllable pauses are detected. In the phase denoted by [13] the tonal change pattern of the three symbolic values is determined. This tonal change pattern is used to identify the vocal information spoken by the source person for causing artificial hands, controlled by the system, to start, for example, their "grasping" 15 action.

Besides controlling the artificial hands, the tonal change pattern may be applied to other uses such as controlling machine tools, financial records, or door movement. The phases [1] to [13] set forth above constitute one information identification cycle after which another information identification cycle will occur following an input-waiting phase [0'].

20 The aforementioned information identification cycle will now be described further in detail.

If the power switch is turned on, a reset pulse is applied from a power-on reset circuit 12a of the system control unit F (Figure 6) through an OR circuit 12b to a 4-bit counter 12c to reset the counter 12c and a circuit control unit 12d. The 4-bit counter 12c indicates the 25 aforementioned 12 phases of the information identification cycle in sequence and is reset to the phase [0] by the reset pulse. The thus reset circuit control unit 12d simultaneously emits pulses (c), (d), (e), (j), (j'), (j''), (O), (O'), (O'') and (ℓ) as shown in Figure 8.

30 The pulse (c) is applied to the reset input terminal of a flip-flop circuit 4a of the gating circuit 4 shown in Figure 2 to provide an output at the output terminal Q thereof. This functions to cut off the square-wave signal (a) from the Schmidt trigger circuit 3 until the entire circuit of the system is set, because an AND circuit 4b is closed and so inhibits the passage of the square wave signal (a) as no output is emitted from the output terminal Q. The pulse (d) resets the 8-bit counter circuit 5 and a flip-flop circuit 4c so that, when the square wave signal (a) is present, an output is emitted from the terminal Q of the flip-flop 4c and so that the 10KHz or 300Hz clock pulses to be described later are fed to the counter 35 circuit 5 through an AND circuit 4d. The pulse (e) sets the flip-flop circuit 4e to provide an output from the terminal Q thereof and to permit AND circuit 4f to pass clock pulses from a 10KHz clock pulse generator circuit 6a forming part of the clock pulse generator circuit 6.

40 Since the AND circuit 4d is also open, the 10KHz clock pulses from the clock pulse generator circuit 6a are fed to the 8-bit counter circuit 5 through the AND circuit 4f and an OR circuit 4g to be counted by the counter circuit 5. The pulses (j), (j'), (j''), (ℓ), (O), (O'), and (O'') reset 8-bit latches 7a, 7b, 7c and 7d, and 9a, 9b, 9c, respectively, into their input-waiting states as shown in Figures 3 and 4.

45 Then the circuit control unit 12d emits a pulse (b) (Figure 2) and so the flip-flop circuit 4a is set to provide an output at the output terminal Q which is applied to the AND circuit 4b and the terminal D of the flip-flop circuit 4c. Under these conditions, if the first square-wave pulse of the square wave signal (a) is applied to the input of the AND circuit 4b, the pulse is passed there-through to be applied to the terminal T of the flip-flop circuit 4c. When this happens, the output is changed over to the terminal Q at the leading edge of the said square wave pulse (a), and a pulse (g) is emitted from said terminal Q (Figure 9). 50 At the same time, an increment pulse is fed to the 4-bit counter 12c to advance the count from phase [0] to phase [1]. Then, the pulse (d) is again fed out from the circuit control unit 12d to reset the 8-bit counter circuit 5. Also, the flip-flop circuit 4c is reset to change its output to the Q side. Although the 8-bit counter 5 starts to count the clock pulses again, it is 55 reset by the next pulse of the pulse (d) without transmitting the output thereof to the succeeding stage when the pulse (g) is produced at the leading edge of the second pulse of the pulse (a), so that the period of the first square wave pulse is not measured. When the second pulse of the square wave signal (a) appears, namely, after the lapse of the building-up period of the first syllable, the output of the circuit 4c is again changed over to the Q side and, therefore, the 8-bit counter circuit 5 starts to count the 10KHz clock pulses 60 again. Also, an increment pulse is applied from the circuit control unit 12d to the 4-bit counter 12c to advance the count from phase [1] to phase [2].

65 Thereafter, when the third pulse of the square wave signal (a) appears and the third pulse of the pulse (g) is generated at the leading edge thereof, the circuit control unit 12d produces a pulse (i), which is applied to the strobe input terminal of the latch 7a shown in 65

Figure 3. Thus, the count signal from the 8-bit counter 5 is stored in the latch 7a. Further, the pulse (d) is emitted subsequently to the pulse (i), and the 8-bit counter circuit 5 is reset again to start the counting operation, which is continued until the fourth pulse of the square wave signal (a) is supplied.

5 When the fourth pulse of the pulse (g) is generated at the leading edge of the fourth pulse, the circuit control unit 12d produces a pulse (i') and, therefore, a function similar to those described above takes place so that the number of clock pulses counted in the period of the third square wave pulse is stored in the latch 7b and the number of clock pulses counted in the period of the fourth square wave pulse is stored in the latch 7c. The fifth pulse (d) 10 generated subsequently thereto causes an increment pulse to be produced to advance the count from phase [2] to phase [3] (in Figure 10).

15 In the third phase [3], the central value is selected out of the numbers stored in the latches 7a, 7b and 7c (the outputs of these latches being designated as A, B, and C) during the second phase [2]. In order to do this, the comparison function as shown in Table 1 is carried out in any array of 8-bit comparators 8a. 15

TABLE 1

20	Comparator output		Selected channel	20
	[A] > [B]	[B] > [C]	[B]	
		[B] ≦ [C]	[A] > [C]	
25			[A] ≦ [C]	25
	[A] ≦ [B]	[B] > [C]	[A] > [C]	
30			[A] ≦ [C]	30
		[B] ≦ [C]	[B]	

35 The selected one of the outputs A, B, and C is fed out from a data multiplexer 8b (Figure 3) having the output terminal thereof opened by a pulse (m) from the circuit control unit 12d and, at the same time, this output is stored in the 8-bit latch 9a (Figure 4) by the action of a pulse (n) fed from the circuit control unit 12d. Since a pulse (f) is fed out from the circuit control unit 12d simultaneously with the pulse (n), the flip-flop circuit 4e shown in Figure 2 is reset and the output is changed over to the  $\bar{Q}$  terminal side. Accordingly, the 300 40 Hz clock pulse is applied from the 300Hz clock pulse generator 6b through the AND circuit 4h, OR circuit 4g and AND circuit 4d to the 8-bit counter 5. Following the pulse (m), the circuit control unit 12d applies an increment pulse to the 4-bit counter 12c to advance the count from phase [3] to the phase [4].

45 In the fourth phase [4] the syllable pause is detected. The 8-bit counter 5 counts the 300 Hz clock pulses for each pulse (d) and stores the number of clock pulses in the latch 7d for each pulse (k) from the circuit control unit 12d. The output of the latch 7d is transmitted to one of the input terminals of a comparator circuit 11a, to the other input terminal of which is applied a signal corresponding to a suitable syllable pause. Assuming that the output A of the latch 7d is larger than the reference signal B ( $A \geq B$ ), that is to say, when the output 7d 50 is larger than the output predetermined as being suitable for the syllable pause, the comparator circuit 11a produces a syllable pause pulse(s), while, in the case of  $A < B$ , the comparator circuit 11a feeds out a voiced pulse (t). 11b and 11c are AND circuits for transmitting the pulse(s) or (t), respectively, on the application of a timing pulse (r) from said circuit control unit 12d. If the voiced pulse (t) is output from the comparator circuit 55 11a, the aforementioned operation is repeated until a syllable pause pulse (s) appears. When the first pulse (a) of the next syllable is detected and the pulse (s) appears, the circuit control unit 12d provides the pulses (e), (e'), (j), (j'), and (j'') (in Figure 6). Accordingly, the flip-flop circuit 4e shown in Figure 2 is set and the sequence is changed over to the 10KHz clock pulse generator circuit 6a and, thereafter, the 8-bit counter 5 counts 10KHz 60 clock pulses. Also, the latches 7a through 7d are cleared. When the first pulse (d) in the second syllable is provided, an increment pulse is applied to the 4-bit counter 12c to advance the count from the fourth phase [4] to the fifth phase [5].

65 Since the operations in the fifth to twelfth phases are identical to those already described, their description is omitted. The latch 9b shown in Figure 4 stores the central value of the second syllable, and the latch 9c stores the central value of the third syllable. Further upon 65

detecting syllable pause pulse (h) (in Figure 11), in the third syllable, the circuit control unit 12d provides the pulse (c) as a word end signal to cut off the square wave signal (a).

The most significant (upper) four bits of the outputs of the latches 9a to 9c are applied to the inputs of adders 10a to 10c, respectively, to be added to the original 8-bits, thereby, to provide the central values with a certain width. This is because the output of latch 9a does not always coincide with the output of latch 9b due to the variations peculiar to the human voice even if the tones are on the same level. Therefore, it is necessary to provide a tonal difference wider than one scale when a tonal change exists in the generated voice.

Then, a first comparator array 10d compares the input [1] from the latch 9a with the input [2] from the adder 10a, and the input [3] from the latch 9b with the input [4] from the adder 10b, while a second comparator array 10e compares the input [1] from the latch 9b with the input [2] from the adder 10b, and the input [3] from the latch 9c with the input [4] from the adder 10c. As a result of these comparison operations, 2-bit data as shown in Table 2 are produced.

For example, if the tone of a word composed of three syllables

TABLE 2

Comparator output	2-bit data
$\begin{array}{c} \text{Input (1)} > \text{Input (3)} \\ \text{Input (1)} > \text{Input (4)} \end{array}$	11 (Tone up)
$\begin{array}{c} \text{Input (1)} \leq \text{Input (4)} \\ \text{Input (1)} \leq \text{Input (3)} \end{array}$	10 (Level tone)
$\begin{array}{c} \text{Input (2)} \geq \text{Input (3)} \\ \text{Input (2)} < \text{Input (3)} \end{array}$	10 (Level tone)
	01 (Tone down)

changes as shown in Figure 12A, outputs (11, 01) will be generated. The tonal change as shown in Figure 12B produces outputs (10, 11). Nine combinations of these outputs are possible in all. These combinations are stored in a 4-bit latch of 10f by the action of a pulse (p) from the circuit control unit 12d and used to control an electric motor, for example, for performing the "grasping" action of the artificial hands. Upon completing such a "grasping" action, the circuit control unit 12d provides the pulse signal (q) for clearing the 4-bit latch 10f and, then, emits pulses (e), (e'), (j), (j'), (j''), (O), (O') and (O'') to return the sequence to the initial state.

Although the foregoing description of the preferred embodiment has been set forth for the recognition of a word composed of three syllables, recognition of a word composed of one or two syllables may be readily performed as well. Hereinafter, the description will be made with reference to Figure 13.

Assuming that only the first syllable is input to the system, the sequence is advanced up to the third phase in the same manner as mentioned hereinbefore to determine the symbolic value which is to be stored in the latch 9a. Then, the sequence is changed over to the 300Hz clock pulse generator circuit 6b, and the 8-bit counter 5 counts the 300Hz clock pulses. In this case, due to the absence of the second syllable, the 8-bit counter 5 overflows to emit the pulse (h), and the circuit control unit 12d transmits the pulse (c) to reset the flip-flop circuit 4a for cutting off the square wave signal (a) thereto. Immediately thereafter, the pulse (q) is provided for clearing the 4-bit latch 10f. At the same time, the circuit control unit 12d emits pulses (u) and (v) in Figure 13 which, after being inverted through inverter circuits 10g and 10h, inhibit the outputs of the comparator arrays 10d and 10e at the AND circuits 10i and 10j. Then the pulse (p) is supplied to the input of the 4-bit latch 10f and the latter produces outputs (00,00) indicating the end of the word as shown in Figure 12C.

In the case where the first and second syllables are input to the system, but the third syllable is not input, the sequence is advanced up to the seventh phase in the manner as described previously. That is to say, the symbolic values of the first and second syllables are stored in the latches 9a and 9b, respectively, and the 8-bit counter 5 counts 300 Hz clock pulses. In this case, since the third syllable does not appear, the 8-bit counter 5 overflows to emit the pulse (h). Therefore, the flip-flop circuit 4a is reset by the pulse (c). Then, 4-bit



a pattern from the rhythm data stacked in the rhythm data stack 73. 75 is a stress data stack for stacking stress data of sound which is provided at the output of the analog-to-digital converter 58. 76 is a stress pattern forming circuit for forming a pattern from the stress data stacked in the stress data stack 75. 77 denotes an instruction word data bank in which predetermined patterns corresponding to instruction words are stored. 78 is a pattern matching circuit which synthesizes patterns fed from the musical interval pattern forming circuit 72, rhythm pattern forming circuit 74 and stress pattern forming circuit 76. The pattern matching circuit 78 also functions to compare the synthesized pattern with the patterns corresponding to instruction words stored in the instruction word data bank. 79 denotes an output converter circuit for converting an output instruction word signal from the matching circuit 78 into a signal which can be used for controlling a machine or other equipment provided externally on the source person.

Description will now be made of the operation of the system having the aforementioned arrangement.

Assuming that the source person has emitted four syllables as shown by the vocal input envelope in Figure 15a, this vocal information is received as a vocal sound wave signal by the microphone 51 provided on the outer wall of the trachea (preferably, at a position directly below the Adam's apple where higher frequency components and vocal tract characteristics have relatively small effect) and then fed as an electric vocal signal to the amplifier 52 to be amplified therein. Components other than the pitch component are removed by the band-pass filter 53 which passes signal components lying in a frequency band of, for example, 50 to 250 Hz. Then, the output of the filter 53 is fed to the Schmidt trigger circuit where components other than the pitch component are further removed due to the hysteresis characteristic of the Schmidt trigger circuit 54 and the signal is converted into a square wave signal (Figure 15c). For a female voice source, the pass band may be shifted into a higher frequency.

The output of the amplifier 52 is also fed to the envelope detector 56 which detects its amplitude. The peak value of the amplitude is detected and retained by the peak detector 57 (Figure 15d). Then, this peak value is converted into a digital value by the analog-to-digital converter 58.

The sequence of operations will now be described with reference to the flow charts shown in Figures 16 to 18.

The sequence controller 55 has three controlling modes which commence respectively at a first start position, a second start position and a third start position. The sequence controller 55 assumes the first start position at the leading edge of each square wave pulse from the Schmidt trigger 54, the second start position when the period counter 60 overflows (for example, by 50ms) between successive syllables, and the third start position when the rhythm counter 62 overflows (for example, by 500ms) at the end of an instruction word. The period counter 60 and the rhythm counter 62 are arranged so that the period counter 60 will overflow but the rhythm counter 62 will not overflow between successive moras.

The system also operates in four operating states, namely, a phase 0 state in which the system is waiting for the beginning of a syllable phase 1 state in which a syllable is being received, a phase 2 state in which the system is waiting for a syllable pause, and a phase 3 state in which an input signal is recognised as not comprising an instruction word and is rejected.

Referring now to Figure 16, when the system is waiting to receive vocal information it will be in the phase 0 state. Then, when vocal information is received, the sequence controller 55 will assume the first start position when the leading edge of the first square wave pulse is fed to it from the Schmidt trigger 54. Because the system is in the phase 0 state, rhythm counter 62 is reset, clock pulse generator 59 starts generating clock pulses, and period counter 60 and rhythm counter 62 start counting. The operating state is then shifted from phase 0 to phase 1.

When the leading edge of the next square wave pulse is fed from the Schmidt trigger circuit 54 into the sequence controller 55, the sequence controller returns to its first start position. Because the operating state is now phase 1, the data unit in the period counter 60 is latched in the counter buffer register 63, the period counter 60 is reset and then starts counting again. Next the data unit latched in the counter buffer register 63 is compared in the digital comparator 64 with the values preset in the upper and lower pitch limit registers 65 and 66, namely, values corresponding to the levels  $\alpha$  and  $\beta$  shown in Figure 15b. If the data unit fed from the register 63 is outside the range defined by the values in the registers 65 and 66, the system waits for the next pulse to be fed from the Schmidt trigger 54. If the data unit fed from the register 63 is in said range defined by the preset values  $\alpha$  and  $\beta$ , the data unit in the counter buffer register 63 is subjected to a logarithmic conversion in the logarithmic converter circuit 67 and the converted data unit is stacked in the data stack 68. It is then determined whether the number of data units in the data stack 68 has reached a

predetermined value or not. If said predetermined value is not reached, the system waits for the next square wave pulse to be fed from the Schmidt trigger 54, and data units are thus successively stored in the data stack 68. When a predetermined number of data units, namely, four data units in the preferred embodiment, are stacked in the data stack 68, the symbolic value evaluating circuit 69 determines the mean value of the data units which is then fed into the interval comparator 70 as a symbolic value. The interval comparator 70 then compares each data unit from the data stack 68 with the symbolic value determined as above. If any one of the data units in the data stack 68 is outside said range, the data stack 68 clears the initially stacked data units, and the system waits for the next square wave pulse to be fed from the Schmidt trigger 54. This operation is repeated until all the data units in the data stack 68 lie with the said range from the symbolic value. When all data units lie within said range from the symbolic value, the symbolic value is fed into the musical interval data stack 71 and stacked therein. Under these conditions, the system takes the syllable-pause waiting state, namely, the phase 2 state and waits for the next square wave pulse to be fed from Schmidt trigger circuit 54.

Then, when the next square wave pulse is fed from the Schmidt trigger circuit 54, the period counter 60 is reset, without the contents being latched in the counter buffer register 63 because the operating state is neither phase 0 nor phase 1. Thereafter, the sequence controller 55 causes the aforementioned operation to be repeated until the sequence controller moves to the second start position.

Referring now to Figure 17, when the period counter 60 overflows, the sequence controller 55 assumes the second start position. If the operating state is phase 1 indicating that stable pitch data have not been obtained, the state is shifted to phase 3 to reject the input words, and the system then waits for the next square wave pulse to be fed from the Schmidt trigger circuit 54. If the state is phase 3, the operation is identical to that described immediately above.

If the state is phase 2, the data unit in the rhythm counter 62 is stacked in the rhythm data stack 73 and the data unit in the A-D converter 58 is stacked in the stress data stack 75. Then, it is determined whether the number of units of the data in each of the musical interval data stack 71, rhythm data stack 73 and stress data stack 75 is four, or more. If it is less than four, the period counter 60 is reset and stops its counting operation, the rhythm counter 62 is reset and starts counting again, and the system waits for the next square wave pulse to be fed from the Schmidt trigger circuit 54, with the operating state returned to phase 0.

Then, the aforementioned operation is repeated again. Referring now to Figure 18, when the rhythm counter 62 overflows, the sequence controller assumes the third start position, and it is determined whether the state is phase 3 or not. If the state is phase 3, the system is generally reset. In other words, all registers are reset and the counter operation is stopped. The state is shifted to phase 0 and the sequence controller 55 stops operating.

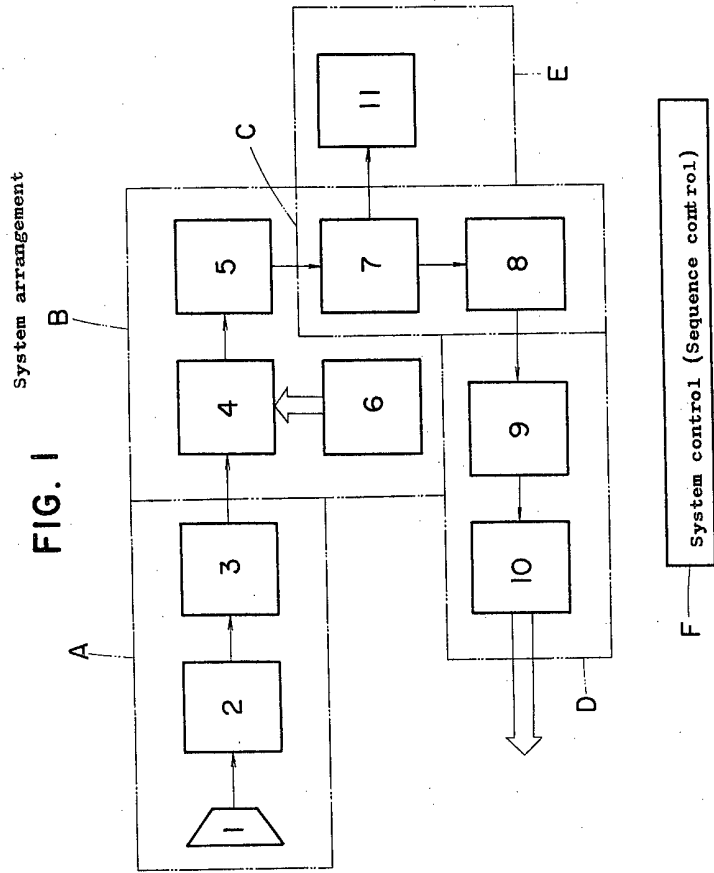
If the state is not phase 3, the data in the musical interval data stack 71, rhythm data stack 73 and the stress data stack 75 are formed into patterns, in the musical interval pattern forming circuit 72, the rhythm pattern forming circuit 74 and stress pattern forming circuit 76 through a relative comparison of the respective data units, and the data units thus formed into patterns are synthesized in the pattern matching circuit 78 into a musical note type pattern as shown in Figure 19. Then this pattern is compared in the pattern matching circuit 78 with the patterns in the instruction word data bank 77 in order to determine which instruction the pattern obtained from the voice sound is, and what kind of function the controlled system such as a machine, should make. This output instruction is converted by the output converter circuit 79 into a signal for causing the controlled machine to make a predetermined function. Upon completing this function, the system is generally reset and the operating state is shifted to phase 0.

In the systems described hereinbefore the microphone is provided advantageously on the outer wall of the trachea of the source person where the noise level is relatively low and the voice is subjected to detection by being divided into syllables. Also, advantageously the initial unstable portion of each sound is removed and the remaining portion is used for determining a value symbolic of the fundamental frequency thereof and the tonal change pattern exhibited by a plurality of thus determined symbolic values is subjected to matching for providing a specific controlling function of a device to be controlled. The system disclosed by way of the second embodiment detects the pitch, duration and stress of the sounds of each instruction word and synthesizes these factors into a musical note type of pattern. Thus, a higher rate of voice recognition can be achieved regardless of whether the voice comprises ordinary sounds or humming sounds. Also, the adoption of a simplified pattern matching system ensures a higher rate of voice recognition. Further, the combination of two or three patterns permits a plurality of instruction words to be properly identified.

**WHAT I CLAIM IS:-**

1. A voiced instruction identification system comprising means for converting a series of human voice sounds uttered separately into a series of electrical signals; means for extracting fundamental waves from the respective electrical signals; means for sampling at least one feature of each fundamental wave to obtain a symbolic value; means for forming an information pattern using relative changes of the symbolic value of each fundamental wave from that of other fundamental waves; and means for identifying a voiced instruction from said information pattern. 5
2. A voiced instruction identification system according to Claim 1, wherein said sampling means includes means for sampling at least one of frequency, amplitude, and duration of the fundamental wave. 10
3. A voiced instruction identification system according to Claim 1, wherein said fundamental wave extracting means includes means for eliminating the unstable initial few periods of the fundamental wave.
4. A voiced instruction identification system substantially as hereinbefore described with reference to and as shown in Figures 1 to 12 of the accompanying drawings. 15
5. A voiced instruction identification system substantially as hereinbefore described with reference to and as shown in Figures 1 to 12 as modified by Figure 13 of the accompanying drawings.
6. A voiced instruction identification system substantially as hereinbefore described with reference to and as shown in Figures 14 to 19 of the accompanying drawings. 20

MARKS & CLERK,  
Alpha Tower,  
ATV Centre,  
Birmingham, B1 1TT.  
Agents for the Applicant.



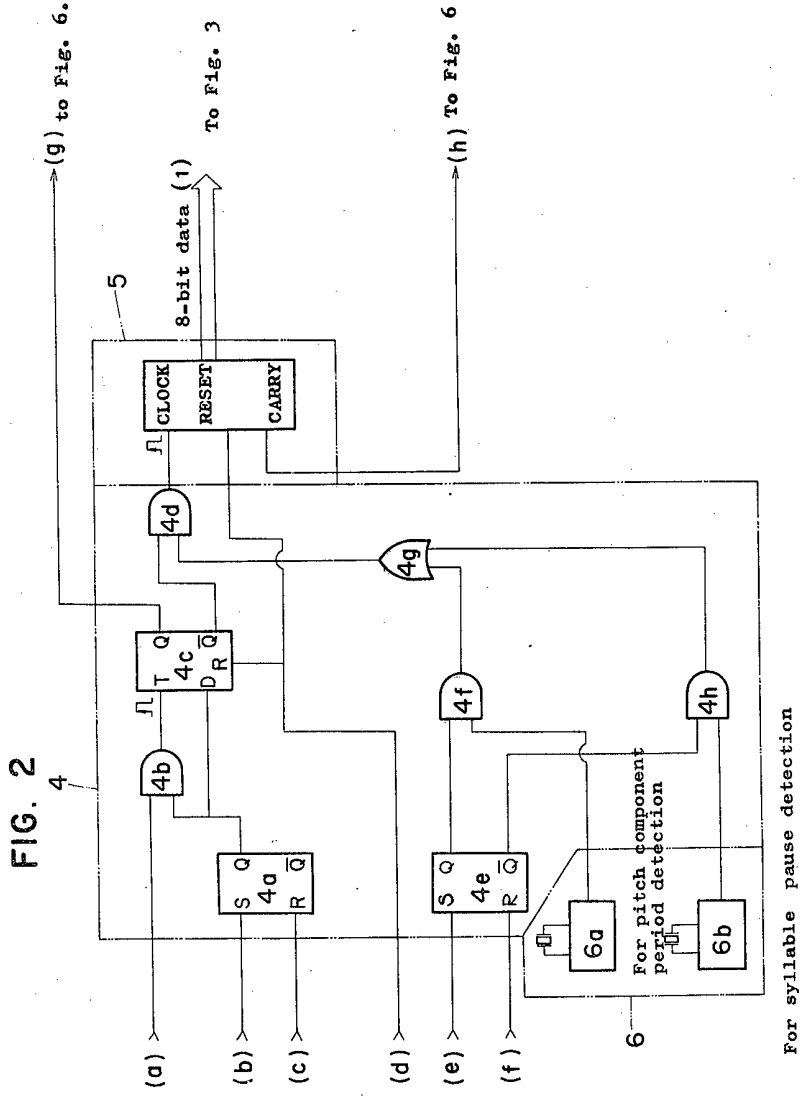


FIG. 3

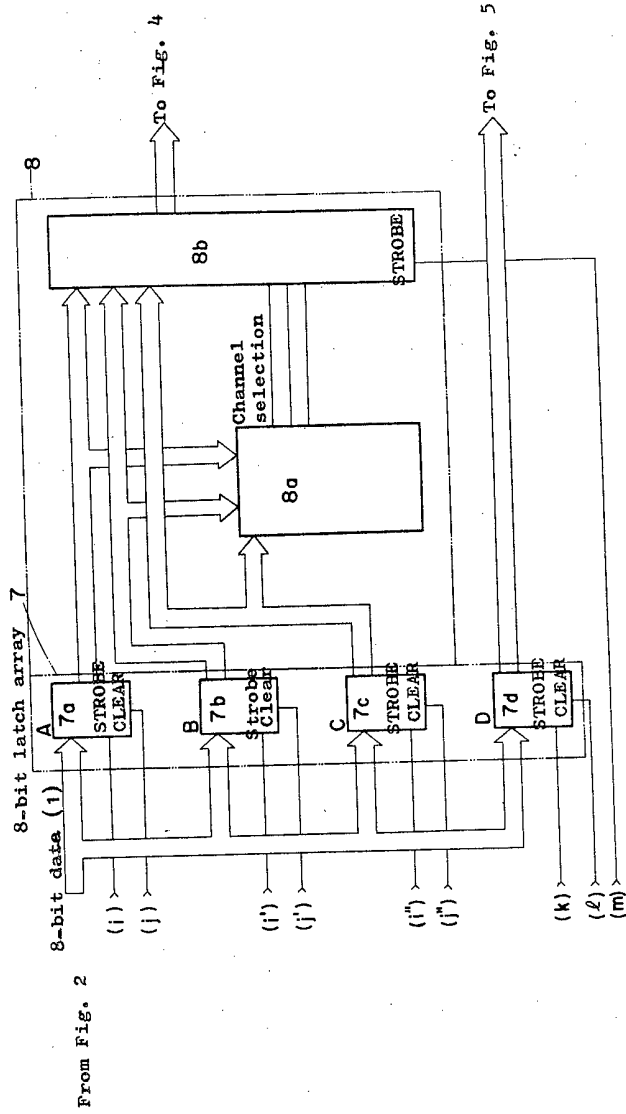
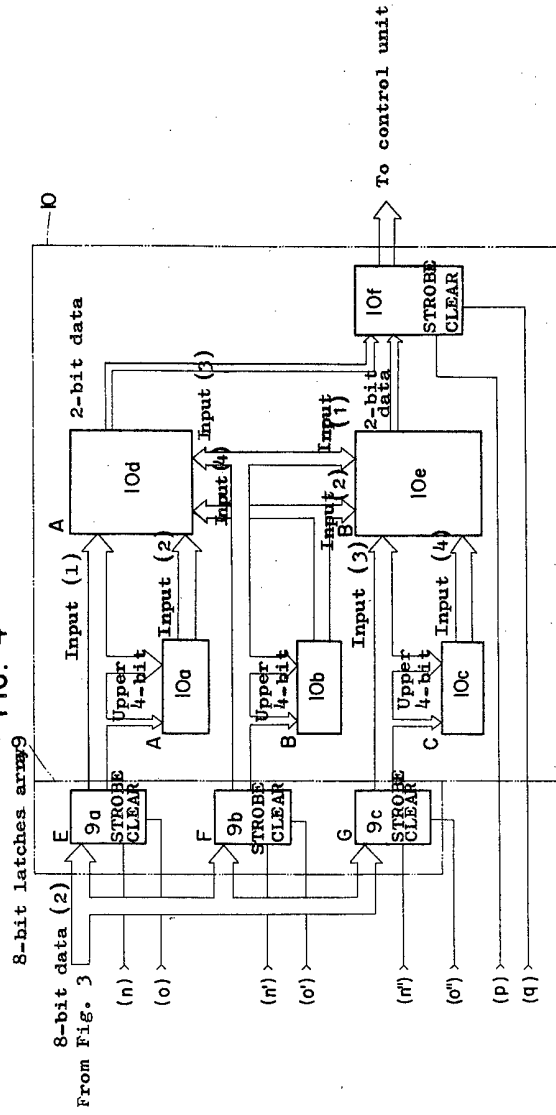
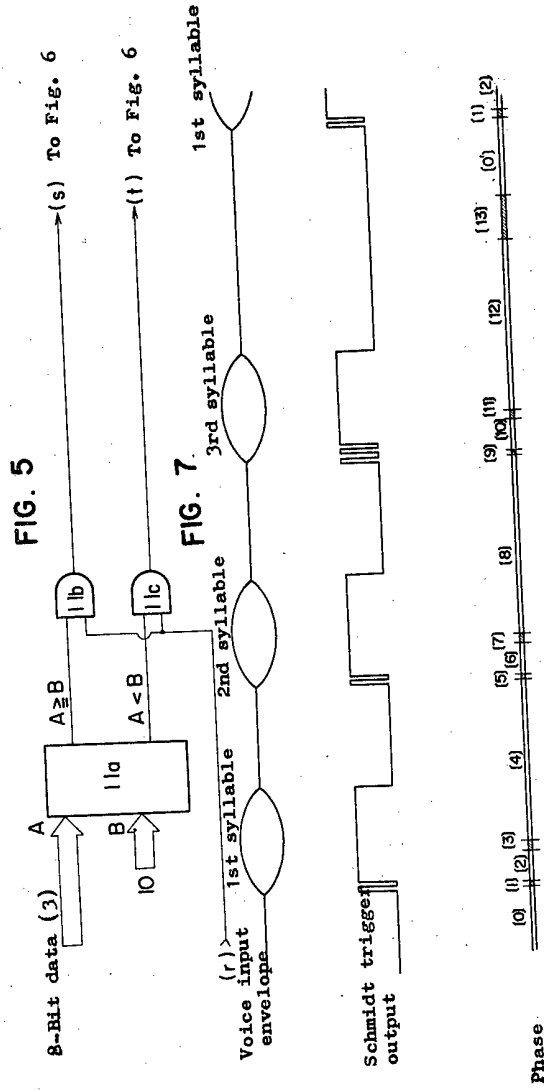


FIG. 4





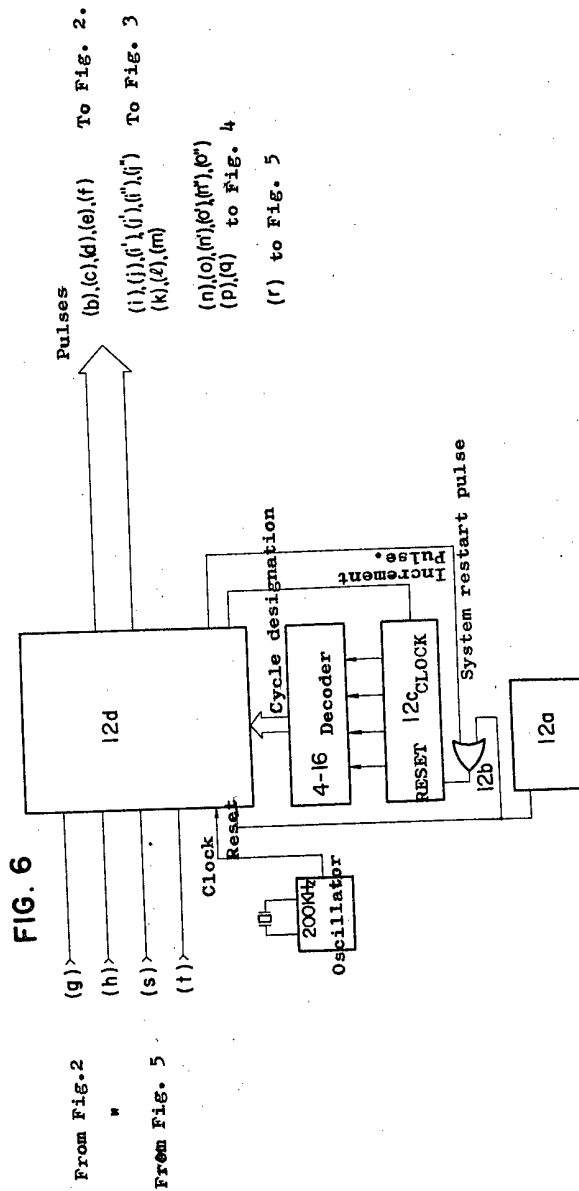


FIG. 8

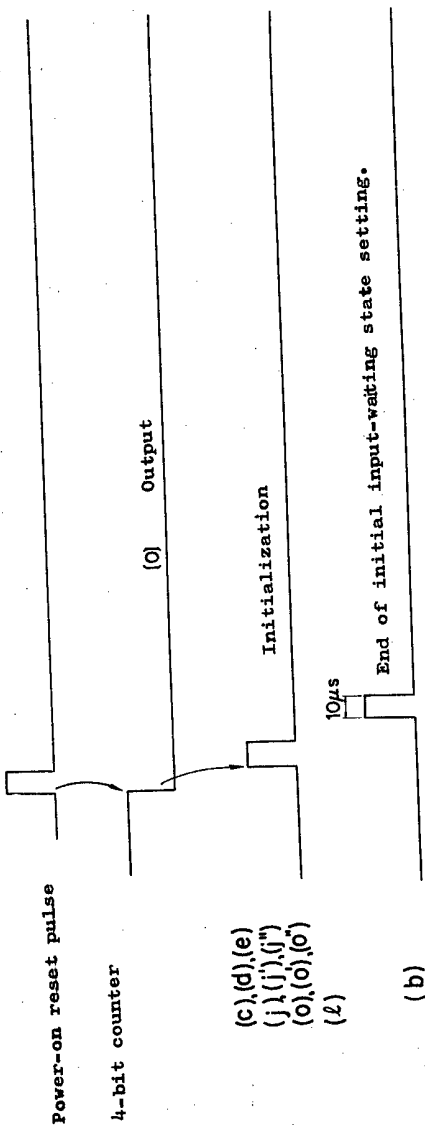


FIG. 9

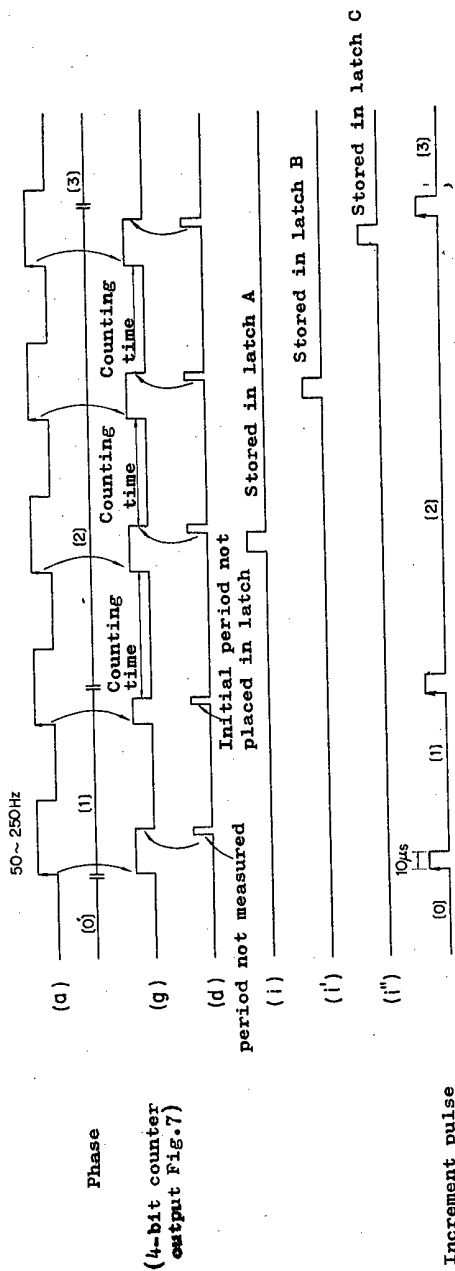


FIG. 10

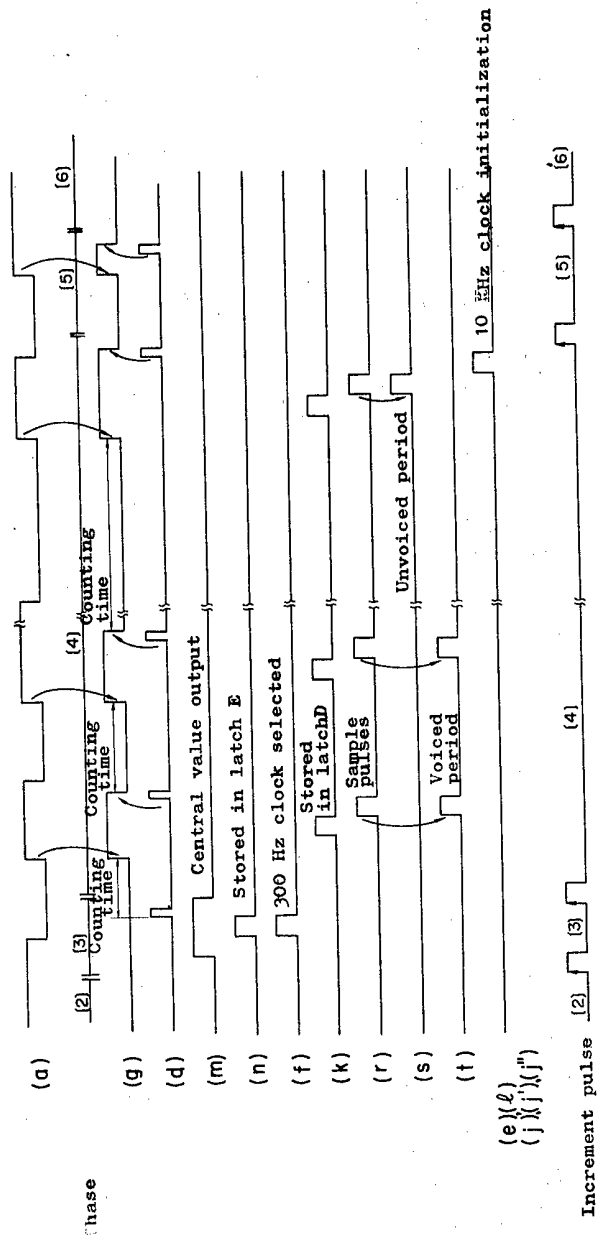
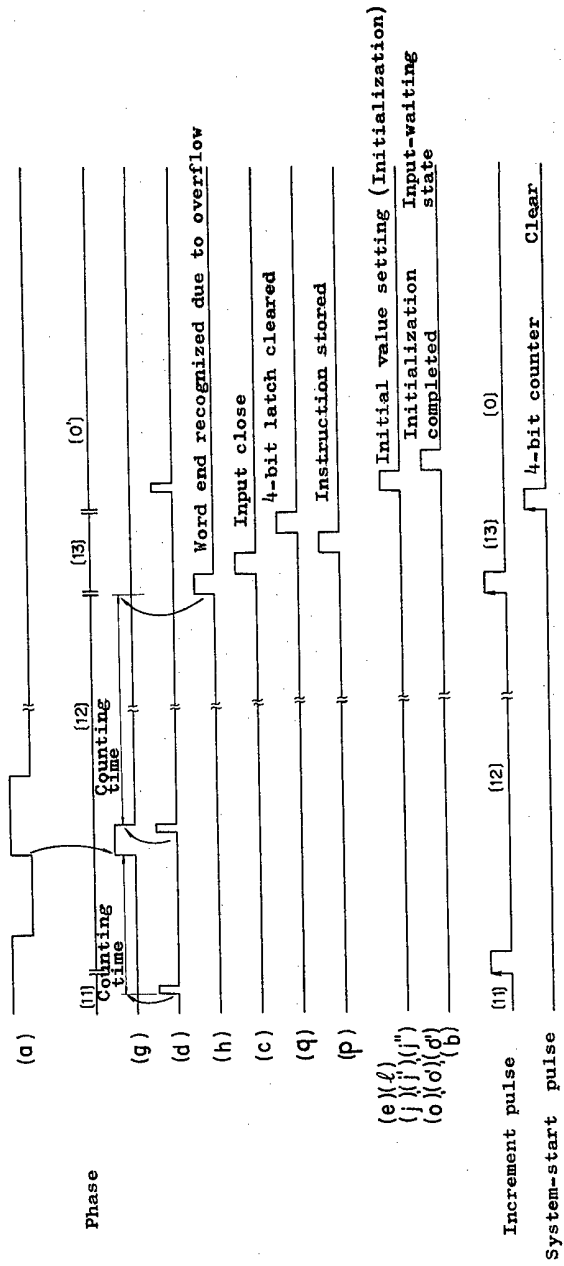


FIG. 11



*This drawing is a reproduction of  
the Original on a reduced scale  
Sheet 11*

FIG. 13

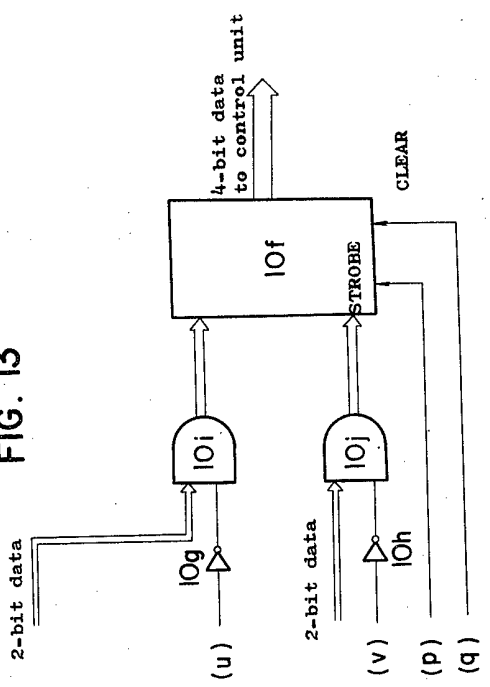


FIG. 12C

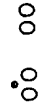


FIG. 12D

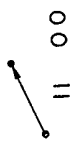


FIG. 12A

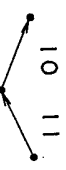


FIG. 12B



FIG. 14

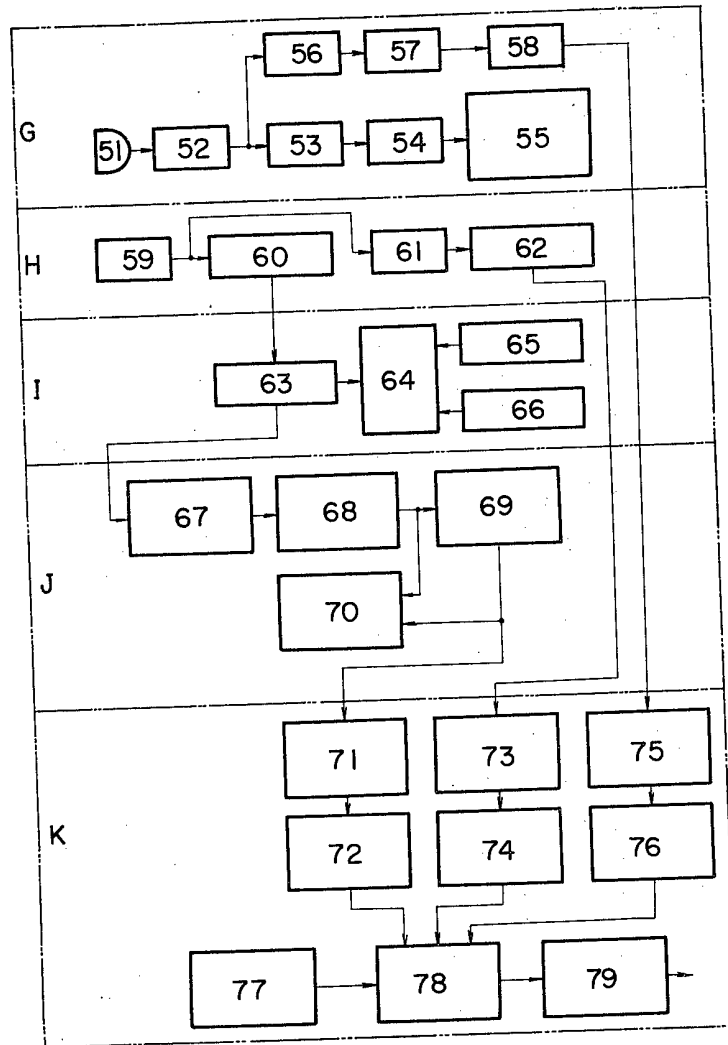




FIG. 15a

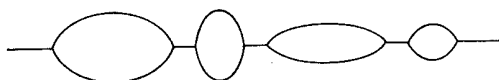


FIG. 15b



FIG. 15c



FIG. 15d

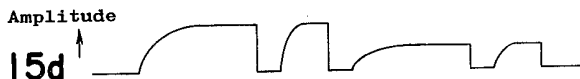


FIG. 16

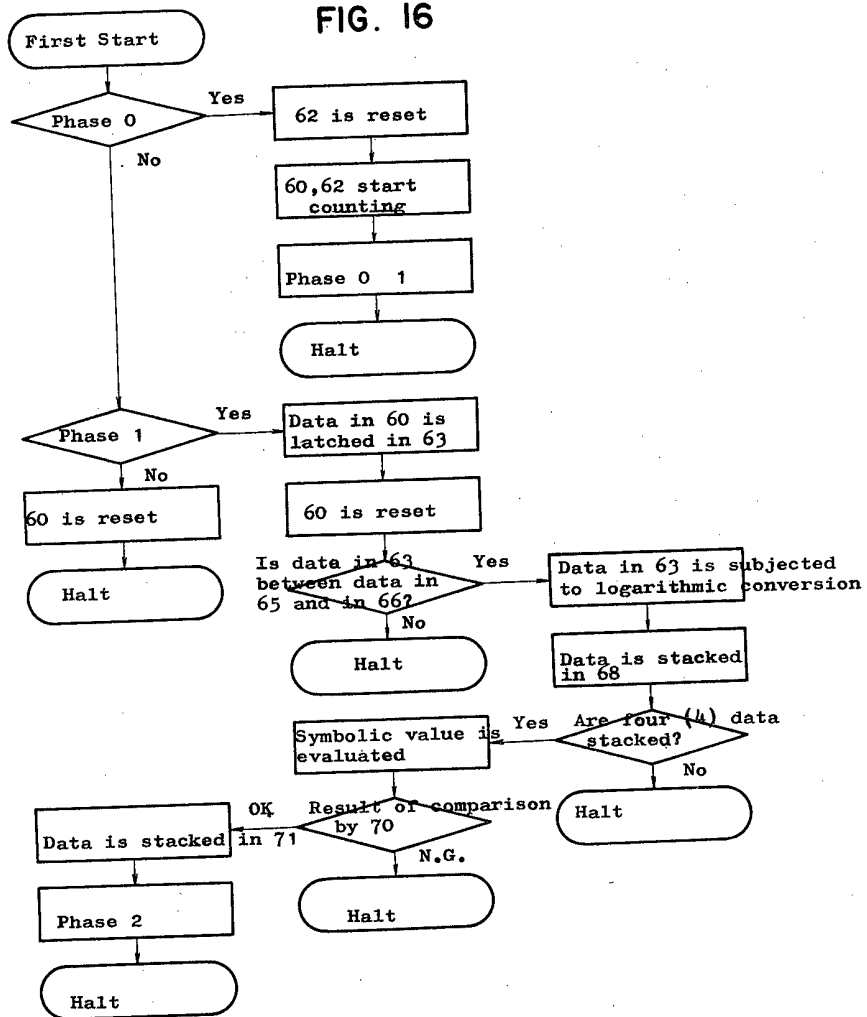


FIG. 17

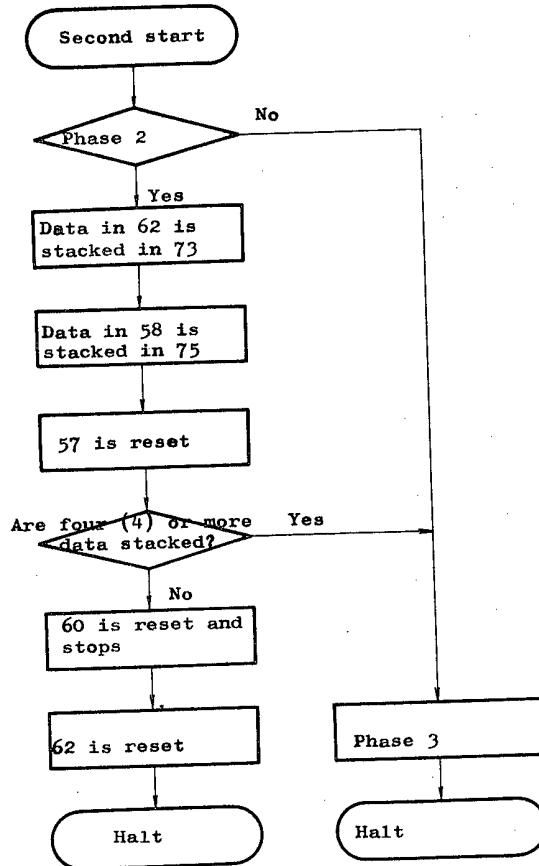


FIG. 18

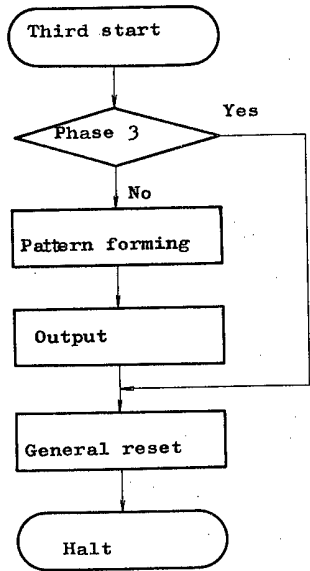


FIG. 19

