

12 **EUROPEAN PATENT APPLICATION**

21 Application number: 81201230.0

51 Int. Cl.³: **G 10 L 1/08**

22 Date of filing: 30.10.81

30 Priority: 31.10.80 NL 8005989

43 Date of publication of application:
12.05.82 Bulletin 82/19

84 Designated Contracting States:
BE DE FR GB SE

71 Applicant: **Staat der Nederlanden (Staatsbedrijf der
Posterijen, Telegrafie en Telefonie)**
P.O. Box 430
NL-2260 AK Leidschendam(NL)

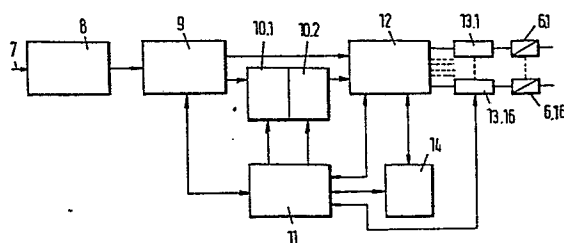
72 Inventor: **Maathuis, Petrus Gerardus Maria**
Edelsteenweg 15
NL-2651 SP Berkel en Rodenrijs(NL)

72 Inventor: **Oberman, Maarten Roelof**
Bergluststraat 6a
NL-3054 BM Rotterdam(NL)

54 **Multichannel digital speech synthesizer employing adjustable parameters.**

57 Device for decoding a number of linear-predictive-coded speech signals. The device comprises a digital noise generator, a pitch generator, a voiced/unvoiced switch, an adjustable ladder filter and a volume control. In order that the device can be used as efficiently as possible for a number of channels simultaneously, the device also comprises a pre-processing unit (9), a store (10) and a computing device (12), permitting the speech signals to be reproduced more accurately according as fewer channels are used at the output. When all channels are used, the quality of speech decreases. This variation of quality according to the number of channels used is achieved by varying the number of adjusting coefficients per channel.

FIG. 2



- 1 -

Multichannel digital speech synthesizer employing adjustable parameters.

The invention relates to a digital speech synthesizer comprising a digital noise generator, an adjustable digital pitch generator, an adjustable digital filter and a digital-to-analog converter, the pitch generator, the multiplier and the filter being adjusted
5 according to parameters derived from the original speech signal by the "linear predictive coding (LPC)" method, and means for computing the adjusting coefficients.

Devices of this type are generally known and have the advantage that for synthesizing the speech signal storage and processing
10 require a considerably smaller number of bits than with other digital speech synthesizing methods.

In known devices of said type, it is a problem that the relatively complex computing equipment required for the interpolations and for adjusting the filters is expensive.

15 The device according to the invention provides a solution for the said problem, because a large number of channels can be handled simultaneously, so that a considerably higher efficiency can be achieved.

Another object of the invention consists in providing a device of
20 the said type, in which the quality of the speech produced depends on the degree of occupation. By making an admissible sacrifice of quality, the degree of occupation can be considerably changed. This can be done, for example, by decreasing or increasing the number of

interpolations between successive speech samples, by changing the adjustment of the adjustable filter or by changing the de-emphasis. In this manner, the quantity of the speech channels handled by the device can be increased, albeit - as has been said - at the expense
5 of the quality. On the other hand, in low-traffic hours, the quality can be improved considerably. Such a device offers great advantages, e.g. with message sources, such as those for time (speaking clock), weather forecasts, cueing systems and so on. When use is made of a transmission medium having a high traffic
10 density, a special embodiment offers the possibility of providing, already at the transmitting end, means for transmitting a reduced quantity of information per channel on the transmission medium in periods of increased traffic density.

The device according to the invention is the synthesizing part and
15 is characterized in that the means for computing the adjusting coefficients can serve a plurality of channels on a time-division basis. Thus, a more efficient use can be made of these relatively expensive means.

Another feature of the device according to the invention consists
20 in that the quality of the reproduced speech signal is adjustable, because this device comprises means for computing the number of bits per adjusting coefficient according to the number of channels for which speech samples have to be computed.

The device may comprise means for computing the number of bits of
25 the adjusting coefficients according to the information content of the signal supplied to the input of the device.

The invention will now be explained in connection with the appended drawing, in which

Fig. 1 is a general block diagram of a speech synthesizer;

30 Fig. 2 is a block diagram of the speech synthesizer according to the invention;

Fig. 3 illustrates a preferred embodiment of the pre-processing unit
of the device according to the invention;

Fig. 4 illustrates a preferred embodiment of the computing unit of
35 the device according to the invention;

Fig. 5 is the flow-chart of the pre-processing unit (9) of the device
of Fig. 2 and

Fig. 6 is the flow-chart of the central processing unit (11) of the device of Fig. 2.

Fig. 1 is a general block diagram of a speech synthesizer. The adjusting parameters for the device are designated by the letters a, b, c and d. The circuit comprises a digital noise source 1, which generates white noise for unvoiced speech components, and a digital pitch generator 2, which generates the fundamental frequency for voiced speech components and is adjusted according to parameter a. The choice between generators 1 and 2 is made by switch 3 as controlled by parameter b. Then the digital signal is applied successively to an adjustable digital ladder filter 4, controlled by parameter c, and a digital volume regulator 5, controlled by parameter d. A digital-to-analog converter 6 converts the digital signal into an analog signal. Fig. 2 is a block diagram of the device according to the invention. A digital input signal incorporating the parameters a, b, c and d is applied to input 7 of the speech synthesizer and led to a buffer 8. The parameters a, b, c and d have been determined by the "linear predictive coding" method and can come from a storage medium, in the case of a message that has to be repeated regularly or from a transmission line. A preprocessing unit 9 ensures the reading of the parameters and their storage in portion 10.1 of store 10, the interpolation of two successive groups of parameters, the transfer of the interpolation results to other parts of the circuit and the passing of control data to the central processing unit 11. The data stored in store portion 10.1 can be transferred to a second store portion 10.2, when the preceding data stored in 10.2 have been processed. Processing takes place in a computing unit 12, which employs the interpolated data for adjusting the ladder filter (Fig. 1;4) incorporated in the computing unit. In the meantime store portion 10.1 is filled again. The computing unit 12 of this embodiment can compute the digital speech signals for 16 speech channels simultaneously. These digital speech signals are stored in "first-in-first-out" buffers 13.1 ... 13.16 (one signal per channel) and then led to digital-to-analog converters 6.1... 6.16, respectively. The computing unit 12 is controlled in conformity with fixed rules by a

control unit 14, which receives its instructions from the central processing unit 11.

Fig. 3 illustrates a preferred embodiment of the pre-processing unit 9 according to the invention, and store portion 10.1. The data coming from the buffer (Fig. 2; 8) are led to a series-to-parallel converter 15. The discriminator 16 infers from the first few bits of a 24-bit frame whether this frame contains speech data or control information, in which cases a data buffer 17 or a control buffer 18 is opened, respectively. The speech data are led from the data buffer 17 via a data bus 19 to a microprocessor 20, which is connected to the central processing unit (Fig. 2; 11) via a control bus 21 and an address bus 22. Store 23 (RAM) and decoding store 24 (ROM) are also connected to this data bus. Further, the circuit comprises an adder-multiplier 25 for carrying out parts of interpolation calculations.

The group of parameters comprises, as has already been observed, the following four:

- a. the pitch, for which the data are led via output 27 to the computing unit (Fig. 2; 12);
- 20 b. the voiced/unvoiced decision, for which the data are also led to said computing unit via output 26;
- c. the reflection coefficients K_1^n to K_M^n , 1st and Mth reflection coefficient, respectively, of the nth group of parameters read, the data of which are led to store portion 10.1, while the data transferred to store portion 10.2 are being handled;
- 25 d. the volume of output signal C^n , the data of which are also led to store portion 10.1.

The data for parameters a and b have no effect upon the ladder filter (Fig. 1; 4), so that they can be handled in a different manner and need not pass the store 10, as will be described hereinafter.

The function of the speech data portion of the circuit of Fig. 3 is described as separating the parameters a, b, c and d and interpolating the parameters c and d. Interpolation is necessary, because the speech information arrives in bursts and because annoying clicks could occur without interpolation. Every time a group of data n is written in

register 23, after which the interpolation with the preceding group of data, $n-1$, is carried out by the microprocessor 20 with the aid of the adder-multiplier 25 and store 24 in accordance with the formula

$$5 \quad K_i^n = K_i^{n-1} \cdot \left[\frac{L-I}{L} \right] + K_i^n \left[\frac{I}{L} \right], \quad (1)$$

in which $I = 0 \dots L-1$, L being the number of interpolations desired. I indicates the number of a group interpolated between the groups $n-1$ ($I=0$) and n ($I=L$). The coefficients $\frac{L-I}{L}$ and $\frac{I}{L}$ are generated by the microprocessor 20.

- 10 The reflection coefficients interpolated on the basis of rule (1) and the interpolated volume are led to store 10.1.

According to the invention, the pre-processing unit comprises means for adjusting the quality of the speech reproduced according to the degree of occupation of the transmission medium. Therefore, at the

- 15 transmitting end, relevant data are sent along with the control signals.

These data are interpreted in the function decoder 28. Further, the circuit comprises a register 29, for recording the number of interpolations to be carried out by the microprocessor 20 on the unvoiced part of the speech, and a register 30, which has an analogous function

- 20 with regard to the voiced part of the speech. Registers 29 and 30 are connected to ROM store 31, which converts the number of interpolations to be carried out into a signal for positioning counter 32, stepping in synchronism with a counter incorporated in microprocessor 20. The position of counter 32 is passed to a fraction table 33 (ROM), connected

- 25 via a selector 34 to control bus 21 and address bus 22. Under the control of the central processing unit (Fig. 2; 11), the number of interpolations to be carried out by the microprocessor 20 can be fixed.

The circuit of Fig. 3 also contains registers 35 and 36 for recording adjusting data for the adjustable filter incorporated in the computing

- 30 unit (Fig. 2; 12). The adjusting data for unvoiced speech are stored in register 35, those for voiced speech in register 36. A ROM 37, converts the adjusting data into positioning data for counter 38.

Via selector 39 the counter position is passed to buses 21 and 22, after which the number of calculations to be carried out by the control

unit (Fig. 2; 14) is fixed under the control of the central processing unit (Fig. 2; 12).

Furthermore, the circuit may contain a register 40 for recording a signal indicating that the next one or two frames contain no speech.

- 5 The relevant data can be passed via selector 41 and buses 21 and 22 to the central processing unit (Fig. 2; 12), so that the computing unit (Fig. 2; 14) can spend the time thus saved in dealing with other channels. The circuit may comprise a register 42 and a selector 43 for recording the signal indicating that one or two new frames contain the
- 10 same information as the preceding frame, so that the new frames need not be transmitted. Because the preceding frame is in the buffer (Fig. 2; 8) for interpolation purposes, repetition will suffice, so that transmission capacity is saved. In an analogous way information concerning the degree of compression and expansion of the speech
- 15 signal can be received and handled.

- Fig. 4 illustrates a preferred elaboration of store 10.2, computing unit 12, buffers 13 and control unit 14. The data stored in 10.1 (Fig. 2) are transferred to store 10.2 under the control of the central processing unit 11. The data stored in 10.2, containing the information
- 20 for computing the digital signal to be supplied to the buffers 13, are led to multipliers 44 and 45 working in parallel, adder-subtractor 46, AND-circuit 47 and D-flip-flop 48. Selector 49 determines the number of bits to be calculated per PCM-word and a round-off factor. D-flip-flop 50 ensures in a well-known manner the adaptation to bus traffic. The
- 25 results of a first calculation are written, for sixteen separate channels, in buffers 51, from which they can be output via D-flip-flops 52. The voiced/unvoiced and pitch data are sent via output 42 to electronic switch 3 and via output 43 to generator 2, respectively, and combined by means of D-flip-flop 53 with the digital signal
- 30 to be calculated. The whole algorithm can be represented by the following formulae:

$$Y_m(i) = Y_{m+1}(i) - K_m b_m(i-1) \quad (2)$$

and

$$b_{m+1}(i) = b_m(i-1) + K Y_m \quad (3)$$

in which $m = M-1, \dots, 0$ (4)

and $b_o(i) = Y_o(i)$ (5)

Multipliers 44 and 45 ensure the multiplications and adder-subtractor 46 carries out the adding and subtracting operations. The intermediate results of the operations are put away, every time, in the 51-buffer associated with the channel dealt with. Every time one sample has been calculated, its value is multiplied by the volume factor C_n . The various operations carried out on the data from memory 10.2 are controlled by a programmable store (PROM) 54, which, under the control of a counter 55, makes a step every time after the calculation of one PCM-sample for each of the 16 channels. The stepping of counter 55 is timed by clock 56. Store 54 supplies the data required for carrying out the various operations via a control bus 57 and the address data for store 10 via address bus 58. The last instruction in store 54 relates to writing the calculated final results in buffers 13 and signalling to the central processing unit 11 (Fig. 2) that the programme has finished. Then, under the control of central processing unit 11 (Fig. 2), a fresh set of data is transferred from store 10.1 to store 10.2, clock 56 being started in order to carry out again the programme contained in store 54. The data produced by the programme will only be stored when the central processing unit 11 (Fig. 2) has found that the buffers 13 are not full. After the data have been stored in buffers 13, the programme is started again under the control of the central processing unit 11. Thus the invention provides a relatively simple device for generating, from an input signal produced by the LPC-method referred to hereinabove, an analog signal for a large number of channels. The pre-processing unit 9 and the central processing unit 11 comprise microcomputers, for which the flow-charts are given in Figs. 5 and 6, respectively. The arrangement is not relevant for a good understanding of the invention, so that the flow-chart need not be described in detail. After a start command the microprogramme stored in the control unit 14 (Fig. 4) is carried out.

As soon as counter 55 (Fig. 4) has reached its final position, the last instruction contained in memory 34 passes the signal "End of programme" to the central processing unit 11, which checks whether buffers 13 are full. In the affirmative case the programme is stopped. If there is space left in buffers 13, the next PCM-word is calculated.

Claims.

1. Digital speech synthesizer comprising a digital noise generator, an adjustable digital pitch generator, an adjustable digital filter and a digital-to-analog converter, the pitch generator, the multiplier and the filter being adjusted according to parameters derived from
5 the original speech signal by the "linear predictive coding (LPC)" method, and means for computing the adjusting coefficients, characterized in that the means for computing the adjusting coefficients can serve a plurality of channels on a time-division basis.
- 10 2. Device according to claim 1, characterized in that the means for computing are adjustable.
3. Device according to claim 1 or 2, characterized by means for computing the number of bits per adjusting coefficient according to the number of channels for which speech samples have to be computed.
- 15 4. Device according to claim 1, characterized in that the means for computing the adjusting coefficients on a time-division basis comprise a computing unit (12) including a digital noise generator (1), a pitch generator (2), a switch (3) for switching from voiced to unvoiced speech and conversely, a digital filter and one store per channel,
20 which computing unit is controlled by a microprocessor circuit (14) controlling the number of computations to be carried out.
5. Device according to claim 2, characterized in that the means for adjusting the coefficients comprises a function decoder (28), which derives the kind of the coefficient from the incoming signal, and a
25 register for each kind, in which the value of the adjusting coefficient is stored.
6. Device according to claim 5, characterized in that the means for adjusting the number of interpolations between successive adjusting coefficients of one or more kinds comprise a store (31) indicating
30 the position of a counter (32) and a fraction table (33) permitting to indicate the accuracy with which a microprocessor (20) incorporated in a pre-processing unit (9) computes the number of interpolations.

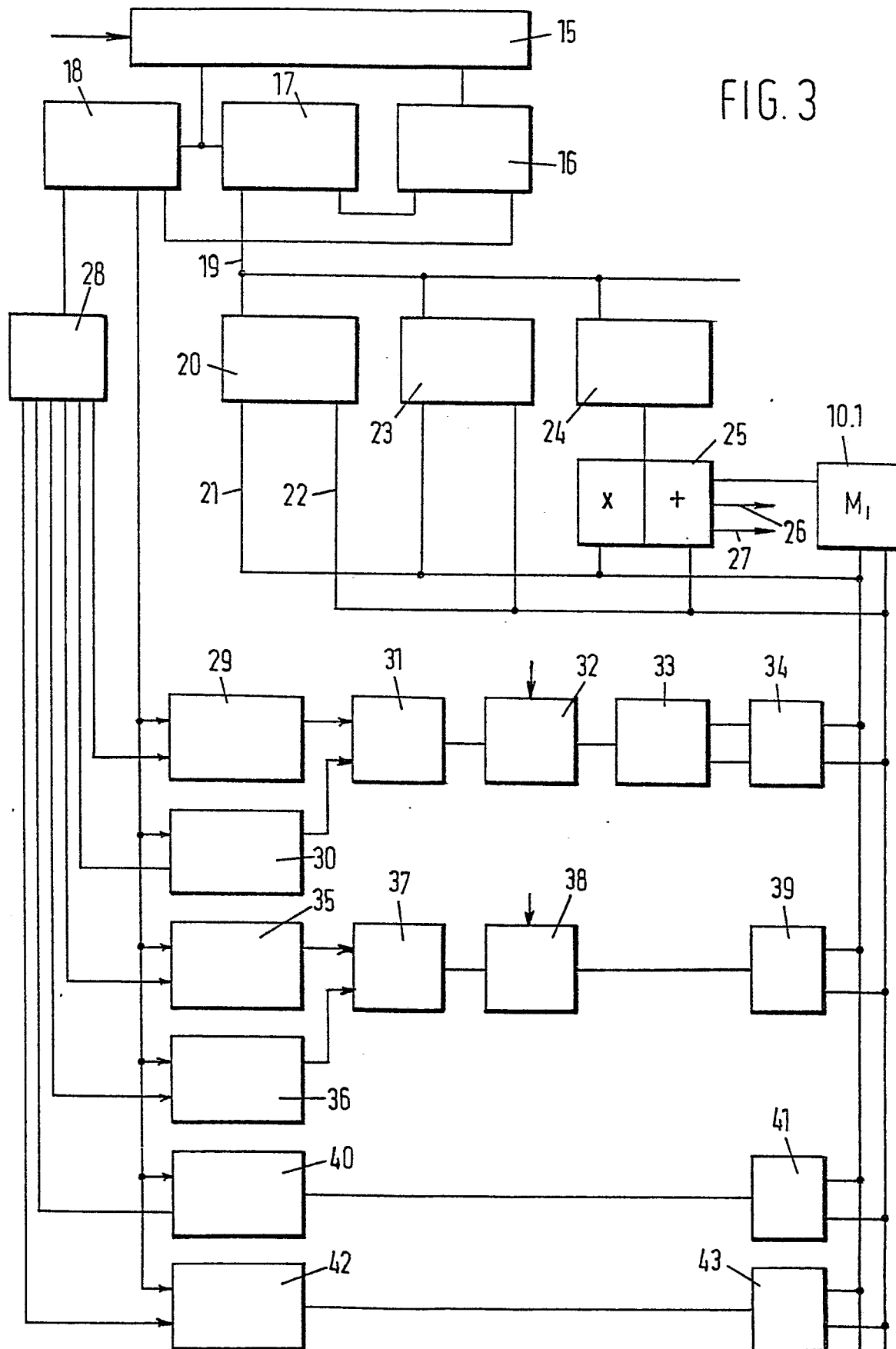
•



•

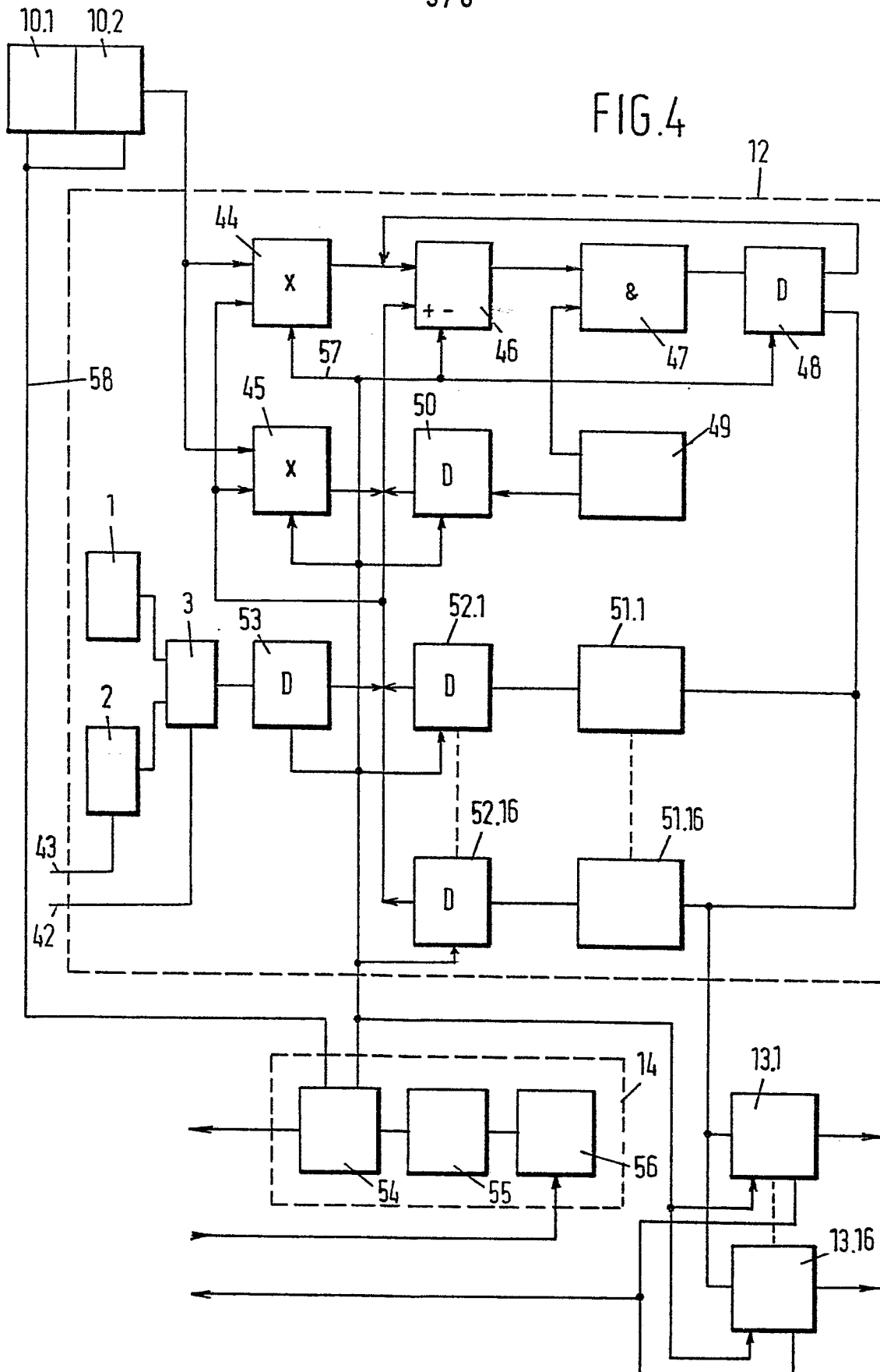


FIG. 3



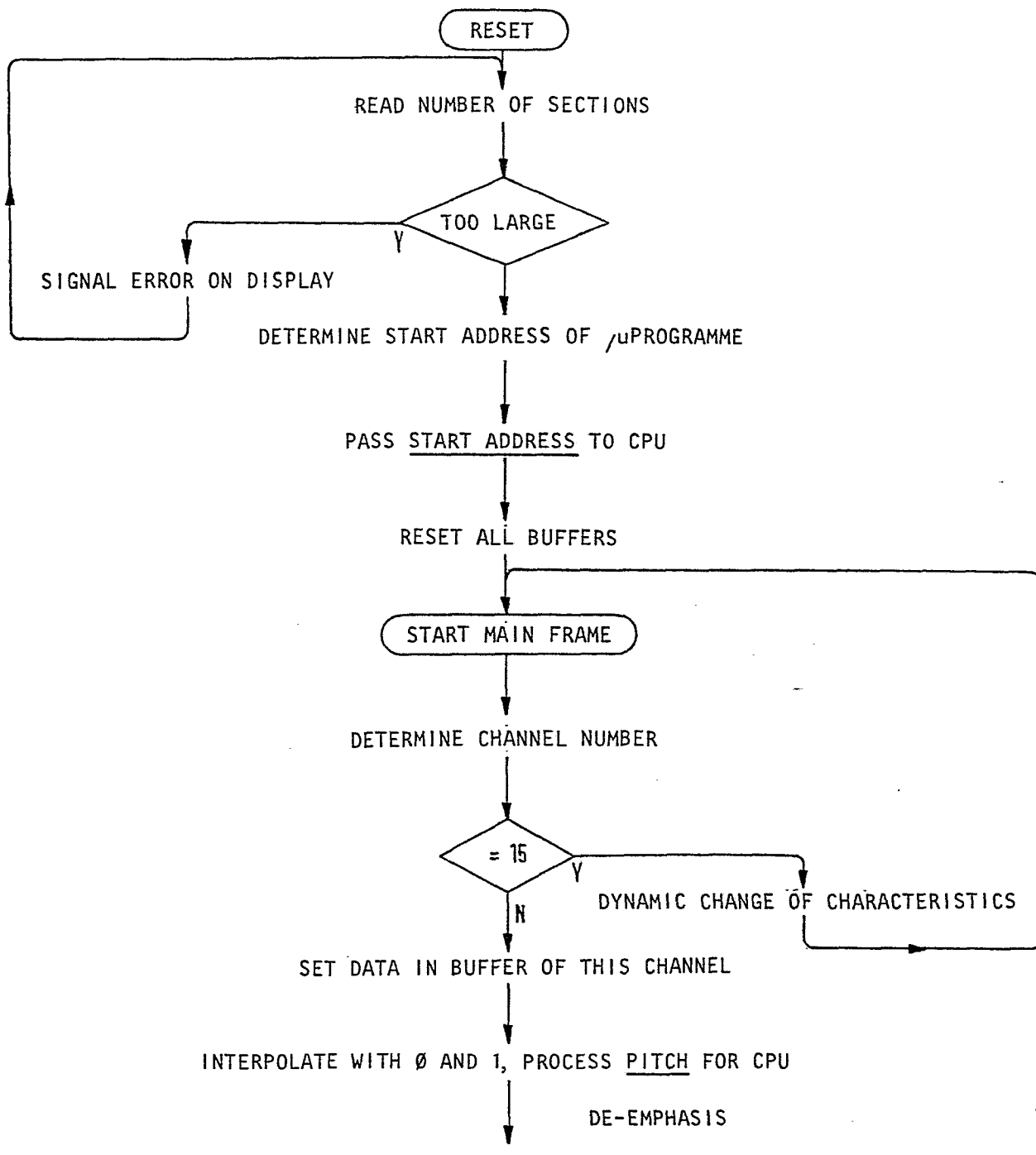
3/6

FIG. 4



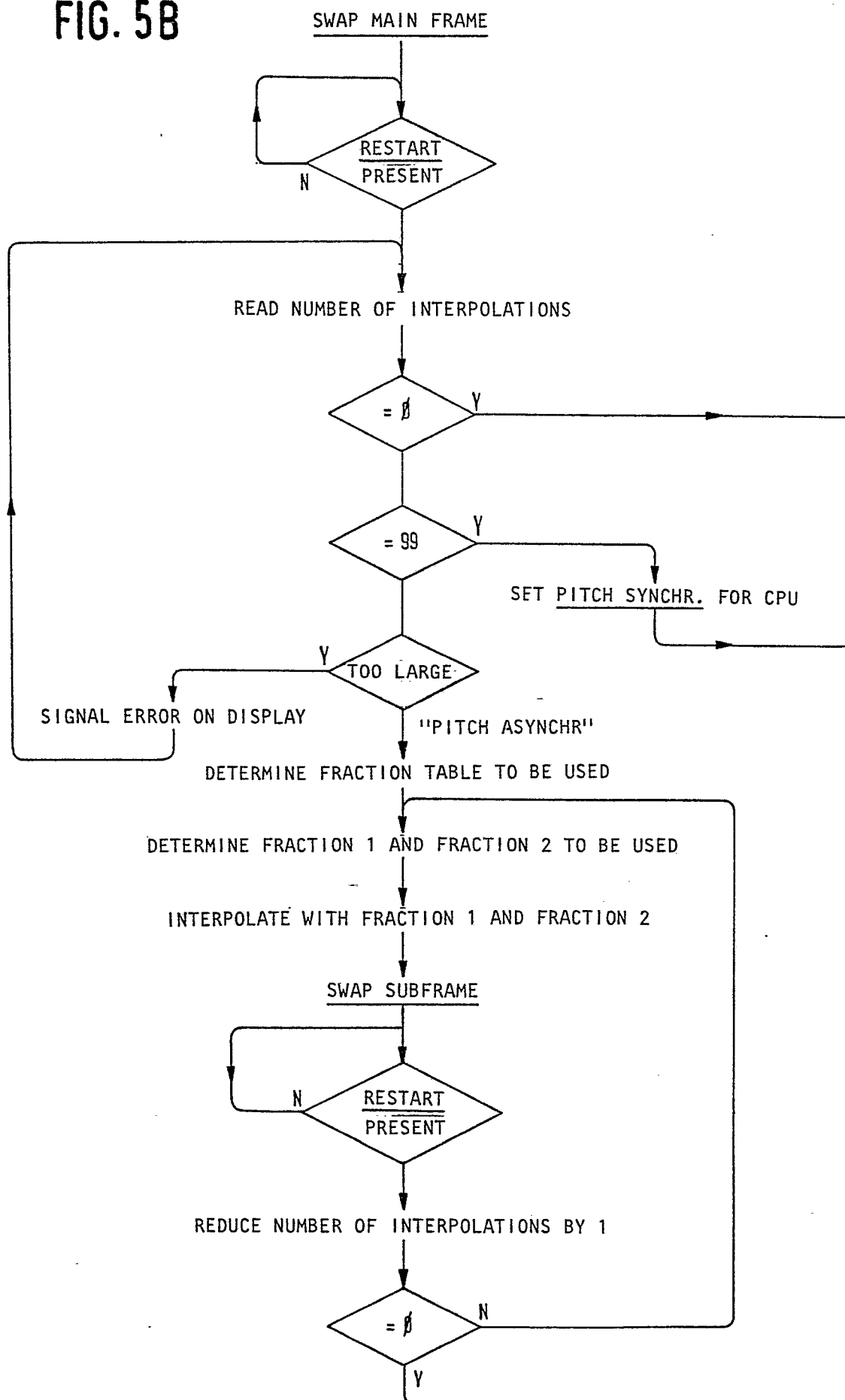
4/6

FIG 5A



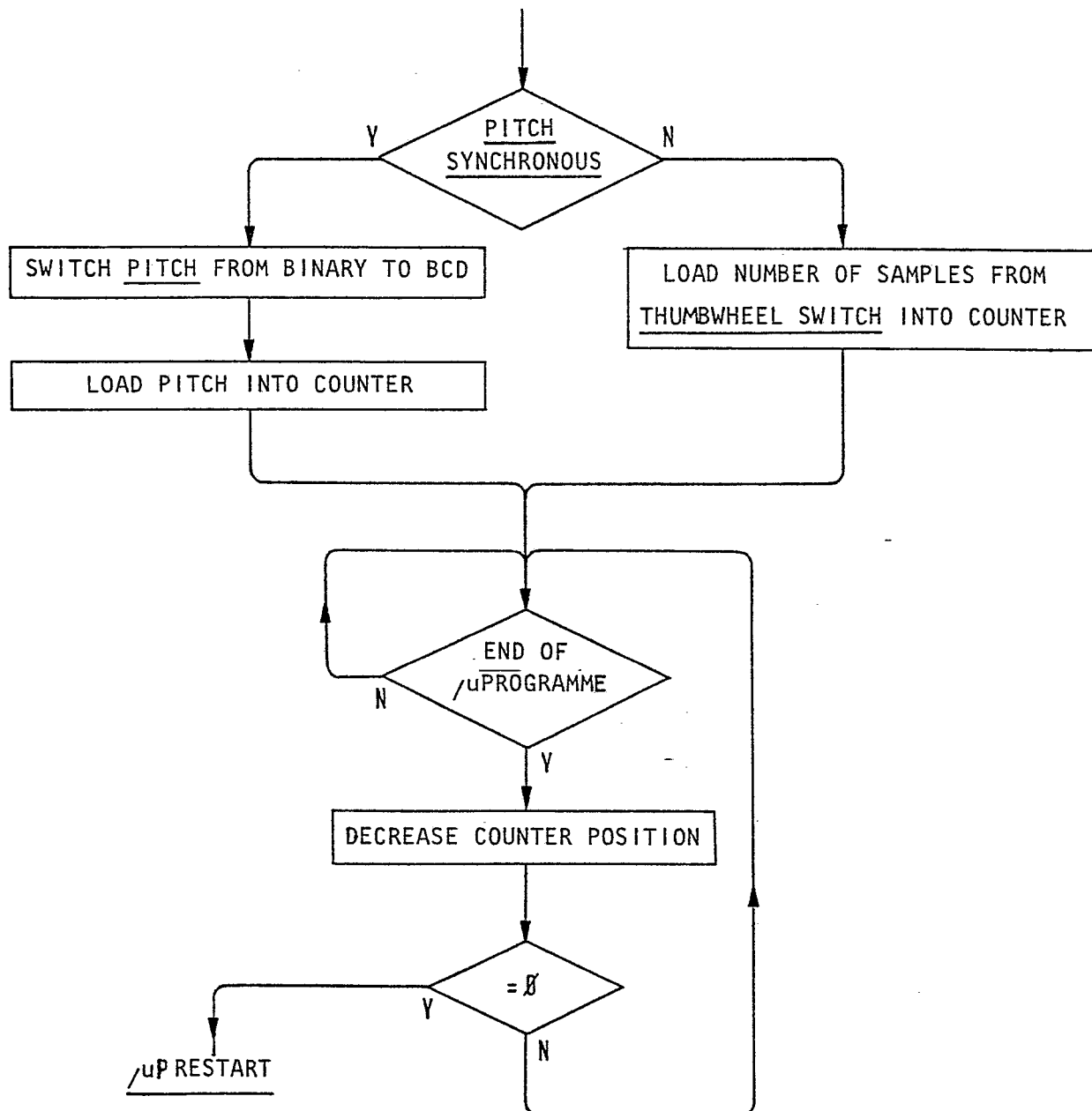
5/6

FIG. 5B



6/6

FIG. 6



COUNTER COUNTS REMAINING NUMBER OF SAMPLES IN THIS FRAME



European Patent
Office

EUROPEAN SEARCH REPORT

0051342

Application number
EP 81 20 1230

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int. Cl. 3)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
X	ICASSP 79, PROCEEDINGS OF A IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (April 2-4, 1979, WASHINGTON) NEW YORK (US) L. NEBBIA et al.: "Eight-channel digital speech synthesizer based on LPC techniques" pages 884-886 * Figures 2,5; page 886, left-hand column * ---	1,4	G 10 L 1/08
X	EP - A - 0 016 427 (CSELT CENTRO STUDI E LABORATORI TELECOMUNICAZIONI) * Figure 1 * ---	1,4	TECHNICAL FIELDS SEARCHED (Int.Cl. 3)
A	US - A - 3 715 512 (J.M. KELLY) * Figure 3 * ----	4,5	G 10 L
A	US - A - 3 975 587 (J.G. DUNN et al.) * Figure 2 * ---	5	
			CATEGORY OF CITED DOCUMENTS
			X: particularly relevant if taken alone Y: particularly relevant if combined with another document of the same category A: technological background O: non-written disclosure P: intermediate document T: theory or principle underlying the invention E: earlier patent document, but published on, or after the filing date D: document cited in the application L: document cited for other reasons
			&: member of the same patent family, corresponding document
☒ The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 17-02-1982	Examiner ARMSPACH