



AU9523948

(12) PATENT ABRIDGMENT (11) Document No. AU-B-23948/95
(19) AUSTRALIAN PATENT OFFICE (10) Acceptance No. 683750

- (54) Title
A MULTI-PULSE ANALYSIS SPEECH PROCESSING SYSTEM AND METHOD
- International Patent Classification(s)
(51)⁶ **G10L 005/00**
- (21) Application No. : **23948/95** (22) Application Date : **27.04.95**
- (87) PCT Publication Number : **WO95/30222**
- (30) Priority Data
- | | | |
|---------------|-----------------|------------------------------------|
| (31) Number | (32) Date | (33) Country |
| 236764 | 29.04.94 | US UNITED STATES OF AMERICA |
- (43) Publication Date : **29.11.95**
- (44) Publication Date of Accepted Application : **20.11.97**
- (71) Applicant(s)
JONATHAN EDWARD SHERMAN; AUDIOCODES LTD.
- (72) Inventor(s)
LEON BIALIK; FELIX FLOMEN
- (74) Attorney or Agent
WATERMARK PATENT & TRADEMARK ATTORNEYS, Locked Bag 5, HAWTHORN VIC 3122
- (56) Prior Art Documents
US 4932061
- (57) Claim

1. A speech processing system comprising:

a short-term analyzer connected to an input and an output line wherein, in response to an input speech signal on said input line, said short-term analyzer generates short-term characteristics of said input speech signal;

a target vector generator for generating a target vector from at least said input speech signal and, optionally, said short-term characteristics; and

a multi-pulse analyzer connected to an output line of said target vector generator, wherein said multi-pulse analyzer generates a plurality of sequences of equal amplitude, variable sign, variably spaced pulses, each of said sequences having a different amplitude value, each of said pulses within each sequence having equal

(11) AU-B-23948/95
(10) 683750

-2-

amplitudes but variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulses which, according to a maximum likelihood criterion, most closely represents said target vector.



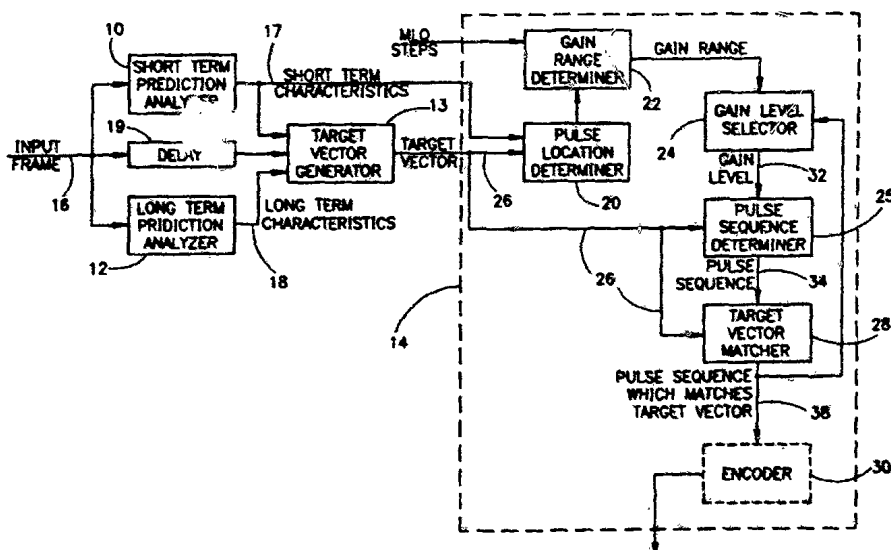
IN

)

(51) International Patent Classification 6 : G10L 5/00		A1	(11) International Publication Number: WO 95/30222
		(43) International Publication Date: 9 November 1995 (09.11.95)	
(21) International Application Number: PCT/US95/05014		(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).	
(22) International Filing Date: 27 April 1995 (27.04.95)			
(30) Priority Data: 08/236,764 29 April 1994 (29.04.94) US			
(71) Applicants: SHERMAN, Jonathan, Edward [US/IL]; 20 Zahal Street, 55905 Kfar Azar (IL). AUDIOCODES LTD. [IL/IL]; 3A Yoni Netanyahu Street, 60256 Or Yehuda (IL).			
(72) Inventors: BIALIK, Leon; 14/3 Shaftei Street, 75234 Rishon LeZion (IL). FLOMEN, Felix; 4/233 Hanarkis Street, 75424 Rishon LeZion (IL).			
(74) Agents: GUNNISON, Forrest, E. et al.; Skjerven, Morrill, MacPherson, Franklin & Friel, Suite 700, 25 Metro Drive, San Jose, CA 95110 (US).			
		<p>Published <i>With international search report.</i> <i>With amended claims.</i></p>	

683750

(54) Title: A MULTI-PULSE ANALYSIS SPEECH PROCESSING SYSTEM AND METHOD



(57) Abstract

A speech processing system and method are disclosed. In one embodiment of the present invention, the system includes at least a maximum likelihood quantization (MLQ) multi-pulse analysis unit (item 21) operating on a target vector (item 26). The MLQ multi-pulse analysis unit typically determines an initial gain level for the multi-pulse sequence and performs single gain multi-pulse analysis (MPA) a number of times, each with a different gain level. The pulse sequence which most closely represents the target vector is provided as an output signal (item 38). In another embodiment, the system includes at least a pulse train multi-pulse analysis unit wherein the target vector is modeled as a series of pulse trains. Each pulse train comprises a plurality of single gain pulses, wherein each pulse is at a position which is a pitch value distance apart from the previous pulse in the pulse train. Combinations of maximum likelihood analysis with pulse trains are also part of the present invention.

**A MULTI-PULSE ANALYSIS SPEECH PROCESSING
SYSTEM AND METHOD**

5

FIELD OF THE INVENTION

The present invention relates to speech processing systems generally and to multi-pulse analysis systems in particular.

BACKGROUND OF THE INVENTION

Speech signal processing is well known in the art and is often utilized to compress an incoming speech signal, either for storage or for transmission. The speech signal processing typically involves dividing the incoming speech signals into frames and then analyzing each frame to determine its components. The components are then stored or transmitted.

Typically, the frame analyzer determines the short-term and long-term characteristics of the speech signal. The frame analyzer can also determine one or both of the short- and long-term components, or "contributions", of the speech signal. For example, linear prediction coefficient analysis (LPC) provides the short-term characteristics and contribution and pitch analysis and prediction provides the long-term characteristics as well as the long-term contribution.

Typically, either, both or neither of the long- and short-term predictor contributions are subtracted from the input frame, leaving a target vector whose shape has to be characterized. Such a characterization can be produced with multi-pulse analysis (MPA) which is described in detail in section 6.4.2 of the book Digital Speech Processing, Synthesis and Recognition by Sadaoki Furui, Marcel Dekker, Inc., New York, NY 1989. The book is incorporated herein by reference.

In MPA, the target vector, which is formed of a multiplicity of samples, is modeled by a plurality of

pulses of equal amplitude (or spikes), of varying location and varying sign (positive and negative). To select each pulse, a pulse is placed at each sample location and the effect of the pulse, defined by passing
5 the pulse through a filter defined by the LPC coefficients, is determined. The pulse which provides most closely matches the target vector is selected and its effect is removed from the target vector, thereby generating a new target vector. The process continues
10 until a predetermined number of pulses have been found. For storage or transmission purposes, the result of the MPA analysis is a collection of pulse locations and a quantized value of the gain.

The gain is typically determined from the first
15 pulse which is determined. This gain is then utilized for the remaining pulses. Unfortunately, the gain value of the first pulse is not always indicative of the overall gain value of the target vector and therefore, the match to the target vector is not always very
20 accurate.

SUMMARY OF THE PRESENT INVENTION

It is therefore an object of the present invention to provide an improved speech processing system. In one embodiment of the present invention, the system includes
25 a short-term analyzer, a target vector generator and a maximum likelihood quantization (MLQ) multi-pulse analysis unit. The short-term analyzer determines the short-term characteristics of an input speech signal. The target vector generator generates a target vector
30 from at least the input signal. The MLQ multi-pulse analysis unit operates on the resultant target vector.

The MLQ multi-pulse analysis unit typically determines an initial gain level for the multi-pulse sequence and performs single gain MPA a number of times,
35 each with a different gain level. The gain levels are within a range above and below the initial gain level.



The resultant pulses can be positive or negative.

Like in other maximum likelihood applications, the quality of the result is measured (in this case, by minimizing the energy of an error vector defined as the difference between the target vector and an estimated vector produced by filtering the single gain pulse sequence through a perceptual weighting filter). The pulse sequence which minimizes the energy of the error vector and its corresponding gain level (or the index for the gain level) is then provided as the output signal of the MLQ multi-pulse analysis unit.

In an alternative embodiment, the system includes a long-term prediction analyzer and replaces the MLQ multi-pulse analysis unit with a pulse train multi-pulse analysis unit. In this embodiment, the pulse train multi-pulse analysis unit utilizes a pitch distance from the long-term analyzer to create a train of equal amplitude, same sign pulses, each the pitch distance apart from the previous pulse in the train. The multi-pulse analysis unit then outputs a signal representing the sequence of pulse trains, including positive and negative pulse trains, which best represents the target vector.

In a further alternative embodiment, the system includes an MLQ pulse train multi-pulse analysis unit which combines the operations of the two previous embodiments. In other words, a range of gains are provided, and for each, a sequence of pulse trains is found. The sequence which represents the closest match to the target vector is provided as the output signal.

In a final further embodiment, the output of the maximum likelihood and pulse train multi-pulse analysis units are compared and the sequence which represents the closest match to the target vector is provided as the output signal.

35

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood and appreciated more fully from the following detailed description taken in conjunction with the drawings in
5 which:

Fig. 1 is a block diagram illustration of a first embodiment of the speech processing system of the present invention;

Fig. 2 is a flow chart illustration of the
10 operations of an Multi-Pulse Maximum Likelihood Quantization (MP-MLQ) block of Fig. 1;

Figs. 3A and 3B are graphical illustrations, useful in understanding the operations of Fig. 2;

Figs. 4A and 4B are graphical illustration
15 describing pulse trains and multi-pulse analysis using pulse trains, respectively;

Fig. 5 is a block diagram illustration of a second embodiment of the speech processing system of the present invention utilizing pulse trains;

Fig. 6 is a flow chart illustration of the
20 operations of the pulse train multi-pulse analysis unit of Fig. 5; and

Fig. 7 is a block diagram illustration of a third embodiment comparing the output of the systems of Figs.
25 1 and 5.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Reference is now made to Figs. 1, 2, 3A and 3B which illustrate a first embodiment of the present invention. The speech processing system of the present invention
30 includes at least a short-term prediction analyzer 10, a long-term prediction analyzer 12, a target vector generator 13 and a maximum likelihood quantization multi-pulse analysis (MP-MLQ) unit 14.

Short-term prediction analyzer 10 receives, on input
35 line 16, an input frame of a speech signal formed of a multiplicity of digitized speech samples. Typically,



there are 240 speech samples per frame and the frame is often separated into a plurality of subframes. Typically, there are four subframes, each typically 60 samples long. The input frame can be a frame of an original speech signal or of a processed version thereof.

Short-term prediction analyzer 10 also receives, on input line 16, the input frame and produces, on output line 17, the short-term characteristics of the input frame. In one embodiment, analyzer 10 performs linear prediction analysis to produce linear prediction coefficients (LPCs) which characterize the input frame.

For the purposes of the present invention, analyzer 10 can perform any type of LPC analysis. For example, the LPC analysis can be performed as described in chapter 6.4.2 of the book Digital Speech Processing, Synthesis and Recognition, as follows: a Hamming window is applied to a window of 180 samples centered on a subframe. Tenth order LPC coefficients are generated, using the Durbin recursion method. The process is repeated for each subframe.

Long-term predictor analyzer 12 can be any type of long-term predictor and operates on the input frame received on line 16. Long-term analyzer 12 analyzes a plurality of subframes of the input frame to determine the pitch value of the speech within each subframe, where the pitch value is defined as the number of samples after which the speech signal approximately repeats itself. Pitch values typically range between 20 and 146, where 20 indicates a high-pitched voice and 146 indicates a low-pitched voice.

For example, for every two subframes, a pitch estimate can be determined by maximizing a normalized cross-correlation function of the subframes $s(n)$, as follows:

$$C_i = \frac{\sum s(k)s(k-i)}{\sum s(k-i)s(k-i)}, \quad 0 \leq k \leq 19, \quad 20 \leq i \leq 146 \quad (1)$$

For this example, long-term analyzer 12 selects the index i which maximizes cross-correlation C_i as the pitch value for the two subframes.

Once the long-term analyzer 12 determines the pitch value, the pitch value is utilized to determine the long-term prediction information for the subframe, provided on output line 18.

The target vector generator 13 receives the output signals of the long-term analyzer 12 and the short-term analyzer 10 as well as the input frame on input line 16, via a delay 19. In response to those signals, target vector generator 13 generates a target vector from at least a subframe of the input frame. The long- and short-term information can be utilized, if desired, or they can be ignored. The delay 19 ensures that the input frame which arrives at the target vector corresponds to the output of the analyzers 10 and 12.

An output line 26 of target vector generator 13, which is connected to the MP-MLQ unit 14, carries the target vector output signal. The MP-MLQ unit 14 is typically also connected to output line 17 carrying the short-term characteristics produced by analyzer 10..

It will be appreciated that, without any loss of generality, the target vector to the MP-MLQ unit 14 can be produced in any other desired manner.

In accordance with the first preferred embodiment of the present invention, the MP-MLQ unit 14 includes an initial pulse location determiner 20, a gain range determiner 22, a gain level selector 24, a pulse sequence determiner 25, a target vector matcher 28 and an optional encoder 30. The specific operations performed by elements 20 - 30 are illustrated in Fig. 2 and are described in detail hereinbelow. The following is a general description of the operation of unit 14.

35

The initial pulse location determiner 20 receives the output signals of the target vector generator 13 and the short-term analyzer 10 along output lines 17 and 26, respectively. It determines the sample location of a first pulse in accordance with multi-pulse analysis techniques.

The gain range determiner 22 receives the first pulse output of unit 20 and determines both an amplitude of the first pulse and a range of quantized gain levels around the absolute value of the determined amplitude. The step size, labeled MLQ_STEPS, for moving through the range of quantized gain levels, typically has a value of 3 separate gain levels. The step size, MLQ_STEPS, is not determined by MP_MLQ unit 14.

The gain level selector 24 receives the gain range produced by gain range determiner 22 and moves through the gain values within the gain range. Its output, on output line 32, is a current gain level for which a sequence of equal amplitude pulses is to be determined.

The pulse sequence determiner 25 receives the target vector, on line 26, and the current gain level, on line 32, and determines therefrom, using multi-pulse analysis techniques as described hereinbelow, a pulse sequence (with both positive and negative pulses) which matches the target vector. The pulse sequence is a series of positive and negative pulses having the current gain level.

The target vector matcher 28 receives the pulse sequence output, on output line 34, of determiner 25, and the target vector, on output line 26. Matcher 28 determines the quality of the match by utilizing a maximum likelihood type criterion.

Since there are a range of gain levels, the matcher 28 returns control to the gain level selector 24 to select the next gain level. This return of control is indicated by arrow 36.



For each gain value, matcher 28 determines the quality of the match, saving the match (gain index and



pulse sequence) only if it provides a smaller value for the criterion than previous matches.

Once gain selector 24 has moved through all of the gain values, the gain index and pulse sequence which is in storage in matcher 28 is the closest match to the target vector. Matcher 28 then outputs the stored pulse sequence and gain index along output line 38 to optional encoder 30.

It will be appreciated that, by determining a pulse sequence for each of a few gain levels, the MP-MLQ unit 14 can select the one which most closely matches the target vector.

Optional encoder 30 encodes the output pulse sequence and gain index for storage or transmission.

The specific operations of the MP-MLQ unit 14 are shown in Fig. 2. In initialization step 40, unit 14 generates the following signals:

a) an impulse response $h[n]$ for the input frame from the short-term characteristics a_i defined as:

$$\begin{aligned} h[n] &= \sum a_i h[n-i] + \delta[n], \quad 0 \leq n \leq N-1, \quad 1 \leq i \leq P \\ h[-n] &= 0, n = 1..P \end{aligned} \quad (2)$$

where P is the number of short-term characteristics and N is the number of speech samples in the subframe

b) the result $r_{hh}[l]$ of an impulse response autocorrelation, for each sample position l , as follows:

$$r_{hh}[l] = \sum h[n] * h[n-l], \quad 0 \leq l \leq N-1, \quad 1 \leq n \leq N-1 \quad (3)$$

and c) the result $r_{th}[l]$ of a cross-correlation between the impulse response $h[n]$ and the target vector $t[n]$, for each sample position l , as follows:

$$r_{th}[l] = \sum t[n] * h[n-l], \quad 0 \leq l \leq N-1, \quad 1 \leq n \leq N-1 \quad (4)$$

It will be appreciated that the impulse response is a function of the short-term characteristics a_i provided along line 17 from analyzer 10. The impulse response generated in initialization step 40 corresponds to the Durbin LPC analysis mentioned hereinabove.

The MP-MLQ unit 14 utilizes a local criterion $LC_{kj}[l]$ to determine a quantitative value for each sample position l , each pulse k and each gain level j . As will be seen hereinbelow, the level of the local criterion is dependent on the value of k (i.e. on the number of pulses already determined).

In step 42, the local criterion $LC_{0,j}[l]$ for the first pulse determination is initialized to the cross-correlation function $r_{th}[l]$, as follows:

$$LC_0[l] = LC_{0,j}[l] = r_{th}[l], \quad 0 \leq l \leq N-1, \quad j_{min} \leq j \leq j_{max} \quad (5)$$

A maximum local value for the local criterion is also set to some negative value. The position index l is also initialized to 0.

In steps 44 - 50 the position l of the first pulse $k = 1$ is determined. To do so, the absolute value of the local criterion $LC_{0,j}[l]$ is compared to the maximum local value (step 44). If $LC_{0,j}[l]$ is larger, the position l is stored, the maximum local value is set to the absolute value of the local criterion $LC_{0,j}[l]$ (step 46) and the position index l is increased by 1 (step 48). The operation is repeated until all the positions l have been reviewed. The sample position l_{opt} which is in storage after all of the positions have been reviewed is the selected sample position l_{opt} . Steps 40 - 50 are performed by the pulse location determiner 20.

Step 52 is performed by the gain range determiner 22. In step 52, maximum amplitude A_{max} of the position

l which produced the largest local criterion $LC_{0,j}[1]$ is generated as follows:

$$A_{\max} = A_{\max_j} = |LC_{0,j}[l_{\text{opt}}]|/r_{\text{hh}}[0], \quad j_{\min} \leq j \leq j_{\max} \quad (6)$$

5

where l_{opt} is the position of the first pulse. The maximum value A_{\max} is then approximated by one of a predetermined set of gain levels. For example, if the expected amplitude levels are in the range of 0.1 - 2.0 units, the gain levels might be every 0.1 units. Thus, 10 if A_{\max} is 0.756, it is quantized to 0.8.

Steps 54 - 58 are performed by the gain selector 24. In step 54, gain selector 24 determines the gain index j associated with the determined gain level as well as a 15 range of gain indices around gain index j . The range of gain levels can be any size depending on the predetermined value of MLQ_STEPS . In step 54, the gain selector 24 sets the gain index to the minimum one. For the previous example, 0.1 might have an index 1 and 20 MLQ_STEPS might be 3. Thus, the determined gain index is 8 and the range is between indices 5 - 11. Step 54 also sets a minimum global value to any very large value, such as 10^{13} .

In the present invention, for each gain index, the 25 first pulse is the location of the pulse determined by the pulse location determiner 20 (in steps 44 - 50). The remaining pulses can be anywhere else within the subframe and can have positive or negative gain values. In step 56, the gain selector 24 stores the first pulse position 30 and its amplitude. In step 58, the local criterion $LC_{k,j}[1]$, for the present pulse index k and gain index j is initialized, typically in accordance with equation 5.

Pulse sequence determiner 25 performs steps 60 - 74. 35 In step 60, determiner 25 sets the maximum local value to a large value, as before, and sets the position index 1

to 0.

In step 62, determiner 25 updates the local criterion with the previous pulse, as follows:

5 $LC_{k,j[l]} = LC_{k-1,j[l]} - A_{k-1,j} * r_{hh}(l-l_{opt_{k-1,j}}),$
 j = gain index (7)
 k = pulse index
 l = position index

10 In the loop of steps 64 - 70, pulse sequence determiner 25 determines the location of a pulse in a manner similar to that performed in steps 44 - 50 and therefore, will not be further described herein. In step 72, determiner 25 stores the selected pulse and in step 15 74, it updates the pulse value. Steps 62 - 74 are repeated for each pulse in the sequence, the result of which is the pulse sequence output of pulse sequence determiner 25. It is noted that step 62 updates the local criterion for each pulse which is found.

20 Figs. 3A and 3B illustrate two examples of different pulse sequence outputs of pulse sequence determiner 25. The sequence of Fig. 3A has a gain index of 7 and the sequence of Fig. 3B has a gain index of 8. Both sequences have the same first sample position 10 but the 25 rest of the pulses are at other positions. It is noted that the pulses can be positive or negative.

In step 76, target vector matcher 28 determines the value of a global criterion GC_j for each gain level j . The global criterion GC_j can be any appropriate 30 criterion and is typically a maximum likelihood type criterion. For example, the global criterion can measure the energy in an error vector defined as the difference between the target vector and an estimated vector produced by filtering the single gain pulse sequence 35 through a perceptual weighting filter, in this case defined by the short-term characteristics. For such a criterion, target vector matcher 28 includes a perceptual



weighting filter.

It will be appreciated that the pulse sequence, per se, does not match the target vector; the pulse sequence represents a function which matches the target vector.

5 As given in equations 8a - 8e hereinbelow, the global criterion GC_j is comprised of two elements, p_j and d_j , both of which are functions of a signal $x_j[n]$ which is the pulse series for the gain level j filtered by the short-term impulse response $h[n]$. p_j is the
10 cross-correlation between the target vector $t[n]$ and $x[n]$ and d_j is the energy of $x_j[n]$.

$$GC_j = -2p_j + d_j \quad (8a)$$

$$p_j = \sum t[n] * x_j[n], \quad 0 \leq n \leq N-1 \quad (8b)$$

$$15 \quad d_j = \sum x_j[n] * x_j[n], \quad 0 \leq n \leq N-1 \quad (8c)$$

$$x_j[n] = \sum v_j[i] * h[i-n], \quad 0 \leq i \leq n, \quad 0 \leq n \leq N-1 \quad (8d)$$

$$v_j[n] = \begin{cases} A_{k,j} & \text{for } n = l_{opt_k,j}, \quad 0 \leq k \leq K-1, \quad 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases} \quad (8e)$$

20

In step 78, the global criterion GC_j for the present gain index j is compared to the present minimum global value. If it is less than the present minimum global value, as checked in step 78, the target vector
25 matcher 28 stores (step 80) the gain index and its associated pulse sequence.

In step 82, the gain level selector 24 updates the gain index and, in step 84 it checks whether or not pulse sequences have been determined for all of the gain
30 levels. If so, the pulse sequence and gain index which are in storage are the ones which best match the target vector in accordance with the global criterion GC_j .

In step 86, optional encoder 30 encodes the pulse sequence and gain index as output signals, for
35 transmission or storage, in accordance with any encoding method. If desired, the target vector can be

reconstructed using $x_{jopt}[n]$, where $jopt$ is the gain index resulting from step 84.

It will be appreciated that the MP-MLQ unit 14 of the present invention provides, as output signals, at least the selected pulse sequence and the gain level.

Reference is now made to Figs. 4A, 4B, 5 and 6 which illustrate an alternative embodiment of the present invention which utilizes pulse trains. A pulse train 83 is illustrated in Fig. 4A. It comprises a series of pulses 81 separated by a distance Q which is the pitch.

In the system shown in Fig. 5, a sequence of pulse trains are found which most closely match a target vector. Fig. 4B illustrates an example sequence of three pulse trains 83a, 83b and 83c which might be found. Each pulse train 83 begins at a different sample position. Pulse train 83a is the first and comprises four pulses. Pulse train 83b begins at a later position and comprises three pulses and pulse train 83c, starting at a much later position, comprises only two pulses.

The system of Fig. 5 is similar to that of Fig. 1; the only differences being that a) the pulse location determiner 20 and pulse sequence determiner 25 of Fig. 1 are replaced by pulse train location determiner 88 and pulse train sequence determiner 89; b) the target vector matcher, labeled 90, operates on pulse train sequences rather than pulse sequences; and c) the determiners 88 and 89 receive the pitch value Q along output line 18. In addition, the output lines 34 and 38 are replaced by output lines 92 and 94 which carry signals representing sequences of pulse trains rather than sequences of pulses.

Pulse train determiner 88 operates similar to pulse determiner 20 except that determiner 88 utilizes a pulse train impulse response $h_T[n]$ rather than the pulse impulse response $h[n]$. $h_T[n]$ is defined as:

$$h_T[n] = \sum h[n-kQ], 0 \leq n \leq N-1, 0 \leq k \leq (N-1)/Q \quad (9)$$

where Q is the pitch value. As can be seen, the pulse trains at later positions typically have fewer pulses.

5 The pulse train impulse response autocorrelation of equation 3 becomes:

$$r_hh[l] = \sum h_T[n] * h_T[n-l], 0 \leq l \leq N-1, 1 \leq n \leq N-1 \quad (10)$$

10 and the cross-correlation $r_th[l]$ between the impulse response $h_T[n]$ and the target vector $t[n]$, for each sample position l , becomes:

$$r_th[l] = \sum t[n] * h_T[n-l], 0 \leq l \leq N-1, 1 \leq n \leq N-1 \quad (11)$$

15

Pulse train sequence determiner 89 operates similarly to pulse sequence determiner 25 but determiner 89 generates pulse train sequences.

20 Target vector matcher 90 operates similarly to target vector matcher 28; however, matcher 90 utilizes the pulse train impulse response function $h_T[n]$ rather than $h[n]$. Thus, equation 8d becomes:

$$x_j[n] = \sum v_j[i] * h_T[i-n], 0 \leq i \leq n, 0 \leq n \leq N-1 \quad (12)$$

25

30 The specific operations of the pulse train multi-pulse analysis unit 86 are shown in Fig. 6. The steps are equivalent to those shown in Fig. 2; however, the equations operate on pulse trains rather than individual pulses. Thus, in equation 9, a pulse train impulse response $h_T[n]$ is defined which has pulses every Q steps. The pulse trains at later positions typically have fewer pulses.

35 The remaining equations are similar except that they operate on the impulse response $h_T[n]$.

If it is desired, the gain range determined by gain

range determiner 22 can have only one gain index. In this embodiment, pulse train multi-pulse analysis unit 86 determines the pulse train sequence which has the gain level of the first pulse train sequence. In this
5 embodiment, the target vector matcher 90 does not operate, nor is there any repeating of the operations of gain level selector 24 and pulse train sequence determiner 89.

It will further be appreciated that the output of
10 target vector matchers 28 and 90 can be compared. This is illustrated in Fig. 7 to which reference is now made. The output signals of matchers 28 and 90, representing the sequences and global criteria, are provided, along output lines 38 and 94 to a comparator 100. Comparator
15 100 compares global criteria GC_jopt from matchers 28 and 90 and selects the lowest one. An output signal representing the resulting sequence, pulse or pulse train, is provided along output line 102.

It will be appreciated that the systems of Figs. 1,
20 5 and 7 can be implemented on a digital signal processing chip or in software. In one embodiment, the software was written in the programming language C++, in another in Assembly language.

It will be appreciated by persons skilled in the art
25 that the present invention is not limited to what has been particularly shown and described hereinabove. Rather the scope of the present invention is defined only by the claims which follow:

CLAIMS

1. A speech processing system comprising:

5 a short-term analyzer connected to an input and an output line wherein, in response to an input speech signal on said input line, said short-term analyzer generates short-term characteristics of said input speech signal;

10 a target vector generator for generating a target vector from at least said input speech signal and, optionally, said short-term characteristics; and

15 a multi-pulse analyzer connected to an output line of said target vector generator, wherein said multi-pulse analyzer generates a plurality of sequences of equal amplitude, variable sign, variably spaced pulses, each of said sequences having a different amplitude value, each of said
20 pulses within each sequence having equal amplitudes but variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulses which, according to a
25 maximum likelihood criterion, most closely represents said target vector.

2. A speech processing system incorporating a short term analyzer for generating short term
30 characteristics utilizing linear prediction coefficient analysis on an input speech signal, comprising:

a target vector generator for generating a target vector from at least said input speech signal and, optionally, the short term
35 characteristics;

an initial pulse location determiner for determining the location of an initial pulse in accordance with multi-pulse analysis techniques,



based on said target vector and the short term characteristics;

5 an amplitude range determiner for determining both an amplitude of said initial pulse and a range of quantized amplitude levels grouped around the absolute value of said amplitude;

10 an amplitude level selector for stepping through said range of quantized amplitude levels in accordance with a predetermined step size, said amplitude level selector outputting a selected quantized amplitude at each step;

15 a pulse sequence determiner for generating, based on said selected quantized amplitude, a sequence of equal amplitude, variable sign, variably spaced pulses which corresponds to said target vector; and

20 a target vector matcher for determining an error vector corresponding to the quality of the match between said sequence of equal amplitude, variable sign, variably spaced pulses and said target vector, for determining said error vector for each of said selected amplitudes, for outputting said sequence of equal amplitude, variable sign, variably spaced pulses that
25 corresponds to a minimum error vector.

3. The system according to claim 2 wherein the initial pulse of each of said sequences of equal amplitude, variable sign, variably spaced pulses is
30 located at the same sample position.

4. The system according to claim 2 wherein said target vector matcher includes a global criterion determiner, said global criterion determiner includes a
35 perceptual weighting filter for filtering said sequence of equal amplitude, variable sign, variably spaced



pulses and a determiner for determining the amount of energy in said error vector, for each of said selected quantized amplitudes, said error vector defined as the difference between said target vector and the output of said filter, said perceptual weighting filter having characteristics corresponding to the short term characteristics.

5. A speech processing system incorporating a short term analyzer for generating short term characteristics utilizing linear prediction coefficient analysis from an input speech signal and incorporating a long term analyzer for determining long term characteristics and a pitch value of speech from the input speech signal, the system comprising:

a target vector generator for generating a target vector from at least said input speech signal and optionally, the short term and long term characteristics;

an initial pulse train location determiner for determining the location of an initial pulse train in accordance with multi-pulse analysis techniques, based on said target vector, the short term characteristics and the pitch value;

a pulse train sequence determiner for generating a plurality of variable sign trains of equal amplitude, uniformly spaced pulses which corresponds to said target vector, said pulses within said trains having a pulse spacing corresponding to the pitch value, said pulses within each train having the same sign, and said pulses of all of said trains having the same amplitude level;

a target vector matcher for comparing the trains with the target vector and outputting a signal which represents the train which best matches the target vector.

6. A speech processing system comprising:

a long term analyzer connected to an input and an output line wherein, in response to an



input speech signal on said input line, said long-term analyzer generates long term characteristics including at least a pitch value of said input speech signal;

5 a short-term analyzer connected to said input line and to an output line wherein, in response to said input speech signal on said input line, said short-term analyzer generates short-term characteristics of said input speech signal;

10 a target vector generator for generating a target vector from at least said input speech signal and, optionally the short term and long term characteristics; and

15 a pulse train multi-pulse analyzer, connected to an output line of said target vector generator for generating a plurality of sequences of variable sign trains of equal amplitude, uniformly spaced pulses, said pulses within each train having the same sign, and each of said sequences
20 of trains of pulses having a different amplitude value said pulse train multi-pulse analyzer outputting a signal corresponding to the plurality of trains of equal amplitude, uniformly spaced pulses which, in accordance with a maximum
25 likelihood criterion, most closely represents said target vector.

7. The system according to claim 6 wherein each of said pulses within each said train of pulses is
30 separated from each other by said pitch value.

8. The system according to claim 6 wherein the initial pulse of the initial train of each said sequence of trains of pulses is located at the same
35 sample position.



9. A speech processing system incorporating a short term analyzer for generating short term characteristics utilizing linear prediction coefficient analysis from an input speech signal and incorporating
5 a long term analyzer for determining long term characteristics including a pitch value of speech from the input speech signal, the system comprising:

a target vector generator for generating a target vector from at least said input speech
10 signal and, optionally, the short term and long term characteristics;

an initial pulse train location determiner for determining the location of an initial pulse train in accordance with multi-pulse analysis
15 techniques, based on said target vector, the short term characteristics and the pitch value;

an amplitude range determiner for determining both an amplitude of said initial pulse train and a range of quantized amplitude levels grouped
20 around the absolute value of said amplitude;

an amplitude level selector for stepping through said range of quantized amplitude levels in accordance with a predetermined step size, said amplitude level selector outputting a selected
25 quantized amplitude at each step;

a pulse train sequence determiner for generating, for each of said selected quantized amplitudes, a plurality of variable sign trains of equal amplitude, uniformly spaced pulses which
30 corresponds to said target vector, said pulses within said trains having a pulse spacing corresponding to the pitch value, said pulses within each train having the same sign, said pulses within each train of pulses having an equal
35 amplitude, said equal amplitude corresponding to said selected quantized amplitude; and



5 a target vector matcher for determining an
error vector corresponding to the quality of the
match between said plurality of sequences of
variable sign trains of equal amplitude, uniformly
spaced pulses and said target vector, for
determining said error vector for each said
selected quantized amplitude, said target vector
matcher for outputting said sequence of trains of
equal amplitude, equal sign, uniformly spaced
10 pulses that corresponds to a minimum error vector.

10. The system according to claim 9 wherein said
target vector matcher includes a global criterion
determiner, said global criterion determiner includes a
15 perceptual weighting filter for filtering said
plurality of variable sign trains of equal amplitude,
uniformly spaced pulses and a determiner for
determining the amount of energy in said error vector,
for each said selected quantized amplitude, said error
20 vector defined as the difference between said target
vector and the output of said filter, said perceptual
weighting filter having characteristics corresponding
to the short term characteristics.

25 11. The system according to claim 10 further
comprising:

a multi-pulse analyzer connected to said
output line of said target vector generator,
wherein said multi-pulse analyzer generates a
30 plurality of sequences of equal amplitude,
variable sign, variably spaced pulses, each of
said sequences having a different amplitude value,
each of said pulses within each sequence having
equal amplitudes but variable signs, said
35 multi-pulse analyzer for outputting a signal
corresponding to the sequence of equal amplitude,



variable signs variably spaced pulses which,
according to a maximum likelihood criterion, most
closely represents said target vector; and

5 a comparator receiving output from both said
pulse train multi-pulse analyzer and said
multi-pulse analyzer for selecting the output
which best matches said target vector.

12. A method of speech process' comprising the
10 steps of:

determining short-term characteristics of an
input speech signal;

generating a target vector from at least said
input speech signal and, optionally from said
15 short-term characteristics;

determining the location of an initial pulse
in accordance with multi-pulse analysis
techniques, based on said target vector and said
short-term characteristics;

20 determining both an amplitude of said initial
pulse and a range of quantized amplitude levels
grouped around the absolute value of said
amplitude;

stepping through said range of quantized
25 amplitude levels in accordance with predetermined
step size and outputting a selected quantized
amplitude at each step;

generating, based on said selected quantized
amplitude, a sequence of equal amplitude, variable
30 sign, variably spaced pulses which corresponds to
said target vector;

comparing each said sequence of equal
amplitude, variable sign, variably spaced pulses
to said target vector; and

35 selecting said sequence of equal amplitude,
variable sign, variably spaced pulses which, in



accordance with a maximum likelihood criterion,
most closely represents said target vector.

13. The method according to claim 12 wherein the
5 initial pulse of each said sequence of equal amplitude,
variable sign, variably spaced pulses is located at the
same sample position.

14. The method according to claim 12 wherein said
10 step of comparing includes the steps of:

filtering the sequence of equal amplitude,
variable sign, variably spaced pulses through a
perceptual weighting filter whose characteristics
are said short-term characteristics; and

15 determining, for each quantized amplitude
level, the amount of energy in an error vector
defined as the difference between said target
vector and the output of said filter.

20 15. A method of speech processing comprising the
steps of:

determining short term characteristics of an
input speech signal;

25 determining long term characteristics of said
input speech signal including at least a pitch
value of said input speech signal;

generating a target vector from at least said
input speech signal, and, optionally from said
short term and long term characteristics;

30 determining the location of an initial pulse
train in accordance with multi-pulse analysis
techniques based on said target vector, said short
term characteristics and said pitch value; and

35 generating a plurality of variable sign
trains of equal amplitude, uniformly spaced pulses
which correspond to said target vector, said



pulses within said trains having a pulse spacing corresponding to said pitch value, said pulses within said trains having the same amplitude level, said pulses within each train having the same sign;

generating a signal representing the train which most closely matches the target vector.

16. A method of speech processing comprising the steps of:

determining short-term characteristics of said input speech signal;

determining long-term characteristics of said input speech signal including at least a pitch value of said input speech signal;

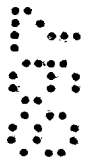
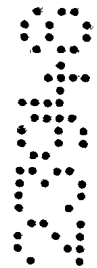
generating a target vector from at least said input speech signal, and, optionally, from said short-term and long term characteristics;

determining the location of an initial pulse train in accordance with multi-pulse analysis techniques, based on said target vector, the short-term characteristics and the pitch value;

determining both an amplitude of said initial pulse train and a range of quantized levels grouped around the absolute value of said amplitude;

stepping through said range of quantized amplitude levels in accordance with a predetermined step size and outputting a selected quantized amplitude at each step;

generating, for each selected quantized amplitude, a plurality of variable sign trains of equal amplitude, uniformly spaced pulses which correspond to said target vector, said pulses within said trains of pulses having a pulse spacing corresponding to said pitch value, said pulses within each said train of pulses having the same amplitude, said same amplitude corresponding



to the selected quantized amplitude, the pulses
within each train having the same sign;

comparing said plurality of variable sign
trains of equal amplitude, uniformly spaced pulses
to said target vector; and

selecting said plurality of variable sign
trains of equal amplitude, uniformly spaced pulses
which, in accordance with a maximum likelihood
criterion, most closely represents said target
vector.

17. The method according to claim 16 wherein the
initial pulse of each said sequence of trains of pulses
is located at the same sample position.



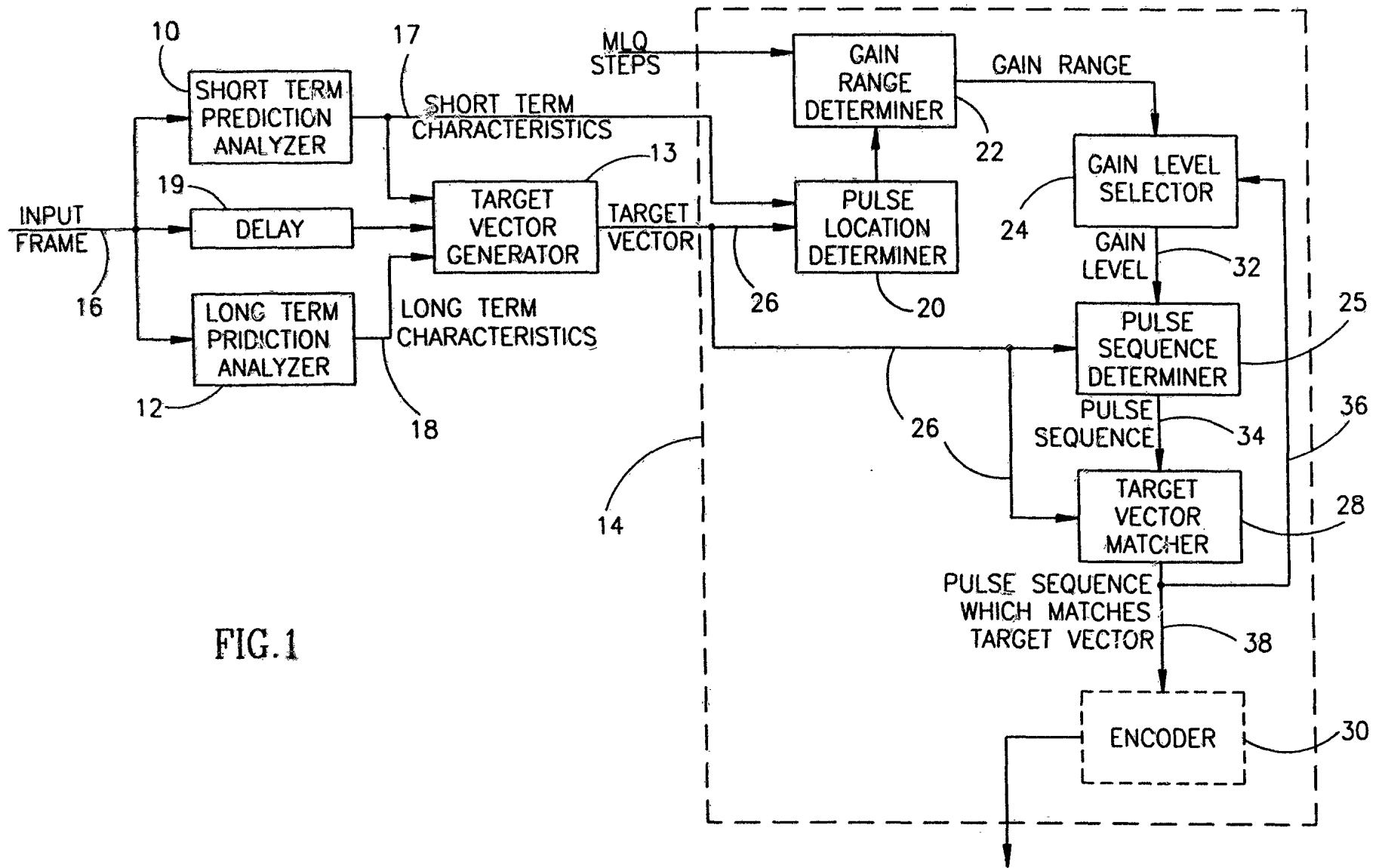


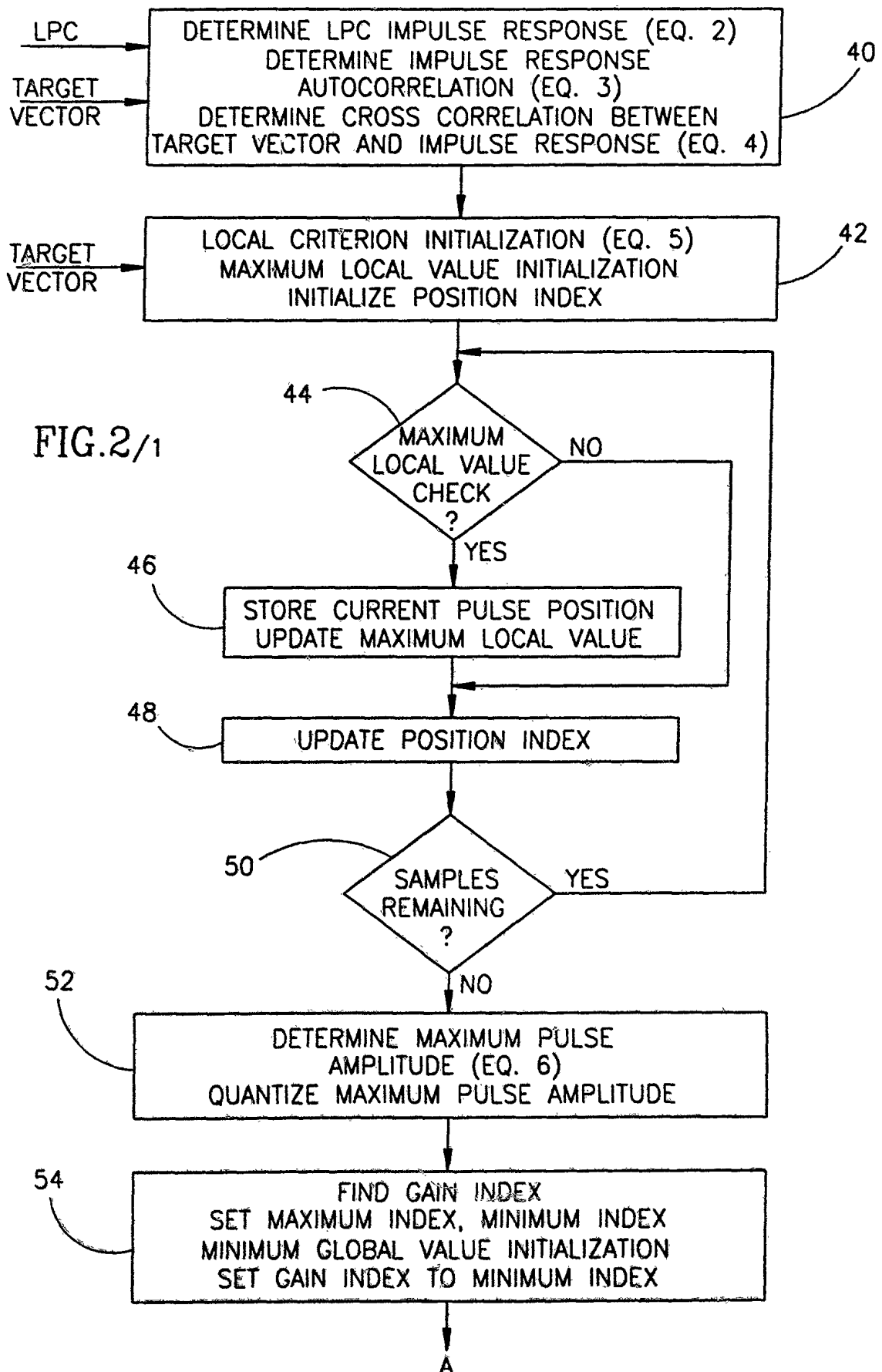
FIG.1

1/10

2,394,815

PCT/US 95/0501
11 Rec'd PCT/PTO 14 NOV 1995

2/10



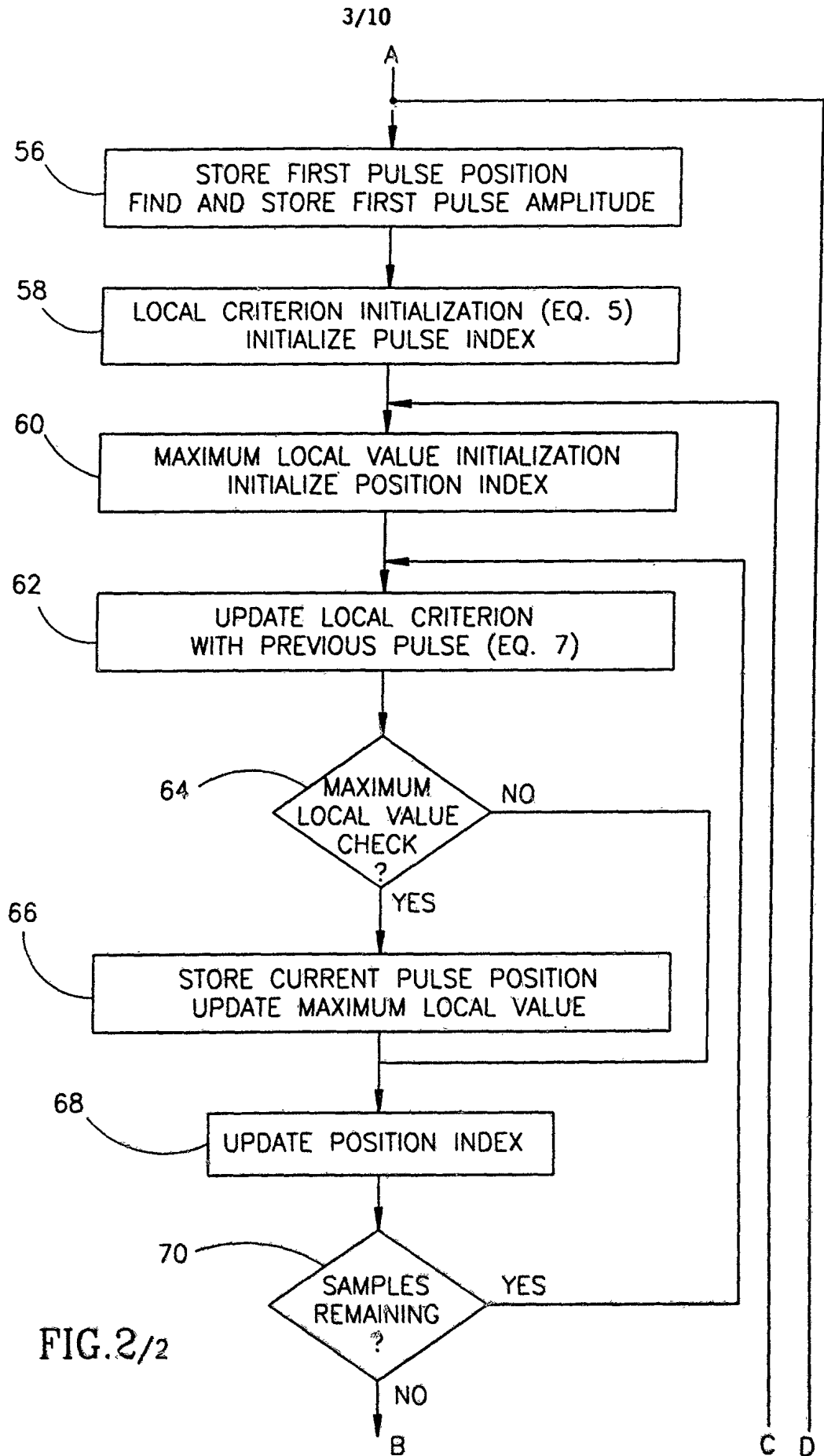
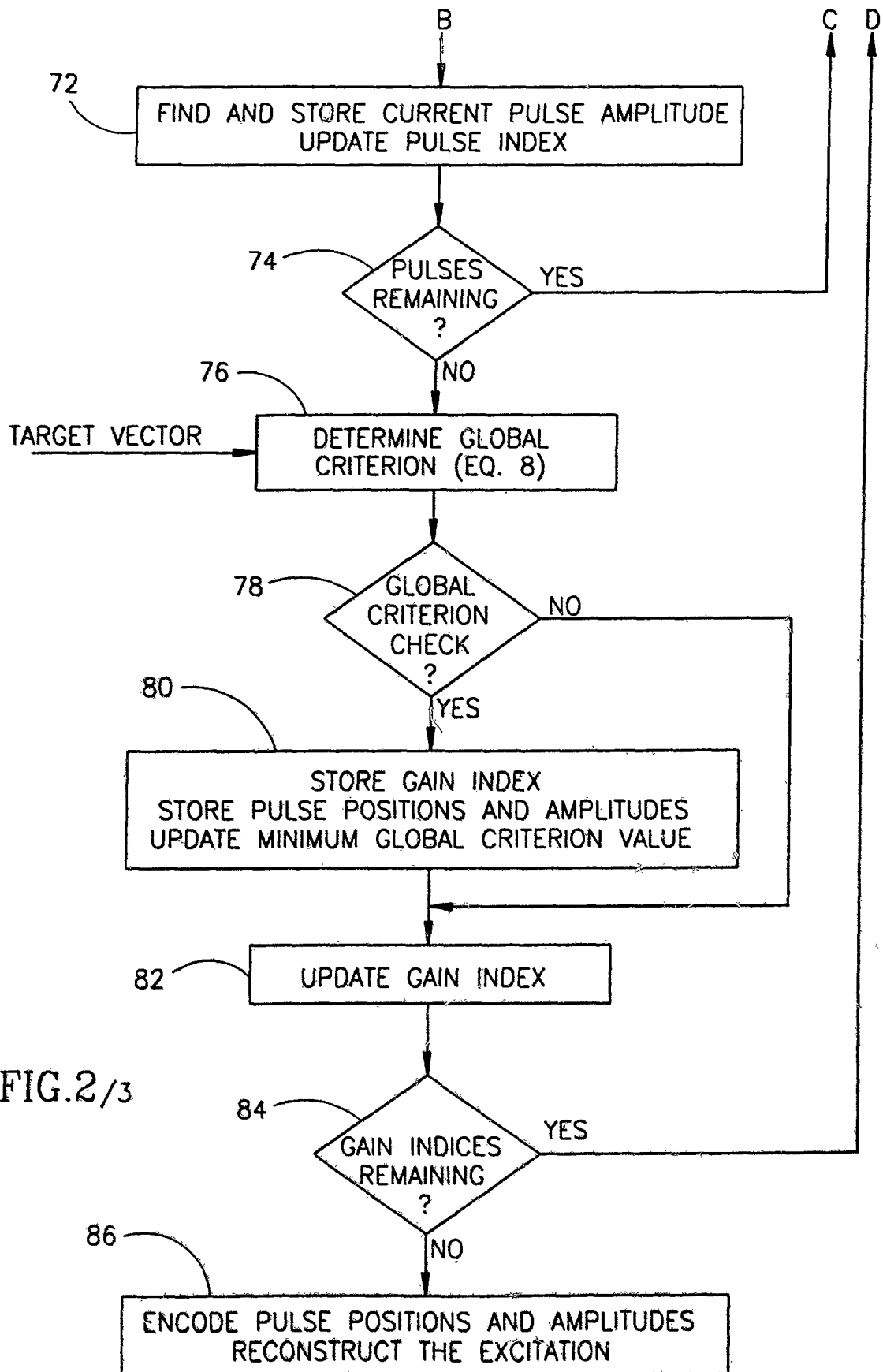
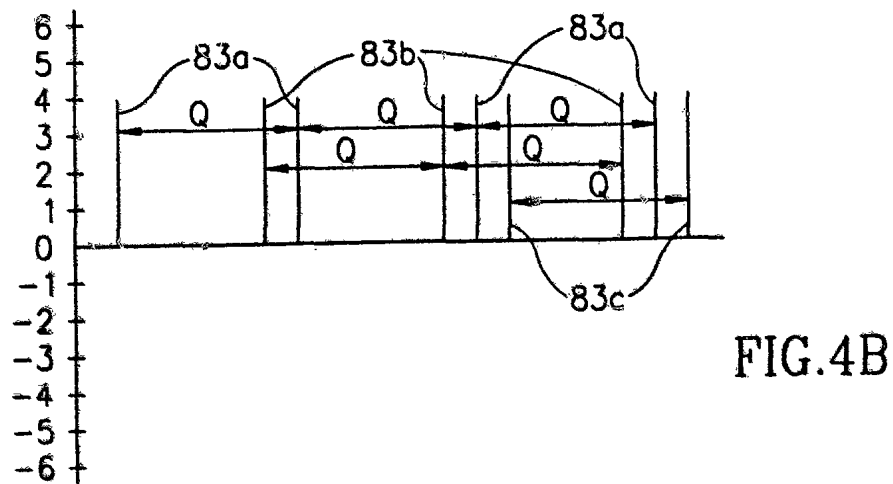
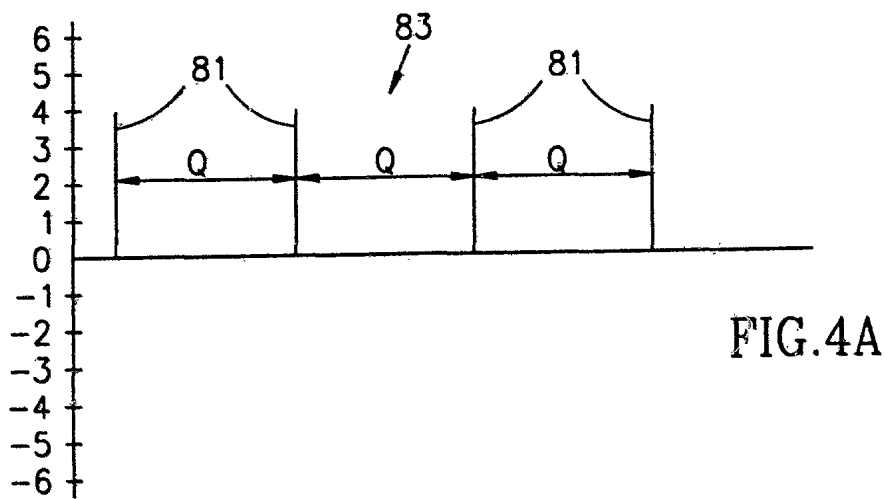
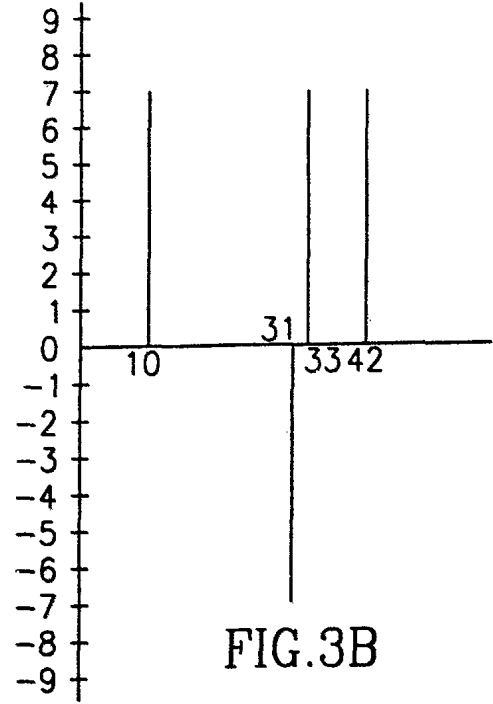
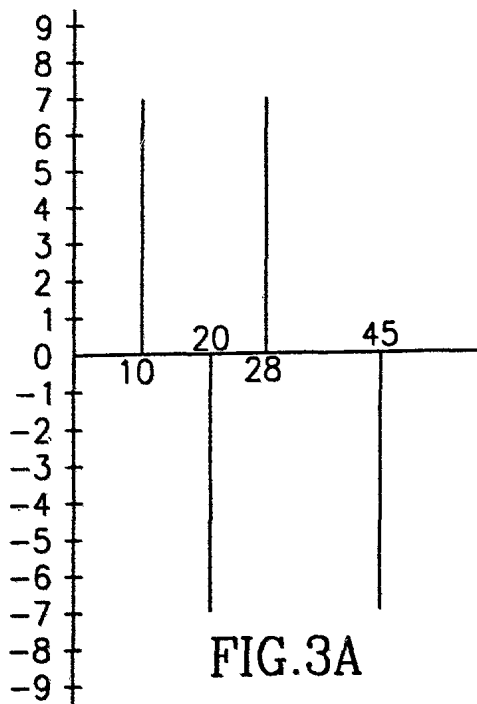


FIG.2/2

4/10



5/10



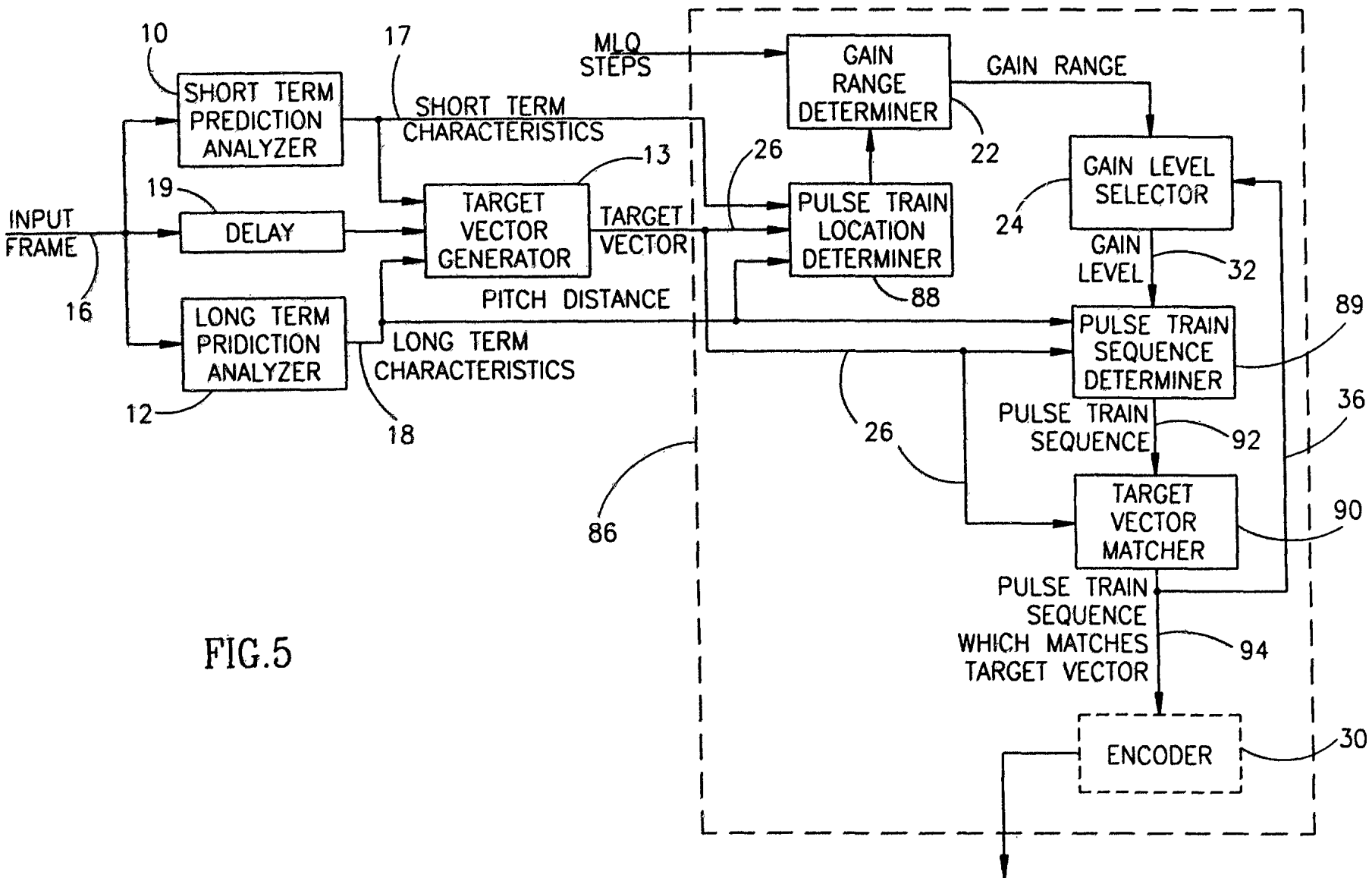
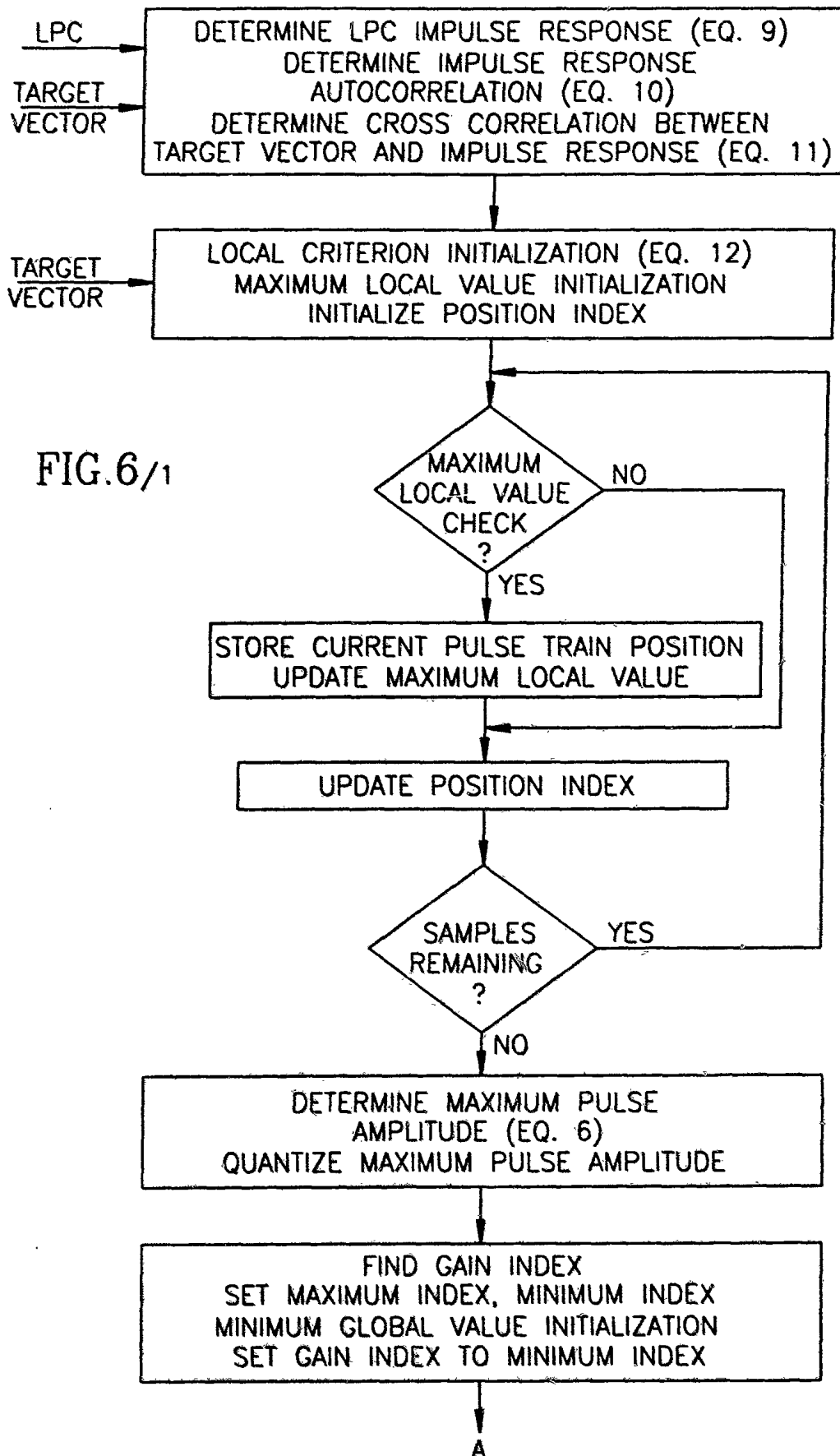


FIG.5

7/10



8/10

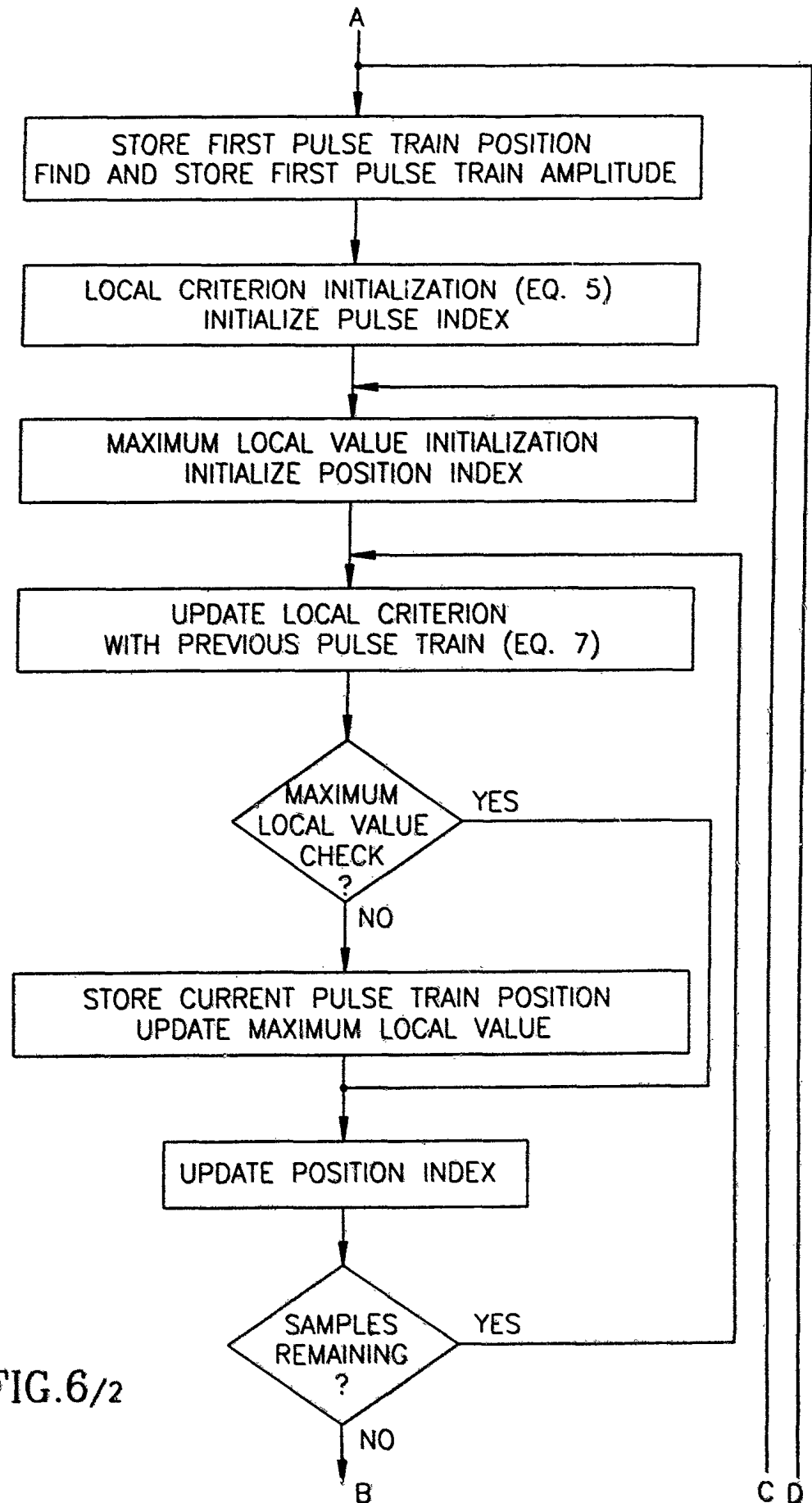


FIG.6/2

9/10

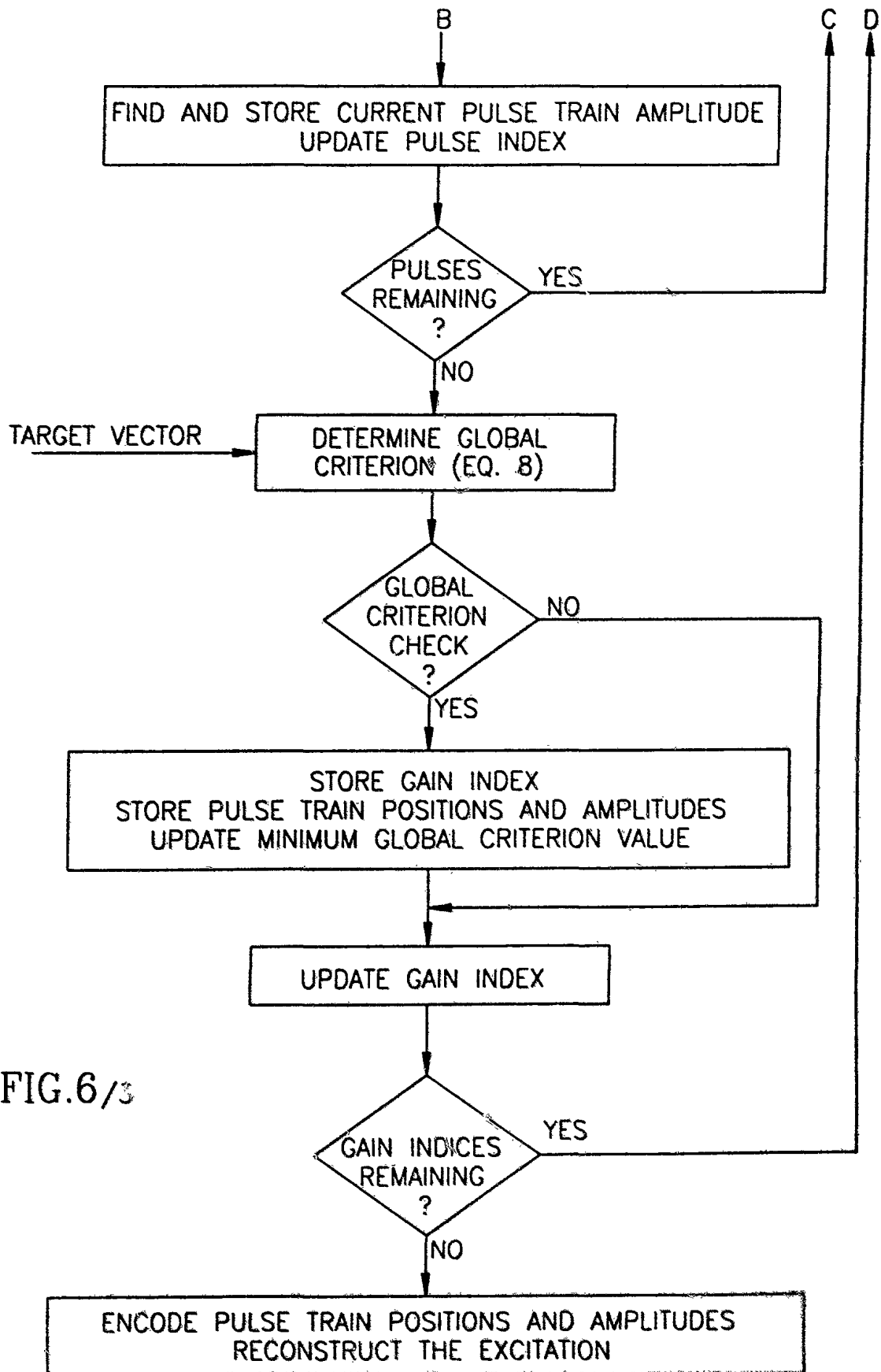
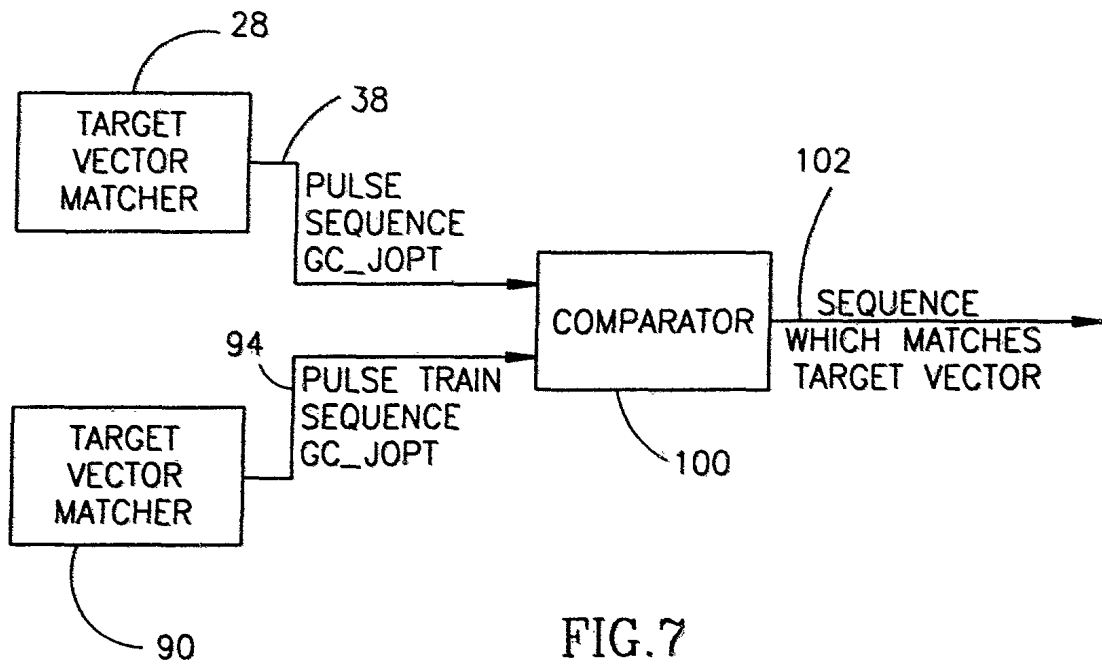


FIG.6/3

10/10



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/05014

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G10L 5/00

US CL :395/2.32

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 395/2.32, 2.10, 2.21, 2.30, 2.31, 2.32, 2.28, 2.29, 2.34

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, IEE, IEEE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, A, 4,932,061 (KROON ET AL) 05 June 1990, Fig. 1 and col. 5, lines 48-63.	1-25



Further documents are listed in the continuation of Box C.



See patent family annex.

<p>* Special categories of cited documents:</p>		<p>* T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p>	
* A*	document defining the general state of the art which is not considered to be part of particular relevance	* X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
* E*	earlier document published on or after the international filing date	* Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
* L*	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* G*	document member of the same patent family
* O*	document referring to an oral disclosure, use, exhibition or other means		
* P*	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

20 JUNE 1995

Date of mailing of the international search report

04 AUG 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer


THOMAS ONKA

Telephone No. (703) 305-9600