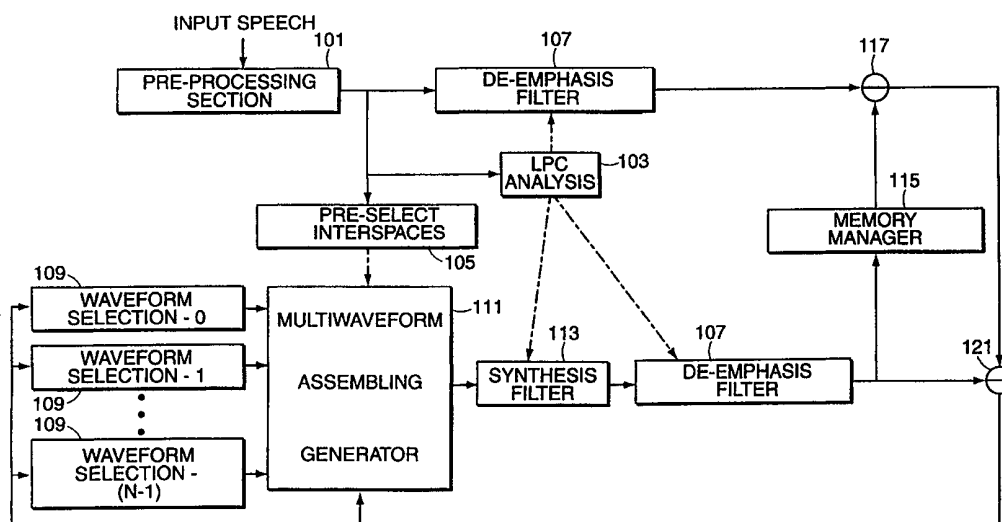




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G10L 9/14	A1	(11) International Publication Number: WO 99/44192 (43) International Publication Date: 2 September 1999 (02.09.99)
(21) International Application Number: PCT/IB99/00392 (22) International Filing Date: 25 February 1999 (25.02.99) (30) Priority Data: 09/031,522 27 February 1998 (27.02.98) US (71) Applicant: LERNOUT & HAUSPIE SPEECH PRODUCTS N.V. [BE/BE]; Flanders Language Valley 50, B- 8900 Ieper (BE). (72) Inventors: ALPUENTE, Manel Guberna; Joan Gamper, 8, 4rt 3a, E-08014 Barcelona (ES). RASAMINJANAHARY, Jean-François; Fiadanana Atsimondrano, 308-Fandriana (MG). FERAHOUI, Mohand; Lot. Cadat, 11, 35300 Rouiba (DZ). VAN COMPERNOLLE, Dirk; Nijvelsebaan 181, B-3060 Korbeek-Dijle (BE). (74) Agent: FROUD, Clive; Elkington and Fife, Prospect House, 8 Pembroke Road, Sevenoaks, Kent TN13 1XR (GB).		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

(54) Title: APPARATUS AND METHOD FOR HYBRID EXCITED LINEAR PREDICTION SPEECH ENCODING

**(57) Abstract**

A method is given of encoding a speech signal using analysis-by-synthesis to perform a flexible selection of the excitation waveforms in combination with an efficient bit allocation. This approach yields improved speech quality compared to other methods at similar bit rates.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

APPARATUS AND METHOD FOR HYBRID EXCITED LINEAR PREDICTION SPEECH ENCODING

5

FIELD OF THE INVENTION

This invention relates to speech processing, and in particular to a method for speech encoding using hybrid excited linear prediction.

10

BACKGROUND OF THE INVENTION

Speech processing systems digitally encode an input speech signal before additionally processing the signal. Speech encoders may be generally classified as either waveform coders or voice coders (also called vocoders). Waveform coders can produce natural sounding speech, but require relatively high bit rates. Voice coders have the advantage of operating at lower bit rates with higher compression ratios, but are perceived as sounding more synthetic than waveform coders. Lower bit rates are desirable in order to more efficiently use a finite transmission channel bandwidth. Speech signals are known to contain significant redundant information, and the effort to lower coding bit rates is in part directed towards identifying and removing such redundant information.

Speech signals are intrinsically non-stationary, but they can be considered as quasi-stationary signals over short periods such as 5 to 30 msec, generally known as a frame. Some particular speech features may be obtained from the spectral information present in a speech signal during such a speech frame.

Voice coders extract such spectral features in encoding speech frames.

It is also well known that speech signals contain an important correlation between nearby samples. This redundant short term correlation can be removed from a speech signal by the technique of linear prediction. For the past 30 years, such linear predictive coding (LPC) has been used in speech coding, in which the coding defines a linear predictive filter representative of the short term spectral information which is computed for each presumed quasi-stationary segment. A

general discussion of this subject matter appears in Chapter 7 of Deller, Proakis & Hansen, *Discrete-Time Processing of Speech Signals* (Prentice Hall, 1987), which is incorporated herein by reference.

A residual signal, representing all the information not captured by the
5 LPC coefficients, is obtained by passing the original speech signal through the linear predictive filter. This residual signal is normally very complex. In early LPC coders, this complex residual signal was grossly approximated by making a binary choice between a white noise signal for unvoiced sounds, and a regularly spaced pulse signal for voiced sounds. Such approximation resulted in a highly
10 degraded voice quality. Accordingly, linear predictive coders using more sophisticated encoding of the residual signal have been the focus of further development efforts.

All such coders could be classified under the broad term of residual excited linear predictive (RELP) coders. The earliest RELP coders used a
15 baseband filter to process the residual signal in order to obtain a series of equally spaced non-zero pulses which could be coded at significantly lower bit rates than the original signal, while preserving high signal quality. Even this signal can still contain a significant amount of redundancy, however, especially during periods of voiced speech. This type of redundancy is due to the regularity of the
20 vibration of the vocal cords and lasts for a significantly longer time span, typically 2.5-20 msec., than the correlation covered by the LPC coefficients, typically <2 msec.

In order to avoid the low speech quality of the original LPC coders and the simple baseband RELP coder's sub-optimal bit efficiency due to the limited
25 flexibility of the residual modeling, many of the more recent speech coding approaches may be considered more flexible applications of the RELP principle, with a long-term predictor also included. Examples of such include the Multi-Pulse LPC arrangement of Atal, U.S. Patent No. 4,701,954, the Algebraic Code Excited Linear Prediction arrangement of Adoul, U.S. Patent No. 5,444,816, and
30 the Regular-Pulse Excited LPC coder of the GSM standard.

SUMMARY OF THE INVENTION

A preferred embodiment of the present invention utilizes a very flexible excitation method suitable for a wide range of signals. Different excitations are used to accurately represent the spectral information of the residual signal, and
5 the excitation signal is efficiently encoded using a small number of bits.

A preferred embodiment of the present invention includes an improved apparatus and method of creating an excitation signal associated with a segment of input speech. To that end, a spectral signal representative of the spectral parameters of the segment of input speech is formed, composed, for instance, of
10 linear predictive parameters. A set of excitation candidate signals is created, the set having at least one member, each excitation candidate signal comprised of a sequence of single waveforms, each waveform having a type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a
15 preceding single waveform. In a further embodiment, selected parameters indicative of redundant information in the segment of input speech may be extracted from the segment of input speech. In such an embodiment, members of the set of excitation candidate signals created may be responsive to such selected parameters.

20 The first single waveform may be positioned with respect to the beginning of the segment of input speech. The relative positions of subsequent waveforms may be determined dynamically or by use of a table of allowable positions. The single waveforms may be glottal pulse waveforms, sinusoidal period waveforms, single pulses, quasi-stationary signal waveforms, non-
25 stationary signal waveforms, substantially periodic waveforms, speech transition sound waveforms, flat spectra waveforms or non-periodic waveforms. The types of single waveforms may pre-selected or dynamically selected, for instance, according to an error signal. The number and length of single waveforms may be fixed or variable. In the event that a single waveform
30 extends beyond the end of the current segment of input speech, the overflowing

portion of the waveform may be applied to the beginning of the current segment, to the beginning of the next segment, or ignored altogether.

A set of error signals is formed, the set having at least one member, each error signal providing a measure of the accuracy with which the spectral signal
5 and a given one of the excitation candidate signals encode the input speech segment. An excitation candidate signal is selected as the excitation signal when the corresponding error signal is indicative of sufficiently accurate encoding. If no excitation signal is selected, a set of new excitation candidate signals is recursively created as before wherein the position of at least one single
10 waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals. Members of the set of new excitation candidate signals are then processed as described above.

A preferred embodiment of the present invention includes another improved apparatus and method of creating an excitation signal associated with
15 a segment of input speech. To that end, a spectral signal representative of the spectral parameters of the segment of input speech is formed, composed, for instance, of linear predictive parameters. The segment of input speech is then filtered according to the spectral signal to form a perceptually weighted segment of input speech. A reference signal representative of the segment of input
20 speech is produced by subtracting from the perceptually weighted segment of input speech a signal representative of any previously modeled excitation sequence of the current segment of input speech. A set of excitation candidate signals is created, the set having at least one member, each excitation candidate signal comprised of a sequence of single waveforms, each waveform having a
25 type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform. In a further embodiment, selected parameters indicative of redundant information in the segment of input speech may be extracted from the segment of input speech. In such an embodiment,
30 members of the set of excitation candidate signals created may be responsive to such selected parameters.

The first single waveform may be positioned with respect to the beginning of the segment of input speech. The relative positions of subsequent waveforms may be determined dynamically or by use of a table of allowable positions. The single waveforms may be glottal pulse waveforms, sinusoidal
5 period waveforms, single pulses, quasi-stationary signal waveforms, non-stationary signal waveforms, substantially periodic waveforms, speech transition sound waveforms, flat spectra waveforms or non-periodic waveforms. The types of single waveforms may pre-selected or dynamically selected, for instance, according to an error signal. The number and length of single
10 waveforms may be fixed or variable. In the event that a single waveform extends beyond the end of the current segment of input speech, the overflowing portion of the waveform may be applied to the beginning of the current segment, to the beginning of the next segment, or ignored altogether.

Members of the set of excitation candidate signals are combined with the
15 spectral signal, for instance in a synthesis filter, to form a set of synthetic speech signals, the set having at least one member, each synthetic speech signal representative of the segment of input speech. Members of the set of synthetic speech signals may be spectrally shaped to form a set of perceptually weighted synthetic speech signals, the set having at least one member. A set of error
20 signals is formed, the set having at least one member, each error signal providing a measure of the accuracy with which the given members of the set of perceptually weighted synthetic speech signals encode the input speech segment. An excitation candidate signal is selected as the excitation signal when the corresponding error signal is indicative of sufficiently accurate encoding. If
25 no excitation signal is selected, a set of new excitation candidate signals is recursively created as before wherein the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals. Members of the set of new excitation candidate signals are then processed as described above.

30 Another preferred embodiment of the present invention includes an apparatus and method of creating an excitation signal associated with a segment

of input speech. To that end, a spectral signal representative of the spectral parameters of the segment of input speech is formed, composed, for instance, of linear predictive parameters. A set of excitation candidate signals composed of elements from a plurality of sets of excitation sequences is created, the set having
5 at least one member, wherein each excitation sequence is comprised of a sequence of single waveforms, each waveform having a type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform. In one embodiment, at least one of the plurality of
10 sets of excitation sequences is associated with preselected redundancy information, for example, pitch related information. In such an embodiment, members of the set of excitation candidate signals created may be responsive to such selected parameters.

The first single waveform may be positioned with respect to the
15 beginning of the segment of input speech. The relative positions of subsequent waveforms may be determined dynamically or by use of a table of allowable positions. The single waveforms may be glottal pulse waveforms, sinusoidal period waveforms, single pulses, quasi-stationary signal waveforms, non-stationary signal waveforms, substantially periodic waveforms, speech
20 transition sound waveforms, flat spectra waveforms or non-periodic waveforms. The types of single waveforms may pre-selected or dynamically selected, for instance, according to an error signal. The number and length of single waveforms may be fixed or variable. In the event that a single waveform extends beyond the end of the current segment of input speech, the overflowing
25 portion of the waveform may be applied to the beginning of the current segment, to the beginning of the next segment, or ignored altogether.

A set of error signals is formed, the set having at least one member, each error signal providing a measure of the accuracy with which the spectral signal and a given one of the excitation candidate signals encode the input speech
30 segment. An excitation candidate signal is selected as the excitation signal when the corresponding error signal is indicative of sufficiently accurate encoding. If

no excitation signal is selected, a set of new excitation candidate signals is recursively created as before wherein the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals. Members of the set of new excitation
5 candidate signals are then processed as described above.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects and advantages of the invention will be appreciated more fully from the following further description thereof with
10 reference to the accompanying drawings wherein:

Fig. 1 is a block diagram of a preferred embodiment of the present invention;

Fig. 2 is a detailed block diagram of excitation signal generation; and

Fig. 3 illustrates various methods to deal with an excitation sequence
15 longer than the current excitation frame.

DETAILED DESCRIPTION AND PREFERRED EMBODIMENTS

A preferred embodiment of the present invention generates an excitation signal which is constructed such that, in combination with a spectral signal that
20 has been passed through a linear prediction filter, it generates an acceptably close recovery of the incoming speech signal. The excitation signal is represented as a sequence of elementary waveforms, where the position of each single waveform is encoded relative to the position of the previous one. For each single waveform, such a relative, or differential, position is quantised using
25 its appropriate pattern which can be dynamically changed in either the encoder or the decoder. The relative waveform position and an appropriate gain value of each waveform in the excitation sequence are transmitted along with the LPC coefficients.

The general procedure to find an acceptable excitation candidate is as
30 follows. Different excitation candidates are investigated by calculating the error caused by each one. The candidate is selected which results in an acceptably

small weighted error. In terms of an analysis-by-synthesis conception, the relative positions (and, optionally, the amplitudes) of a limited number of single waveforms are determined such that the perceptually weighted error between the original and the synthesized signal is acceptably small. The method used to
5 determine the amplitudes and positions of each single waveform determines the final signal-to-noise ratio (SNR), the complexity of the global coding system, and, most importantly, the quality of the synthesized speech.

In a preferred embodiment, excitation candidates are generated as a sequence of single waveforms of variable sign, gain, and position where the
10 position of each single waveform in the excitation frame depends on the position of the previous one. That is, the encoding uses the differential value between the "absolute" position for the previous waveform and the "absolute" position for the current one. Consequently, these waveforms are subjected to the absolute position of the first single waveform, and to the sparse relative positions allowed
15 to subsequent single waveforms in the excitation sequence. The sparse relative positions are stored in a different table for each single waveform. As a result, the position of each single waveform is constrained by the positions of the previous ones, so that positions of single waveforms are not independent. The algorithm used by a preferred embodiment allows the creation of excitation
20 candidates in which the first waveform is encoded more accurately than subsequent ones, or, alternatively, the selection of candidates in which some regions are relatively enhanced with respect to the rest of the excitation frame.

FIG 1 illustrates a speech encoder system according to a preferred embodiment of the present invention. The input speech is pre-processed at the
25 first stage 101, including acquisition by a transducer, sampling by an analog-to-digital sampler, partitioning the input speech into frames, and removing of the DC signal using a high-pass filter.

In the particular case of speech, the human voice is physically generated by an excitation sound passing through the vocal chords and the vocal-tract. As
30 the properties of the vocal chords and tract change slowly in time, some kind of redundancy appears on the speech signal. The redundancy in the neighborhood

of each sample can be subtracted using a linear predictor **103**. The coefficients for this linear predictor are computed using a recursive method in a manner known in the art. These coefficients are quantised and transmitted as a spectral signal that is representative of spectral parameters of the speech to a decoder. For

5 quasi-stationary signals other redundancies can be present, and in particular, for speech signals a pitch value represents well the redundancy introduced by the vibration of the vocal chords. In general, for a quasi-stationary signal, several inter-space parameters are extracted which indicate the most critical redundancies found in this signal, and its evolution, in interspace parameter

10 extractor **105**. This information is used afterwards to generate the most likely train of waveforms matching this incoming signal. The high-pass filtered signal is de-emphasized by filter **107** to change the spectral shape so that the acoustical effect introduced by the errors in the model is minimized. The best excitation is selected using a multiple stage system. Several waveforms (WF) are selected in

15 waveform selectors **109**, from a bank of different types of waveforms, for example, glottal pulses, sinusoidal periods, single pulses. and historical waveform data or any subset of the types of waveforms. One subset, for example, may be simple pulse and historical waveform data. However, a larger variety of waveform types may assist in achieving more accurate encoding,

20 although at potentially higher bit rates. Of course, other waveform types in addition to those mentioned may also be employed. Fig. 2 shows the detailed structure for blocks **109** and **111**.

Thus, we define N different sets of waveforms, the k th set being WF_k , $0 \leq k \leq N-1$. As an example, where we set $N = 3$ and define three different sets of

25 waveforms: a first set of waveforms can model the quasi-stationary excitations where the signal is basically represented by some almost periodic waveforms, encoded using the relative position mechanism; a second set could be defined for non-stationary signals representing the beginning of a sound or a speech burst, being the excitation modeled with a single waveform or a small number of

30 single pulses locally concentrated in time, and thus encoded with the benefit of this knowledge using the relative position method; in general a third set may be

defined for non-stationary signals where the spectra are almost flat, and a large number of sparse single pulses can represent this sparse energy for the excitation signal, and they can be efficiently encoded using the relative position system.

Each one of these waveform sets contains M different single waveforms, where
 5 wf_{ik} represents the i th single waveform included in the k th set of waveforms in
 201 and:

$$wf_{ik} \in WF_k, 0 \leq I \leq M - 1, 0 \leq k \leq N - 1.$$

For example, in the third set of waveforms, three different single waveforms may be defined: the first one consisting of three samples, wherein
 10 the first one has a unity weight, the second one has a double weight, and the third one has also a double weight; the second single waveform consisting of two samples, the first one being a unity pulse, and the second one a "minus one" pulse; and finally, a third single waveform may be defined by a single pulse. The best single waveforms are either pre-selected or dynamically selected as a
 15 function of the feedback error caused by the excitation candidate in 203. The selected single waveforms pass through the multiple stage train excitation generator 111. To simplify, we can consider the case in which only one set of waveforms WF enters this block. This set is formed by M different single waveforms,

$$20 \quad wf_i \in WF, 0 \leq I \leq M - 1.$$

To create the current excitation candidate for the current excitation frame some single waveforms are assembled to form a sequence. Each single waveform is affected by a gain, and the distances between them (for simplicity, only the "relative" distances between successive single waveforms are
 25 considered) are constrained to some sparse values. The length for each of the single waveforms is variable. For this reason, the sequence of single waveforms may go beyond the end of the current excitation frame. Fig. 3 shows different solutions to this problem in the case of only two single waveforms. In the first case 301, the "overflowing" part of the signal is placed at the beginning of the
 30 current excitation frame and added to the existing signal. In a second case 303, the excitation frame continues and the overflowing part of the signal is stored to

be applied in the next excitation frame. Finally, in 305, the overflowing part of the signal is discarded and not taken into account in creating the excitation candidate for the current excitation frame.

Thus, the expression for the excitation signal $\hat{s}_k(n)$ may be simplified by
 5 considering only the case, as in 305, in which the overflowing part of the signal in the excitation frame is discarded, and also by requiring that the number of single waveforms admitted in the excitation frame is not variable, but limited to j single waveforms in 203. Then, the gain g_i affecting the i th single waveform of the train may be defined. Moreover, Δ_i is defined as the constrained "relative"
 10 distance between the i th single waveform and the $(i-1)$ th single waveform, and for simplicity, Δ_0 is considered an "absolute" position. Due to the fact that the number of single waveforms has been limited, the constraints in the "relative" positions for the j single waveforms may be represented by j different tables, each one having a different number of elements. Thus, the i th quantisation table
 15 defined as QT_i in 205 has NB_POS_i different sparse "relative" values, and Δ_i is constrained to satisfy the condition $\Delta_i \in QT_i[NB_POS_i]$, $0 \leq i \leq j-1$. Therefore, the "absolute" positions generated in 207 where the single waveforms can be placed are constrained following the recursion:

$$\begin{aligned}
 p_0 &= \Delta_0 \\
 20 \quad p_1 &= (\Delta_0 + \Delta_1) \\
 p_2 &= (\Delta_0 + \Delta_1 + \Delta_2) \\
 &\dots \\
 p_{i-1} &= (\Delta_0 + \Delta_1 + \Delta_2 + \dots + \Delta_{i-1}) \\
 &\dots \\
 25 \quad p_{j-1} &= (\Delta_0 + \Delta_1 + \Delta_2 + \dots + \Delta_{j-1}).
 \end{aligned}$$

Now, the excitation signal $\hat{s}_k(n)$ may be expressed as a function of the single waveforms wf_i . Each single waveform is delayed by 209 to its "absolute" position in the excitation frame basis and for each single waveform, a gain and a windowing process is applied by 211. Finally, all the single waveform
 30 contributions are added in 213. Mathematically, this concept is expressed:

$$\hat{s}_k(n) = \prod (n) \cdot \sum_{q=0}^{j-1} g_{qwf_{i_q}}(n-p_q) \cdot \prod (n-p_q), \text{ where } wf_{i_q} \in WF, \quad 0 \leq i_q \leq M-1$$

and where $\prod(n)$ is the rectangular window defined by:

$$\prod(n) = \begin{cases} 1 & \text{for } 0 \leq n \leq \text{length}-1, \\ 0 & \text{otherwise} \end{cases} \text{ and } \text{length} \text{ is the length of the excitation frame}$$

basis.

- 5 Nevertheless, in general there may be N sets of waveforms, which means there may be N different excitation signals. Among them, T excitation signals are selected in 215, that are mixed in 217, being $T < N$. Thus, the mixed excitation signal for a generic excitation frame is:

$$10 \quad \hat{s}(n) = \sum_{k=0}^{T-1} \hat{s}_k(n)$$

- where $\hat{s}_k(n)$ corresponds to the k th excitation generated from one set of waveforms. Each mixed excitation candidate passes through the synthesis LPC filter 113, then it is spectrally shaped by the de-emphasis filter 107 obtaining a
 15 new signal $\tilde{s}(n)$, and compared with a reference signal, called $\tilde{s}(n)$, in 121:

$$e(n) = \tilde{s}(n) - \tilde{s}(n).$$

This reference signal $\tilde{s}(n)$ is obtained after subtracting in 117 the contribution of the previous modeled excitation during the current excitation frame, managed in 115.

- 20 The criteria to select the best mixed excitation sequence is to minimize $e(n)$ using, for example, the least mean squared criteria.

From the above, it can be seen how an excitation signal is produced in accordance with various embodiments of the invention. This excitation signal is combined with the spectral signal referred to above to produce encoded speech
 25 in accordance with various embodiments of the invention. The encoded speech may thereafter be decoded in a manner analogous to the encoding, so that the

spectral signal defines filters that are used in combination with the excitation signal to recover an approximation of the original speech.

Although various exemplary embodiments of the invention have been disclosed, it should be apparent to those skilled in the art that various changes
5 and modifications can be made which will achieve some of the advantages of the invention without departing from the true scope of the invention. These and other obvious modifications are intended to be covered by the appended claims.

WHAT IS CLAIMED IS:

1. A method of creating an excitation signal associated with a segment of input speech, the method comprising:
 - 5 a. forming a spectral signal representative of the spectral parameters of the segment of input speech;
 - b. creating a set of excitation candidate signals, the set having at least one member, each excitation candidate signal comprised of a sequence of single waveforms, each waveform having a type, the sequence having at
10 least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform;
 - c. forming a set of error signals, the set having at least one member, each error signal providing a measure of the accuracy with which the
15 spectral signal and a given one of the excitation candidate signals encode the input speech segment;
 - d. selecting as the excitation signal an excitation candidate for which the corresponding error signal is indicative of sufficiently accurate encoding; and
 - 20 e. if no excitation signal is selected, recursively creating a set of new excitation candidate signals according to step (b) wherein the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals, and repeating steps (c)-(e).
- 25 2. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein step (a) further includes composing the spectral signal of linear predictive coefficients.
- 30 3. A method of creating an excitation signal associated with a segment of input speech according to claim 1, further including extracting from the segment

of input speech selected parameters indicative of redundant information present in the segment of input speech.

4. A method of creating an excitation signal associated with a segment of
5 input speech according to claim 3, wherein in step (b), at least one excitation candidate is further responsive to the selected parameters indicative of redundant information present in the segment of input speech.

5. A method of creating an excitation signal associated with a segment of
10 input speech as in claim 1, wherein in step (b), the first single waveform in a given one of the excitation candidate signals is positioned with respect to the beginning of the segment of input speech.

6. A method of creating an excitation signal associated with a segment of
15 input speech as in claim 1, wherein in step (b), the relative positions of subsequent single waveforms are determined dynamically.

7. A method of creating an excitation signal associated with a segment of
input speech as in claim 1, wherein in step (b), the relative positions of
20 subsequent single waveforms are determined by use of a table of allowable positions.

8. A method of creating an excitation signal associated with a segment of
input speech as in claim 1, wherein in step (b), the single waveforms include at
25 least one of: glottal pulse waveforms, sinusoidal period waveforms, and single pulses.

9. A method of creating an excitation signal associated with a segment of
input speech as in claim 1, wherein in step (b), the single waveforms include at
30 least one of: quasi-stationary signal waveforms and non-stationary signal waveforms.

10. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein in step (b), the single waveforms include at least one of: substantially periodic waveforms, speech transition sound
5 waveforms, flat spectra waveforms and non-periodic waveforms.

11. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein in step (b), the types of single waveforms are pre-selected.
10

12. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein in step (b), the types of single waveforms are dynamically selected.

15 13. A method of creating an excitation signal associated with a segment of input speech as in claim 12, wherein the dynamic selection of the types of single waveforms is a function of the set of error signals.

14. A method of creating an excitation signal associated with a segment of
20 input speech as in claim 1, wherein in step (b), the single waveforms are variable in length.

15. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein in step (b), the single waveforms are fixed in
25 length.

16. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein in step (b), the number of single waveforms in the sequence is variable.
30

17. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein in step (b), the number of single waveforms in the sequence is fixed.
- 5 18. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein step (b) further includes applying any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the current segment of input speech.
- 10 19. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein step (b) further includes applying any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the next segment of input speech.
- 15 20. A method of creating an excitation signal associated with a segment of input speech as in claim 1, wherein step (b) further includes ignoring any portion of a single waveform extending beyond the end of the current segment of input speech.
- 20 21. A method of creating an excitation signal associated with a segment of input speech according to claim 1, wherein in step (b) at least one single waveform is modulated in accordance with a gain factor.
22. A method of creating an excitation signal associated with a segment of
25 input speech as in claim 1, wherein step © employs a synthesis filter.
23. An excitation signal generator for use in encoding segments of input speech, the generator comprising:
- a. a spectral signal analyzer for forming a spectral signal
30 representative of the spectral parameters of the segment of input speech;

- b. an excitation candidate generator for creating a set of excitation candidate signals, the set having at least one member, each excitation candidate signal comprised of a sequence of single waveforms, each waveform having a type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform;
- c. an error signal generator for forming a set of error signals, the set having at least one member, each error signal providing a measure of the accuracy with which the spectral signal and a given one of the excitation candidate signals encode the input speech segment;
- d. an excitation signal selector for selecting as the excitation signal an excitation candidate signal for which the corresponding error signal is indicative of sufficiently accurate coding; and
- e. a feedback loop including the excitation candidate generator and the error signal generator configured so that the excitation candidate generator, if no excitation signal is selected, recursively creates a set of new excitation candidate signals such that the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals.

24. An excitation signal generator as in claim 23, wherein the spectral signal analyzer forms the spectral signal with linear predictive coefficients.

25. An excitation signal generator as in claim 23 further including an extractor for extracting from the segment of input speech selected parameters indicative of redundant information present in the segment of input speech.

26. An excitation signal generator as in claim 25, wherein the excitation candidate generator is responsive to the selected parameters indicative of redundant information present in the segment of input speech.

27. An excitation signal generator as in claim 23, wherein the excitation candidate generator positions the first single waveform in at least one excitation candidate signal with respect to the beginning of the segment of input speech.
- 5
28. An excitation signal generator as in claim 23, wherein the excitation candidate generator determines the relative positions of subsequent single waveforms dynamically.
- 10 29. An excitation signal generator as in claim 23, wherein the excitation candidate generator determines the relative positions of subsequent single waveforms by use of a table of allowable positions.
30. An excitation signal generator as in claim 23, wherein the excitation candidate generator uses single waveforms including at least one of: glottal pulse waveforms, sinusoidal period waveforms, and single pulses.
- 15
31. An excitation signal generator as in claim 23, wherein the excitation candidate generator uses single waveforms including at least one of: quasi-stationary signal waveforms and non-stationary signal waveforms.
- 20
32. An excitation signal generator as in claim 23, wherein the excitation candidate generator uses single waveforms including at least one of: substantially periodic waveforms, speech transition sound waveforms, flat spectra waveforms and non-periodic waveforms.
- 25
33. An excitation signal generator as in claim 23, wherein the excitation candidate generator pre-selects the types of single waveforms.
- 30 34. An excitation signal generator as in claim 23, wherein the excitation candidate generator dynamically selects the types of single waveforms.

35. An excitation signal generator as in claim 34, wherein the dynamic selection of the types of single waveforms is a function of the set of error signals.
- 5 36. An excitation signal generator as in claim 23, wherein the excitation candidate generator uses variable length single waveforms.
37. An excitation signal generator as in claim 23, wherein the excitation candidate generator uses fixed length single waveforms.
- 10 38. An excitation signal generator as in claim 23, wherein the excitation candidate generator uses a variable number of single waveforms.
39. An excitation signal generator as in claim 23, wherein the excitation
15 candidate generator uses a fixed number of single waveforms.
40. An excitation signal generator as in claim 23, wherein the excitation candidate generator applies any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the current
20 segment of input speech.
41. An excitation signal generator as in claim 23, wherein the excitation candidate generator applies any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the next
25 segment of input speech.
42. An excitation signal generator as in claim 23, wherein the excitation candidate generator ignores any portion of a single waveform extending beyond the end of the current segment of input speech.

43. An excitation signal generator as in claim 23, wherein the excitation candidate generator modulates at least one single waveform in accordance with a gain factor.
- 5 44. A method of creating an excitation signal associated with a segment of input speech, the method comprising:
- a. forming a spectral signal representative of the spectral parameters of the segment of input speech;
 - b. filtering the segment of input speech according to the spectral
10 signal to form a perceptually weighted segment of input speech;
 - c. producing a reference signal representative of the segment of input speech by subtracting from the perceptually weighted segment of input speech a signal representative of any previous modeled excitation sequence of the current segment of input speech;
 - 15 d. creating a set of excitation candidate signals, the set having at least one member, each excitation candidate signal comprised of a sequence of single waveforms, each waveform having a type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position
20 of a preceding single waveform;
 - e. combining a given one of the excitation candidate signals with the spectral signal to form a set of synthetic speech signals, the set having at least one member, each synthetic speech signal representative of the segment of input speech;
 - 25 f. spectrally shaping each synthetic speech signal to form a set of perceptually weighted synthetic speech signals, the set having at least one member;
 - g. determining a set of error signals by comparing the reference signal representative of the segment of input speech to each member of
30 the set of perceptually weighted synthetic speech signals;

- h. selecting as the excitation signal an excitation candidate signal for which the corresponding error signal is indicative of sufficiently accurate encoding; and
 - i. if no excitation signal is selected, recursively creating a set of new excitation candidate signals according to step (d) wherein the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals, and repeating steps (e)-(I).
45. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein step (a) further includes composing the spectral signal of linear predictive coefficients.
46. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein step © further includes subtracting a contribution due to previously modeled excitation in the current segment of input speech.
47. A method of creating an excitation signal associated with a segment of input speech according to claim 44, further including extracting from the segment of input speech selected parameters indicative of redundant information present in the segment of input speech.
48. A method of creating an excitation signal associated with a segment of input speech according to claim 47, wherein in step (d), the set of excitation candidate signals is further responsive to the selected parameters indicative of redundant information present in the segment of input speech.
49. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein in step (d), the first single waveform in a

given one of the excitation candidate signals is positioned with respect to the beginning of the segment of input speech.

50. A method of creating an excitation signal associated with a segment of
5 input speech as in claim 44, wherein in step (d), the relative positions of subsequent single waveforms are determined dynamically.

51. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein in step (d), the relative positions of
10 subsequent single waveforms are determined by use of a table of allowable positions.

52. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein in step (d), the single waveforms include at
15 least one of: glottal pulse waveforms, sinusoidal period waveforms, and single pulses.

53. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein in step (d), the single waveforms include at
20 least one of: quasi-stationary signal waveforms and non-stationary signal waveforms.

54. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein in step (d), the single waveforms include at
25 least one of: substantially periodic waveforms, speech transition sound waveforms, flat spectra waveforms and non-periodic waveforms.

55. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein in step (d), the types of single waveforms
30 are pre-selected.

56. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein in step (d), the types of single waveforms are dynamically selected.
- 5 57. A method of creating an excitation signal associated with a segment of input speech as in claim 55, wherein the dynamic selection of the types of single waveforms is a function of the set of error signals.
58. A method of creating an excitation signal associated with a segment of
10 input speech as in claim 44, wherein in step (d), the single waveforms are variable in length.
59. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein in step (d), the single waveforms are fixed
15 in length.
60. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein in step (d), the number of single waveforms in the sequence is variable.
20
61. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein in step (d), the number of single waveforms in the sequence is fixed.
- 25 62. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein step (d) further includes applying any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the current segment of input speech.
- 30 63. A method of creating an excitation signal associated with a segment of input speech as in claim 44, wherein step (d) further includes applying any

portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the next segment of input speech.

64. A method of creating an excitation signal associated with a segment of
5 input speech as in claim 44, wherein step (d) further includes ignoring any portion of a single waveform extending beyond the end of the current segment of input speech.

65. A method of creating an excitation signal associated with a segment of
10 input speech as in claim 44, wherein in step (d) at least one single waveform is modulated in accordance with a gain factor.

66. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein step (e) employs a synthesis filter.

15

67. A method of creating an excitation signal associated with a segment of
input speech as in claim 44, wherein step (f) employs a de-emphasis filter.

68. An excitation signal generator for use in encoding segments of input
20 speech, the generator comprising:

- a. a spectral signal analyzer for forming a spectral signal representative of the spectral parameters of the segment of input speech;
- b. a de-emphasis filter which filters the segment of input speech according to the spectral signal to form a perceptually weighted segment
25 of input speech;
- c. a reference signal generator which produces a reference signal representative of the segment of input speech by subtracting from the perceptually weighted segment of input speech a signal representative of any previously modeled excitation sequence of the current segment of
30 input speech;

- d. an excitation candidate generator for creating a set of excitation candidate signals, the set having at least one member, each excitation candidate signal comprised of a sequence of single waveforms, each waveform having a type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform;
- e. a synthesis filter which combines a given one of the excitation candidate signals with the spectral signal to form a set of synthetic speech signals, the set having at least one member, each synthetic speech signal representative of the segment of input speech;
- f. a spectral shaping filter which shapes each synthetic speech signal to form a set of perceptually weighted synthetic speech signals, the set having at least one member;
- g. a signal comparator which determines a set of error signals by comparing the reference signal representative of the segment of input speech to each member of the set of perceptually weighted synthetic speech signals;
- h. an excitation signal selector for selecting as the excitation signal an excitation candidate signal for which the corresponding error signal is indicative of sufficiently accurate encoding; and
- i. a feedback loop including the excitation candidate generator and the error signal generator configured so that the excitation candidate generator, if no excitation signal is selected, recursively creates a set of new excitation candidate signals such that the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals.

69. An excitation signal generator as in claim 68, wherein the spectral signal analyzer forms the spectral signal with linear predictive coefficients.

70. An excitation signal generator as in claim 68, wherein the reference signal generator further includes means for subtracting a contribution due to previously modeled excitation in the current segment of input speech.
- 5 71. An excitation signal generator as in claim 68 further including an extractor for extracting from the segment of input speech selected parameters indicative of redundant information present in the segment of input speech.
72. An excitation signal generator as in claim 71, wherein the excitation
10 candidate generator is responsive to the selected parameters indicative of redundant information present in the segment of input speech.
73. An excitation signal generator as in claim 68, wherein the excitation candidate generator positions the first single waveform in a given one of the
15 excitation candidate signals with respect to the beginning of the segment of input speech.
74. An excitation signal generator as in claim 68, wherein the excitation candidate generator determines the relative positions of subsequent single
20 waveforms dynamically.
75. An excitation signal generator as in claim 68, wherein the excitation candidate generator determines the relative positions of subsequent single waveforms by use of a table of allowable positions.
25
76. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses single waveforms including at least one of: glottal pulse waveforms, sinusoidal period waveforms, and single pulses.

77. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses single waveforms including at least one of: quasi-stationary signal waveforms and non-stationary signal waveforms.
- 5 78. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses single waveforms including at least one of: substantially periodic waveforms, speech transition sound waveforms, flat spectra waveforms and non-periodic waveforms.
- 10 79. An excitation signal generator as in claim 68, wherein the excitation candidate generator pre-selects the types of single waveforms.
80. An excitation signal generator as in claim 68, wherein the excitation candidate generator dynamically selects the types of single waveforms.
- 15 81. An excitation signal generator as in claim 80, wherein the dynamic selection of the types of single waveforms is a function of the set of error signals.
82. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses variable length single waveforms.
- 20 83. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses fixed length single waveforms.
- 25 84. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses a variable number of single waveforms.
85. An excitation signal generator as in claim 68, wherein the excitation candidate generator uses a fixed number of single waveforms.

86. An excitation signal generator as in claim 68, wherein the excitation candidate generator applies any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the current segment of input speech.

5

87. An excitation signal generator as in claim 68, wherein the excitation candidate generator applies any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the next segment of input speech.

10

88. An excitation signal generator as in claim 68, wherein the excitation candidate generator ignores any portion of a single waveform extending beyond the end of the current segment of input speech.

15 89. An excitation signal generator as in claim 68, wherein the excitation candidate generator modulates at least one single waveform in accordance with a gain factor.

90. A method of creating an excitation signal associated with a segment of
20 input speech, the method comprising:
a. forming a spectral signal representative of the spectral parameters of the segment of input speech;
b. creating a set of excitation candidate signals, the set having at least one member, each excitation candidate signal composed of members from
25 a plurality of sets of excitation sequences, wherein each excitation sequence is comprised of a sequence of single waveforms, each waveform having a type, the sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform;
30 c. forming a set of error signals, the set having at least one member, each error signal providing a measure of the accuracy with which the

spectral signal and a given one of the excitation candidate signals encode the input speech segment;

d. selecting as the excitation signal an excitation candidate signal for which the corresponding error signal is indicative of sufficiently accurate encoding; and

e. if no excitation signal is selected, recursively creating a set of new excitation candidate signals according to step (b) wherein the position of at least one single waveform in at least one of the excitation sequences is modified in response to the error signal, and repeating steps (c)-(e).

10

91. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein step (a) further includes composing the spectral signal of linear predictive coefficients.

15 92. A method of creating an excitation signal associated with a segment of input speech according to claim 90, further including extracting from the segment of input speech selected parameters indicative of redundant information present in the segment of input speech.

20 93. A method of creating an excitation signal associated with a segment of input speech according to claim 92, wherein in step (b), at least one of the excitation sequences is further responsive to the selected parameters indicative of redundant information present in the segment of input speech.

25 94. A method of creating an excitation signal associated with a segment of input speech as in claim 90; wherein step (b) further includes positioning the first single waveform in each excitation sequence with respect to the beginning of the segment of input speech.

30 95. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein in step (b), in at least one excitation

sequence the relative positions of subsequent single waveforms are determined dynamically.

96. A method of creating an excitation signal associated with a segment of
5 input speech as in claim 90, wherein in step (b), in at least one excitation sequence the relative positions of subsequent single waveforms are determined by use of a table of allowable positions.

97. A method of creating an excitation signal associated with a segment of
10 input speech as in claim 90, wherein in step (b), the single waveforms include at least one of: glottal pulse waveforms, sinusoidal period waveforms, and single pulses.

98. A method of creating an excitation signal associated with a segment of
15 input speech as in claim 90, wherein in step (b), the single waveforms include at least one of: quasi-stationary signal waveforms and non-stationary signal waveforms.

99. A method of creating an excitation signal associated with a segment of
20 input speech as in claim 90, wherein in step (b), the single waveforms include at least one of: substantially periodic waveforms, speech transition sound waveforms, flat spectra waveforms and non-periodic waveforms.

100. A method of creating an excitation signal associated with a segment of
25 input speech as in claim 90, wherein in step (b), the types of single waveforms are pre-selected for at least one of the excitation sequences.

101. A method of creating an excitation signal associated with a segment of
input speech as in claim 90, wherein in step (b), the types of single waveforms
30 are dynamically selected for at least one of the excitation sequences.

102. A method of creating an excitation signal associated with a segment of input speech as in claim 101, wherein the dynamic selection of the types of single waveforms is a function of the set of error signals.

5 103. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein in step (b), the single waveforms are variable in length.

104. A method of creating an excitation signal associated with a segment of
10 input speech as in claim 90, wherein in step (b), the single waveforms are fixed in length.

105. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein in step (b), the number of single waveforms
15 in at least one of the excitation sequences is variable.

106. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein in step (b), the number of single waveforms in at least one of the excitation sequences is fixed.

20

107. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein, for at least one of the excitation sequences, step (b) further includes applying any portion of a single waveform extending beyond the end of the current segment of input speech to the beginning of the
25 current segment of input speech.

108. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein, for at least one of the excitation sequences, step (b) further includes applying any portion of a single waveform extending
30 beyond the end of the current segment of input speech to the beginning of the next segment of input speech.

109. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein, for at least one of the excitation sequences, step (b) further includes ignoring any portion of a single waveform extending
5 beyond the end of the current segment of input speech.

110. A method of creating an excitation signal associated with a segment of input speech according to claim 90, wherein in step (b) at least one of the plurality of sets of excitation sequences is associated with preselected
10 redundancy information.

111. A method of creating an excitation signal associated with a segment of input speech according to claim 110, wherein the preselected redundancy information is pitch related information.

15

112. A method of creating an excitation signal associated with a segment of input speech according to claim 90, wherein in step (b) at least one single waveform is modulated in accordance with a gain factor.

20 113. A method of creating an excitation signal associated with a segment of input speech as in claim 90, wherein step (c) employs a synthesis filter.

114. An excitation signal generator for use in encoding segments of input speech, the generator comprising:

- 25 a. a spectral signal analyzer for forming a spectral signal representative of the spectral parameters of the segment of input speech;
- b. an excitation candidate generator for creating a set of excitation candidate signals, the set having at least one member, each excitation candidate signal composed of members from a plurality of sets of
30 excitation sequences, wherein each excitation sequence is comprised of a sequence of single waveforms, each waveform having a type, the

sequence having at least one waveform, wherein the position of any single waveform subsequent to the first single waveform is encoded relative to the position of a preceding single waveform;

- 5 c. an error signal generator for forming a set of error signals, the set having at least one member, each error signal providing a measure of the accuracy with which the spectral signal and a given one of the excitation candidate signals encode the input speech segment;
- d. an excitation signal selector for selecting as the excitation signal an
10 excitation candidate signal for which the corresponding error signal is indicative of sufficiently accurate encoding; and
- e. a feedback loop including the excitation candidate generator and the error signal generator configured so that the excitation candidate generator, if no excitation signal is selected, recursively creates a set of
15 new excitation candidate signals such that the position of at least one single waveform in the sequence of at least one excitation candidate signal is modified in response to the set of error signals.

115. An excitation signal generator as in claim 114, wherein the spectral signal analyzer forms the spectral signal with linear predictive coefficients.

20

116. An excitation signal generator as in claim 114 further including an extractor for extracting from the segment of input speech selected parameters indicative of redundant information present in the segment of input speech.

25 117. An excitation signal generator as in claim 114, wherein the excitation candidate generator is responsive in at least one of the excitation sequences to the selected parameters indicative of redundant information present in the segment of input speech.

118. An excitation signal generator as in claim 114, wherein the excitation candidate generator positions the first single waveform in each excitation sequence with respect to the beginning of the segment of input speech.
- 5 119. An excitation signal generator as in claim 114, wherein the excitation candidate generator determines the relative positions of subsequent single waveforms in at least one of the excitation sequences dynamically.
120. An excitation signal generator as in claim 114, wherein the excitation
10 candidate generator determines the relative positions of subsequent single waveforms in at least one of the excitation sequences by use of a table of allowable positions.
121. An excitation signal generator as in claim 114, wherein the excitation
15 candidate generator uses single waveforms including at least one of: glottal pulse waveforms, sinusoidal period waveforms, and single pulses.
122. An excitation signal generator as in claim 114, wherein the excitation candidate generator uses single waveforms including at least one of: quasi-
20 stationary signal waveforms and non-stationary signal waveforms.
123. An excitation signal generator as in claim 114, wherein the excitation candidate generator uses single waveforms including at least one of:
substantially periodic waveforms, speech transition sound waveforms, flat
25 spectra waveforms and non-periodic waveforms.
124. An excitation signal generator as in claim 114, wherein the excitation candidate generator pre-selects the types of single waveforms for at least one of the excitation sequences.

125. An excitation signal generator as in claim 114, wherein the excitation candidate generator dynamically selects the types of single waveforms for at least one of the excitation sequences.
- 5 126. An excitation signal generator as in claim 125, wherein the dynamic selection of the types of single waveforms is a function of the set of error signals.
127. An excitation signal generator as in claim 114, wherein the excitation candidate generator uses variable length single waveforms.
- 10 128. An excitation signal generator as in claim 114, wherein the excitation candidate generator uses fixed length single waveforms.
129. An excitation signal generator as in claim 114, wherein the excitation
15 candidate generator uses a variable number of single waveforms in at least one of the excitation sequences.
130. An excitation signal generator as in claim 114, wherein the excitation candidate generator uses a fixed number of single waveforms in at least one of
20 the excitation sequences.
131. An excitation signal generator as in claim 114, wherein the excitation candidate generator in at least one of the excitation sequences applies any portion of a single waveform extending beyond the end of the current segment
25 of input speech to the beginning of the current segment of input speech.
132. An excitation signal generator as in claim 114, wherein the excitation candidate generator in at least one of the excitation sequences applies any portion of a single waveform extending beyond the end of the current segment
30 of input speech to the beginning of the next segment of input speech.

133. An excitation signal generator as in claim 114, wherein the excitation candidate generator in at least one of the excitation sequences ignores any portion of a single waveform extending beyond the end of the current segment of input speech.

5

134. An excitation signal generator as in claim 114, wherein in the excitation candidate generator at least one of the plurality of sets of excitation sequences is associated with preselected redundancy information.

10 135. An excitation signal generator as in claim 134, wherein the preselected redundancy information is pitch related information.

136. An excitation signal generator as in claim 132, wherein the excitation candidate generator modulates at least one single waveform in accordance with
15 a gain factor.

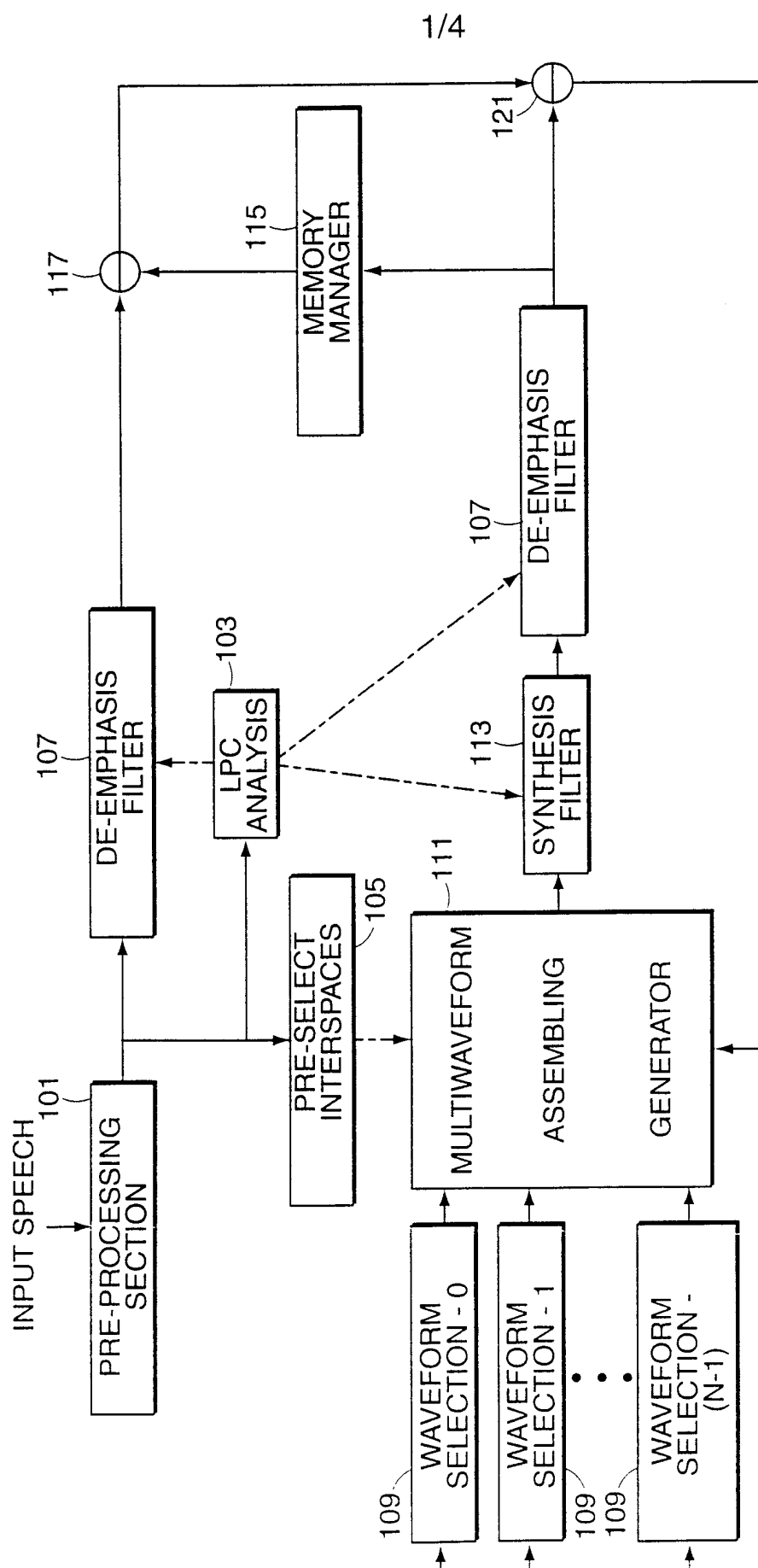


FIG. 1

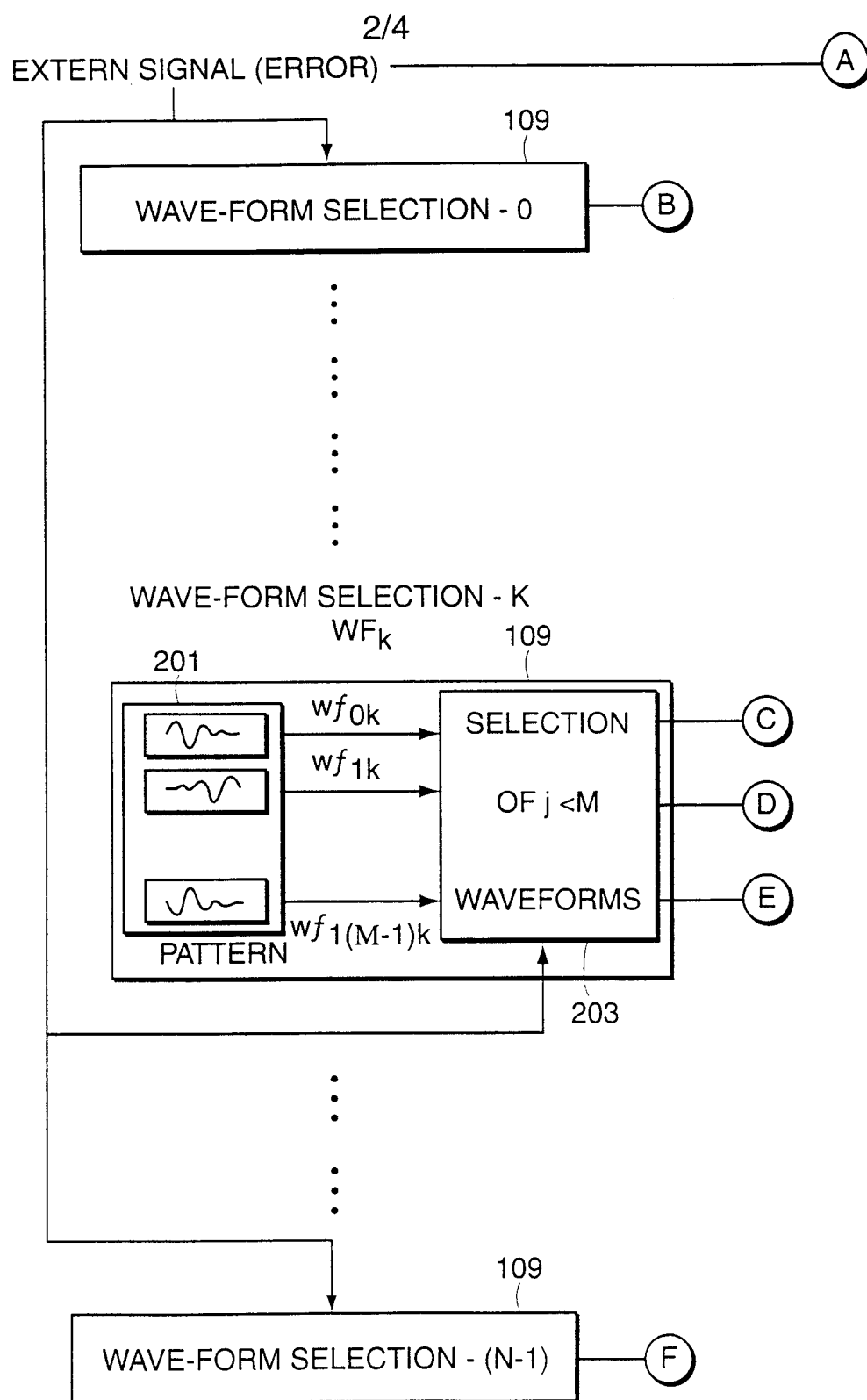


FIG. 2A

FIG. 2A	FIG. 2B
------------	------------

FIG. 2

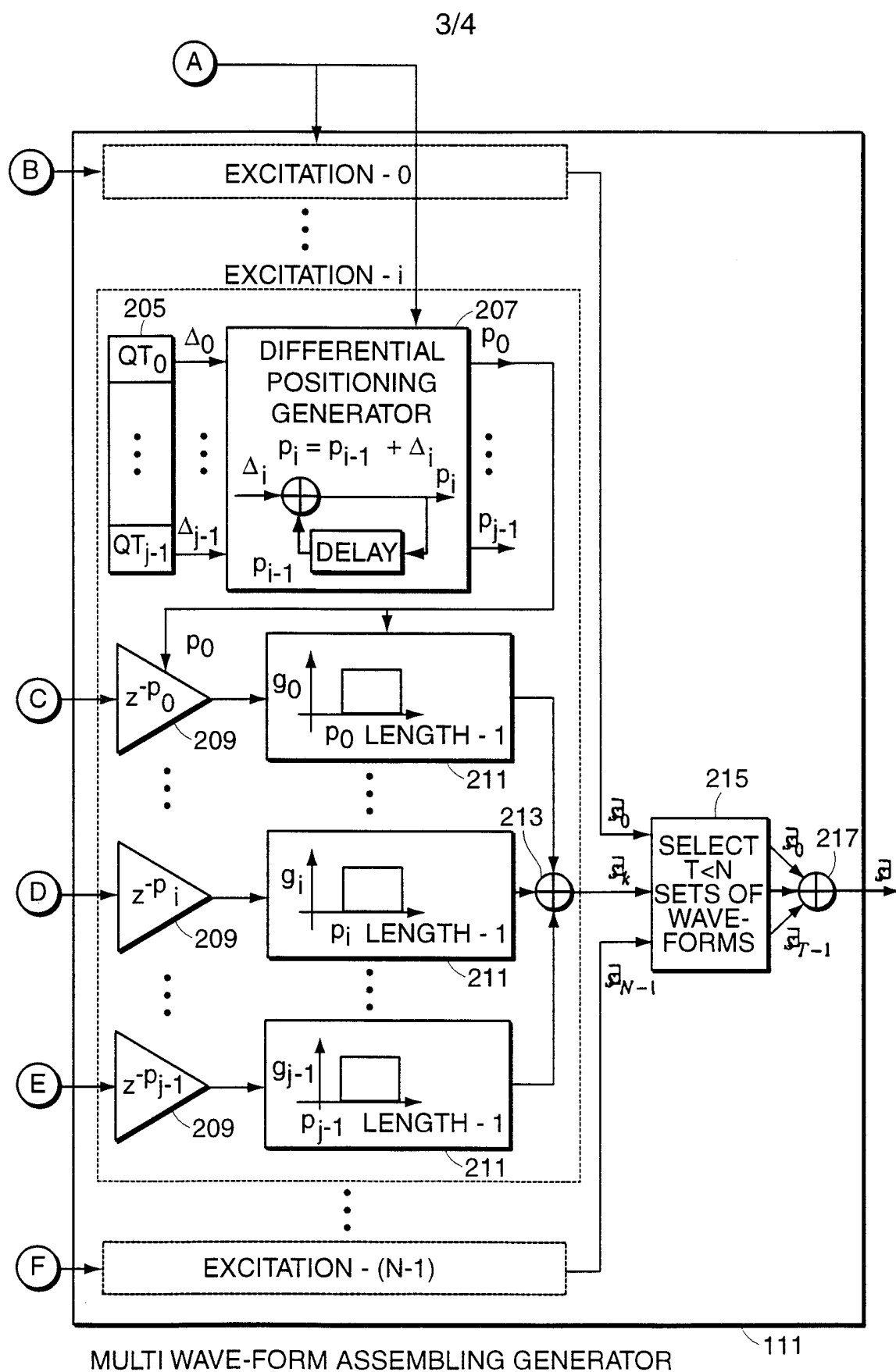


FIG. 2B

4/4

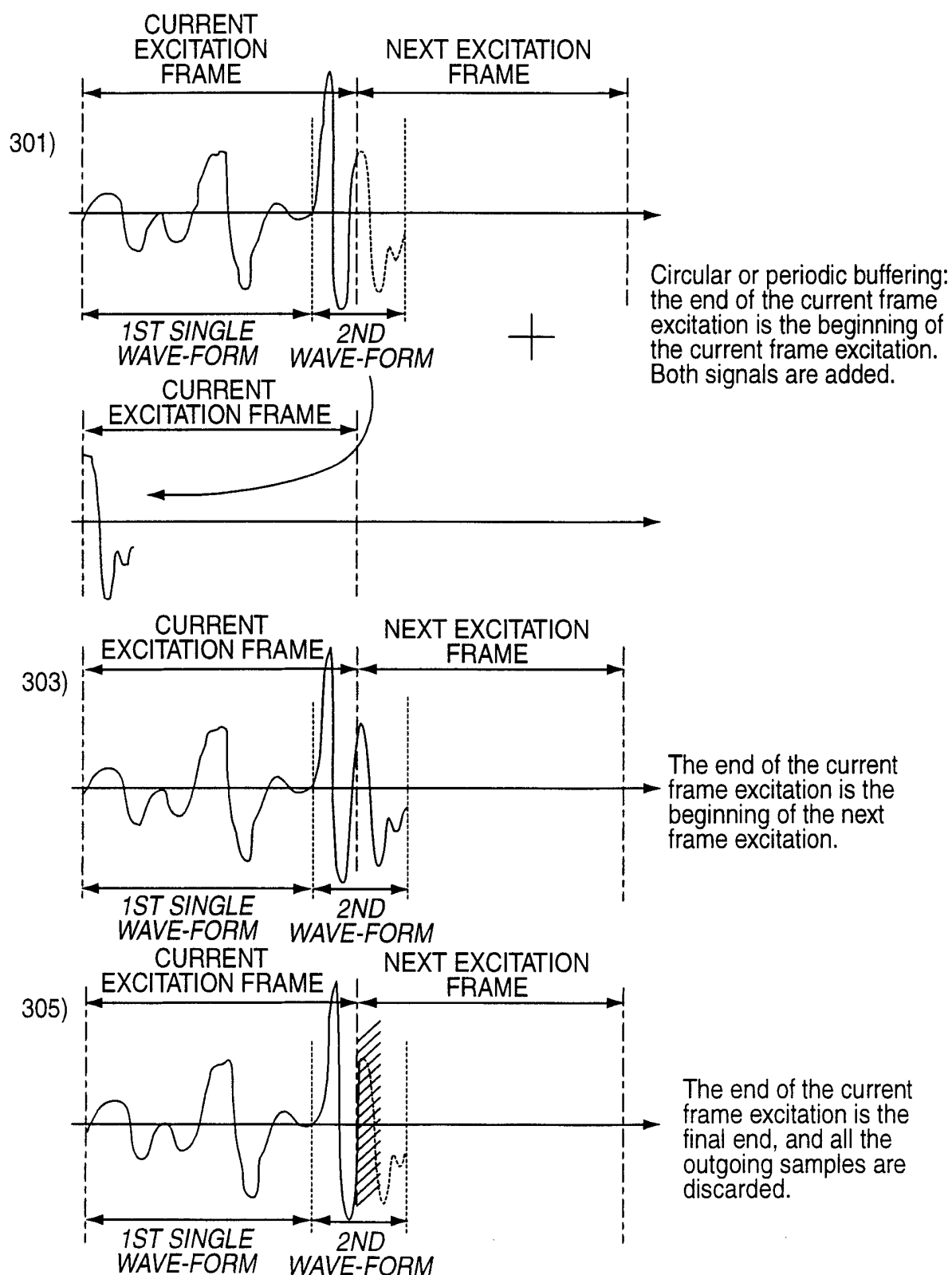


FIG. 3

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 99/00392

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G10L9/14

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)

IPC 6 G10L

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 practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 95 21443 A (QUALCOMM INC) 10 August 1995	1,2, 8-10,12, 15,21, 23,24, 30-32, 34,37, 43,90, 91, 97-99, 101,104, 112,114, 115
X		121-123, 125,128, 136
A	see abstract; claims 1-10	44,45, 52-54, 56,59,
	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"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)

"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

"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

"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

"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.

"&" document member of the same patent family

Date of the actual completion of the international search

14 June 1999

Date of mailing of the international search report

22/06/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

VAN DOREMALEN, H

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 99/00392

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	----- US 5 444 816 A (J. ADOUL, C.LAFLAMME) 22 August 1995 see abstract; claims 1-3,7 -----	65, 68, 69, 76-78, 80, 83, 89 1, 23, 44, 68, 90, 114

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 99/00392

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9521443 A	10-08-1995	AU 693519 B	02-07-1998
		AU 1739895 A	21-08-1995
		BR 9506574 A	23-09-1997
		CA 2181456 A	10-08-1995
		CN 1139988 A	08-01-1997
		EP 0744069 A	27-11-1996
		FI 962968 A	24-09-1996
		JP 9508479 T	26-08-1997
		US 5621853 A	15-04-1997
US 5444816 A	22-08-1995	CA 2010830 A	23-08-1991
		AT 164252 T	15-04-1998
		AU 6632890 A	18-09-1991
		WO 9113432 A	05-09-1991
		DE 69032168 D	23-04-1998
		DE 69032168 T	08-10-1998
		EP 0516621 A	09-12-1992
		ES 2116270 T	16-07-1998
		US 5699482 A	16-12-1997
		US 5754976 A	19-05-1998
		US 5701392 A	23-12-1997