(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 355 298 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
22.10.2003 Bulletin 2003/43

(51) Int Cl.7: G10L 19/12

(21) Application number: 03013629.5

(22) Date of filing: 10.06.1993

(84) Designated Contracting States:
DE FR GB SE

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
93913500.0 / 0 654 909

(71) Applicant: Oki Electric Industry Company, Limited
Tokyo 105 (JP)

(72) Inventors:
- Hosoda, Kenichiro
  Minato-ku, Tokyo 105 (JP)
- Aoyaki, Hiromi
  Minato-ku, Tokyo 105 (JP)
- Katsuragawa, Hiroshi
  Minato-ku, Tokyo 105 (JP)
- Ariyama, Yoshihiro
  Minato-ku, Tokyo 105 (JP)

(74) Representative: Williams, Ceili
Stevens Hewlett & Perkins
Halton House
20/23 Holborn
London EC1N 2JD (GB)

Remarks:
This application was filed on 16 - 06 - 2003 as a divisional application to the application mentioned under INID code 62.

(54) **Code Excitation linear prediction encoder and decoder**

(57) A code excitation linear predictive (CELP) coding or decoding apparatus is provided in which a code vector, which is provided by a stochastic codebook (108), is converted adaptively in accordance with vocal tract analysis information (LPC) so that a high quality reproduction speech is obtained at a low coding rate. Further, in order to obtain a similar effect, a pulse-like excitation codebook formed of an isolated impulse is provided in addition to the adaptive excitation codebook (107) and stochastic excitation codebook (108) so that either the stochastic excitation codebook or the pulse-like excitation codebook is selectively used to provide a vocal tract parameter as a linear spectrum pair parameter.

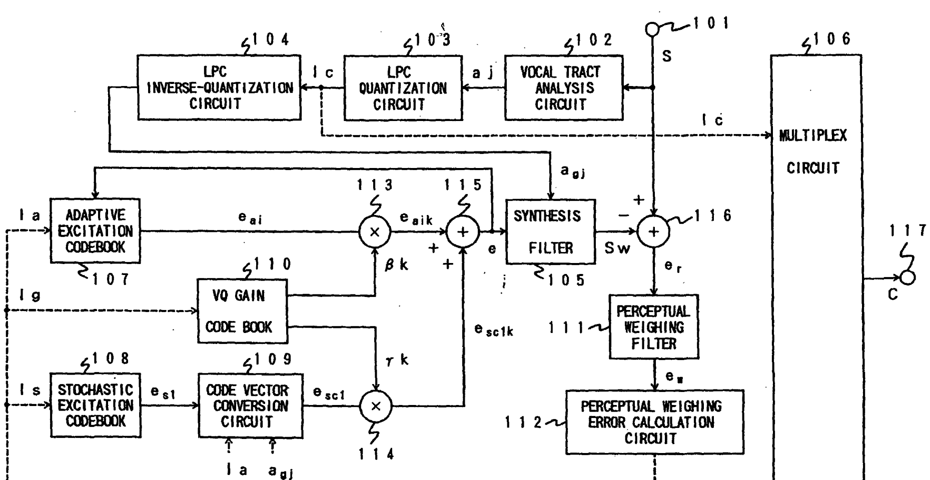


Fig. 1

EP 1 355 298 A2

Printed by Jouve, 75001 PARIS (FR)

Description

TECHNICAL FIELD OF THE INVENTION

[0001] This invention relates to an encoder and a decoder based on the code excitation linear predictive coding (CELP) system.

BACKGROUND OF THE INVENTION

[0002] Conventionally, as a high efficient coding system for speech signal including audible signal in a field of digital transportable communication system, a code excitation linear predictive coding and its modification, that is, a vector sum excitation linear predictive coding system (VSELP) have been used. The coding apparatus which uses the code excitation linear predictive coding (CELP) is disclosed in, for example, N.S. Jayant and J.H.Chen, "Speech Coding with Time-varying Bit Allocation to Excitation and LPC Parameters", Proc. ICASSP, pp65-68, 1989.

[0003] A fundamental construction of the coding system relative to the speech signal is to obtain vocal tract parameters representing vocal tract properties and excitation source parameters representing excitation source information. In the recent CELP system, an excited signal as a excitation source information is encoded by means of both an adaptive excitation codevectors, which contribute to stochastically stronger periodic excitation signal and stochastic excitation codevectors which contribute to stochastic less periodic random excitation signal, and then the coded excitation signals are stored in a codebook, and an optimum adaptive excitation codevectors and stochastic excitation codevectors are found out in each codebook so that weighted error power sum between an input speech vector and synthetic speech vector becomes minimum. Then, whatever it is of a forward-type coding system which obtains vocal tract parameters from an input speech vector or of a backward-type coding system which obtains vocal tract parameters from synthetic speech vectors, at least the excitation source parameters, that is, adaptive excitation code and stochastic excitation code information are transmitted.

[0004] By utilizing the code excitation linear predictive (CELP) system as described above, it is known that a high quality regenerated speech signals are obtained at a coding rate of 6 kbit/s to 8 kbit/s.

[0005] However, some communication systems require lower coding rate, for example 4kbit/s or less. In such a lower coding rate, regardless of being the forward type which transmits both vocal tract parameters and excitation source parameters or being the backward type which transmits excitation source parameters, the number of coded bits which are assigned to the excitation source parameters is smaller and the number of adaptive excitation codevectors stored in the adaptive excitation codebook and the number of stochastic excitation codevectors stored in the stochastic excited codebook become smaller. Consequently, the quality of the regenerated speech signal inevitably degrades at the lower coding rate as described above.

[0006] Besides, the adaptive excited codebook are adaptively renewed by synthetic codevectors of optimum adaptive excitation codevectors and stochastic excitation codevectors and, accordingly, it can be determined that the adaptive excitation codevectors are formed on the basis of the stochastic excitation codevectors. Therefore, the current CELP coding has a poor tracking capability for a voice signal having a nature of strong periodicity. Consequently, generated speech signal lacks clearness.

[0007] A speech coding and decoding system that attempts to realise a higher compression of speech information is described in EP 476614. Here, a sparse adaptive codebook is used in association with a time-reversed perceptual weighting filter.

SUMMARY OF THE INVENTION

[0008] The present invention is based upon the foregoing problems and an object of the present invention is to provide code excitation linear predictive coding encoder and decoder which can provide a high quality regenerated speech signal even when pulse-like noise components are contained in the input speech vectors.

[0009] Another object of the present invention is to provide code excitation linear predictive coding encoder and decoder which can provide high-quality regenerated speech signal even when a lower coding rate is employed.

[0010] According to the present invention, there is provided a code excitation linear predictive coding apparatus which uses, as a speech excitation source information, excitation signals in the form of excitation codebook, wherein the apparatus is provided with a codevectors conversion circuit which converts the frequency characteristics of fixed codevectors such as stochastic excitation codevectors transmitted from the excitation codebook into the predetermined frequency characteristics at the time of output of the excitation codevectors. A primary reason for providing the codevectors conversion circuit is as set forth below. Conventionally, the frequency characteristics of an excitation signal is modelled as "theoretically white" and yet it actually is not "white" but is recognized by examinations to have a characteristic which is near to a frequency characteristics of an input speech vectors. Therefore, the nearer the fixed codevectors frequency characteristics is set to the frequency characteristics of the input speech vectors, the higher the quality of the synthetic speech vector is obtained and, moreover, an effective frequency component of the excitation codevectors becomes much larger than a quantization error vectors so that a masking effect of the quantization error vector can be obtained. As an information representing frequency characteristics of the code conversion circuit, parameters of LPC (linear predictive coefficient) and optimum adaptive excitation code information which means pitch predictive information (which includes VQ gains) are used. Thus, the codevectors conversion circuit controls the frequency characteristics of the stochastic excitation codevectors and so forth, in accordance with these information.

[0011] Further, in the present invention, there is provided a code excitation linear predictive decoding apparatus which has codevectors conversion circuit which forces the fixed codevector frequency characteristics near to the input speech vector frequency characteristics in accordance with the respective code excitation linear predictive coding system.

[0012] In the codevector converter circuit, an impulse response determined by the following formula (1) as filter transfer function H(Z) according to the vocal tract parameters,

$$H(Z)= (1-\Sigma A^{j}ajZ^{-j}) / (1-\Sigma B^{j}aj^{-j}) \qquad (1)$$

or an impulse response determined by the following formula (2) in accordance with a excited pitch lag,

$$H(Z) = 1/ (1-\varepsilon\ Z^{-L}) \qquad (2)$$

or an impulse response which is cascade-connected filter represented by formulas (1) and (2) is used to proceed a convolution treatment to the stochastic excitation codevectors and thereafter a adaptive excitation codevectors are added to produce excitation codevectors. Here, aj(j=1 to p) represents a parameter of LPC and p represents the order of LPC analysis. A, B and E are constants which are determined in the range of $0<A<1$, $0<B<1$ and $0<\varepsilon\leq 1$, respectively, and L represents a pitch lag.

[0013] Further, the present invention provides a code excitation linear predictive coding or decoding apparatus which is provided, as an excitation codebook, with a adaptive excitation codebook and stochastic excitation codebook, in which pulse-like excitation codebook storing a pulse-like excitation codevector which consists of isolated impulse in addition to the adaptive excitation codebook and stochastic excitation codebook is provided so that the current CELP coding has a good tracking capability for a speech signal having a nature of strong periodicity. Thus, clear regenerated speech signal can be obtained.

[0014] Further, in the code excited linear predictive coding apparatus, excitation codevectors from the stochastic excitation codebook or pulse-like excitation codebook are selectively used, and this selected information is transmitted to the code excitation linear predictive decoder apparatus. In this code excitation linear predictive decoder apparatus, the excitation codevectors from the stochastic excitation codebook or pulse-like excitation codebook are selected in accordance with the information transmitted from the code excitation linear predictive coding apparatus.

[0015] In addition, in each of the above-described code excitation linear predictive encoders, the output of vocal tract parameters are assigned to be LSP (linear spectral pair) parameters and this linear spectral pair parameters are utilized for the speech regeneration in the code excitation linear predictive decoder so that the regeneration speech quality at the lower coding rate can be improved from a viewpoint of vocal tract parameters. The reasons for using LSP parameters as the vocal tract parameters reside in that an interpolation characteristics relative to the frequency characteristics of the vocal tract are improved, that the LSP parameters provides less distortion to the vocal tract spectral than LPC parameters even when the LSP parameters are coded by smaller number of code bits, and that an effective coding can be obtained by combination with vector quantization.

BRIEF DESCRIPTION OF THE DRAWING

[0016]

Fig. 1 is a block diagram of a code excitation linear predictive encoder (coding apparatus) according to a first and a second embodiments of the present invention.

Fig. 2 is a block diagram of a code excitation linear predictive decoder in correspondence with the code excitation linear predictive encoder shown in Fig. 1.

Fig. 3 is a block diagram of a code excitation linear predictive encoder (coding apparatus) according the a third embodiment of the invention.

Fig. 4 is a block diagram of a code excitation linear predictive decoder in correspondence with the code excitation linear predictive encoder shown in Fig. 3.

Fig. 5 is a detailed block diagram of a codevector conversion circuit shown in Figs. 3 and 4.

BEST MODE FOR CARRYING OUT THE INVENTION

[0017] Preferred embodiments of the code excitation linear predictive coding apparatus (encoder) and the code excitation linear predictive decoding apparatus (decoder) according to the present invention will be described with reference to the figures of the drawing attached herewith.

[0018] Referring to Fig. 1 which shows a code excitation linear predictive encoder (coding apparatus) according the a first embodiment of the present invention, an input speech vector S which has been inputted in each frame from an input terminal 101 is first transmitted to a vocal tract analysis circuit 102 to obtain a vocal tract parameter aj (linear predictive coefficient).

[0019] An LPC (linear predictive coefficient) quantization circuit 103 quantizes vocal tract predictive parameter aj and transmits its code Ic (quantized LPC code) to an LPC inverse-quantization circuit 104 and a multiplex circuit 106.

[0020] The LPC inverse-quantization circuit 104 serves to convert the LPC code Ic into vocal tract predictive parameter aqj and transmits the same to a synthesis filter 105.

[0021] Then, an adaptive excitation codevector e ai (i=1 to n) is outputted from a adaptive excitation codebook 107 and similarly, a stochastic excitation codevector e sl (l=1 to m) is from a stochastic excitation codebook 108. Similarly, an excitation gains βk and γk (k=1 to r) are outputted from a VQ gain codebook 110.

[0022] A codevector conversion circuit 109, which has an impulse response of filter transfer function H(Z) represented by the following formula (3), performs convolutional computation with stochastic excitation codevector e sl from a stochastic excitation codebook 108, and transmits a converted stochastic excitation codevector e scl.

$$H(Z)=(1-\sum_{j=1}^{P} 0.4^{j}aqjZ^{-j})/(1-\sum_{j=1}^{P} 0.9^{j}aqjZ^{-j})\ldots(3)$$

wherein aqj represents an output of LPC inverse quantization circuit 104 and p represents vocal tract analysis order.

[0023] The adaptive excitation codevector e ai is multiplied by the gain βk by means of a multiplier 113 to produce a vector e aik and, on the other hand, the converted stochastic excitation codevector e scl is multiplied by the gain γk by means of a multiplier 114 to produce a vector e sclk.

[0024] An adder 115 adds the components of vector e alk and vector e sclk and produces an excitation codevector e.

[0025] The synthesis filter 105 calculates synthetic speech vector Sw corresponding to the excitation codevector e and transmits it to a subtracter 116.

[0026] The subtracter 116 performs the subtraction between the synthesized speech vector Sw and the input speech vector S, and the obtained error vector between Sw and S is transmitted to a perceptual weighting filter 111.

[0027] The perceptual weighting filter 111 transmits a perceptual weighting error vector ew corresponding to the error vector er to a perceptual weighting error calculation circuit 112.

[0028] The perceptual weighting error calculation circuit 112 calculates a mean square value of each component of the perceptual weighting error vector ew, and determines the excitation codevector (i.e., combination of i, l and k) to minimize the mean square error power of ew for the input speech vector at the present time. Indexes Ia, Is and Ig of each codebook at this moment are transmitted to each of the adaptive excitation codebook 107, stochastic excitation codebook 108, VQ gain codebook 110 and multiplex circuit 106.

[0029] The adaptive excitation codebook 107 outputs an optimum adaptive excitation codevector ea0 assigned by index Ia, the stochastic excitation codebook 108 outputs an optimum stochastic excitation codevector es0 assigned by index Is, and the VQ gain codebook 110 transmits optimum VC gain β_0 and γ_0 assigned by index Ig. A codevector conversion circuit 109 converts the stochastic codevector es0 which has been transmitted from the stochastic excitation codebook in accordance with the index Is into an optimum converted stochastic excitation codevector e sc0 and then outputs it to the multiplier 114.

[0030] The optimum excitation codevector $e_0$pt composed by the ea_0, esc_0, β_0 and γ_0 is transmitted to the adaptive excitation codebook 107 and updates the content of the adaptive excitation codebook 107.

[0031] The multiplex circuit 106 multiplexes Ic, Ia, Is and Ig, as a total code C, and transmits it to the receiver through an output terminal 117.

[0032] Fig. 2 is a block diagram of a code excitation linear predictive decoder corresponding to the code excitation linear predictive encoder.

[0033] In Fig. 2 the total code C from an input terminal 201 is separated by a demultiplex circuit 212 into LPC code Ic, adaptive excitation code index Ia, stochastic excitation code index Is, and VQ gain code index Ig and they are transmitted, respectively, to LPC inverse quantization circuit 202, adaptive excitation codebook 204, stochastic excitation codebook 205 and VQ gain codebook 207.

[0034] The LPC inverse quantization circuit 202 converts the LPC code Ic into vocal tract predictive parameter aj and transmits to a synthesis filter 203. The adaptive excitation codebook 204 outputs adaptive excitation codevector ea assigned by the index Ia, the stochastic excitation codebook 205 outputs a stochastic excitation codevector es assigned by the index Is, and a VQ gain codebook 207 outputs excitation gains β and γ, assigned by index Ig.

[0035] A codevector conversion circuit 206 converts the vector es into vector e sc and outputs it as similar as the aforementioned code excitation linear predictive coding apparatus (encoder).

[0036] The adaptive excitation codevector ea is multiplied by gain β by means of multiplier 208, and the vector e sc is multiplied by gain γ by means of multiplier 209. These multiplied vector components are added by adder 210, and final excitation codevector e for synthesis filter is obtained.

[0037] A synthesis filter 203 calculates a synthesized speech vector S corresponding to the excitation codevector e and outputs to an output terminal 211. At the same time, the content of the adaptive excitation codebook 204 is updated by vector e.

[0038] The code excitation linear predictive encoder according to the second embodiment of the invention will be explained with reference to Fig. 1 again.

[0039] This code excitation linear predictive encoder according the a second embodiment has the similar construction as that of the first embodiment except the codevector conversion circuit 109 and, therefore, an operational mode of the codevector conversion circuit 109 will be explained presently.

[0040] The codevector conversion circuit 109, which has an impulse response of filter transfer function H(Z) shown by the following formula (4) performs convolutional computation with the vector e sl and results in vector e scl.

$$H(Z)=1/(1-\varepsilon Z^{-L}) \tag{4}$$

[0041] Where ε is $\varepsilon \leq 1.0$, and L is a pitchlag obtained from index of the adaptive excitation code.

[0042] Incidentally, in the codebook of a shift-type adaptive excitation codebook, the index of the adaptive excitation code corresponds with the pitch lag index as below.

Adaptive code index	**0**	**1**	**2......**	**i**
	↓	↓	↓	↓
Pitch lag	**20**	**21**	**22.....**	**20+i**

[0043] The convolutional processing of the aforementioned code excitation linear predictive coding apparatus (encoder) are represented by the following formula (5), provided that the e sl is an output stochastic excitation codevector of the stochastic excitation codebook, e scl is a stochastic excitation codevector after the conversion, and h is an impulse response of conversion circuit.

$$e\ scl = e\ sl \times h \tag{5}$$

wherein:

e scl = $[x_0, x_1 ,,,,,, x_{n-1}]$, e sl=$[y_0, y_1,,,y_{n-1}]$,
h=$[h_0, h_1,,,,,h_{n-1}]$ (The bracket [ ] is column vector.),
x, y and h are elements, and n is subframe length (or frame length).

[0044] A transfer function composed of a vocal tract parameter, or a transfer function composed of the pitch lag can be used for the impulse response of code conversion circuit, alternatively, said two transfer functions can be cascaded to form the impulse response.

[0045] Fig. 3 is a block diagram of a code excitation linear predictive encoder according to the third embodiment of the invention. In Fig. 3 this code excitation linear predictive encoder is primarily composed of a input speech process portion 301, optimum synthesized speech search portion 302 and multiplex circuit 303.

[0046] The input speech process 301 has LSP parameter analysis circuit 311, LSP parameter coding circuit 312, LSP parameter decoding circuit 313, LPC conversion circuit 314, perceptual weighting filter 315, synthesis filter zero input response generation circuit 316, perceptual weighting filter zero input response generation circuit 317, and subtracters 318 and 319. When an input vector is given, a speech parameter which is to be transmitted to the decoder is obtained and, target speech vector for a synthesized speech vector which is formed by local reproduction.

[0047] In the code excitation linear predictive encoder, digitalized discrete input speech vector series are stored as much as the time which corresponds to an analysis frame length for obtaining a vocal tract parameter and, this analysis frame length is separated into several subframes and processed by input speech processing portion 301.

[0048] The input speech vector is given to the LSP parameter analysis circuit 311, analyzed by the LSP analysis circuit 311, and converted to LSP parameter as vocal tract parameter. This LSP parameter is coded (for example, to be vector quantized) by LSP parameter coding circuit 312 and given to the multiplex circuit 303 and transmitted to the code excitation linear decoder. The coded LSP parameter is decoded (vector quantized) by LSP parameter decoding circuit 313 and converted to LPC by the LPC conversion circuit 314. The thus converted LPC is used as a tap coefficient for perceptual weighting filter 315, synthesis filter zero input response generation circuit 316, perceptual weighting filter zero input generation circuit 317 and a synthesis filter 329 which will be described presently, and given also to a code vector conversion circuit 328. The quantized LSP parameter is converted into LPC.

[0049] Next, an operation for forming a target speech vector relative to synthesized speech vector which is locally reproduced from the input speech vector will be explained.

[0050] The input speech vector described above is given to the perceptual weighting filter 315 and after the weighing processing in consideration of human perceptual characteristics, the input speech vector is given to a subtracter 318 to be subtracted. Further, a zero input response vector in relation to a synthesis filter 329, is given for input of subtracter 318. Thus, a speech vector, from which an influence of the synthesis filter 329 in the immediately before analysis frame is excluded, is given to subtracter 319. Further, a zero input response vector in relation to a perceptual weighting filter 315, is given for input of subtracter 139. Thus, a speech vector, from which an influence of the weighted filter 315 in the immediately before analysis frame is obtained, is given to subtracter 330.

[0051] The optimum synthesizedtic speech search portion 302 serves to search a excitation source parameter in which the synthesis speech vector in the local reproduction is most similar to the target speech vector, and is composed of adaptive excitation codebook 320, stochastic excitation codebook 321, pulse-like excitation codebook 322, VQ gain codebook 323, VQ gain controllers 324 and 327, adder 325, fixed codebook selection switch 326, codevector conversion circuit 328, synthesis filter 329, subtracter 330, error power sum computing circuit 331 and code selection circuit 332.

[0052] Each of the adaptive excitation codebook 320, stochastic excitation codebook 321 and pulse-like excitation codebook 322 stores adaptive excitation codevector, which is a waveform code in relation to an excitation signal, stochastic excitation codevector and pulse-like excitation codevector, respectively, and VQ gain codebook 323 stores VQ gain code which is related to adaptive excitation codevector and fixed codevector (which generally represents stochastic excitation codevector and pulse-like excitation codevector).

[0053] The adaptive excitation code vector contributes to the voiced speech signal having stochastically periodicity, while the stochastic excitation codevector contributes to the unvoiced speech signal having stochastically less periodicity. The adaptive excitation codevector of the adaptive excitation codebook 320 is adaptively updated as described presently.

[0054] The pulse-like excitation codevector is a waveform excitation codevector consisting of an unit impulse and is considered to contribute to the steady portion of the voiced speech signal having a strong periodicity.

[0055] The VQ gain code is vector-quantized, for example, and one component of the vector relates to VQ gain for adaptive excitation code vector and the other component relates to VQ gain for the fixed code vector.

[0056] Pulse-like excitation code vector is a periodic simple signal which can be generated by means of a pulse signal generating circuit but, it can preferably be generated by coding and reading out from the codebook 322 as this code excitation linear predictive encoder, the reason of which will be explained presently. Namely, it is easy to synchronize the excitation vector with an output from the adaptive excitation codebook 320. The same processing for selecting the stochastic excitation codebook can be pulse-like excitation codevector search by constituting the excitation code vector to have the same codebook construction with the codebook 321.

[0057] By utilizing said various codebook to obtain an optimum code so that the locally synthesized speech vector becomes the most similar to the target speech vector, and its indices are given to the multiplex circuit 303 and are transmitted to the code excitation linear predictive decoder portion.

[0058] In case of the search of an optimum code including a selection of the stochastic excitation code vector or the pulse-like excitation code vector as described above, the searching is carried out with respect to the adaptive excitation code, stochastic excitation code, pulse-like excitation code and VQ gain code, in turn, in this code excitation linear predictive encoder.

[0059] In case of searching an optimum adaptive excitation code vector, an output from the stochastic excitation codebook 321 and the pulse-like excitation codebook 322 are assigned to be zero (0), and the VQ gain controller 324 multiply a suitable value of VQ coefficient ("1", for example). In this state, the adaptive excitation codebook 320 outputs all of the stored adaptive excitation code vector sequentially or in parallel, and gives it as an excitation code vector to the synthesis filter 329 through the VQ gain controller 324 and the adder 325. The synthesis filter 329 carries out a convolutional computing relative to the excitation code vector, by utilizing, as a tap coefficient, the LPC which is given from the LPC conversion circuit 314, and a synthesized speech vectors, which are synthesized only by the content of the adaptive excitation code vector as the excitation source signal, are obtained with respect to all the adaptive excitation code vector.

[0060] The subtracter 330 obtains, with respect to all of the adaptive excitation code vector, an error vector between the synthesized speech vector on which only the content of the adaptive excitation code vector is effected and the target speech vector, and then gives it to an error power sum calculation circuit 331. The error power sum calculation circuit 331 obtains square sum (error power sum) of the error vector, with respect to all the adaptive code vector, and gives it to a code selection circuit 332. The code selection circuit 332 determines the the adaptive excitation code vector to minimize the error power sum.

[0061] Next, an optimum stochastic excitation code vector searching is carried out and in the searching of this, a fixed codebook selection switch 326 is driven to the side of the stochastic excitation codebook 321 the output from adaptive excitation codebook is set to zero (0) or to the previously obtained optimum adaptive excitation code vector. In the state as this, the stochastic excitation codebook 321 outputs sequentially or in parallel, all the stored stochastic excitation code vectors,and inputs them into the code vector conversion circuit 328 through the fixed codebook selection switch 326 and VQ controller 324.

[0062] The code vector conversion circuit 328 proceeds the conversion of the frequency characteristics of inputted stochastic excitation code vector so that it is moved to close to frequency characteristics of an input speech vector in correspondence with time-length of the stochastic excitation code vector. As described above, all the stochastic exited code vector with its frequency characteristics being conversion-processed is given, as an excitation code vector, to a synthetic filter 329. Thereafter, it is processed as similar as the searching of the optimum adaptive excitation code vector, and the code selection circuit 332 determines an optimum stochastic excitation code vector.

[0063] After the searching of the optimum stochastic excitation code vector is finished as described above, a searching of an optimum pulse-like excitation code vector is carried out. At this searching, the fixed codebook selection switch 326 is driven to the side of the pulse-like excitation codebook 322 the output from adaptive excitation codebook 326 is set to zero (0) or to the previously obtained optimum adaptive excitation code vector. In this state, the pulse-like excitation codebook 322 outputs sequentially or in parallel, all the stored pulse-like excitation code vectors. Processings thereafter will be substantially similar with those of the moment when an optimum stochastic excitation code vector is searched and, accordingly, more detailed explanation will not be necessary.

[0064] As described above, when the optimum pulse-like excitation code vector is determined, the code selection circuit 332 compares the error power sum of the selected code vector in the stochastic excitation code vector search with the error power sum of the selected code vector in the pulse-like excitation code vector search to obtain smallest error power sum, and determin a fixed code to be transmitted to the code excitation linear predictive decoder.

[0065] Thereafter, a searching of an optimum VQ gaincode is carried out. At the searching of this VQ gain code, an optimum (selected) adaptive excitation code vector is transmitted from the adaptive excitation codebook 320, and the fixed codebook selection switch 326 is switched to either the selected stochastic excitation codebook 321 or pulse-like excitation codebook 322, and an optimum (selected) fixed code vector is outputted from the selected fixed codebook 321 or 322. A VQ gain codebook 323 is composed of VQ gain for an adaptive excitation code vector and VQ gain for the fixed code vector. The VQ gain for the adaptive excitation code vector is given to a VQ gain controller 324 and the VQ gain for the fixed code vector is given to a VQ gain controller 327. Thus, both the VQ gain-controlled optimum adaptive excitation code vector and the optimum fixed code vector, which have been processed with respect to a frequency characteristic operation and VQ gain control, are added by an adder 325 and then given to a synthesis filter as an excitation code vector. This processing is carried out sequentially or in parallel, relative to all the VQ gain codes in the VQ gain codebook 323.

[0066] After an optimum adaptive excitation code, optimum fixed code and optimum VQ gain code are selected, the code selection circuit 332 gives the indexes of these codes to a multiplex circuit 303 and, a fixed codebook selection switching information which one of the stochastic excitation code vector and the pulse-like excitation code vector is selected actually, is given to the multiplex circuit 303. The multiplex circuit 303 multiplexes said indexes with LSP parameter given from the LSP parameter coding circuit 312 and transmits it to the code excitation linear predictive decoder. Incidentally, in case of utilizing a vector quantization for a VQ gain coding method, the transmitted index is vector number.

[0067] The coding processings described above is repeated with respect of each subframe, and the coded speech information is transmitted in turn to the code excitation linear predictive decoder.

[0068] Fig. 5 shows in detail the specific structure of the code vector conversion circuit 328. In Fig. 3, the code vector conversion circuit 328 has two cascaded filters 328a and 328b, and a pitch lag decision circuit 328c.

[0069] The fixed code vector is given to a first filter 328a. An impulse response H1(Z) of the first filter 328a is set as shown by formula (6), by which the frequency conversion processing is carried out relative to the fixed vector.

$$H1(Z)=(1-\Sigma A^{j}ajZ^{-j}) / (1-\Sigma B^{j}ajZ^{-j}) \qquad (6)$$

wherein aj(j is 1 to p) is a tap coefficient relative to a synthesis filter 329 which is supplied from the LPC conversion circuit 324, and p is vocal tract analysis order. Further, A and B are constants which are determined in the ranges of $0<A \leq 1$, and $0 < B \leq 1$.

[0070] The code vector which was processed in its frequency characteristics by the first filter 328a is transmitted to the second filter 328b. The pitch lag decision circuit 328c obtains a pitch lag L from the index of the optimum adaptive excitation code relative to the adaptive excitation codebook 320 and then gives the pitch lag L to the second filter 328b. An impulse response H2(Z) of the second filter 328b is determined as shown by formula (7), by which a frequency ' conversion is carried out relative to the inputted fixed code vector.

$$H2(Z)= 1/(1-\varepsilon Z^{-L}) \qquad (7)$$

wherein ε is a constant determined in the range of $0<\varepsilon\leq 1$. An output of the second filter 328b is given to VQ gain controller 327 shown in Fig. 3.

[0071] By the code vector conversion circuit 328 as described above, the frequency characteristics of inputted fixed code vector can be made closer to the frequency characteristics of the input speech vector, in accordance with a time length of the fixed code vector.

[0072] Accordingly, the code excited linear predictive coding apparatus (encoder) can provide a high quality regenerated speech signal.

[0073] Next, a code excitation linear predictive decoder in correspondence with the code excitation linear predictive coding apparatus (encoder) shown in Fig. 3 will be described with reference to the accompanying drawing.

[0074] Fig. 4 is a block diagram of code excitation linear predictive decoder which corresponds to the code excitation linear predictive coding apparatus (encoder) shown in Fig. 3. In Fig. 4, the code excitation linear predictive decoder has demultiplex circuit 440, LSP parameter decoding circuit 441, LPC conversion circuit 442, adaptive excitation codebook 443, stochastic excitation codebook 444, pulse-like excitation codebook 445, VQ gain codebook 446, VQ gain controller 447, VQ gain controller 449, fixed codebook selection switch 448, code vector conversion circuit 450, adder 451 and synthesis filter 452.

[0075] The coded speech information given from the code excitation linear predictive encoder is inputted to the demultiplex circuit 440. The demultiplex circuit 440 separates the coded speech information into LSP parameter code, index of the optimum adaptive excitation code, index of the optimum fixed code, index of the optimum VQ gain codebook and fixed code selection switch information.

[0076] Then, LSP parameter code is given to the LSP parameter decoding circuit 441 and the index of the optimum adaptive excitation code is given to the adaptive excitation codebook 443. Further, the index of optimum VQ gain code is given to the VQ gain codebook 446 and the fixed codebook selection switch information is given to the fixed codebook selection switch 448.

[0077] The index of the optimum fixed code 443 is given to a pulse-like excitation codebook 445 or a stochastic excitation codebook 444 which are determined by the fixed code selection switching information. The adaptive excitation codebook outputs an adaptive excitation code vector which is determined by a given index, and this adaptive excitation code vector is VQ gain-controlled through VQ gain controller 447 and given to an adder 451. Further, the adaptive excitation codebook 443 gives adaptive excitation code vector to a code vector conversion circuit 450.

[0078] The stochastic excitation codebook 444 or pulse-like excitation codebook 445 gives a stochastic excitation code vector or pulse-like excitation code vector, which corresponds to the given index, to a code vector conversion circuit 450 through'a fixed codebook selection switch 448.

[0079] The code vector conversion circuit 450 operates so that the frequency characteristics become closer to a frequency characteristics of the input speech vector in accordance with the index of the LPC and adaptive excitation code vector. A specific structure of the code vector conversion circuit 450 will be the same as that of the structure shown in Fig. 5. Thus, the frequency-processed fixed code vector is VQ gain-controlled by a VQ gain controller and then given to an adder 451.

[0080] The adder 451 adds the given adaptive excitation code vector and the fixed code vector together, and the added vector is assigned to be an excitation code vector, which is then given to a synthesis filter 452. The synthesis filter 452 outputs a synthesized speech vector.

[0081] The code excitation linear predictive decoder conducts the above-described processes every time when a decoded speech vector is given or, in other words, for each subframe.

[0082] Important features of the present invention are that the LSP parameter is used and transmitted as a vocal tract parameter; pulse-like excitation codebook is provided for giving an excitation source parameter; and a frequency characteristic of fixed code vector is controlled. These features can be independently provided to each of the coding apparatus and decoding apparatus without failure of the advantages and effects thereof.

[0083] In addition, the coding apparatus and decoding apparatus described above are related primarily to the forward-type code excitation linear predictive encoder and decoder, respectively, but the present invention is not limited thereto but applicable to backward-type code excitation linear predictive encoder and decoder, respectively.

[0084] The above-described encoder and decoder were intentionally designed under the technological basis for seeking to solve the problems induced from the low rate coding of 4-bit/s or less. However, more favorable sound reproduction can be realized if they are adapted to encoders and decoders of high rate coding. If the higher coding rate is allowable, both of the stochastic excitation codebook and pulse-like excitation codebook can be co-operated effectively rather than selectively operating either the stochastic excitation codebook or the pulse-like excitation codebook.

INDUSTRIAL APPLICABILITY

[0085] According to the present invention, it is considered that a frequency characteristic of actual excitation code vector is relatively close to that of an input speech vector and, in order to make it closer the frequency of the excitation code vector to a frequency of the input speech vector, the stochastic excitation code vector is convolutionaly computed with utilizing a specific impulse response. Thereafter, an adaptive excitation code vector is added to produce excitation code vector and, therefore, an excitation code vector which is well adaptive to an input speech vector by a small number of vector can be obtained and, at the same time, quantization error can be masked with conversion operation of an excitation code vector, thereby improving a reproduction quality.

[0086] Further, in addition to the adaptive excitation codebook and stochastic excitation codebook, pulse-like excitation codebook is disposed which stores therein pulse-like excitation code vector composed of unit impulse and, accordingly, a rapid tracking to a speech signal having periodicity can be realized, and a clear pulse-like excitation code vector can be formed at a steady portion of the speech signal.

[0087] Besides, since the pulse-like excitation code vector and the stochastic excitation code vector are switched over, the apparatus of the present invention can be adapted to low rate coding, and a favorably reproduced speech can be realized at the time , for example of a transitional period of the speech in which there are random signals and pulse-like signals together.

[0088] In addition, according to the code excitation linear coding apparatus and decoding apparatus, an excitation code vector is selected and used from either stochastic excitation codebook or pulse-like excitation codebook and, therefore, a favorable reproduction speech sound can be realized with the condition that the number of coded bit of the excitation source parameter is small.

[0089] Further, the vocal tract parameter for sound synthecization is used as ISP parameter which gives less distortion to the vocal tract vector than LPC when it is coded with a smaller number of code bit and, therefore, reproduction quality at a lower coding rate can be improved from a vocal tract parameter viewpoint.

Claims

1. A speech coding apparatus comprising:

an adaptively-renewable first codebook means (107) for selectively outputting a first signal;
a first gain controller means (110,113) for controlling a value of the first signal and outputting a second signal;
a second codebook means (108) for selectively outputting a third signal;
a signal conversion circuit means (109) for converting the third signal into a frequency characteristic and outputting a fourth signal;
a second gain controller means (110, 114) for controlling a value of the fourth signal and outputting a fifth signal;
an adder means (115) for adding the second signal and the fifth signal and thereby obtaining an excitation signal for use in speech synthesis; wherein
the first codebook means (107) is adaptively renewed on the basis of the excitation signal; and

the signal conversion circuit means (109) is arranged to generate an impulse response of a transfer function which is determined in accordance with pitch information relative to the first signal and to obtain the fourth signal by convolving the third signal with this impulse response.

2. A speech decoding apparatus comprising:

an adaptively-renewable first codebook means (204) for selectively outputting a first signal;
a first gain controller means (207, 208) for controlling a value of the first signal and outputting a second signal;
a second codebook means (205) for selectively outputting a third signal;
a signal conversion circuit means (206) for converting the third signal into a frequency characteristic and outputting a fourth signal;
a second gain controller means (207, 209) for controlling a value of the fourth signal and outputting a fifth signal;
an adder means (210) for adding the second signal and the fifth signal and thereby obtaining an excitation signal for use in speech synthesis; wherein
the first codebook means (204) is adaptively renewed on the basis of the excitation signal; and
the signal conversion circuit means (206) is arranged to generate an impulse response of a transfer function which is determined in accordance with pitch information relative to the first signal and to obtain the fourth signal by convolving the third signal with this impulse response.

3. A code excitation linear predictive coding apparatus which uses an excitation signal of an excitation codebook (108) as an excitation source information of a speech signal, the apparatus being **characterised in that** it comprises:-

a code vector conversion circuit means (109) for converting an excitation code vector selected from the excitation codebook (108) into a frequency characteristic which is determined at the time of output of said excitation code vector, said frequency characteristic serving as the input of a synthesis filter (105).

4. A code excitation linear predictive decoding apparatus which uses an excitation signal of an excitation codebook (205) as an excitation source information of a speech signal, the apparatus being **characterised in that** it comprises:-

a code vector conversion circuit means (206) for converting an excitation code vector selected from the excitation codebook (205) into a frequency characteristic which is determined at the time of output of said excitation code vector, said frequency characteristic serving as the input of a synthesis filter (203).

5. A coding or decoding apparatus according to claims 3 or 4, wherein the code vector conversion circuit means (109, 206) generates an impulse response of a transfer function which is determined in accordance with a vocal tract parameter of a speech signal input, and convolutionally computes the excitation code vector with the impulse response.

6. A coding or decoding apparatus according to claims 1, 2 or 5, wherein the impulse response of the transfer function is represented by:-

$$H(Z) = \frac{1 - \sum_{j=1}^{p} A^j a_j Z^{-j}}{1 - \sum_{j=1}^{p} B^j a_j Z^{-j}}$$

where a_j are linear predictive coefficients; p is a vocal tract analysis order; and A and B are within the range: 0 < A < 1 and 0 < B < 1.

7. A coding or decoding apparatus according to claims 3 or 4, wherein the code vector conversion circuit means (109, 206) generates an impulse response of a transfer function which is determined in accordance with an excited pitch lag, and convolutionally computes the excitation code vector with the impulse response.

8. A coding or decoding apparatus according to claims 1, 2 or 7, wherein the impulse response of the transfer function which is determined in accordance with the excited pitch lag is represented by:-

$$H(Z) = \frac{1}{1 - \varepsilon Z^{-L}}$$

where ε is a constant within the range $0 < \varepsilon \leq 1$; and L is pitch lag signal.

9. A coding or decoding apparatus according to claim 1 or 2, wherein the codebook vector conversion circuit means (109, 206) convolutionally computes the excitation code vector with the impulse response of the transfer function which is determined in accordance with transfer functions represented by:-

$$H(Z) = \frac{1 - \sum_{j=1}^{p} A^j a_j Z^{-j}}{1 - \sum_{j=1}^{p} B^j a_j Z^{-j}}$$

and

$$H(Z) = \frac{1}{1-\varepsilon Z^{-L}}$$

where a_j are linear predictive coefficients; p is a vocal tract analysis order, A, B and ε are within the range: $0 < A < 1$, $0 < B < 1$ and $0 < \varepsilon \leq 1$; and L is pitch lag signal.

10. A coding or decoding apparatus according to claim 4 or 4, wherein the impulse response of the transfer function is determined in accordance with transfer functions represented by:-

$$H(Z) = \frac{1 - \sum_{j=1}^{p} A^j a_j Z^{-j}}{1 - \sum_{j=1}^{p} B^j a_j Z^{-j}}$$

and

$$H(Z) = \frac{1}{1-\varepsilon Z^{-L}}$$

where a_j are linear predictive coefficients; p is a vocal tract analysis order, A, B and ε are within the range: $0 < A < 1$, $0 < B < 1$ and $0 < \varepsilon \leq 1$; and L is pitch lag signal.

11. A coding or decoding apparatus according to any one of claims 1 to 4 wherein said excitation code book or second codebook means is a pulse-like excitation code book (322).

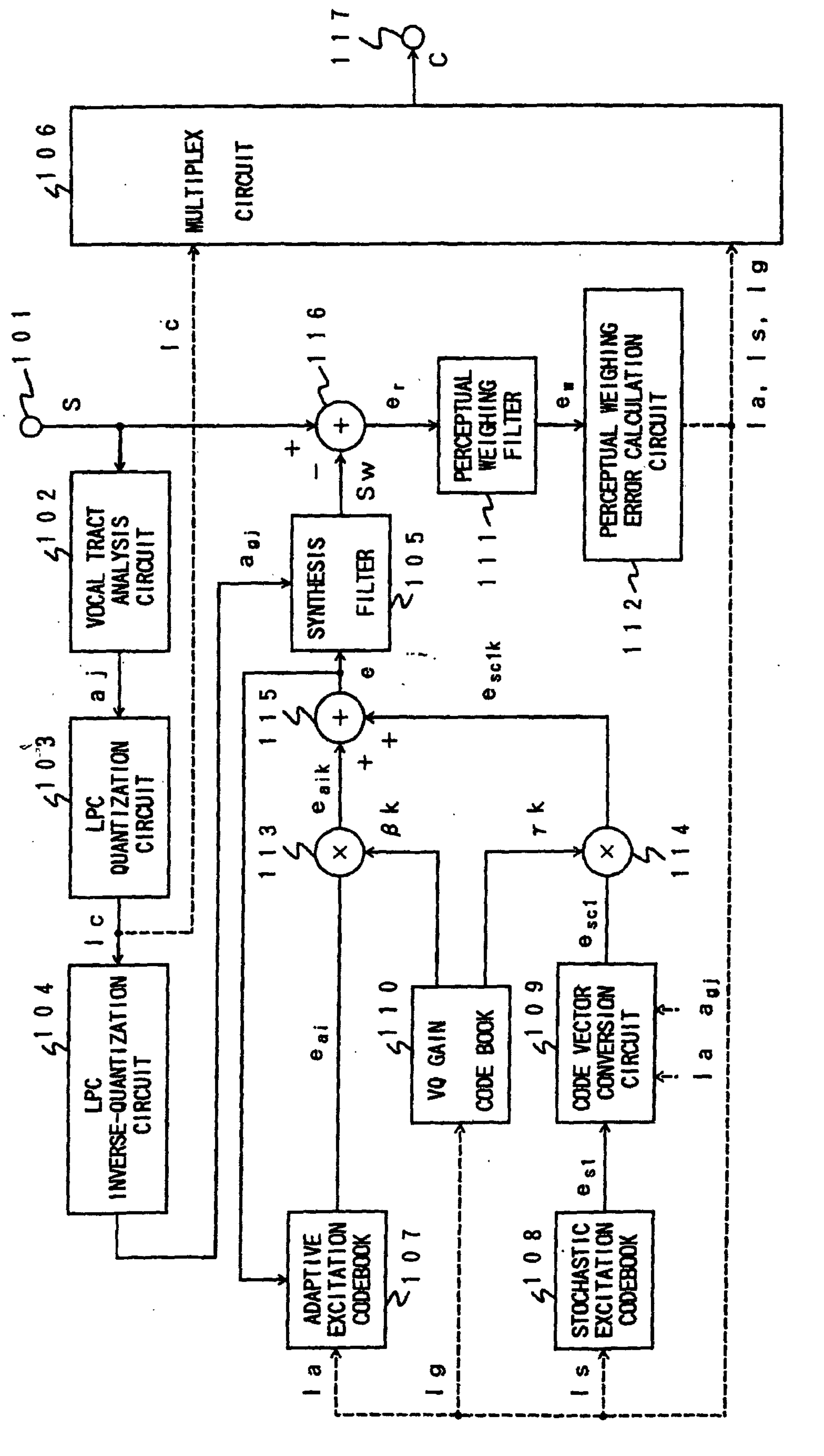

101
S
102
VOCAL TRACT ANALYSIS CIRCUIT
a j
103
LPC QUANTIZATION CIRCUIT
I c
104
LPC INVERSE-QUANTIZATION CIRCUIT
106
MULTIPLEX CIRCUIT
117
C
116
e r
Sw
111
PERCEPTUAL WEIGHING FILTER
e w
112
PERCEPTUAL WEIGHING ERROR CALCULATION CIRCUIT
I a, I s, I g
105
SYNTHESIS FILTER
a qj
e
115
e sclk
e aik
113
114
β k
γ k
110
VQ GAIN CODE BOOK
e ai
109
CODE VECTOR CONVERSION CIRCUIT
e scl
I a
a qj
107
ADAPTIVE EXCITATION CODEBOOK
108
STOCHASTIC EXCITATION CODEBOOK
e sl
I a
I g
I s

Fig. 1

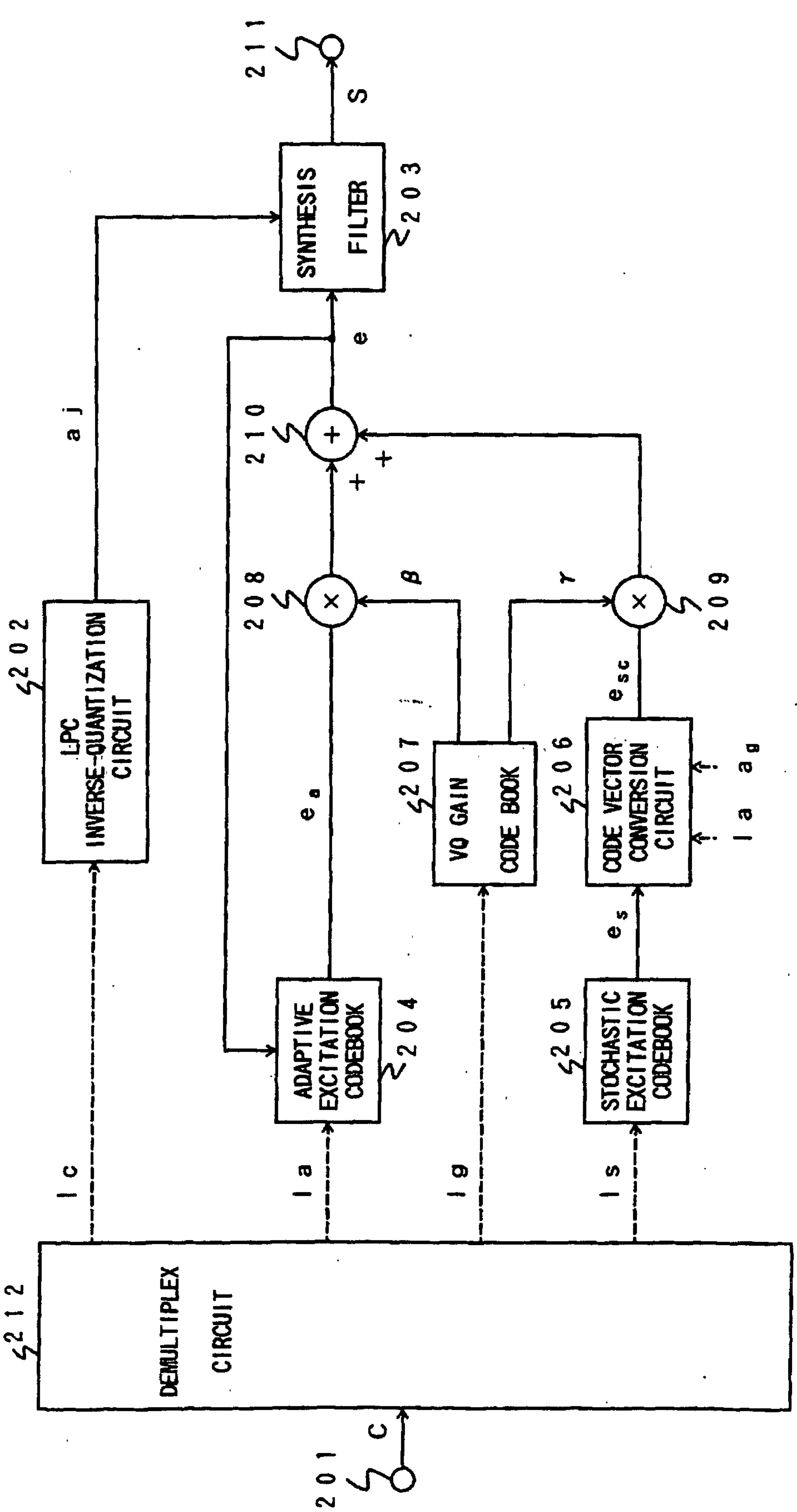

DEMULTIPLEX CIRCUIT
212
201
C
Ic
Ia
Ig
Is
LPC INVERSE-QUANTIZATION CIRCUIT
202
aj
ADAPTIVE EXCITATION CODEBOOK
204
STOCHASTIC EXCITATION CODEBOOK
205
VQ GAIN CODE BOOK
207
CODE VECTOR CONVERSION CIRCUIT
206
Ia
ag
ea
es
esc
β
γ
208
209
210
+
+
e
SYNTHESIS FILTER
203
S
211

Fig. 2

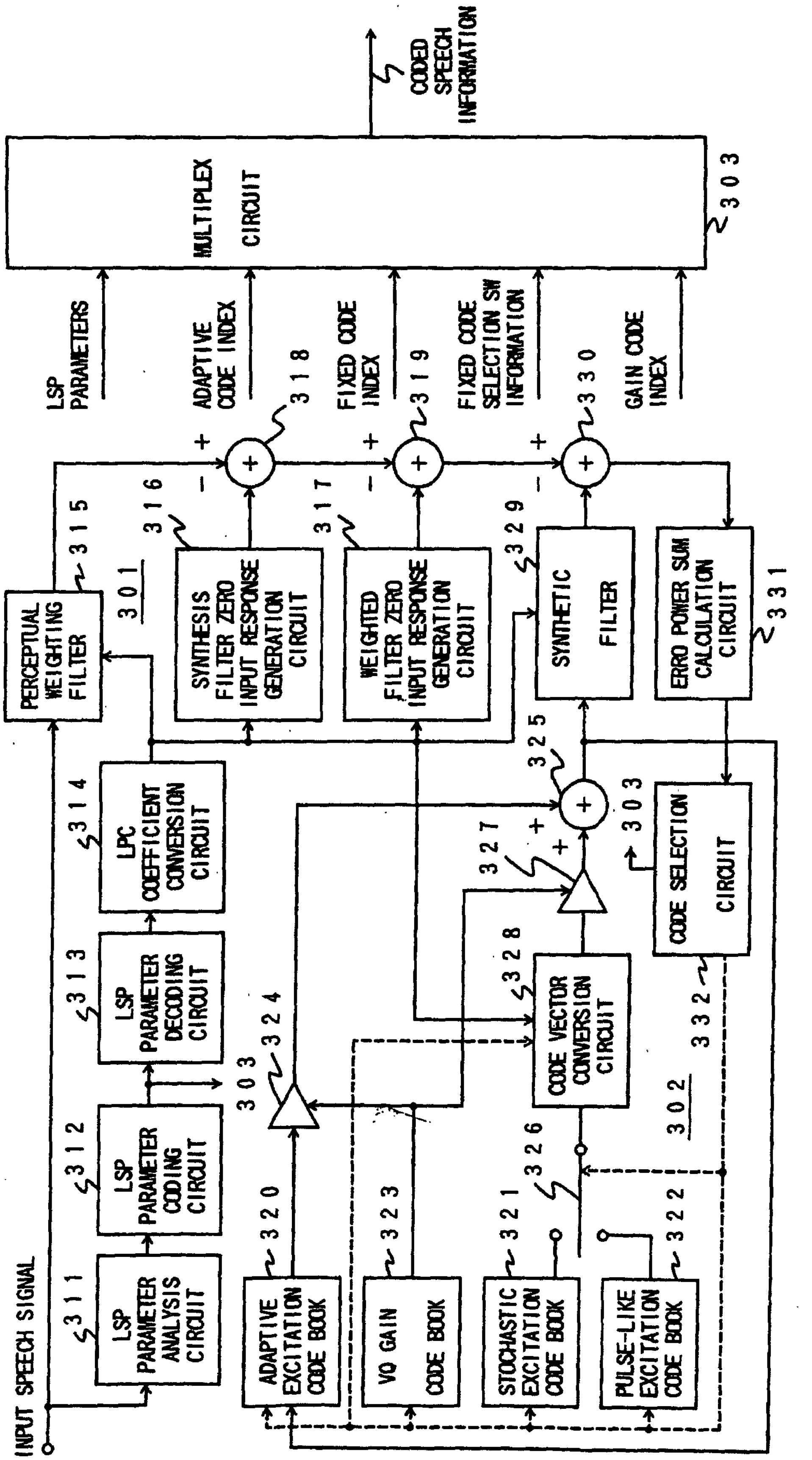
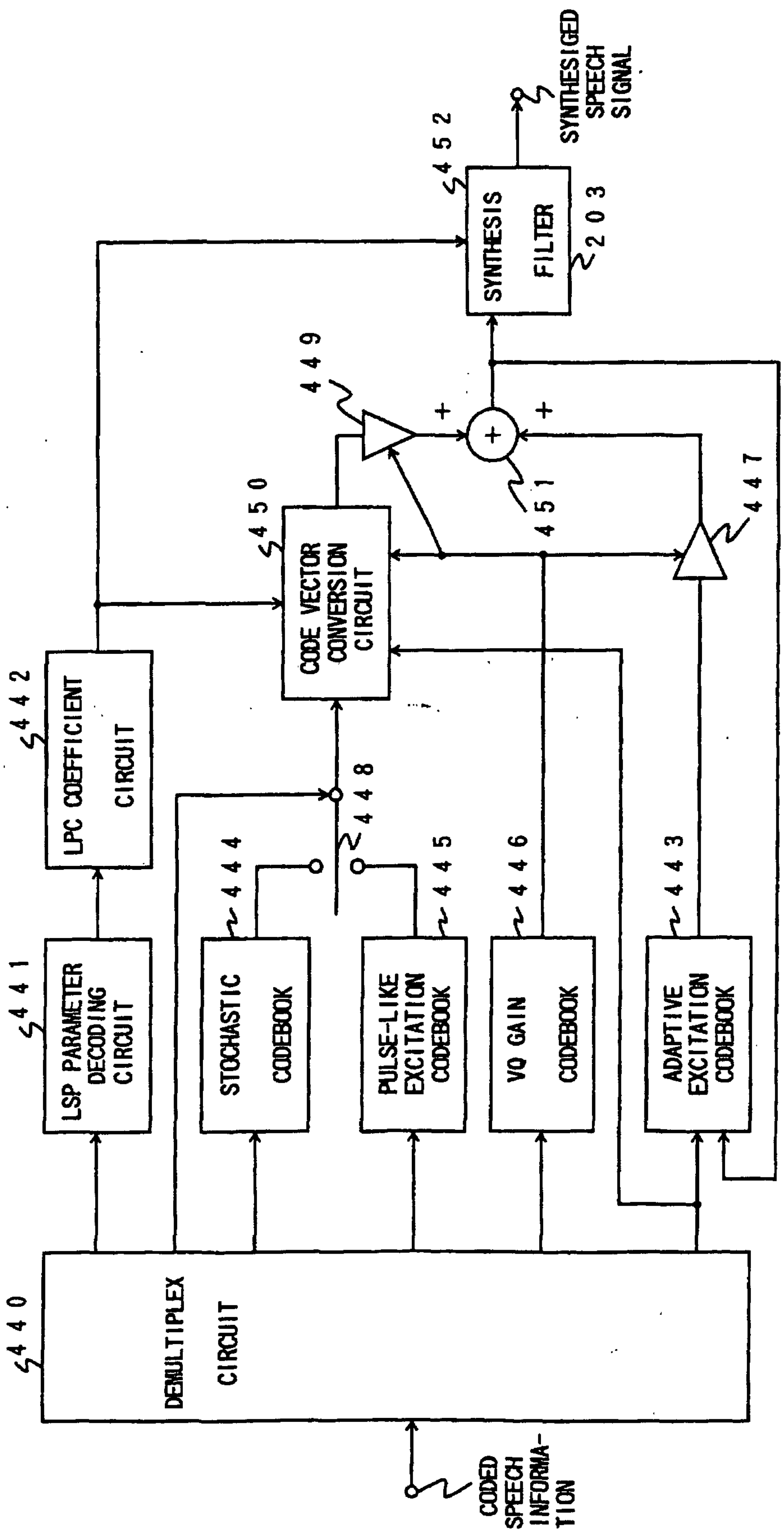

INPUT SPEECH SIGNAL
LSP PARAMETER ANALYSIS CIRCUIT
311
LSP PARAMETER CODING CIRCUIT
312
LSP PARAMETER DECODING CIRCUIT
313
LPC COEFFICIENT CONVERSION CIRCUIT
314
PERCEPTUAL WEIGHTING FILTER
315
301
SYNTHESIS FILTER ZERO INPUT RESPONSE GENERATION CIRCUIT
316
WEIGHTED FILTER ZERO INPUT RESPONSE GENERATION CIRCUIT
317
SYNTHETIC FILTER
329
ERRO POWER SUM CALCULATION CIRCUIT
331
CODE SELECTION CIRCUIT
332
303
ADAPTIVE EXCITATION CODE BOOK
320
VQ GAIN CODE BOOK
323
STOCHASTIC EXCITATION CODE BOOK
321
PULSE-LIKE EXCITATION CODE BOOK
322
326
CODE VECTOR CONVERSION CIRCUIT
328
327
324
325
302
318
319
330
MULTIPLEX CIRCUIT
303
LSP PARAMETERS
ADAPTIVE CODE INDEX
FIXED CODE INDEX
FIXED CODE SELECTION SW INFORMATION
GAIN CODE INDEX
CODED SPEECH INFORMATION

Fig. 3



CODED SPEECH INFORMATION
440
DEMULTIPLEX CIRCUIT
441
LSP PARAMETER DECODING CIRCUIT
442
LPC COEFFICIENT CIRCUIT
444
STOCHASTIC CODEBOOK
445
PULSE-LIKE EXCITATION CODEBOOK
446
VQ GAIN CODEBOOK
443
ADAPTIVE EXCITATION CODEBOOK
448
450
CODE VECTOR CONVERSION CIRCUIT
449
447
451
+
+
452
SYNTHESIS FILTER
203
SYNTHESIGED SPEECH SIGNAL

Fig. 4

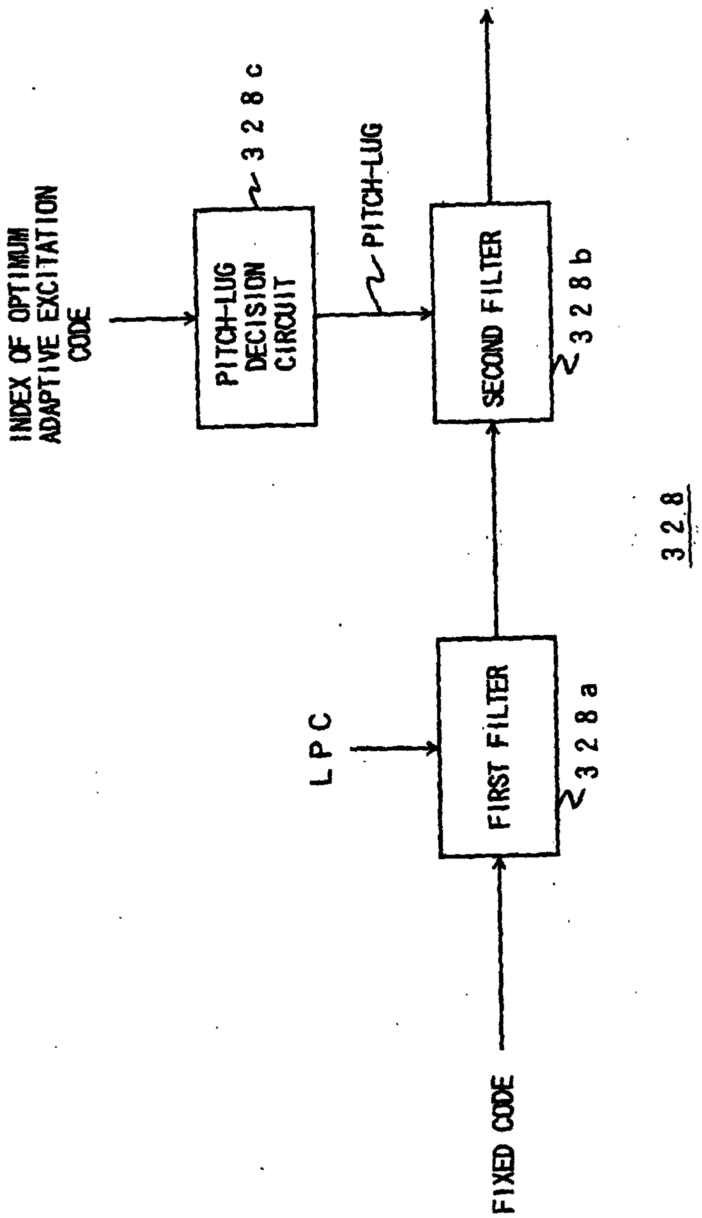

INDEX OF OPTIMUM ADAPTIVE EXCITATION CODE
PITCH-LUG DECISION CIRCUIT
328c
PITCH-LUG
SECOND FILTER
328b
LPC
FIRST FILTER
328a
FIXED CODE
328

Fig. 5