

SPECIFICATION

TITLE OF THE INVENTION

VECTOR CODING METHOD, ENCODER USING THE SAME AND DECODER
5 THEREFOR

BACKGROUND OF THE INVENTION

The present invention relates to a vector coding
method that is used to encode speech, images and various
10 other pieces of information and is particularly suited to
encoding of information that is transmitted over an error-
prone channel such as a mobile radio channel and encodes
an input vector through the use of a plurality of
codebooks each composed of plural representative vectors.
15 The invention also pertains to a vector encoder using the
above-mentioned vector coding method and a vector decoder
for decoding codes encoded by the vector encoder.

Methods that have been proposed to transmit vectors
over channels prone to channel errors are to set
20 representative vectors in anticipation of possible channel
errors and to take into account the channel errors when
labeling representative vectors. These methods are
disclosed in Kumazawa, Kasahara and Namekawa, "A
Communication of Vector Quantizers for Noisy Channels,"
25 Transactions of the Institute of Electronics, Information
and Communication Engineers of Japan , Vol. J67-B, No. 1,
pp. 1-8, 1984, Zeger and Gersho, "Pseudo-Gray Coding,"
IEEE Trans. on Comm., Vol. 38, No. 12, pp. 2147-2158, 1990,
and other literature. These methods hold all
30 representative vectors directly in one codebook, and hence
require large storage capacity for storing the codebook.

As a method that does not need large storage capacity
in transmitting vectors over channels prone to channel

errors, there has been proposed to transmit vectors after quantizing them through the use of two structured codebooks. This is disclosed in Moriya, "Two-Channel Vector Quantizer Applied to Speech coding," Transactions of the Institute of Electronics, Information and Communication Engineers of Japan, IT87-106, pp. 25-30, 1987 and other literature. This method has two small-scale codebooks and uses two representative vectors in combination to reduce the storage capacity needed and transmits two labels indicative of the two representative vectors to lessen the influence of channel errors. This method will be described with reference to Figs. 1A and 1B. The representative vectors of the codebooks are generated beforehand by learning, for instance. In an encoder depicted in Fig. 1A, one representative vector z_{1i} is fetched from a codebook CB1 and one representative vector z_{2j} from a codebook CB2, then they are added together in a vector combining part 3 to generate a vector sum $y_{ij}=z_{1i}+z_{2j}$, and the distance, $d(X, y_{ij})$, between the combined representative vector y_{ij} and an input vector X via an input terminal 4 is calculated, as distortion, in a distance calculating part 5. A control part 6 controls representative vector select switches 7 and 8 for the codebooks CB1 and CB2 and searches them for the representative vectors z_{1i} and z_{2j} that minimize the output $d(X, y_{ij})$ from the distance calculating part 5. The control part 6 provides, as an encoded outputs to an output terminal 9, labels i and j of the representative vectors z_{1i} and z_{2j} that provides minimum distance.

30 In a decoder shown in Fig. 1B, the control part 6 controls representative vector select switches 13 and 14 in accordance with the labels i and j in the input code via an input terminal 11 and reads out representative

vectors z_{1i} and z_{2j} from codebooks CB3 and CB4, respectively. The thus read-out representative vectors z_{1i} and z_{2j} are combined in a vector combining part 17 into a reconstructed vector $y_{ij}=z_{1i}+z_{2j}$, which is provided to an output terminal 18. Incidentally, the codebooks CB3 and CB4 are identical with those CB1 and CB2, respectively.

The method described above in respect of Figs. 1A and 1B saves the storage capacity of the codebooks for storing the representative vectors and lessens the influence of channel errors by combining the vectors in the decoder through utilization of the two labels corresponding thereto.

With this method, however, if an error arises in the labels during transmission over the channel, then distortion will occur in all elements of the received vector. According to the circumstances, the error will cause an abnormally large amount of distortion in the decoded output.

Another problem of this method is that the amount of processing required is very large because it involves the calculation of the distance $d(X, y_{ij})$ for every combination of representative vectors of the two codebooks in search for the pair of representative vectors that minimizes the distance.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a vector coding method that prevents an error in the input code to the decoder from causing serious distortion of its output.

Another object of the present invention is to provide a vector coding method that prevents an error in the input code to the decoder from causing serious distortion of its

output and permits reduction of the amount of processing required.

Another object of the present invention is to provide a vector encoder that embodies the above-mentioned vector
5 coding method.

Still another object of the present invention is to provide a vector decoder that decodes a vector encoded by the vector coding method that serves the above-mentioned objects.

10 According to the vector coding method and the encoder of the first aspect of the present invention, representative vectors from respective codebooks are combined and the distance between the combined
representative vector and the input vector is calculated;
15 in this instance, the representative vectors to be combined are those multiplied by predetermined different weighting coefficient vectors, each of which is composed of the same number of components. At least one of the
20 components in each weighting coefficient vector assumes a maximum value, and the positions of the maximum components in the respective weighting coefficient vectors differ with the codebooks. The multiplication of each
representative vector by the weighting coefficient vector is done on the representative vector read out of each
25 codebook, or weighted representative vectors respectively premultiplied by the weighting coefficient vectors are prestored in each codebook.

According to the vector coding method and the vector encoder of a second aspect of the present invention, in
30 the first aspect, the distribution of the set of weighted representative vectors multiplied by the weighting coefficient vectors for each codebook are approximated with straight lines, then the input vector is projected on

each straight line and a plurality of weighted representative vectors present around the projection are chosen for each codebook. The thus chosen weighted representative vectors of the respective codebooks are
5 combined in pairs and that one of the combined vectors which has the minimum distance to the input vector is selected as the combined representative vector.

In a third aspect, the present invention relates to a decoder for codes encoded into vectors according to the
10 first or second aspect. The representative vectors read out of respective codebooks are multiplied by weighting coefficient vectors by multiplying means and the multiplied weighted representative vectors are combined into a reconstructed vector; the weighting coefficient
15 vectors are selected in the same fashion as in the first aspect of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A is a block diagram showing the configuration of an encoder embodying a conventional vector coding
20 method;

Fig. 1B is a block diagram showing a conventional decoder for use with the encoder of Fig. 1A;

Fig. 2A is a block diagram illustrating an example of the encoder embodying the present invention;
25

Fig. 2B is a block diagram illustrating an embodiment of the decoder according to the present invention;

Fig. 3A is a graph showing each representative vector z_{1i} of a codebook CB1;

30 Fig. 3B is a graph showing a vector obtained by multiplying each representative vector z_{1i} by a weighting coefficient vector w_1 ;

Fig. 3C is a graph showing each representative vector z_{2j} of a codebook CB2;

Fig. 3D is a graph showing a vector obtained by multiplying each representative vector z_{2j} by a weighting
5 coefficient vector w_2 ;

Fig. 3E is a graph showing examples of a combined vector and an erroneous combined vector in the present invention;

Fig. 3F is a graph showing examples of a combined
10 vector and an erroneous combined vector in the prior art;

Fig. 4 is a graph showing a set of weighted representative vectors and a straight line for approximation, for explaining the vector coding method of the present invention;

Fig. 5 is a block diagram illustrating an example of
15 the encoder of the present invention applied to the CELP scheme;

Fig. 6A is a table showing a first codebook with weighted gain vectors stored therein;

Fig. 6B is a table showing a second codebook with
20 weighted gain vector stored therein;

Fig. 7 is a graph showing weighted gain vectors of Figs. 6A and 6B on a coordinate system;

Fig. 8 is a block diagram illustrating the encoder of
25 the present invention applied to the quantization of random excitation vectors;

Fig. 9 is a block diagram illustrating the encoder of the present invention applied to the quantization of pitch excitation vectors;

Fig. 10 is a block diagram illustrating the encoder
30 of the present invention applied to the VSELP scheme;

Fig. 11 is a graph showing segmental SN ratio of reconstructed speech with respect to a channel error rate

in the cases of using one gain codebook and two gain codebooks for vector encoding of gains g_p and g_c in Fig. 5;

Fig. 12 is a graph showing, as an equivalent Q value, the MOS of the reconstructed speech with respect to the channel error rate in the two cases of Fig. 11; and

Fig. 13 is a graph showing the segmental SN ratio of reconstructed speech with respect to the channel-error rate in the case of the present invention that uses two weighted gain codebooks for vector encoding of the gains g_p and g_c in Fig. 5.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

In Figs. 2A and 2B there are illustrated in block form an embodiment of the present invention, in which the parts corresponding to those in Figs. 1A and 1B are identified by the same reference numerals. In the encoder of Fig. 2A, multipliers 21 and 22 are provided between the representative selection switches 7, 8 and the vector combining part 3, by which components of L-dimensional (where L is an integer equal to or greater than 2) representative vectors $z_{1i}=(z_{1i1}, z_{1i2}, \dots, z_{1iL})$ and $z_{2j}=(z_{2j1}, z_{2j2}, \dots, z_{2jL})$ selected from the codebooks CB1 and CB2 are multiplied by the corresponding components of L-dimensional weighting coefficient vectors $w_1=(w_{11}, w_{12}, \dots, w_{1L})$ and $w_2=(w_{21}, w_{22}, \dots, w_{2L})$, respectively. At least one of the L components forming each of the weighting coefficient vectors w_1 and w_2 assumes a maximum value and the positions of the maximum components in the respective weighting coefficient vectors w_1 and w_2 differ with the codebooks CB1 and CB2. According to the present invention, letting the weighting coefficient vectors w_1 and w_2 be represented by the following weighting

coefficient matrixes W_1 and W_2 that have, as diagonal elements, the values $w_{11}, w_{12}, \dots, w_{1L}$ of the components by which the respective components of the representative vectors are multiplied:

5

$$W_1 = \begin{bmatrix} w_{11} & 0 & \dots & \dots & 0 \\ 0 & w_{12} & 0 & \dots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & w_{1L} \end{bmatrix} \quad (1)$$

$$W_2 = \begin{bmatrix} w_{21} & 0 & \dots & \dots & 0 \\ 0 & w_{22} & 0 & \dots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & w_{2L} \end{bmatrix} \quad (2)$$

10

the weighting coefficient vectors w_1 and w_2 may preferably be selected so that the sum of the weighting coefficient matrixes W_1 and W_2 of the codebooks CB1 and CB2 becomes a constant multiple of the unit matrix as follows:

$$\begin{bmatrix} w_{11} & 0 & \dots & \dots & 0 \\ 0 & w_{12} & 0 & \dots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & w_{1L} \end{bmatrix} + \begin{bmatrix} w_{21} & 0 & \dots & \dots & 0 \\ 0 & w_{22} & 0 & \dots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & w_{2L} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \times K \quad (3)$$

where K is a predetermined constant. Vectors $w_1 z_{1i}$ and $w_2 z_{2j}$, obtained by multiplying the representative vectors z_{1i} and z_{2j} by the weighting coefficient vectors w_1 and w_2 , respectively, are combined in the vector combining part 3, and the codebooks CB1 and CB2 are searched for representative vectors z_{1i} and z_{2j} that minimize the distance between the combined vector y_{ij} and the input vector X .

With such a configuration as described above, for example, when $L=2$, the representative vectors z_{1i} and z_{2j} are expressed by two-dimensional vectors $z_{1i}=(z_{1i1}, z_{1i2})$ and $z_{2j}=(z_{2j1}, z_{2j2})$, respectively. Suppose that $k=2$ and that the weighting coefficients which satisfy Eq. (3) are $w_1=(w_{11}=1.8, w_{12}=0.2)$ and $w_2=(w_{21}=0.2, w_{22}=1.8)$. Assuming that representative vectors $z_{11}, z_{12}, \dots$ of the codebook CB1 are distributed substantially uniformly all over a plane of a certain two-dimensional range defined by z_{1i1} in the first-dimensional direction and z_{1i2} in the second-dimensional direction as shown in Fig. 3A, weighted representative vectors z_{11}' and $z_{12}', \dots$, obtained by multiplying each representative vector $z_{1i}=(z_{1i1}, z_{1i2})$ by the weighting coefficient vector $w_1=(1.8, 0.2)$, are concentrated close to the first-dimensional axis as shown in Fig. 3B. Similarly, assuming that representative vectors $z_{21}, z_{22}, \dots$ of the codebook CB2 are distributed substantially uniformly all over the plane of a certain two-dimensional range defined by two axes as depicted in Fig. 3C, weighted representative vectors, obtained by multiplying the representative vectors $z_{21}, z_{22}, \dots$ by the weighting coefficient vectors $w_2=(0.2, 1.8)$, are concentrated close to the second-dimensional axis as shown in Fig. 3D.

Suppose, for example, that when it is judged at the transmitting side that the weighted combined vector y_{ij} of the representative vectors z_{21} and z_{1i} has the minimum distortion with respect to the input signal X , the label of the one weighted representative vector z_{1i} becomes z_{1i}' because of a channel error as shown in Fig. 3E. In this instance, the combined vector y_{ij} changes to y_{ij}' at the receiving side. There is a possibility that the weighted representative vector z_{1i} changes to any other weighted representative vectors z_{1i}' , but since the vector z_{1i} has a biased distribution, the value of the second-dimensional component of an error vector, $\Delta y = \Delta y_{ij} - y_{ij}'$, between the combined vectors y_{ij} and y_{ij}' is relatively small, no matter how much the vector z_{1i} may change. In contrast to this, in the case where the combined vector is not multiplied by the weighting coefficient, if the one representative vector z_{1i} changes to a representative vector z_{1i}' , combined vectors of these vectors z_{1i} and z_{1i}' and the other representative vector z_{2j} become y_{ij} and y_{ij}' , respectively, as shown in Fig. 3F. Since there is a likelihood of the representative vector z_{1i} changing to any of the representative vectors of the codebook CB1 and since the representative vectors $z_{11}, z_{12}, \dots$ are distributed over a wide range, the error vector Δy between the combined vector y_{ij} and the changed combined vector y_{ij}' is likely to have appreciably large first- and second-dimensional components.

In other words, in the example of Fig. 3E, when the weighted representative vector $w_1 z_{1i} = (w_{11} z_{1i1}, w_{12} z_{1i2})$ becomes $w_1 z_{1i}' = (w_{11} z_{1i1}', w_{12} z_{1i1}')$ because of a channel error, distortion is concentrated on the first-dimensional component $w_{11} z_{1i1}'$ to keep down distortion of the second-

dimensional component $w_{12}z_{1i}z_{2j}$, by which distortion is reduced as a whole.

Fig. 2B illustrates in block form an embodiment of the decoder of the present invention, which is supplied with the labels i and j and the weighting coefficient vectors w_1 and w_2 from the encoder of Fig. 2A and decodes the code y_{ij} . The decoder has the same codebooks CB3 and CB4 as those CB1 and CB2 in Fig. 2A, read out representative vectors of the labels i and j of the inputted code from the codebooks CB3 and CB4 and combines them as is the case with Fig. 1B. In this embodiment, multipliers 24 and 25 are provided between the switches 13, 14 and the vector combining part 17, by which representative vectors z_{1i} and z_{2j} read out of the codebooks CB3 and CB4 are multiplied by the same weighting coefficient vectors w_1 and w_2 as those used by the corresponding multipliers 21 and 22 of the Fig. 2A encoder. The thus multiplied representative vectors w_1z_{1i} and w_2z_{2j} are combined in the vector combining part 17 into the reconstructed vector y_{ij} . As will be evident from the above, it is also possible to omit the multipliers 21, 22, 24 and 25 by prestoring in the codebooks CB1, CB3 and CB2, CB4 in Figs. 2A and 2B weighted representative vectors obtained by multiplying the representative vectors z_{1i} and z_{2j} by the weighting coefficient vectors w_1 and w_2 , respectively.

In the encoder of Fig. 2A, the combined vector y_{ij} is determined for every combination of representative vectors z_{1i} and z_{2j} prestored in the codebooks CB1 and CB2, then the distortion of each combined vector is calculated with respect to the input signal vector X , and a decision is made of which pair of representative vectors z_{1i} and z_{2j} provides the minimum distortion. With this method,

however, the number of calculations increases sharply as the codebooks CB1 and CB2 increase in size. Next, a description will be given of a scheme which pre-selects small numbers of representative vectors z_{1i} and z_{2j} and determines the pair of representative vectors of the minimum distortion among them, thereby reducing the computational complexity and hence shortening the operation time.

Let it be assumed, for example, that the codebooks CB1 and CB2 in Fig. 2A have 8 and 16 representative vectors, respectively, and that the vectors z_{1i} and z_{2j} are all two-dimensional. In Fig. 4, eight weighted representative vectors, obtained by multiplying the eight representative vectors z_{1i} of the codebook CB1 by the weighting coefficient vector $w_1=(w_{11}=1.8, w_{12}=0.2)$, are indicated by crosses; 16 weighted representative vectors, similarly obtained by multiplying the 16 representative vectors z_{2j} of the codebook CB2 by the weighting coefficient vectors $w_2=(w_{21}=0.2, w_{22}=1.8)$, are indicated by white circles. The input signal vector is indicated by X , which is composed of a predetermined number of signal samples of each frame, two samples in this example.

The two-dimensional weighting coefficient vectors w_1 and w_2 for the two-dimensional vectors z_{1i} and z_{2j} are determined such that they satisfy Eq. (3); $w_{11}+w_{21}=w_{12}+w_{22}=2$ in this example. As depicted in Fig. 4, the weighted representative vectors marked with the white circles and the weighted representative vectors marked with the crosses are distributed separately on the sides opposite with respect to a straight line of a 45°-gradient and passing through the origin (0, 0). The following description will be given on the assumption that there are stored such weighted representative vectors in the

codebooks CB1 and CB2 in Fig. 2A and in the CB3 and CB4 in Fig. 2B, with the multipliers 21, 22, 24 and 25 left out.

With this scheme, the set of weighted representative vectors (indicated by the crosses) of the codebook CB1 is approximated with a straight line 27. That is, the straight line 27 is determined so that the sum of distances, $D_{11}, D_{12}, \dots, D_{18}$, between it and the respective crosses (or distances in the second-dimensional axis direction) is minimum. Likewise, the set of weighted representative vectors (indicated by the white circles) of the codebook CB2 is approximated with a straight line 28. The straight line 28 is also determined so that the sum of distances, $D_{21}, D_{22}, \dots, D_{216}$, between it and the respective white circles (or distances in the first-dimensional axis direction) is minimum.

The input vector X is projected on the approximating straight lines 27 and 28 and pluralities of weighted representative vectors present around the projections are selected. That is, a calculation is made of the value on the abscissa and hence a first-dimensional value p_{1x} at the intersection P_1 of a straight line 29 passing through the input signal vector X and parallel to the approximating straight line 28 and the approximating straight line 27, then the value P_{1x} and first-dimensional values (values of first components) of the cross-marked weighted representative vectors having a wide first-dimensional distribution are compared, and a predetermined number of, for example, three, weighted representative vectors are selected, as a subgroup H_1 , in increasing order of the difference between the value p_{1x} and the first-dimensional value of the respective weighted representative vector. In this way, the weighted representative vectors are pre-selected for the codebook

CB1. Similarly, a calculation is made of a value on the ordinate and hence a second-dimensional value p_{2y} at the intersection P_2 of a straight line 31 passing through the input signal vector X and parallel to the approximating straight line 27 and the approximating straight line 28, then the second-dimensional value p_{2y} and second-dimensional values (values of second components) of the white-circled weighted representative vectors having a wide second-dimensional distribution, and a predetermined number of, for example, three, weighted representative vectors are selected, as a subgroup H_2 , in the order of increasing differences between the value p_{2y} and the second-dimensional values of the weighted representative vectors. This is the pre-selection of the weighted representative vectors for the codebook CB2.

Only the weighted representative vectors thus pre-selected from the codebooks CB1 and CB2 are searched for a pair of weighted representative vectors that provides the minimum distance between their combined vector and the input signal vector. In this example, since three weighted representative vectors are pre-selected from each of the codebooks CB1 and CB2, the number of their combinations is nine, and hence the number of combined vectors is nine. When the pre-selection scheme is not adopted, the number of combinations of the weighted representative vectors (the number of combined representative vectors) is $8 \times 16 = 128$, and when the pre-selection is made, the number of calculations for the distance to the input signal vector X is reduced down to $9/128$ of the number of calculations needed when no pre-selection takes place. In this pre-selection scheme, when M codebooks are used, the number of dimensions of representative vectors is also set to M . The M weighting

coefficient vectors each have at least one maximum component at a different component position (that is, in a different dimension), and by multiplying the representative vector by the weighting coefficient vector, that dimension is emphasized more than the other dimensions.

Fig. 5 illustrates an embodiment of the coding method of the present invention applied to speech coding of the CELP (Code-Excited Linear Prediction Coding) system. In the CELP system, as disclosed by M.R. Schroeder and B.S. Atal in "Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates", Proc. ICASSP'85, pp. 937-940, 1985, for instance, pitch excitation vectors read out of a pitch excitation source codebook and random excitation vectors read out of a random excitation source codebook are respectively provided with gains and are combined in pairs, then the combined vectors are each fed as an excitation signal to a synthesis filter to obtain synthesized speech, then two vectors and two gains therefor are determined which minimize distortion of the synthesized speech with respect to input speech, and labels of these vectors and labels of the gains are outputted as encoded results of the input speech, together with the filter coefficients of the synthesis filter. By applying the vector coding method of the present invention to the encoding of the gains for the two vectors in the CELP system, it is possible to prevent the occurrence of gross distortion in the decoded or reconstructed speech by a channel error in the codes representing the gains.

The input speech signal X fed via an input terminal 34 is sampled with a fixed period and is expressed as a vector that is provided as a sequence of digital sample values for each frame period. The input signal vector X

of every frame is subjected to, for example, an LPC analysis in a filter-coefficient determining part 35, from which linear predictive coefficients are provided. The linear predictive coefficients are used to calculate a spectrum envelope parameter, which is quantized in a filter coefficient quantizing part 36, and the quantized value is set as the filter coefficient of a synthesis filter 37. In a pitch excitation source codebook 39 there are stored sequences of sample values of waveforms respectively containing different pitch period components and labeled as pitch excitation vectors. In a random-excitation source codebook 43 there are stored sequences of sample values of various random waveforms respectively labeled as random-excitation vectors. The pitch-excitation vectors and the random-excitation vectors stored in the pitch-excitation source codebook 39 and the random-excitation source codebook 43 are each composed of components of the same number as that of samples of one frame. A selection switch 38 is controlled by the control part 6 to select one of the pitch-excitation vectors in the pitch-excitation source codebook 39, and the selected pitch-excitation vector is multiplied by a given gain in a gain providing part 41, thereafter being applied to the synthesis filter 37. The difference between synthesized speech X_p from the synthesis filter 37 and the input speech signal X is calculated by a subtractor 48, and in the distortion calculating part 5 the difference is used to calculate distortion D as $D=||X-X_p||^2$. Similarly, the other pitch-excitation vectors are sequentially taken out from the pitch-excitation source codebook 39 via the switch 38 under the control of the control part 6, then the above-mentioned distortion is calculated for each pitch-excitation vector, and the pitch-excitation vector

of the minimum distortion is determined. Next, one of the random-excitation vectors stored in the random-excitation source codebook 43 is taken out through a switch 42 and amplified by a given gain in a gain providing part 46, thereafter being fed to an adder 47 wherein it is combined to the already determined pitch-excitation vector into an excitation signal vector E. The excitation signal vector E is provided to the synthesis filter 37 to generate synthesized speech and its distortion with respect to the input speech signal is similarly calculated. Likewise, such distortion is also calculated for each of the other remaining random-excitation vectors in the random-excitation source codebook 43 and the random excitation vector of the minimum distortion is determined.

After the selection of the pitch excitation vector and the random excitation vector as mentioned above, gains g_p and g_c of the gain providing parts 41 and 46 are so determined as to minimize the distortion as described hereafter. In gain codebooks CB1 and CB2 there are stored gain vectors z_{1i} (where $i=1, \dots, a$) and z_{2j} (where $j=1, \dots, b$), respectively. The gain vectors z_{1i} and z_{2j} are each composed of two components and expressed as $z_{1i}=(z_{1i1}, z_{1i2})$ and $z_{2j}=(z_{2j1}, z_{2j2})$, respectively. The gain vectors z_{1i} and z_{2j} taken out of the gain codebooks CB1 and CB2 are multiplied by weighting coefficient vectors $w_1=(w_{11}, w_{12})$ and $w_2=(w_{21}, w_{22})$ by multipliers 21 and 22, respectively, from which weighted gain vectors $y_i=(y_{i1}, y_{i2})$ and $y_j=(y_{j1}, y_{j2})$ are provided. Here,

$$y_{i1}=z_{1i1}w_{11}, y_{i2}=z_{1i2}w_{12}, y_{j1}=z_{2j1}w_{21}, y_{j2}=z_{2j2}w_{22} \quad (4)$$

In a vector combining part 3 the weighted gain vectors y_i and y_j are combined into a combined gain vector $G=(g_p, g_c)$ which is composed of such components as follows:

$$g_p=y_{i1}+y_{j1}, g_c=y_{i2}+y_{j2} \quad (5)$$

The first and second components g_p and g_c of the combined gain vector G are provided as first and second gains to the gain providing parts 41 and 46, wherein they are used to multiply the pitch excitation vector C_p and the random excitation vector C_R from the pitch excitation source codebook 39 and the random excitation source codebook 46, respectively.

The pitch excitation vector $g_p C_p$ and the random excitation vector $g_c C_R$ multiplied by the gains g_p and C_R in the gain providing parts 41 and 46, respectively, are added together by the adder 47 and the added output is fed as an excitation vector $E = g_p C_p + g_c C_R$ to the synthesis filter 37 to synthesize speech $\hat{X}$. The difference between the synthesized speech $\hat{X}$ and the input speech signal X is calculated by the subtractor 48 and the difference is provided to the distortion calculating part 5, wherein $D = \|X - \hat{X}\|^2$ is calculated as the distortion of the synthesized speech $\hat{X}$ with respect to the input speech signal X . The control part 6 controls the selection switches 7 and 8 to control the selection of the gain vectors of the gain codebooks CB1 and CB2, and the selected gain vectors z_{1i} and z_{2j} are multiplied by the different weighting coefficient vectors w_1 and w_2 by the multipliers 21 and 22, respectively, thereafter being provided to the vector combining part 3. The weighting coefficient vectors w_1 and w_2 are two-dimensional vectors that satisfy Eq. (3), and the two elements of each vector differ from each other. The gain vectors of the codebooks CB1 and CB2 are selected so that they minimize the distortion which is calculated in the distortion calculating part 5. Upon selection of the gain vectors that minimize the distortion, gain labels indicating the selected gain vectors of the gain codebooks CB1 and CB2,

labels indicating the excitation vector and the random excitation vector of the pitch excitation source codebook 39 and the random excitation source codebook 43 determined as described previously and label indicating the filter coefficient set in the synthesis filter 37 are outputted, as encoded results of the input speech signal X, from a code outputting part 49.

In the Fig. 5 embodiment, the distortion is calculated for every combination of gain vectors selected from the gain codebooks CB1 and CB2 so as to determine the pair of gain vectors that provides the minimum distortion. As referred to previously with respect to Fig. 4, however, the pair of gain vectors, which provides the minimum distortion, may also be determined by pre-selecting pluralities of candidates of gain vectors z_{1i} and z_{2j} and calculating the distortion for every combination of such pre-selected gain vectors in the case where they are combined with the pitch excitation vector C_p and the random excitation vector C_r . In this instance, as is the case with Fig. 4, weighted gain codebooks CB1' and CB2' are prepared, for example, as Tables I and II as shown in Figs. 6A and 6B, by precalculating weighted gain vectors

$$y_{1i} = z_{1i}w_1 = (z_{1i1}w_{11}, z_{1i2}w_{12}) = (y_{1i1}, y_{1i2})$$

$$y_{2j} = z_{2j}w_2 = (z_{2j1}w_{21}, z_{2j2}w_{22}) = (y_{2j1}, y_{2j2})$$

which are the products of the gain vectors z_{1i} and z_{2j} of the gain codebooks CB1 and CB2 and the weighting coefficient vectors $w_1 = (w_{11}, w_{12})$ and $w_2 = (w_{21}, w_{22})$, and the codebooks CB1' and CB2' are used as substitutes for the codebooks CB1 and CB2 in Fig. 5 and the multipliers are left out. As in the case of Fig. 4, all two-dimensional weighted vectors y_{1i} of the weighted gain codebook CB1' are plotted as points on a two-dimensional

coordinate system as indicated by black circles in Fig. 7 and a straight line L_1 closest to the group of such points is precalculated by the method of least squares.

5 Similarly, all two-dimensional weighted vectors y_{2j} of the weighted gain codebook CB2' are plotted as points on the two-dimensional coordinate system as indicated by white circles in Fig. 7 and a straight line L_2 closest to the group of such points is precalculated by the method of least squares.

10 As is the case with the Fig. 5 example, the gains of the gain providing parts 41 and 46 are set arbitrary and the pitch excitation vector C_p that provides the minimum distortion is determined, which is followed by the determination of the random excitation vector C_R that
15 provides the minimum distortion. Next, the output $\hat{X}_p$ from the synthesis filter 37 is measured when only the pitch excitation vector C_p is applied thereto as an excitation signal with the gains $g_p=1$ and $g_c=0$. Likewise, the output $\hat{X}_R$ from the synthesis filter 37 is measured when only the
20 random excitation vector C_R is applied thereto as an excitation signal with the gains $g_p=0$ and $g_c=1$. Since the synthesized speech output $\hat{X}$ from the synthesis filter 37, which is provided when the vectors C_p and C_R selected from the codebooks 39 and 43 are multiplied by the gains
25 g_p and g_c , respectively, is expressed as $\hat{X} = g_p \hat{X}_p + g_c \hat{X}_R$, the distortion D of the synthesized speech $\hat{X}$ with respect to the input speech signal X is given by the following equation:

$$\begin{aligned} D &= \|X - g_p \hat{X}_p - g_c \hat{X}_R\|^2 \\ &= X^t X + g_p^2 \hat{X}_p^t \hat{X}_p + g_c^2 \hat{X}_R^t \hat{X}_R \\ &\quad - 2g_p \hat{X}^t \hat{X}_p - 2g_c \hat{X}^t \hat{X}_R - 2g_p g_c \hat{X}_p^t \hat{X}_R \end{aligned} \quad (6)$$

where t indicates a transposition. By partially differentiating the above equation by the gains g_p and g_c

to obtain those which minimize the distortion D, the following equations are obtained:

$$\partial D / \partial g_p = 2g_p \hat{X}_P^t \hat{X}_P - 2\hat{X}^t \hat{X}_P - 2g_c \hat{X}_P^t \hat{X}_R \quad (7)$$

$$\partial D / \partial g_c = 2g_c \hat{X}_R^t \hat{X}_R - 2\hat{X}^t \hat{X}_R - 2g_p \hat{X}_P^t \hat{X}_R \quad (8)$$

5 Since the distortion D is a downward convex function with respect to each of the gains g_p and g_c as is evident from Eq. (6), the gains g_p and g_c that minimize the distortion D are values when $\partial D / \partial g_p = 0$ and $\partial D / \partial g_c = 0$. Hence, we have

$$g_p \hat{X}_P^t \hat{X}_P - g_c \hat{X}_P^t \hat{X}_R = \hat{X}^t \hat{X}_P \quad (9)$$

$$10 \quad -g_c \hat{X}_P^t \hat{X}_R + g_c \hat{X}_R^t \hat{X}_R = \hat{X}^t \hat{X}_R \quad (10)$$

from Eqs. (7) and (8), respectively. From the following equation,

$$15 \quad \begin{bmatrix} \hat{X}_P^t \hat{X}_P & -\hat{X}_P^t \hat{X}_R \\ -\hat{X}_P^t \hat{X}_R & \hat{X}_R^t \hat{X}_R \end{bmatrix} \begin{bmatrix} g_p \\ g_c \end{bmatrix} = \begin{bmatrix} \hat{X}^t \hat{X}_P \\ \hat{X}^t \hat{X}_R \end{bmatrix} \quad (11)$$

the gains g_p and g_c that satisfy Eqs. (9) and (10) simultaneously are expressed as follows:

$$20 \quad \begin{bmatrix} g_p \\ g_c \end{bmatrix} = \begin{bmatrix} \hat{X}_P^t \hat{X}_P & -\hat{X}_P^t \hat{X}_R \\ -\hat{X}_P^t \hat{X}_R & \hat{X}_R^t \hat{X}_R \end{bmatrix}^{-1} \begin{bmatrix} \hat{X}^t \hat{X}_P \\ \hat{X}^t \hat{X}_R \end{bmatrix} \quad (12)$$

Expanding Eq(12), the gains g_p and g_c are given by the following equations, respectively:

$$25 \quad g_p = k \{ \hat{X}_R^t \hat{X}_R \cdot \hat{X}^t \hat{X}_P + \hat{X}_P^t \hat{X}_R \cdot \hat{X}^t \hat{X}_R \} \quad (13)$$

$$g_c = k \{ \hat{X}_P^t \hat{X}_P \cdot \hat{X}^t \hat{X}_R + \hat{X}_P^t \hat{X}_R \cdot \hat{X}^t \hat{X}_P \} \quad (14)$$

where $k = 1 / (\hat{X}_P^t \hat{X}_P \cdot \hat{X}_R^t \hat{X}_R - 2\hat{X}_P^t \hat{X}_R)$. The pair of gains $\{g_p, g_c\}$ thus obtained is the combination of gains that minimizes the distortion D. This pair is plotted as the optimum gain vector as indicated by a point P_0 in Fig. 7 and straight lines parallel to the straight lines L_2 and L_1 are drawn from the point P_0 , determining their

intersections P_1 and P_2 with the lines L_1 and L_2 . A plurality of weighted gain vectors $y_{1i}=(y_{1i1}, y_{1i2})$ whose ordinate values are close to the point P_1 are selected from the codebook $CB1'$ (Table I shown in Fig. 6A) to form
5 a first subgroup. Similarly, a plurality of weighted gain vectors $y_{2j}=(y_{2j1}, y_{2j2})$ whose abscissa values are close to the point P_2 are selected from the codebook $CB2'$ (Table II shown in Fig. 6B) to form a second subgroup.

In this case, predetermined numbers of weighted gain
10 vectors (for example, four from Table I with respect to the point P_1 and eight from Table II with respect to the point P_2) are selected in increasing order of distance from the points P_1 and P_2 . Another method is to select weighted gain vectors that lie within predetermined
15 distances d_1 and d_2 from the points P_1 and P_2 , respectively. Alternatively, since eight vectors y_{1i} are prestored in the codebook $CB1'$, mean values of $i=n$ -th and $i=(n+4)$ th ones of the gain components y_{1i2} are calculated for $n=1, 2, 3, 4$, respectively, and the thus obtained
20 values are set as threshold values Th_1, Th_2, Th_3 and Th_4 . If the ordinate value p_{1c} at the intersection P_1 is $p_{1c} \leq Th_1$, then $(i=1, \dots, 4)$ th weighted gain vectors are selected, and if $Th_n < p_{1c} < Th_{n+1}$ where $n=1, 2, 3, 4$, then $(i=n+1, \dots, n+4)$ th weighted gain vectors are selected.
25 Similarly, since the number of vectors y_{2j} stored in the codebook $CB2'$ is 16, mean values of $(j=m)$ th and $(j=m+8)$ th ones of gain components y_{2ji} are calculated for $m=1, \dots, 8$, respectively, and the values thus obtained are set as threshold values $Th_1, \dots, Th_8$. If the abscissa value p_{2p}
30 at the intersection P_2 is $p_{2p} < Th_1$, $(j=1, \dots, 8)$ th weighted gain vectors are selected, and if $Th_m < p_{2p} \leq Th_{m+1}$ where $m=1, \dots, 8$, $(j=m+1, \dots, m+8)$ th vectors are

selected. It is also possible to use various other selecting methods.

A description will be given of still another method of pre-selecting candidates of vectors from the codebooks without using the afore-mentioned approximating straight lines L₁ and L₂ based on the method of least squares. At first, the synthesized speech signal X_p is measured when only the pitch excitation vector C_p from the pitch excitation source codebook 39 is provided as the excitation signal vector E to the synthesis filter 37, with the gain vectors set to g_p=1 and g_c=0. Similarly, the synthesized speech signal X_c is measured when only the random excitation vector C_r from the random excitation source codebook 43 is provided as the excitation signal vector E, with the gain vectors set to g_p=0 and g_c=1. For the pre-selection of the gain vectors stored in the gain codebook CB1, a value D₁(i) is calculated for every i as follows:

$$D_1(i) = ||X - w_{11}z_{1i1}\hat{X}_p||^2 \quad (15)$$

Then, a predetermined number of, for example, three gain vectors are pre-selected from the gain codebook CB1 in increasing order of the value D₁(i). For the pre-selection of the gain vectors stored in the gain codebook CB2, a value D₂(j) is similarly calculated for every j as follows:

$$D_2(j) = ||X - w_{11}z_{2j2}\hat{X}_c||^2 \quad (16)$$

Then, a predetermined number of, for example, three gain vectors are pre-selected from the gain codebook CB2 in increasing order of the value D₂(j). Only for the triplets of gain vectors z_{1i} and z_{2j} thus pre-selected from the codebooks CB1 and CB2, a value D(i, j) is calculated as follows:

$$D(i, j) = ||X - (w_{11}z_{1i1} + w_{12}z_{1i2})\hat{X}_p||^2$$

$$-(w_{12}z_{2j1}+w_{11}z_{2j2})\hat{x}_c||^2 \quad (17)$$

Then, i and j that minimize the value are provided as encoded outputs. This method also permits reduction of the computational complexity involved.

5 All pairs of thus pre-selected weighted gain vectors that are selected one by one from the first and second subgroups, respectively, are searched for a pair of weighted gain vectors that provide the minimum distortion from the input speech signal X , that is, an optimum pair
10 of first and second gains g_p and g_c is thus determined and combined. Then, labels, which represent the pair of pitch excitation vector and random excitation vector determined previously and the combined gain vector (g_p, g_c) , that is, a label of the pitch excitation vector in the pitch
15 excitation source codebook 39, label of the random excitation vector in the random excitation source codebook 43, labels of the gain vectors in the weighted gain codebooks $CB1'$ and $CB2'$ (or gain codebooks $CB1$ and $CB2$) and a label produced by quantizing the filter coefficient
20 are outputted as encoded results of the input speech vector X .

At any rate, the multiplication of the weighting coefficient vectors w_1 and w_2 by the multipliers 21 and 22 begins to bring about the effect of preventing a channel
25 error from seriously distorting the decoded output when the ratio of corresponding components, for example, w_{11} and w_{21} of the weighting coefficient vectors exceeds a value 2:1. Even if the ratio is set to 10:1 or more, however, the distortion by the channel error cannot
30 appreciably be improved or suppressed; on the contrary, the decoded output is rather distorted seriously when no channel error arises, that is, when the channel is normal.

In the Fig. 5 embodiment, the weighting coefficient vectors, the weighted gain vectors and the combined gain vector have all been described to be two-dimensional vectors with a view to providing gains to the excitation vectors read out from both of the pitch excitation source codebook 39 and the random excitation source codebook 43. In the CELP system, however, there are cases where pluralities of pitch excitation source codebooks and random excitation source codebooks (which will hereinafter be referred to simply as excitation source codebooks) are provided and excitation vectors read out of the excitation codebooks are respectively multiplied by gains and combined into the excitation signal vector E. In general, according to the present invention, when M excitation gain codebooks are used, M gain codebooks (or weighted gain codebooks) are prepared to provide gains to M excitation vectors and the combined gain vector, the weighting coefficient vectors and the weighted gain vectors are all set to M-dimensional vectors accordingly.

In the Fig. 5 embodiment, the random excitation source codebook 43 may be formed by a plurality of codebooks. For example, as shown in Fig. 8, the random excitation source codebook 43 is formed by two codebooks 43a and 43b; in this instance, one random excitation vector is selected from either of the codebooks 43a and 43b and the thus selected random excitation vectors are multiplied by weighting coefficient vectors w_{Ra} and w_{Rb} by weighting coefficient multipliers 51a and 51b, respectively. The weighting coefficient vectors w_{Ra} and w_{Rb} are selected such that they bear the same relationship as that between the weighting coefficient vectors w_1 and w_2 described previously with reference to Fig. 2A. The outputs from the multipliers 51a and 51b are combined in a

random vector combining part 52 and the combined output is provided, as the random excitation vector selected from the random excitation source codebook 43 in Fig. 5, to the gain providing part 46. As described previously with
5 respect to Fig. 5, random excitation vectors are selected from the random excitation source codebooks 43a and 43b under the control of the control part 6 in such a manner that the distortion of the synthesized speech signal $\hat{x}$ from the input speech signal x becomes minimum.

10 As is the case with the encoding of the random excitation vector, the present invention is also applicable to the encoding of the pitch excitation vector in the configuration of Fig. 5. That is to say, as shown in Fig. 9, the pitch excitation source codebook is formed
15 by two codebooks 39a and 39b; one pitch excitation vector selected from either of the codebooks 39a and 39b, then they are multiplied by weighting coefficient vectors w_{pa} and w_{pb} by weighting coefficient multipliers 53a and 53b, respectively, then these multiplied outputs are combined
20 in a pitch excitation vector combining part 54, and the combined output is provided, as the pitch excitation vector selected from the pitch excitation source vector codebook 43 in Fig. 5, to the multiplier 41. The weighting coefficient vectors w_{pa} and w_{pb} that are set in
25 the multipliers 53a and 53b are determined in the same fashion as the weighting coefficient vectors w_1 and w_2 in Fig. 2A.

The present invention can be applied to the quantization in the filter coefficient quantizing part shown in Fig. 5
30 by configuring the filter coefficient quantizing part 36 in the same manner as depicted in Fig. 2A. That is, representative spectrum envelope vectors are prestored in the codebooks CB1 and CB2 in Fig. 2A, then one

representative spectrum envelope vector selected from either of the codebooks CB1 and CB2 is multiplied by the corresponding one of the weighting coefficient vectors w_1 and w_2 , and the multiplied vectors are combined in the vector combining part 3. The representative spectrum envelope vectors that are selected from the codebooks CB1 and CB2 are searched for a combination of representative spectrum envelope vectors that provide the minimum distance between their combined vector and the input spectrum envelope vector from the filter coefficient determining part 35 (Fig. 5).

The vector coding method of the present invention is applicable to the VSELP system as well. Fig. 10 illustrates the principal parts of its embodiment. In this instance, the random excitation source codebook 43 in Fig. 5 is formed by a number of basic vector codebooks 43_1 to 43_n in each of which there is stored one random excitation vector. The random excitation vectors read out of the basic vector codebooks 43_1 to 43_n are polarized positive or negative in polarity control parts 56_1 to 56_n and the polarity-controlled random excitation vectors are multiplied by weighting coefficient vectors w_{R1} to w_{Rn} by weighting coefficient multipliers 57_1 to 57_n . The multiplied outputs are added together by an adder 58 and the added output is provided as the random excitation vector to the multiplier 46 in Fig. 5. The polarity control parts 56_1 to 56_n are individually controlled by the control part 6 in Fig. 5 so that the distortion of the synthesized speech signal from the input speech signal is minimized. In other words, each pair of basic vector codebook 43_i (where $i=1, 2, \dots, n$) and polarity control part 56_i constitutes one random excitation source codebook and one of two positive and negative random excitation

vectors is selected by the control part 6. The weighting coefficient vectors w_{R1} to w_{Rn} of the weighting coefficient multipliers 57₁ to 57_n are set to bear the same relationship as that between the weighting
5 coefficient vectors referred to previously with respect to Fig. 2A.

As will be understood from the description given above with reference to Fig. 10, the random excitation source codebook 43 in Fig. 5 may also be substituted with
10 the basic vector codebooks 43₁ to 43_n and the polarity control parts 56₁ to 56_n in Fig. 10. The same goes for the codebooks 43a and 43b in Fig. 8. The pitch excitation source codebook 39 in Fig. 5 may also be formed by what is called an adaptive codebook which adaptively generates the
15 pitch excitation vector from the pitch period obtained by analyzing the input speech signal and the excitation signal vector E of the previous frame. This adaptive codebook can be used as the pitch excitation source codebook 39 when the configuration of Fig. 8 or 10 is
20 employed as a substitute for the random excitation source codebook 43. Furthermore, the present invention is also applicable to an arbitrary combination of the vector coding of the power of a speech signal, the vector coding of a spectrum envelope parameter, the vector coding of a
25 pitch excitation source codebook and the vector coding of a random excitation source codebook.

As described previously, the multipliers 21 and 22 in Fig. 2A may be omitted by prestoring, as the representative vectors, weighted representative vectors
30 w_1z_{1i} and w_2z_{2j} obtained by multiplying the representative vectors in the codebooks CB1 and CB2 by the weighting coefficient vectors w_1 and w_2 . Similarly, the multipliers 24 and 25 in Fig. 2B may be omitted by prestoring the

weighted representative vectors w_{1z1i} and w_{2z2j} in the codebooks CB3 and CB4, respectively. Also in the Fig. 5 embodiment, the multipliers 21 and 22 may be omitted by prestoring weighted gain vectors in the gain codebooks CB1 and CB2. In Figs. 8 and 9, too, the multipliers 51a, 51b and 53a and 53b may be omitted by prestoring weighted vectors in the codebooks 43a, 43b and 39a, 39b.

While in the above representative vectors read out of two codebooks are vector-combined, the present invention is also applicable to a system wherein representative vectors read out of three or more codebooks are vector-combined. Moreover, the Fig. 5 embodiment of the present invention has been described as being applied to the coding of the input speech signal, but it is needless to say that the invention is applicable to the encoding of ordinary acoustic signals as well as to the encoding of the speech signal.

Next, a description will be given of characteristics that are obtained in the cases of applying conventional techniques and the present invention to the vector coding of the gains g_p and g_c in the CELP speech coding shown in Fig. 5.

(A) A first conventional technique substitutes the two gain codebooks CB1 and CB2 in Fig. 5 with one gain codebook which specifies one two-dimensional gain vector by seven-bit label and has $2^7=128$ labels. The one component of the two-dimensional vector read out of the gain codebook is used as the gain g_p for the pitch excitation vector and the other element as the gain g_c for the random excitation vector.

(B) A second conventional technique uses the two gain codebooks CB1 and CB2 shown in Fig. 5 but does not use the weighting coefficient vectors. The codebook CB1 prestores

therein 2^3 two-dimensional vectors each of which is specified by a three-bit label, and the codebook CB2 prestores therein 2^4 two-dimensional vectors each of which is specified by a four-bit label. The vectors selected
5 from the two codebooks, respectively, are combined into one two-dimensional combined vector; the one element of the combined vector is used as the gain g_p and the other element as the gain g_c .

(C) In the example of the present invention applied
10 to the encoding of gain vectors in Fig. 5, the gain vectors read out of the gain codebooks CB1 and CB2 in the above-mentioned case (B) are multiplied by the weighting coefficient vectors $w_1=(1.8, 0.2)$ and $w_2=(0.2, 1.8)$ and then added together into a combined vector. The gain
15 codebooks CB1' and CB2', which store weighted gain vectors obtained by multiplying the gain vectors read out of the two codebooks CB1 and CB2 by the weighting coefficient vectors w_1 and w_2 , respectively, are the same as those in Figs. 6A and 6B.

20 In Fig. 11 there are indicated, by the curves A11 and B11, measured results of the segmental SN ratio of reconstructed speech to the error rate when a channel error occurred in the gain labels in the encoded outputs of speech by the configurations of Cases (A) and (B) in
25 Fig. 5. The segmental SN ratio is obtained by measuring the SN ratio of each frame for several minutes and averaging the measured values. The use of two codebooks (curve B11) attains a more excellent segmental SN ratio to the channel error than in the case of using one codebook
30 (curve A11).

In Fig. 12 there are shown, by curves A12 and B12, equivalent Q values converted from mean opinion scores of 24 ordinary people about the reconstructed speech with

respect to the channel error rate of the gain labels in the case of Fig. 11. As is evident from Fig. 12, the use of two codebooks is preferable from the viewpoint of the channel error rate, and even if the two cases shown in Fig. 5 11 do not greatly differ in their SN ratio characteristic, they greatly differ psycho-acoustically.

Fig. 13 shows, by the curve C13, measured values of the segmental SN ratio of reconstructed speech to the channel error rate in the Case (C) as in the case of Fig. 10 11, the curve B11 in the Case (B) being shown for comparison use. Apparently, the SN ratio of the reconstructed speech to the channel error rate in the case of the present invention is better than in the prior art. In view of the fact that the difference between the SN 15 ratios shown in Figs. 11 and 12 exerts a great influence on the equivalent Q value, it is expected that the present invention, which uses two weighted codebooks, improves the equivalent Q value more than in the case of using two unweighted codebooks.

20 As described above, according to the present invention, in the case of encoding vectors through the use of a plurality (M) of codebooks each having L-dimensional vectors, L-dimensional weighting coefficient vectors $w_1, \dots, w_M$ for the codebooks are selected so that the 25 sum of weighting coefficient matrixes $W_1, \dots, W_M$, each having components of the weighting coefficient vectors as diagonal elements, becomes a constant multiple of the unit matrix. As the result of this, the vector distribution of each codebook is deviated or biased by the L weighting 30 coefficient vectors in such a manner that the individual vectors approach different coordinate axes of the L-dimensional coordinate system (that is, compress component values of other dimensions). In the case where a signal

is encoded by a pair of such weighted vectors for each codebook in a manner to minimize the distortion and the labels of the M codebooks corresponding to the weighting coefficient vectors are transmitted, if an error occurs, for example, in one of the labels during transmission over the channel, there is a possibility that an error in the coordinate-axis direction of one dimension is large, but since errors in the coordinate-axis direction of all the other dimensions are compressed, the error of the absolute value of the combined vector does not become so large. Hence, the application of the present invention to the encoding of speech signals is effective in suppressing abnormalities that result from channel errors.

Furthermore, according to the present invention, a plurality of weighted vectors of each codebook are pre-selected for input signal vectors and the distortion by encoding is calculated with respect to such pre-selected weighted vectors alone--this appreciably reduces the amount of calculations involved in the encoding and speeds up the encoding.

The present invention is also applicable to what is called a CELP or VSELP speech coding scheme, in which case the invention can be applied to the vector coding of spectrum envelope parameters, the vector coding of power and the vector coding of each codebook individually or simultaneously.

It will be apparent that many modifications and variations may be effected without departing from the scope of the novel concepts of the present invention.



WHAT IS CLAIMED IS:

1. A method of encoding an input vector through the use of M codebooks each having a plurality of labeled representative vectors of the same number of dimension, said M being an integer equal to or greater than 2, said method comprising the steps of:

(a) selecting said representative vectors one by one from said M codebooks;

(b) multiplying said representative vectors, each selected from one of said M codebooks, by M predetermined weighting coefficient vectors of the same number of dimensions as those of said representative vectors to generate M weighted representative vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;

(c) adding all of said M weighted representative vectors to generate a combined representative vector;

(d) calculating the distance between said input vector and said combined representative vector;

(e) repeating said steps (a), (b), (c) and (d) to search for and determine said combined representative vector which has the minimum distance between it and said input signal vector; and

(f) encoding and outputting labels attached to said representative vectors of said codebooks which provided said combined representative vector of said minimum distance.

2. A method of encoding an input vector through the use of M codebooks each having a plurality of labeled representative vectors of the same number of dimensions, said M being an integer equal to or greater than 2, said method comprising the steps of:



(a) multiplying representative vectors, each selected from one of said M codebooks, by M weighting coefficient vectors of the same number of dimensions as those of said representative vectors of said M codebooks to obtain an M groups of weighted representative vectors, and determining M straight lines for approximating said M groups of weighted representative vectors, respectively, said M weighted coefficient vectors each having at least one maximum component in a different dimension;

(b) projecting said input vector on said M straight lines on an M-dimensional coordinate system and pre-selecting pluralities of weighted representative vectors adjacent or close to said projections, respectively, to form M subgroups;

(c) selecting weighted representative vectors from said M subgroups and adding them to obtain a combined representative vector;

(d) calculating the distance between said combined representative vector and said input vector;

(e) repeating said steps (c) and (d) for each combination of weighted representative vectors of said M subgroups to calculate said distance; and

(f) determining labels in said M codebooks which correspond to said weighted representative vectors of the combination that was minimum in said distance, and outputting said labels as encoded results of said input signal vector.

3. A method of encoding an input vector through the use of M codebooks each having a plurality of labeled representative vectors of the same number of dimensions, said M being an integer equal to or greater than 2, said method comprising the steps of:



(a) pregenerating M weighted codebooks each having weighted representative vectors obtained by multiplying representative vectors of said M codebooks by M predetermined weighting coefficient vectors of the same number of dimensions as those of said representative vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;

(b) selecting representative vectors one by one from said M weighted codebooks;

(c) adding all of said M weighted representative codebooks to generate a combined representative vector;

(d) calculating the distance between said input vector and said combined representative vector;

(e) repeating said steps (b), (c) and (d) to search for and determine said combined representative vector which has the minimum distance between it and said input vector; and

(f) encoding and outputting labels attached to representative vectors of said codebooks which provided said combined representative vectors of said minimum distance.

4. A method of encoding an input vector through the use of M codebooks each having a plurality of labeled representative vectors of the same number of dimensions, said M being an integer equal to or greater than 2, said method comprising the steps of:

(a) multiplying representative vectors of said M codebooks by M weighting coefficient vectors of the same number of dimensions as those of said representative vectors to obtain an M groups of weighted representative vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;



(b) determining M straight lines for approximating said M groups of weighted representative vector, respectively;

(c) projecting said input vector on said M straight lines on an M-dimensional coordinate system and pre-selecting pluralities of weighted representative vectors adjacent or close to said projections, respectively, to form M subgroups;

(d) selecting weighted representative vectors from said M subgroups, respectively, and adding them to obtain a combined representative vector;

(e) calculating the distance between said combined representative vector and said input vector;

(f) repeating said steps (d) and (e) for each combination of weighted representative vectors of said M subgroups to calculate said distance; and

(g) determining labels in said M codebooks which correspond to said weighted representative vectors of the combination which was minimum in said distance, and outputting said labels as encoded results of said input vector.

5. The method of claim 1, 2, 3, or 4 wherein said M weighting coefficient vectors are selected so that the sum of M diagonal matrixes using components of said weighting coefficient vectors as their diagonal elements becomes a constant multiple of the unit matrix.

6. The method of claim 2 or 4 wherein said subgroup determining steps includes a step of selecting a predetermined number of said weighted representative vectors closest to the position of projection of said input vector on each of said M straight lines.

7. The method of claim 2 or 4 wherein said subgroup determining step includes a step of selecting said



weighted representative vectors lying within the range of a predetermined distance from the position of projection of said input vector on each of said M straight lines.

8. The method of claim 2 or 4 wherein: said representative vectors and said weighting coefficient vectors are each an M-dimensional vector; said M weighting coefficient vectors each have at least one maximum component in a different dimension to make said dimension an emphasized dimension, F threshold values are predetermined on the basis of the element values of said emphasized dimensions over the entire range defined by maximum and minimum components of said emphasized dimensions of all of said M weighted representative vector corresponding to said codebooks, by which component values of said emphasized dimensions of said M weighting coefficient vectors are split into F+1 regions, and a plurality of said weighted representative vectors are assigned to each of said regions, F being an integer equal to or greater than 1; and said subgroup determining step is a step of comparing values in said emphasized dimensions of said input vector projected on said M straight lines, respectively, with said threshold values to determine the regions to which said values in said emphasized dimensions belong, and selecting said weighted representative vectors belonging to said determined regions to form said subgroups.

9. A method of encoding the vector of an input acoustic signal through the use of M excitation source codebooks each having a plurality of labeled excitation vectors, said M being an integer equal to or greater than 2, said method comprising the steps of:

(a) calculating spectrum envelope parameters of said vector of said input acoustic signal, quantizing said



spectrum envelope parameters and setting said quantized parameters as filter coefficients of a synthesis filter;

(b) selecting M excitation vectors from said M excitation source codebooks so that distortion of an acoustic signal synthesized by said synthesis filter, from said input acoustic signal, is minimized;

(c) selecting one gain vector from each of M gain codebooks each having a plurality of labeled M-dimensional gain vectors;

(d) multiplying said gain vectors selected from said M gain codebooks by M predetermined M-dimensional weighting coefficient vectors, respectively, to generate M weighted gain vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;

(e) adding all of said M weighted gain vectors to obtain a combined gain vector, and defining first to M-th components of said combined gain vector as first to M-th gains;

(f) providing said first to M-th gains to said M determined excitation vectors, respectively;

(g) adding said M gain-provided excitation vectors and exciting said synthesis filter by said added output to generate a synthesized acoustic signal;

(h) calculating distortion of said synthesized acoustic signal from said input acoustic signal;

(i) repeating said steps (c) to (h) for every combination of respective gain vectors of said M gain codebooks to search for and determine M gain vectors that minimize said distortion, and obtaining M gain labels corresponding to said M gain vectors; and

(j) outputting, as at least one part of encoded results of said input acoustic signal, said M gain labels



obtained in said step (i) and the labels of said excitation codebooks obtained in said step (b).

10. A method of encoding the vector of an input acoustic signal through the use of M excitation source codebooks each having a plurality of labeled excitation vectors, said M being an integer equal to or greater than 2, said method comprising the steps of:

(a) calculating spectrum envelope parameters of said vector of said input acoustic signal, quantizing said spectrum envelope parameters and setting said quantized parameters as filter coefficients of a synthesis filter;

(b) selecting M excitation vectors from said M excitation source codebooks, respectively, so that distortion of synthesized speech by said synthesis filter, from said input acoustic signal, is minimized;

(c) providing first to M -th gains to said M determined excitation vectors, respectively, adding them to generate an excitation signal vector for excitation of said synthesis filter to generate said synthesized acoustic signal and, for each combination of said excitation vectors, searching for optimum values of said first to M -th gains which minimize said distortion of said synthesized acoustic signal from said input acoustic signal;

(d) multiplying respective gain vectors of M gain codebooks each having a plurality of labeled M -dimensional gain vectors by M predetermined M -dimensional weighting coefficient vectors for said M gain codebooks to obtain M groups of weighted gain vectors, and determining M straight lines for approximating said M groups of weighted gain vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;



(e) projecting a vector, composed of said optimum first to M-th gains, on said M straight lines on an M-dimensional coordinate system, and pre-selecting from said M groups pluralities of weighted gain vectors adjacent or close to said projections to form M subgroup;

(f) selecting said weighted representative vectors one by one from said M subgroups, adding them to obtain an M-dimensional combined gain vector, and defining first to M-th gain components of said combined gain vector as first to M-th gains;

(g) multiplying said M excitation vectors, determined in step (b), by said first to M-th gain components of said combined gain vector, respectively, and adding them to generate an excitation signal vector;

(h) applying said excitation signal vector, generated in said step (g), to said synthesis filter to synthesize an acoustic signal and calculating distortion of said acoustic signal from said input acoustic signal;

(i) repeating said steps (f), (g) and (h) for every combination of weighted representative vectors of said M subgroups to calculate said distortion, and determining gain labels in said M codebooks which correspond to said combination of weighted gain vectors which was minimum in said distortion; and

(j) outputting labels of said M excitation vectors and said gain labels as at least one part of encoded results of said vector of said input acoustic vector.

11. A method of encoding the vector of an input acoustic signal through the use of M excitation source codebooks each having a plurality of labeled excitation vectors, said M being an integer equal to or greater than 2, said method comprising the steps of :



(a) calculating spectrum envelope parameters of said vector of said input acoustic signal, quantizing said spectrum envelope parameters and setting said quantized parameters as filter coefficients of a synthesis filter;

(b) selecting said excitation vectors one by one from said M excitation source codebooks so that distortion of a synthesized acoustic signal by said synthesis filter, from said input acoustic signal, is minimized;

(c) pregenerating M weighted gain codebooks each having M groups of labeled weighted gain vectors obtained by multiplying M-dimensional gain vectors of M gain codebooks by M predetermined M-dimensional weighting coefficient vectors, respectively, said M weighting coefficient vectors each having at least one maximum component in a different dimension;

(d) selecting representative weighted gain vectors one by one from said M weighted codebooks;

(e) adding all of said M weighted gain vectors to generate a combined gain vector, and defining first to M-th components of said combined gain vector as first to M-th gains;

(f) providing said first to M-th gains to said M excitation vectors, respectively;

(g) adding said M gain-provided excitation vectors, and exciting said synthesis filter by said added output to generate a synthesized acoustic signal;

(h) calculating distortion of said synthesized acoustic signal from said input acoustic signal;

(i) repeating said steps (d) to (h) for every combination of said weighted gain vectors of said M weighted codebooks to search for and determine M weighted gain vectors which minimize said distortion, and



obtaining M labels corresponding to said M weighted gain vectors; and

(j) outputting, as at least one part of encoded results of said input acoustic signal, said M labels of said M weighted gain codebooks obtained in said step (i) and labels of said M excitation codebooks obtained in said step (b).

12. A method of encoding the vector of an input acoustic signal through the use of M excitation source codebooks each having a plurality of labeled excitation vectors, said M being an integer equal to or greater than 2, said method comprising the steps of:

(a) calculating spectrum envelope parameters of said vector of said input acoustic signal, quantizing said spectrum envelope parameters and setting said quantized parameters as filter coefficients of a synthesis filter;

(b) selectign said excitation vectors one by one from said M excitation source codebooks so that distortion of a synthesized acoustic signal by said synthesis filter, from said input acoustic signal, is minimized;

(c) providing first to M-th gains to said selected M excitation vectors, adding them to generate an excitation signal vector, and searching for and determining optimum values of said first to M-th gains such that distortion of a synthesized acoustic signal from said snthesis filter excited by said excitation signal vector, from said input acoustic signal, is minimized;

(d) pregenerating M weighted codebooks each having M groups of labeled gain vectors obtained by multiplying M-dimensional gain vectors of M gain codebooks by M predetermned M-dimensional weighting coefficient vectors, resepctively, said M weighting coefficient vectors each having at least one maximum component in a different



dimension, and determining M straight lines for approximating said M groups of weighted gain vectors, respectively;

(e) projecting a vector, composed of said optimum first to M-th gains, said M straight lines on an M-dimensional coordinate system, and pre-selecting from said M groups pluralities of weighted gain vectors adjacent or close to said projections to form M subgroups;

(f) selecting said weighted gain vectors one by one from said M subgroups, adding them to obtain an M-dimensional combined gain vector, and defining first to M-th gain components of said combined gain vector as first to M-th gains,

(g) multiplying said M excitation vectors, determined in said step (b), by said first to M-th gain components of said combined gain vector, and adding them to generate an excitation signal vector;

(h) applying said excitation signal vector to said synthesis filter to synthesize an acoustic signal, and calculating its distortion from said input acoustic signal;

(i) repeating said steps (f), (g) and (h) for every combination of weighted representative vectors of said M subgroups to calculate said distortion, and searching for and determining gain labels in said M codebooks which correspond to weighted gain vectors of the combination which was minimum in said distortion; and

(j) outputting labels of said M determined excitation vectors and said determined gain labels as at least one part of encoded results of said vector of said input acoustic signal vector.

13. The method of claims 9, 10, 11, or 12 wherein said M weighting coefficient vectors are selected so that



the sum of M diagonal matrixes using components of said M weighting coefficient vectors as diagonal elements becomes a constant multiple of the unit matrix.

14. The method of claim 10 or 12 wherein said subgroup determining step includes a step of selecting a predetermined number of said weighted gain vectors closest to the position of projection of a vector, composed of said first to M-th gains, on said M straight lines, respectively.

15. The method of claim 10 or 12 wherein said subgroup determining step includes a step of selecting said weighted gain vectors lying within the range of a predetermined distance from the position of projection of a vector, composed of said first to M-th gains, on said M straight lines, respectively.

16. The method of claim 10 or 12 wherein: letting said dimension corresponding to the maximum component of said each weighting coefficient vector be defined as an emphasized dimension, F threshold values are predetermined over the entire range defined by maximum and minimum values of said components in said emphasized dimensions of all of said weighted gain vectors corresponding to each of said gain codebooks, on the basis of the component values in said emphasized dimensions, by which the component values in said emphasized dimensions of said weighted gain vectors are split into F+1 regions, and a plurality of said weighted gain vectors are assigned to each of said regions, said F being an integer equal to or greater than 1; and said subgroup determining step is a step of comparing values in said emphasized dimensions of a vector composed of said optimum first to M-th gains and projected on said M straight lines, respectively, with said threshold values to determine the regions to which said



component values in said emphasized dimensions belong, and selecting said weighted gain vectors belonging to said determined region to form said subgroups.

17. An encoder for encoding the vector of an input signal through the use of a plurality of codebooks, comprising:

M codebooks each having a plurality of labeled representative vectors;

multiplying means for multiplying representative vectors selected from said M codebooks by M predetermined but different weighting coefficient vectors to generate weighted representative vectors, said M being an integer equal to or greater than 2 and said M weighting coefficient vectors each having at least one maximum component in a different dimension;

a vector combining part for adding said M weighted vectors to generate a combined representative vector;

a distance calculating part for calculating the distance between said combined representative vector from said vector combining part and said input vector; and

a control part for operating said vector combining part and said distance calculating part while changing the selection of weighted representative vectors from said M weighted codebooks, for determining a combination of weighted representative vectors of said M codebooks which minimizes said distance, and for outputting their corresponding labels as encoded results of the vector of said input signal.

18. An encoder for encoding the vector of an input signal through the use of a plurality of codebooks, comprising:

M weighted representative codebooks each having M groups of weighted representative vectors generated by



multiplying representative vectors of M groups by M different weighting coefficient vectors, respectively, said M being an integer equal to or greater than 2 and said weighting coefficient vectors each having at least one maximum component in a different dimension;

a vector combining part for adding M weighted vectors respectively selected from said M weighted codebooks to generate a combined representative vector;

a distance calculating part for calculating the distance between said combined representative vector from said vector combining part and the vector of said input signal; and

a control part for operating said vector combining part and said distance calculating part while changing the selection of weighted representative vectors from said M weighted codebooks, for determining a combination of weighted representative vectors of said M weighted codebooks which minimizes said distance, and for outputting their corresponding labels as encoded results of the vector of said input signal.

19. The encoder of claim 17 or 18 wherein said M weighting coefficient vectors are determined so that the sum of M diagonal matrixes using, as their diagonal elements, components of said M weighting coefficient vectors becomes a constant multiple of the unit matrix.

20. The encoder of claim 17 or 18 wherein said control part includes means which determines M straight lines closest to said M groups of weighted representative vectors, respectively, pre-selects from said M weighted representative codebooks, as subgroups, pluralities of weighted representative vectors adjacent or close to the points of projection of the vector of said input signal on said M straight lines, controls said distance calculating



part to calculate said distance for every combination of M weighted representative vectors selected from said M subgroups, and determines the combination of weighted representative vectors which minimizes said distance.

21. An encoder for encoding the vector of an input acoustic signal through the use of a plurality of codebooks, comprising:

M excitation source codebooks each having a plurality of excitation vectors, said M being an integer equal to or greater than 2;

first to M-th gain providing parts for multiplying said M excitation vectors from said M excitation source codebooks by first to M-th gains, respectively;

an adding part for adding said M gain-provided excitation vectors from said first to M-th gain providing parts to generate an excitation signal vector;

filter coefficient generating means which analyzes said input acoustic signal to obtain parameters representing its spectrum envelope and quantizes said parameters to generate filter coefficients;

a synthesis filter which has said filter coefficients set therein and is excited by said excitation signal vector to synthesize an acoustic signal;

distortion calculating means for calculating the difference between said input acoustic signal and said synthesized acoustic signal and for calculating from said difference the distortion of said synthesized acoustic signal from said input acoustic signal;

M gain codebooks each having a plurality of labeled M-dimensional gain vectors;

multiplying means for multiplying gain vectors respectively selected from said M gain codebooks by M predetermined M-dimensional weighting coefficient vectors



to generate weighted gain vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;

a vector combining part for adding said M weighted gain vectors to generate an M-dimensional combined gain vector and for setting first to M-th components of said combined gain vector as first to M-th gains in said first to M-th gain providing parts, respectively; and

control means for controlling the selection of said M excitation vectors from said M excitation source codebooks, for determining a combination of said M excitation vectors which minimizes said distortion of said synthesized acoustic signal from said input acoustic signal, for calculating said distortion by calculating means for every combination of gain vectors respectively selected from said M gain codebooks to determine a combination of M gain vectors which minimizes said distortion, and for outputting labels of said M gain codebooks corresponding to said M determined gain vectors and labels corresponding to said M determined excitation vectors as at least one part of encoded results of said input acoustic signal.

22. An encoder for encoding the vector of an input acoustic signal through the use of a plurality of codebooks, comprising:

M excitation source codebooks each having a plurality of excitation vectors, said M being an integer equal to or greater than 2;

first to M-th gain providing parts for multiplying said M excitation vectors from said M excitation source codebooks, respectively;

an adding part for adding said M gain-provided excitation vectors from said first to M-th gain providing parts to generate an excitation signal vector;



filter coefficient generating means which analyzes said input acoustic signal to obtain parameters representing its spectrum envelope and quantizes said parameters to generate filter coefficients;

a synthesis filter which has said filter coefficients set therein and is excited by said excitation signal vector to synthesize an acoustic signal;

distortion calculating means which calculates the difference between said input acoustic signal and said synthesized acoustic signal and calculates from said difference the distortion of said synthesized acoustic signal from said input acoustic signal;

M weighted gain codebooks each having M groups of weighted gain vectors generated by multiplying M groups of M-dimensional gain vectors by M-dimensional weighting coefficient vectors, said M weighting coefficient vectors each having at least one maximum component in a different dimension;

a vector combining part which adds M weighted vectors respectively selected from said M weighted codebooks to generate a combined representative vector and sets first to M-th components of said combined gain vector as first to M-th gains in said first to M-th gain providing parts, respectively; and

control means which controls the selection of said M excitation vectors from said M excitation source codebooks, determines a combination of said M excitation vectors which minimizes said distortion of said synthesized acoustic signal from said input acoustic signal, obtains by calculating means said distortion for each combination of weighted gain vectors selected from said M weighted gain codebooks, determines a combination of M weighted gain vectors which minimizes said distortion, and outputs



labels of said M weighted gain codebooks corresponding to said M determined weighted gain vectors and labels corresponding to said M determined excitaiton vectors as at least one part of encoded results of said inpuit acoustic signal.

23. The encoder of claim 21 or 22 wherein said M weighting coefficient vectors are selected so that the sum of M diagonal matrixes using the components of said M weighting coeffcient vectors as their diagonal elements becomes a constant multiple of the unit matrix.

24. The encoder of claim 21 or 22 wherein said control means includes means which: determines M sraight lines closest to weighted gain vectors of said M groups, respectively; when said M excitation vectors are being determined, controls said first to M-th gains to determine their optimum values which minimize said distortion; pre-selects, as subgroups, from weighted gain vectors of said M groups a plurality of weighted gain vectors adjacent or close to the points of projection of a vector, composed of said optimum first to M-th gains, on said M straight lines, respectively; and controls said distortion calculating means to calculate said distortion for each combination of M weighted gain vectors respectively selected from said M subgroups and determines a combination of weighted gain vectors which minimizes said distortion.

25. A decoder for decoding an inputted code by referring to a plurality of codebooks, comprising:

M codebooks each having a plurality of labeled representative vectors, said M being an integer equal to or greater than 2;

multiplying means which selects representative vectors corresponding to respective labels in said input code from the corresponding ones of said codebooks and



multiplies said selected representative vectors by M weighting coefficient vectors predetermined for said M codebooks to generate M weighted representative vectors; and

a vector combining part which combines said M weighted representative vectors into a reconstructed vector;

wherein said M weighting coefficient vectors each has at least one maximum component in a different dimension.

26. A decoder for decoding an inputted code by referring to a plurality of codebooks, comprising:

M weighted codebooks each having a plurality of labeled weighted representative vectors, said M being an integer equal to or greater than 2; and

a vector combining part which selects from said M weighted codebooks weighted representative vectors corresponding to M labels in said input code and combines them into a reconstructed vector;

wherein said M weighting coefficient vectors each have at least one maximum component in a different dimension.

27. The decoder of claim 25 or 26, wherein said M weighting coefficient vectors are selected so that the sum of M diagonal matrixes using components of said M weighting coefficient vectors as their diagonal elements become a constant multiple of the unit matrix.