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

(11) EP 1 020 848 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
19.07.2000 Bulletin 2000/29

(51) Int. Cl.[7]: **G10L 19/00**

(21) Application number: **00300042.9**

(22) Date of filing: **06.01.2000**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **11.01.1999 US 228102**

(71) Applicant:
LUCENT TECHNOLOGIES INC.
Murray Hill, New Jersey 07974-0636 (US)

(72) Inventors:
- **Benno, Steven A.**
  Dover, New Jersey 07801 (US)
- **Recchione, Michael Charles**
  Nutley, New Jersey 07110 (US)

(74) Representative:
Johnston, Kenneth Graham et al
Lucent Technologies (UK) Ltd,
5 Mornington Road
Woodford Green Essex, IG8 OTU (GB)

(54) Method for transmitting auxiliary information in a vocoder stream

(57) Non-speech information is sent in the bits allocated to one or both of a vocoder's codebooks' output by selling the gain for the corresponding codebook to zero. By selling the gain to zero, the codebook output will not be interpreted by the receiving vocoder. In this way, it is possible to transmit additional information in a way that is totally transparent to the vocoder. Applications for this technique of sending "secret" messages include, but is not limited to, transmitting parameters for generating non-speech signals. As an example, information to generate call waiting tones, DTMF, or TTY/TDD characters can be clandestinely embedded in the compressed bit stream so that these non-speech tones can be regenerated.

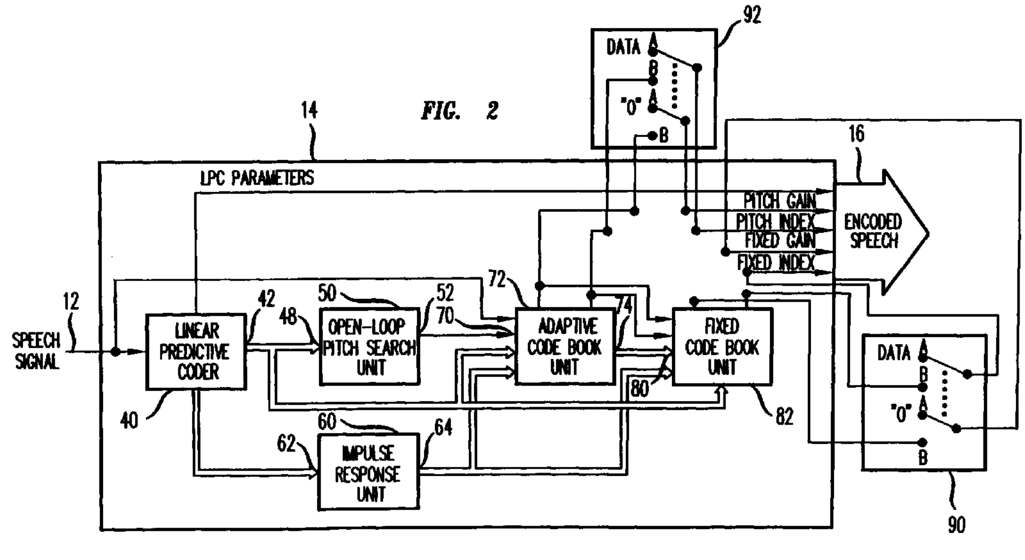

FIG. 2

EP 1 020 848 A2

Description

Background of the Invention

1. Field of the Invention

[0001] The present invention relates to telecommunications; more particularly, to transmitting data in wireless speech channels.

2. Description of the Prior Art

[0002] A voice encoder/decoder (vocoder) is used to compress voice signals so as to reduce the transmission bandwidth over a communications channel. By reducing the bandwidth per call, it becomes possible to place more calls over the same channel. There exists a class of vocoders known as code excited linear prediction (CELP) vocoders. In these vocoders, the speech is modeled by a series of filters. The parameters to these filters can be transmitted with much fewer bits than the original speech. It is also necessary to transmit the input (or excitation) to these filters in order to reconstruct the original speech. Because it would require too much bandwidth to transmit the excitation directly, a crude approximation is made by replacing the excitation by a few non-zero pulses. The locations of these pulses can be transmitted using very few bits and this crude approximation to the original excitation is adequate to reproduce high quality speech. The excitation is represented by a fixed codebook contribution and an associated gain. Also the quasi-periodicity found in speech is represented by an adaptive codebook output and an associated gain. The fixed codebook output and its associated gain, the adaptive codebook output and its associated gain, and filter parameters (also known as linear predictive coder parameters) are transmitted to represent the encoded speech signal.

[0003] The vocoders were initially designed to compress speech by modeling its characteristics and transmitting the parameters of that model in much fewer bits than transmitting the speech itself. As wireless phones become more commonplace, people are increasingly expecting to use them for the same range of non-speech applications as they have used traditional land-line phones, such as accessing voice mail and receiving call waiting tones. Recently, the FCC has mandated that text-telephones for the hearing impaired (TTY/TDD) work with digital cellular phones. The problem with non-speech applications is that they do not fit the vocoder's speech model. When non-speech signals are passed through the vocoder, the decoded result it not always acceptable. The problem is further exacerbated by the fact that wireless phones operate in an error prone environment. In order to recover from transmission errors, the vocoder depends on a speech model to recover from random errors. Once again, non-speech signals do not match this model and so the reconstruction is inadequate.

Summary of the Invention

[0004] The present invention sends information in the bits allocated to one or both of the codebooks' output by setting the gain for the corresponding codebook to zero. By setting the gain to zero, the codebook output will not be interpreted by the receiving vocoder. In this way, it is possible to transmit additional information in a way that is totally transparent to the vocoder. Applications for this technique of sending "secret" messages include, but is not limited to, transmitting parameters for generating non-speech signals. As an example, information to generate call waiting tones, DTMF, or TTY/TDD characters can be clandestinely embedded in the compressed bit stream so that these non-speech tones can be regenerated.

Brief Description of the Drawings

[0005]

FIG. 1 is a block diagram of a typical vocoder;
FIG. 2 illustrates the major functions of encoder 14 of vocoder 10; and
FIG. 3 is a functional block diagram of decoder 20 of vocoder 10.

Detailed Description

[0006] FIG. 1 illustrates a block diagram of a typical vocoder. Vocoder 10 receives digitized speech on input 12. The digitized speech is an analog speech signal that has been passed through an analog to digitized converter, and has been broken into frames where each frame is typically on the order of 20 milliseconds. The signal at input 12 is passed to encoder section 14 which encodes the speech so as decrease the amount of bandwidth used to transmit the speech. The encoded speech is made available at output 16. The encoded speech is received by the decode section of a similar vocoder at the other end of a communication channel. The decoder at the other end of the communication channel is similar or identical to the decoder portion of vocoder 10. Encoded speech is received by vocoder 10 through input 18, and is passed to decoder section 20. Decoder section 20 uses the encoded signals received from the transmitting vocoder to produce digitized speech at output 22.

[0007] Vocoders are well known in the communications arts. For example, vocoders are described in "Speech and audio coding for wireless and network applications," edited by Bishnu S. Atal, Vladimir Cuperman, and Allen Gersho, 1993, by Kluwer Academic Publishers. Vocoders are widely available and manufactured by companies such as Qualcomm Incorporated of San Diego, California, and Lucent Technologies Inc., of

Murray Hill, New Jersey.

[0008] FIG. 2 illustrates the major functions of encoder 14 of vocoder 10. A digitized speech signal is received at input 12, and is passed to linear predictive coder 40. Linear predictive coder 40 performs a linear predictive analysis of the incoming speech once per frame. Linear predictive analysis is well known in the art and produces a linear predictive synthesis model of the vocal tract based on the input speech signal. The linear predictive parameters or coefficients describing this model are transmitted as part of the encoded speech signal through output 16. Coder 40 uses this model to produce a residual speech signal which represents the excitation that the model uses to reproduce the input speech signal. The residual speech signal is made available at output 42. The residual speech from output 42 is provided to input 48 of open-loop pitch search unit 50 to an input of adaptive codebook unit 72 and to fixed codebook unit 82.

[0009] Impulse response unit 60 receives the linear predictive parameters from coder 40 and generates the impulse response of the model generated in coder 40. This impulse response is used in the adaptive and fixed codebook units.

[0010] Open loop pitch search unit 50 uses the residual speech signal from coder 40 to model its pitch and provides a pitch, or what is commonly called the pitch period or pitch delay signal, at output 52. The pitch delay signal from output 52 and the impulse response signal from output 64 of impulse response unit 60 are received by input 70 of adaptive codebook unit 72. Adaptive codebook unit 72 produces a pitch gain output and a pitch index output which become part of encoded speech output 16 of vocoder 10. Output 74 of adaptive codebook 72 also provides the pitch gain and pitch index signals to input 80 of fixed codebook unit 82. Additionally, adaptive codebook 72 provides an excitation signal and an adaptive codebook target signal to input 80.

[0011] The adaptive codebook 72 produces its outputs using the digitized speech signal from input 12 and the residual speech signal produced by linear predictive coder 40. Adaptive codebook 72 uses the digitized speech signal and linear predictive coder 40's residual speech signal to form an adaptive codebook target signal. The adaptive codebook target signal is used as an input to fixed codebook 82, and as an input to a computation that produces the pitch gain, pitch index and excitation outputs of adaptive codebook unit 72. Additionally, the adaptive codebook target signal, the pitch delay signal from open loop pitch search unit 50, and the impulse response from impulse response unit 60 are used to produced the pitch index, the pitch gain and excitation signals which are passed to fixed codebook unit 82. The manner in which these signals are computed is well known in the vocoder art.

[0012] Fixed codebook 82 uses the inputs received from input 80 to produce a fixed gain output and a fixed index output which are used as part of the encoded speech at output 16. The fixed codebook unit attempts to model the stochastic part of the linear predictive coder 40's residual speech signal. A target for the fixed codebook search is produced by determining the error between the current adaptive codebook target signal and the residual speech signal. The fixed codebook search produces the fixed gain and fixed index signal for excitation pulses so as to minimize this error. The manner in which the fixed gain and fixed index signals are computed using the outputs from adaptive codebook unit 72 are well known in the vocoder art.

[0013] Switches 90 and 92 are used to send data in place of the bits that are used to send the fixed codebook output and the adaptive codebook output, respectively. When the contacts of the switches are in position "A", the associated codebook output is replaced by data or other information and the associated codebook gain is set to zero or substantially zero. As a result, the scaled codebook output or excitation produced at a receiver will be zero or substantially zero and therefore will not have an adverse affect on the filter being used by the receiving vocoder to model the speech that is normally transmitted.

[0014] FIG. 3 illustrates a functional block diagram of decoder 20 of vocoder 10. Encoded speech signals are received at input 18 of encoder 20. The encoded speech signals are received by decoder 100. Decoder 100 produces fixed and adaptive code vectors corresponding to the fixed index and pitch index signals, respectively. These code vectors are passed to the excitation construction portion of unit 110 along with the pitch gain and the fixed gain signals. The pitch gain signal is used to scale the adaptive vector which was produced using the pitch index signal, and the fixed gain signal is used to scale the fixed vector which was obtained using the fixed index signal. Decoder 100 passes the linear predictive code parameters to the filter or model synthesis section of unit 110. Unit 110 then uses the scaled vectors to excite the filter that is synthesized using the linear predictive coefficients produced by linear predictive coder 40, and produces an output signal which is representative of the digitized speech originally received at input 12. Optionally, post filter 120 may be used to shape the spectrum of the digitized speech signal that is produced at output 20.

[0015] When data rather than speech information is being transmitted, the pitch index (adaptive codebook output) and/or the fixed index (the fixed codebook output) are used to receive the data. The affect of non-data signals on the filter synthesize by unit 110 are eliminated because the gain value associated with the pitch or code index is zero.

[0016] The functional block diagrams can be implemented in various forms. Each block can be implemented individually using microprocessors or microcomputers, or they can be implemented using a single microprocessor or microcomputer. It is also possible to implement each or all of the functional blocks using programmable digital signal processing devices or specialized devices received from the aforementioned manufacturers or other semiconductor manufacturers.

Claims

1. A method for transmitting non-speech information over a speech channel, CHARACTERIZED BY the steps of:

   transmitting non-speech information in place of pitch index information; and
   transmitting a pitch gain value having a value of substantially zero.

2. The method of claim 1, CHARACTERIZED IN THAT the non-speech information is DTMF information.

3. The method of claim 1, CHARACTERIZED IN THAT the non-speech information is TTY/TDD information.

4. A method for transmitting non-speech information over a speech channel, CHARACTERIZED BY the steps of:

   transmitting first non-speech information in place of fixed index information; and
   transmitting a index gain value having a value of substantially zero.

5. The method of claim 4, further CHARACTERIZED BY the steps of:

   transmitting second non-speech information in place of pitch index information; and
   transmitting a pitch gain value having a value of substantially zero.

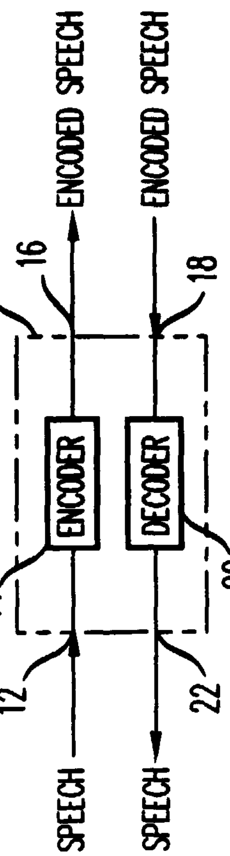
SPEECH
12
14
ENCODER
10
16
ENCODED SPEECH
SPEECH
22
DECODER
20
18
ENCODED SPEECH

FIG. 1

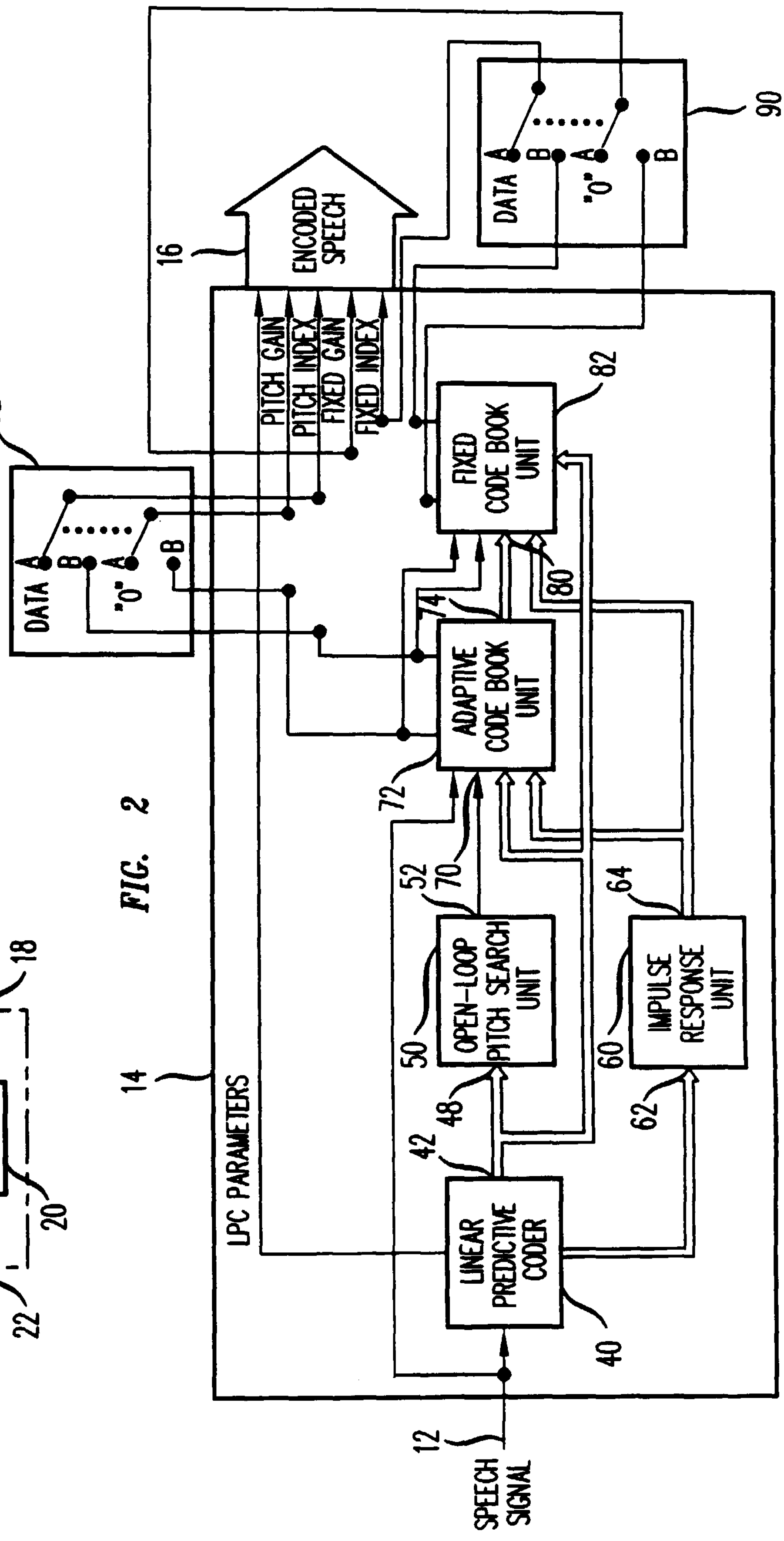
SPEECH SIGNAL
12
14
LPC PARAMETERS
LINEAR PREDICTIVE CODER
40
42
48
OPEN-LOOP PITCH SEARCH UNIT
50
52
70
IMPULSE RESPONSE UNIT
60
62
64
ADAPTIVE CODE BOOK UNIT
72
74
FIXED CODE BOOK UNIT
80
82
PITCH GAIN
PITCH INDEX
FIXED GAIN
FIXED INDEX
ENCODED SPEECH
16
DATA
A
B
"0"
A
B
92
DATA
A
B
"0"
A
B
90

FIG. 2

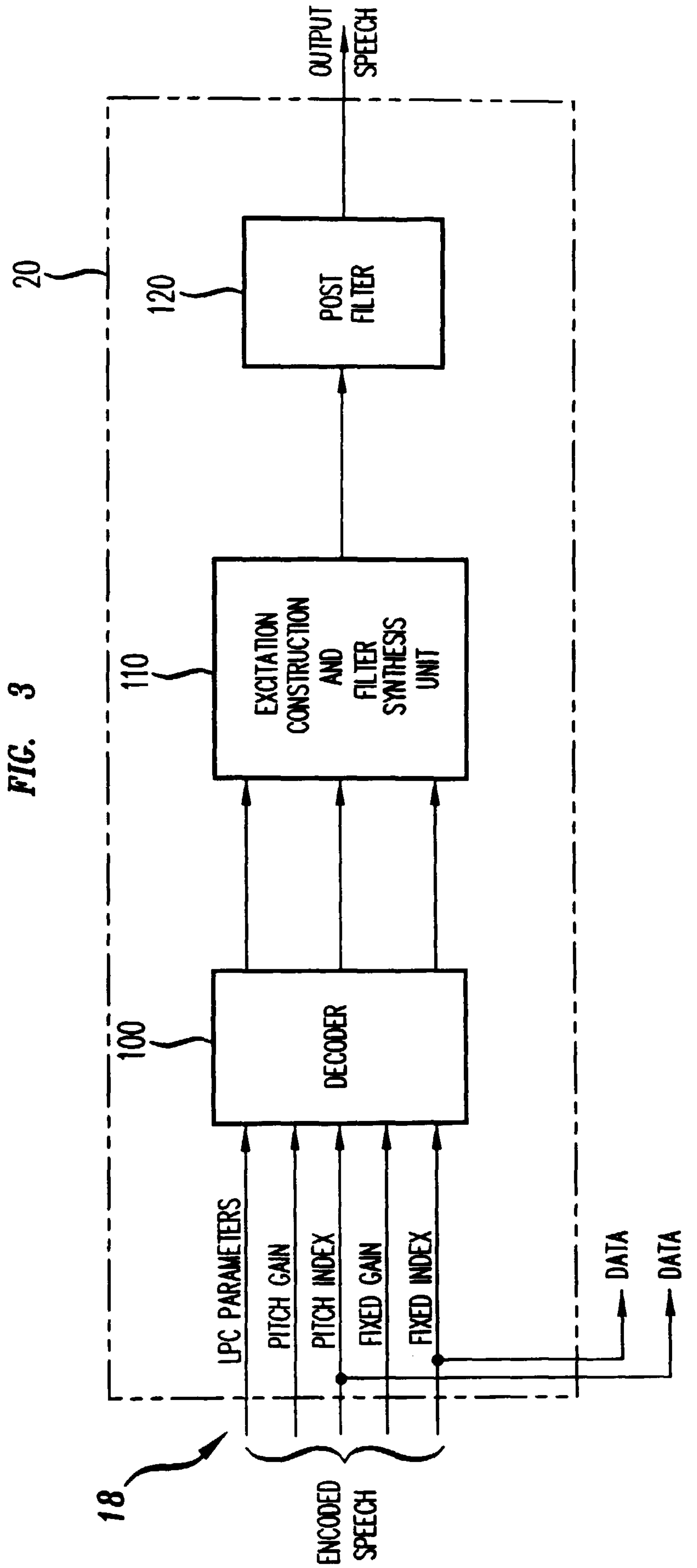
18
ENCODED SPEECH
LPC PARAMETERS
PITCH GAIN
PITCH INDEX
FIXED GAIN
FIXED INDEX
DATA
DATA
100
DECODER
110
EXCITATION CONSTRUCTION AND FILTER SYNTHESIS UNIT
120
POST FILTER
20
OUTPUT SPEECH

FIG. 3