

Aug. 15, 1961

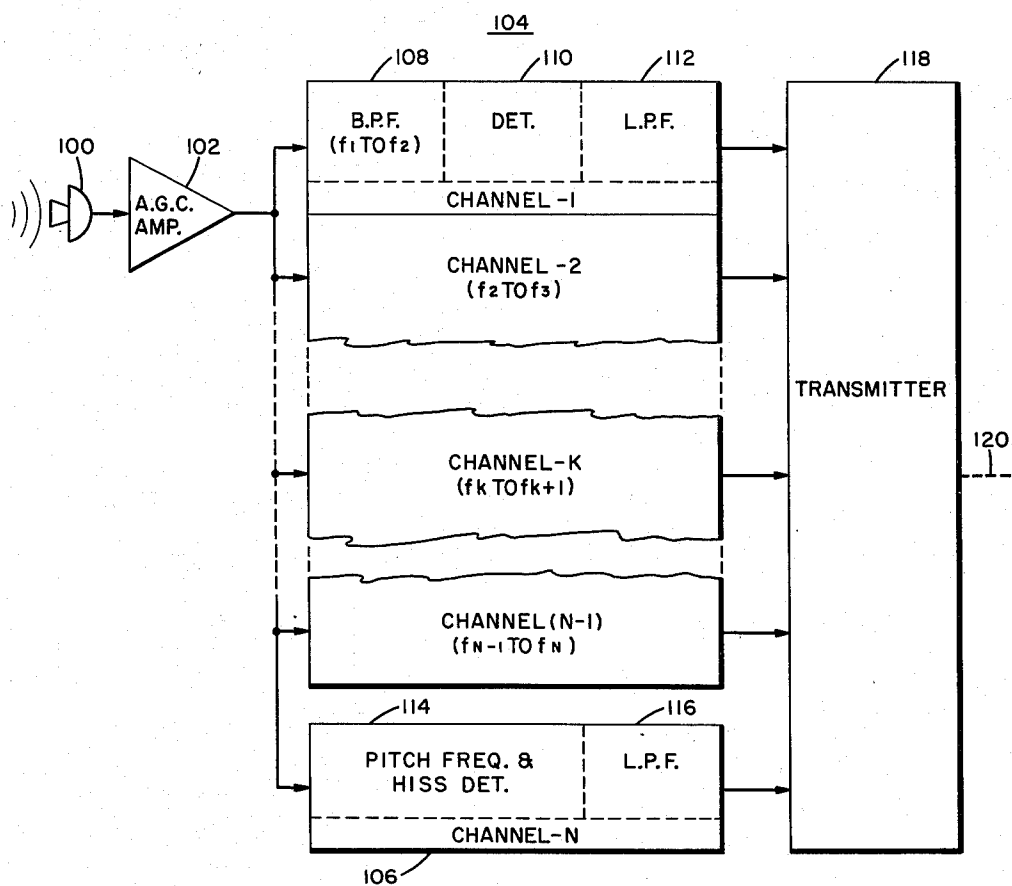
F. H. SLAYMAKER

2,996,579

FEEDBACK VOCODER

Filed Jan. 13, 1960

3 Sheets-Sheet 1



ANALIZING & TRANSMITTING LOCATION

Fig. 1A

INVENTOR.
FRANK H. SLAYMAKER
BY *George J. Seligson*

ATTORNEY

Aug. 15, 1961

F. H. SLAYMAKER
FEEDBACK VOCODER

2,996,579

Filed Jan. 13, 1960

3 Sheets-Sheet 2

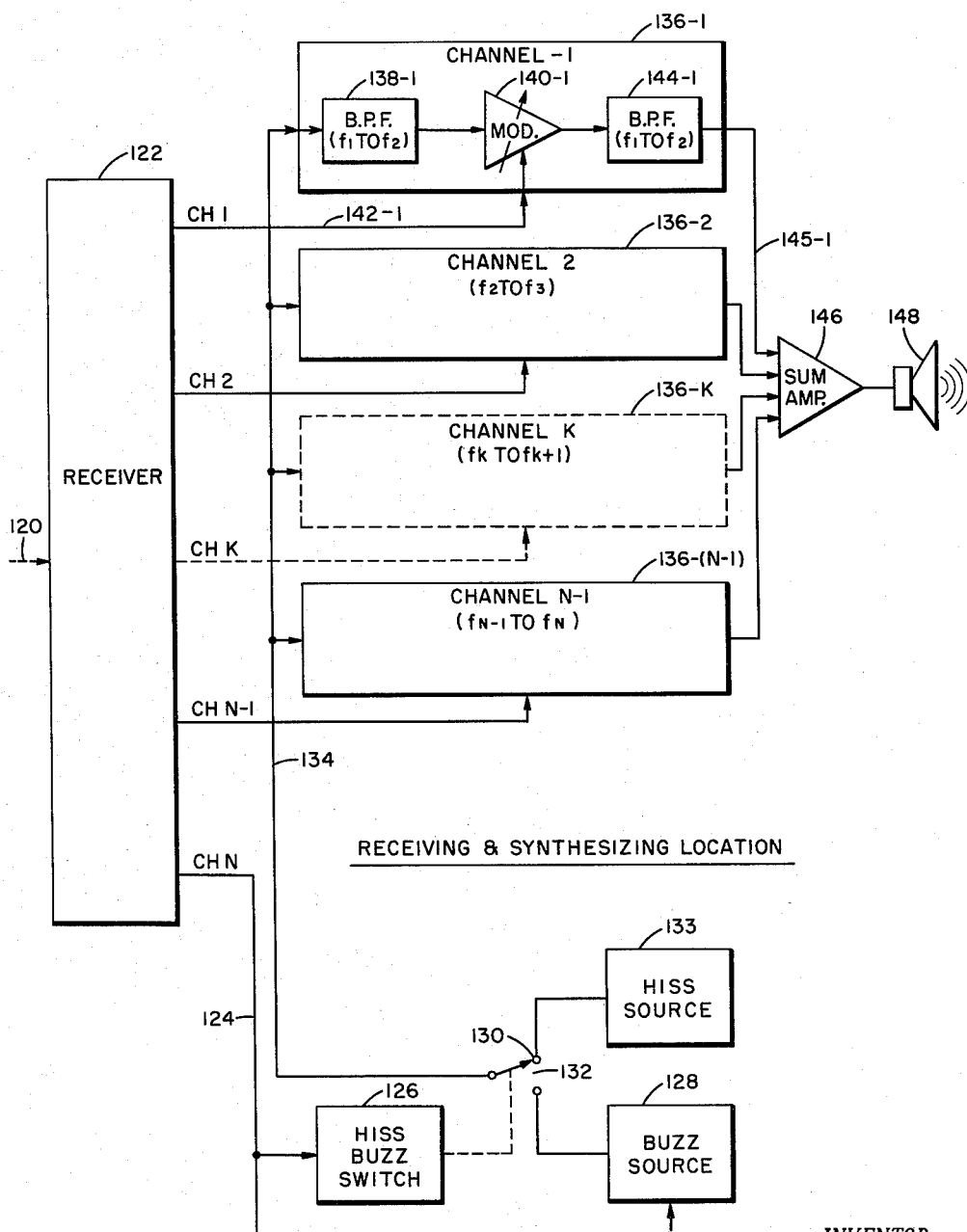


Fig. 1B

INVENTOR.
FRANK H. SLAYMAKER
BY *George J. Seligsohn*

ATTORNEY

Aug. 15, 1961

F. H. SLAYMAKER

2,996,579

FEEDBACK VOCODER

Filed Jan. 13, 1960

3 Sheets-Sheet 3

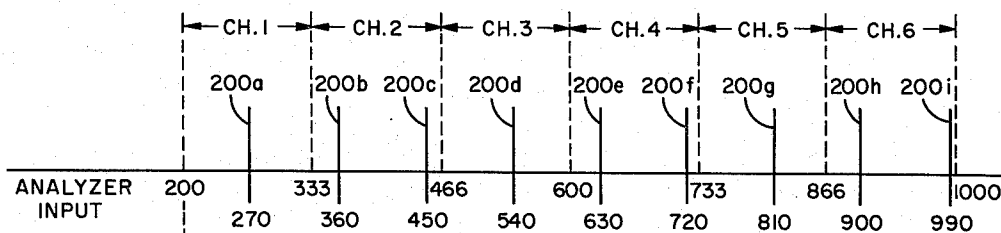


Fig. 2A

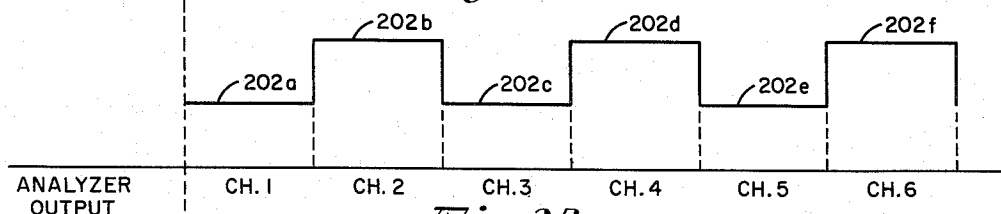


Fig. 2B

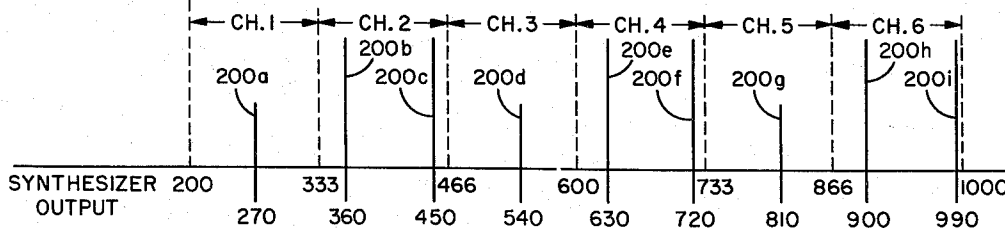


Fig. 2C

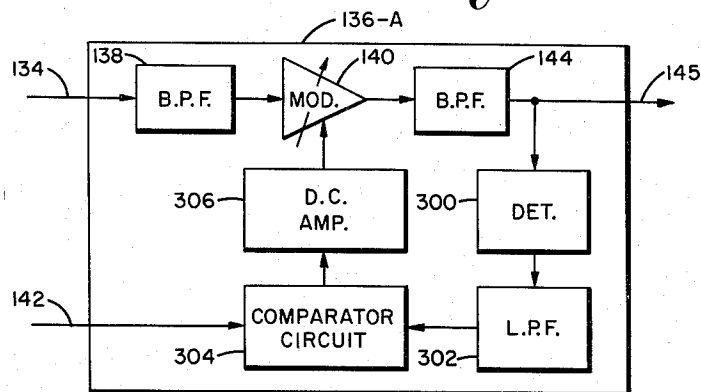


Fig. 3

INVENTOR.
FRANK H. SLAYMAKER
BY *George J. Seligsohn*
ATTORNEY

1

2,996,579

FEEDBACK VOCODER

Frank H. Slaymaker, Rochester, N.Y., assignor to General Dynamics Corporation, Rochester, N.Y., a corporation of Delaware

Filed Jan. 13, 1960, Ser. No. 2,251

10 Claims. (Cl. 179—1)

This invention relates to vocoders and, more particularly, to an improved vocoder employing feedback means to minimize distortion and enhance the naturalness of the vocoder synthesized speech output.

Vocoders are old in the art and are discussed, for instance, in Patent No. 2,194,298, issued to H. W. Dudley, March 19, 1940. Briefly, a vocoder is a device for analyzing speech to obtain the pertinent information contained therein, transmitting this information to a distant point, and at the distant point utilizing the transmitted information to synthesize the original speech. The advantage of a vocoder is that the transmission bandwidth necessary to transmit the pertinent information contained in speech is much smaller than the bandwidth necessary to transmit the speech itself.

In analyzing speech, the vocoder separates the speech frequency spectrum into a preselected number of contiguous frequency bands and then derives a signal for each frequency band which is an analog of the energy contained therein. In addition, a signal is derived which is an analog of the fundamental frequency of the applied speech. These analog signals are then transmitted either by wire or radio to a synthesizer located at a distant point. The synthesizer utilizes these analog signals to recreate the original speech.

The number of contiguous frequency bands and the respective width of each of these bands is fixed by the design of the vocoder. The applied speech, however, which is composed of a fundamental pitch frequency and the various harmonics thereof, is a variable which depends upon the speaker. If the speaker is a man with a low pitched voice, it is possible that more than one harmonic will fall within a single frequency band. Since the magnitude of a signal obtained from any frequency band is merely an analog of the total speech energy within that band, it cannot be determined from the magnitude of the analog signal alone whether this magnitude was derived from a single relatively high level harmonic or from two relatively low level harmonics.

One way of overcoming this problem is to provide a sufficient number of frequency bands of narrow enough bandwidths, so that even for a speaker with a very low pitched voice, no more than one harmonic will fall within any frequency band. However, such a solution would greatly increase the number of analog signals to be transmitted, and, therefore, tend to defeat the advantage of a vocoder over the direct transmission of speech.

The present invention provides novel means incorporated in the synthesizer for discriminating between an analog signal of any given magnitude which was derived from a single relatively high level harmonic and an analog signal of the same given magnitude which was derived from two or more relatively low level harmonics, thereby enhancing the naturalness and fidelity of the synthesized speech. Briefly, this is accomplished by providing feedback means which derive a second analog signal from the portion of the synthesized speech within each frequency band and then compares the second analog signal with the corresponding analog signal received from the analyzer.

It is therefore an object of this invention to provide a vocoder which enhances the naturalness and fidelity of synthesized speech.

2

It is more specific object of this invention to provide a vocoder employing feedback means in the synthesizer thereof to enhance the naturalness and fidelity of the synthesized speech.

These and other objects, features and advantages of the present invention will become more apparent from the following detailed description taken together with the accompanying drawing, in which:

FIGS. 1A and 1B show in block diagram the analyzer and synthesizer, respectively, of a vocoder known to the prior art,

FIGS. 2A, 2B and 2C are graphs showing, respectively, the analyzer input, analyzer output, and synthesizer output as a function of frequency for the vocoder shown in FIG. 1, and

FIG. 3 is a block diagram of a channel of the synthesizer of a vocoder employing the feedback means of the present invention.

Referring now to FIGS. 1A and 1B, there is shown a vocoder of the prior art which includes a first portion at an analyzing and transmitting location, shown in FIG. 1A, and a second portion remote therefrom at a receiving and synthesizing location, shown in FIG. 1B. The first portion of the vocoder comprises microphone 100 having its output applied as an input to automatic gain control amplifier 102. The output of automatic gain control amplifier 102 is applied in parallel to the input of each of channels 1 to N, inclusive, the first (N-1) channels, shown in block 104, are signal level channels, while channel N, shown in block 106, is a pitch signal channel. As shown, signal level channel 1 consists of a band pass filter 108, covering a frequency band from frequency f_1 to f_2 , detector 110 and low pass filter 112. The output from automatic gain control amplifier 102 is applied to the input of band pass filter 108, the output of band pass filter 108 is applied to the input of detector 110 and the output from detector 110 is applied to the input low pass filter 112.

Each of signal level channels 2 to (N-1) is identical in structure to that of channel 1, except for the frequency band covered by the band pass filter thereof. More specifically, the frequency band covered by each signal level channel is above, but contiguous with, the frequency band of the immediately preceding signal level channel. Thus, the band pass filter of channel 2 covers a frequency band from frequency f_2 to f_3 , the band pass filter of channel 3 covers a frequency band from frequency f_3 to f_4 , etc.

In one typical vocoder, the frequency spectrum of the speech is broken up into 16 frequency bands by sixteen signal level channels. The first six of these signal level channels have equal bandwidths of approximately 133 cycles extending from a low frequency f_1 for channel 1 of 200 cycles to a high frequency for channel 6 of 1,000 cycles. The remaining ten channels have progressively wider bandwidths, the frequency band covered by the sixteenth channel being from 3,330 cycles to 3,820 cycles.

The output of the band pass filter of each signal level channel is rectified by the detector thereof, and the output of the detector is smoothed by the low pass filter thereof. In one typical vocoder, the cutoff frequency of the low pass filter of each signal level channel is 25 cycles. It will be seen that the output of each signal level channel, obtained at the output of the low pass filter thereof is a D.C. voltage having a magnitude which is an analog of the speech energy passed by the band pass filter thereof.

The speech of any speaker will have a certain pitch or fundamental frequency for all voiced sounds, such as sound of the letters B and V. However, the pitch frequency will be substantially absent for unvoiced sounds, such as the sound of the letters P and F. Chan-

nel N, shown in block 106, includes pitch frequency and hiss detector 114 for determining the pitch frequency of the applied speech input, and whether or not this pitch frequency is present at any given time.

Since the pitch frequency and hiss detector is well known in the art, and does not form part of the present invention, it will not be described in detail. Briefly, however, it may consist of a variable band pass filter having a nominal uncorrected band width below the lowest pitch frequency which may be encountered, so that the pitch frequency occurs somewhere on the sloping upper cutoff edge of the filter. The output of the filter is applied to a frequency to analog signal converter which consists of a clipper, a differentiator and a rectifier to provide pulses of a given polarity which occur at the zero axis crossings of the pitch frequency. These pulses are utilized to trigger a monostable multivibrator to provide a series of constant width pulses which occur at a repetition rate equal to the pitch frequency. The output of the monostable multivibrator is applied to an averaging integrator to provide a D.C. output signal which has a magnitude which is an analog of the pitch frequency. This signal is fed back to the variable band pass filter to effect an increase in the upper cutoff of the filter in accordance with the magnitude of the signal.

Pitch frequency and hiss detector 114 also may include a switch controlled by the output signal of the averaging integrator for passing this output signal to low pass filter 116 only in response to the magnitude of the output signal being above a given threshold value. An output signal having a magnitude below this given threshold value is indicative of an unvoiced speech sound, while an output signal having a magnitude above this given threshold of value is indicative of a voiced speech sound.

The output of each of the N channels is applied as a separate input to transmitter 118. Transmitter 118 either time or frequency multiplexes the various analog signals applied as an input thereto and then transmits the multiplexed signals, either by wire or radio, over transmission link 120 to receiver 122 located at the receiving synthesizing location, shown in FIG. 1B. Receiver 122 detects and demultiplexes the various analog signals and applies them to the respective output thereof.

The pitch frequency analog signal, derived from analyzer channel N, shown in block 106, is applied over conductor 124 to both the input of hiss-buzz switch 126 and the input to buzz source 128.

Buzz source 128 includes means for generating a series of sharp pulses which occur periodically at a rate determined by the magnitude of the pitch frequency analog signal applied to buzz source 128. More specifically, the series of periodic sharp pulses produced by buzz source 128 occur substantially at the pitch frequency of the speech signals applied to pitch frequency and hiss-detector 114. Since these pulses are very sharp, they include all the harmonics of the pitch frequency and the amplitude of each of the harmonics is substantially equal to the amplitude of the fundamental pitch frequency.

Hiss-buzz switch 126 is operated by the input applied thereto only if the magnitude of the pitch frequency analog signal is above the threshold value indicative of a voiced speech sound. When hiss-buzz switch 126 is operated, it opens normally closed contacts 130 thereof and closes normally opened contacts 132 thereof. When normally closed contacts 130 remain closed, indicative of an unvoiced speech sound, the output of hiss-source 133, which may be a thermal noise source, is applied over conductor 134 to the input of each of synthesizer channels 1 to (N-1), shown in blocks 136-1 to 136-(N-1), as shown. When normally open contacts 132 are closed, indicative of a voiced speech sound, the output of buzz source 128 is applied over conductor 134 to the input of each of synthesizer channels 136-1 to 136-(N-1), as shown.

As shown in detail for synthesizer channel 1, each of the synthesizer channels consists of a first band pass

filter, such as band pass filter 138-1, covering a frequency band identical to that covered by the band pass filter of the corresponding channel of the analyzer, to the input of which conductor 134 is connected.

The output of each synthesizer channel first band pass filter, such as band pass filter 138-1, is applied as the first input to a modulator included in each synthesizer channel, such as modulator 140-1 of channel 1. Each of these modulators is in effect a variable gain amplifier. The output from receiver 122, manifesting the analog signal of signal level channel 1 of the analyzer, is applied over conductor 142-1 as a second input to modulator 140-1 to control the gain thereof in accordance with the magnitude of that analog signal. In a similar manner, each of the other outputs of receiver 122 manifesting the respective analog signals of the other signal level channels of the analyzer is applied as a second input to the modulator of the corresponding synthesizer channel to control the gain thereof in accordance with the magnitude of that analog signal.

The output of each modulator, such as modulator 140-1, is applied as an input to a second synthesizer band pass filter, such as band pass filter 144-1 which covers a frequency band identical to that of the band pass filter of the corresponding analyzer channel.

The outputs of the respective second synthesizer band pass filters, such as band pass filter 144-1, are applied as separate inputs to summing amplifier 146, over conductors, such as conductor 145-1. The output of summing amplifier 146 is applied to a transducer, such as loud-speaker 148 or an earphone, to reproduce the original speech impinging on microphone 100.

The operation of the vocoder shown in FIG. 1, as it relates to the present invention will now be discussed. In this discussion reference will be made to FIGS 2A, 2B and 2C, wherein the analyzer input, analyzer output and synthesizer output, respectively, of the first six channels is shown. For the purpose of this discussion, it will be assumed that the bandwidth of each of these six channels is 133 cycles extending from a lower cutoff for channel 1 of 200 cycles to an upper cutoff for channel 6 of 1,000 cycles. It will also be assumed that the pitch frequency of the applied speech is 90 cycles, and, in order to simplify the discussion, that all the harmonics of the speech pitch frequencies are equal in amplitude, although this is not normally the case.

Referring now to FIG. 2A, which shows the analyzer input as a function of frequency, the frequency band covered by channel 1 extends from 200 to 333 cycles, the frequency band covered by channel 2 extends from 333 to 466 cycles, the frequency band covered by channel 3 extends from 466 to 600 cycles, the frequency then covered by channel 4 extends from 600 to 733 cycles, the frequency band covered by channel 5 extends from 733 to 866 cycles and the frequency band covered by channel six extends from 866 to 1,000 cycles. Lines 200a to 200i, inclusive, represent the third to the eleventh harmonic, which have been assumed to be of equal amplitude, of the applied speech pitch frequency. Since the applied speech pitch frequency has been assumed to be 90 cycles, it will be seen that the third harmonic 200a is 270 cycles, the fourth harmonic 200b is 360 cycles, the fifth harmonic 200c is 450 cycles, the sixth harmonic 200d is 540 cycles, the seventh harmonic 200e is 630 cycles, the eighth harmonic 200f is 720 cycles, the ninth harmonic 200g is 810 cycles, the tenth harmonic 200h is 900 cycles and the eleventh harmonic 200i is 990 cycles. Therefore, as shown in FIG. 2A, only a single harmonic falls within each of channels 1, 3 and 5, while two harmonics fall within each of channels 2, 4 and 6.

Referring now to FIG. 2B, there is shown the analyzer outputs 202a to 202f, inclusive, for each of the first six channels. As shown, the magnitude of each of the analyzer outputs 202b, 202d and 202f of channels 2, 4 and 6 are higher than the magnitudes of the outputs

202a, 202b and 202c of each of channels 1, 3 and 5, since the magnitude of the analyzer output of each channel is an analog of the total speech energy within the frequency band of that channel. Therefore, although all the speech harmonics have been assumed to have equal amplitude, the analyzer output of each of channels 2, 4 and 6, which each includes two harmonics within the frequency band thereof, will be higher than the analyzer output of each of channels 1, 3 and 5, which each includes only a single harmonic within the frequency band thereof.

Referring again for a moment to FIG. 1, it will be remembered that buzz-source 128, in response to the analog signal from channel N, recreates the speech pitch frequency and all harmonics thereof and applies them in parallel to the input of each of the synthesizer channels; that the frequency band passed by each of the synthesizer channels is equal to the corresponding analyzer channel; and that the relative amplitude of the harmonics appearing at the synthesizer output of each channel is determined by the magnitude of the analyzer output of the corresponding signal-level channel applied to the modulator of that synthesizer channel.

Referring now to FIG. 2C, which shows the synthesizer output of the first six channels a function of frequency, it will be seen that the harmonics 204a to 204i, inclusive, present in the synthesizer output are identical in frequency to the corresponding harmonics 200a to 200i, inclusive, of the analyzer input, shown in FIG. 2A. However, although the amplitudes of each of harmonics 200a to 200i, inclusive, are assumed to be equal, the relative amplitude of harmonics 204a to 204i, inclusive, of the synthesizer output will not be equal. More specifically, the relative amplitudes of harmonics 204a, 204d and 204g, falling within the frequency band of channels including only a single harmonic, are lower than the relative amplitudes of harmonics 204b, 204c, 204e, 204f, 204h and 204i, falling within the frequency band of channels including two harmonics. Thus, it will clearly be seen by comparing FIG. 2C to FIG. 2A, the synthesizer output has been distorted relative to the analyzer input. It is the purpose of the present invention to prevent this type of distortion in vocoders.

Referring now to FIG. 3, there is shown a modified synthesizer channel 136-A, incorporating a preferred embodiment of the present invention, which may be substituted for each of the synthesizer channels shown in FIG. 1. Modified synthesizer channel 136-A is similar to each of synthesizer channels 136-1 to 136-(N-1) of FIG. 1 in that the output of hiss source 133 or buzz source 128 is applied over conductor 134 to the input of band pass filter 138, which determines the frequency band covered by the channel. Further, the output of band pass filter 138 is applied as a first input to modulator 140 and the output of modulator 140 is applied to the input of band pass filter 144 covering the same frequency band covered by band pass filter 138. Also, the output of band pass filter 144 is applied to summing amplifier 146 over conductor 145. However, modified synthesizer channel 136-A differs from the synthesizer channel of FIG. 1 in that the associated signal level analog signal present on conductor 142 is not applied directly as a second input to modulator 140, as is the case in the synthesizer channels of FIG. 1, but is applied as a first input to comparator circuit 304. Further, the output of band pass filter 144 present on conductor 145 is applied as an input to detector 300, the output of detector 300 is applied as an input to low pass filter 302 and the output of low pass filter 302 is applied as a second input to comparator circuit 304. The output from comparator circuit 304, which is equal to the difference in magnitudes between the first and second inputs applied thereto, is applied as an input to D.C. amplifier 306, and the output of D.C. amplifier 306 is applied as the second input to modulator 140.

Considering now the operation of modified synthesizer

channel 136-A, shown in FIG. 3, it will be seen that detector 300 and low pass filter 302 provides an output signal, which is applied as the second input to comparator circuit 304, which has a magnitude which is an analog of the speech energy in the synthesizer output of the channel. This is compared with the analyzer output signal present on conductor 142, which is applied as the first input to comparator circuit 304.

Since, as shown in FIGS. 2A and 2C, the number of harmonics falling within a corresponding channel of the analyzer and synthesizer is equal, a difference in the magnitudes of the first and second inputs applied to comparator circuit 304 must be due to an error in the amplitude of the synthesizer output harmonics of a channel with respect to the amplitude of the harmonics of the analyzer input to the corresponding channel. This difference, after amplification by D.C. amplifier 306, varies the gain of modulator 140 to reduce this error to a minimum. Therefore, by means of the feedback circuit incorporated in modified synthesizer channels, such as channel 136-A, the gain of modulator 140 is made independent of the number of harmonics falling within a frequency band, and is dependent solely on the amplitude of the harmonics. Thus, distortion due to different numbers of harmonics falling within the respective frequency bands of different channels, illustrated in FIG. 2C, is eliminated.

Although, only a preferred embodiment of the invention has been described in detail herein, it is not intended that the invention be restricted thereto, but that it be limited only by the true spirit and scope of the appended claims.

What is claimed is:

1. In a vocoder, a synthesizer channel comprising a band pass filter having a bandwidth which covers a given preselected portion of the frequency spectrum of speech, modulating means, detecting means, a comparator circuit, first means for applying a D.C. signal having a magnitude which is an analog of the speech energy of particular speech falling within said preselected portion as a first input to said comparator circuit, second means for applying a signal including harmonics contained in said particular speech as an input to said band pass filter, third means for applying the output of said band pass filter as a first input to said modulating means, fourth means for applying the output of said modulating means as an input to said detecting means, fifth means for applying the output of said detecting means as a second input to said comparator circuit, the magnitude of the output of said comparator circuit manifesting the difference in magnitudes between the first and second inputs applied thereto, and sixth means for applying the output of said comparator circuit as a second input to said modulating means to control the gain thereof in accordance with the magnitude thereof.

2. The vocoder defined in claim 1, wherein said modulating means includes a modulator, and a second band pass filter having the same bandwidth as said first-mentioned band pass filter, the output of said modulator being applied as an input to said second band pass filter and the output of said modulating means being obtained at the output of said second band pass filter.

3. The vocoder defined in claim 1, wherein said detecting means includes a detector and a low pass filter, the output of said detector being applied as an input to said low pass filter and the output of said detecting means being obtained at the output of said low pass filter.

4. The vocoder defined in claim 1, wherein said sixth means includes a D.C. amplifier having the output of said comparator means applied as an input thereto and the output thereof applied as the second input to said modulating means.

5. The vocoder defined in claim 1, wherein the output of said synthesizer channel is obtained at the output of said modulating means.

6. In a vocoder comprising an analyzer including a

7

pitch channel for producing a pitch D.C. signal only in response to voiced sounds applied thereto which has a magnitude which is an analog of the pitch of the applied voiced sound, and a plurality of analyzer signal level channels for dividing the speech frequency spectrum into contiguous preselected frequency bands, each of said signal level channels producing a signal level D.C. signal which has a magnitude which is an analog of the speech energy contained within its associated frequency band, a synthesizer, and means for transmitting said signals to said synthesizer, said synthesizer comprising a hiss source, a buzz source, means for applying said pitch D.C. signal to said buzz source to control the frequency thereof in accordance with the magnitude of said pitch D.C. signal, a plurality of synthesizer signal level channels each of which covers the same frequency band as a corresponding analyzer signal level channel, a hiss-buzz switch having said pitch D.C. signal applied as an input thereto for selectively applying the output of said hiss source as an input to said synthesizer signal level channels in response to the magnitude of said pitch signal being below a given threshold value and for applying the output of said buzz source as an input to said synthesizer signal level channels in response to the magnitude of said pitch signal being at least equal to said given threshold value, and a summing amplifier for summing the respective outputs of said synthesizer signal level channels; wherein each of said synthesizer signal level channels includes a band pass filter having a bandwidth covering the preselected frequency band of the corresponding analyzer signal level channel, modulating means, detecting means, a comparator circuit, first means for applying the signal level D.C. signal of the corresponding analyzer channel as a first input to said comparator circuit, second means for applying said input to said synthesizer channels as an

8

input to said band pass filter, third means for applying the output of said band pass filter as a first input to said modulating means, fourth means for applying the output of said modulating means as an input to said detecting means, fifth means for applying the output of said detecting means as a second input to said comparator circuit, the magnitude of the output of said comparator circuit manifesting the difference in magnitudes between the first and second inputs applied thereto, and sixth means for applying the output of said comparator circuit as a second input to said modulating means to control the gain thereof in accordance with the magnitude thereof.

7. The vocoder defined in claim 6, wherein said modulating means includes a modulator, and a second band pass filter having the same bandwidth as said first-mentioned band pass filter, the output of said modulator being applied as an input to said second band pass filter and the output of said modulating means being obtained at the output of said second band pass filter.

8. The vocoder defined in claim 6, wherein said detecting means includes a detector and a low pass filter, the output of said detector being applied as an input to said low pass filter and the output of said detecting means being obtained at the output of said low pass filter.

9. The vocoder defined in claim 6, wherein said sixth means includes a D.C. amplifier having the output of said comparator means applied as an input thereto and the output thereof applied as the second input to said modulating means.

10. The vocoder defined in claim 6, wherein the output of each synthesizer channel which is applied to said summing amplifier is obtained at the output of said modulating means.

No references cited.