

April 30, 1968

J. L. FLANAGAN

3,381,093

SPEECH CODING USING AXIS-CROSSING AND AMPLITUDE SIGNALS

Filed Aug. 4, 1965

2 Sheets-Sheet 1

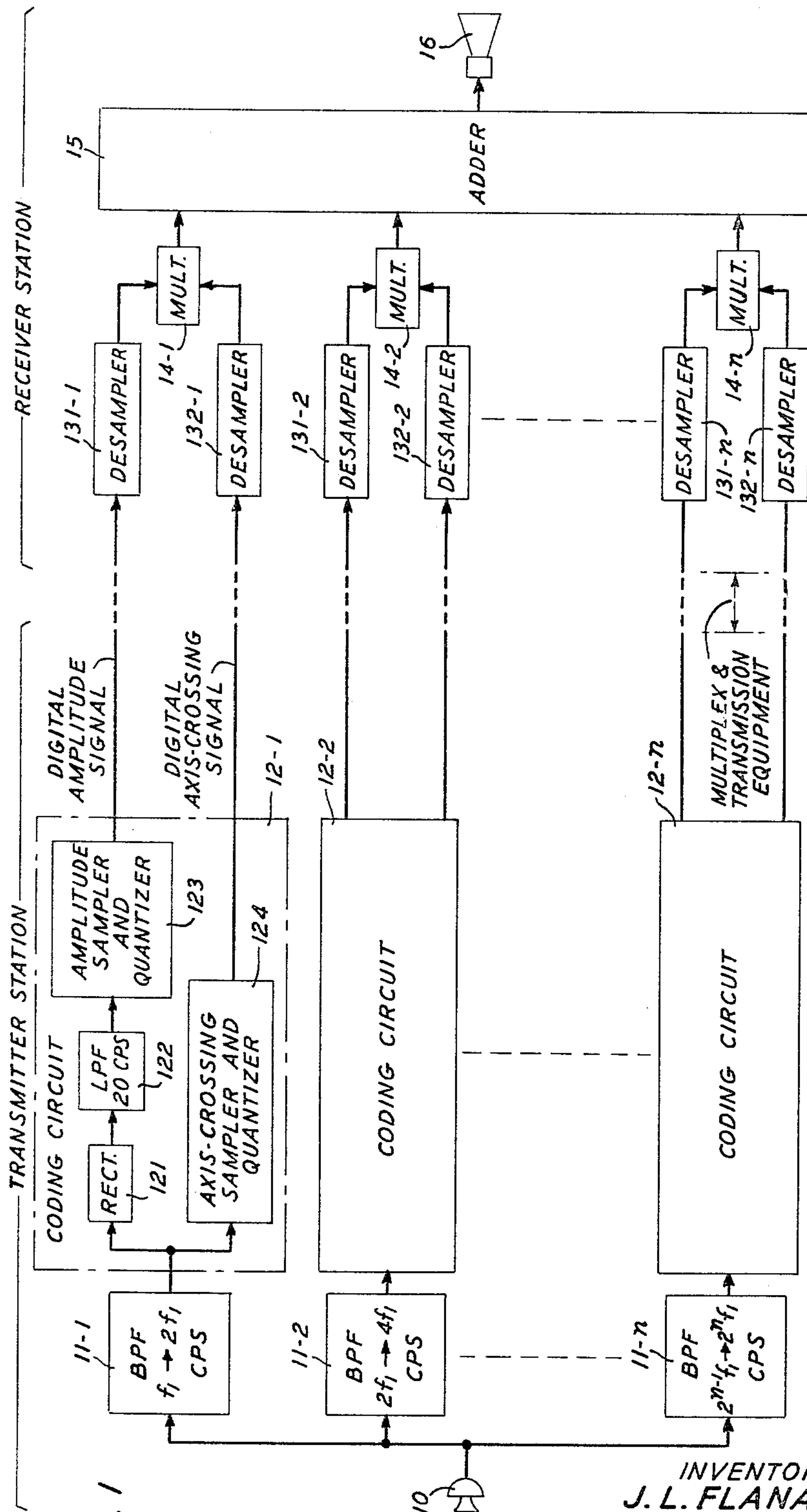


FIG. 1

INVENTOR
J. L. FLANAGAN
BY *Julius F. Kirsch*
ATTORNEY

April 30, 1968

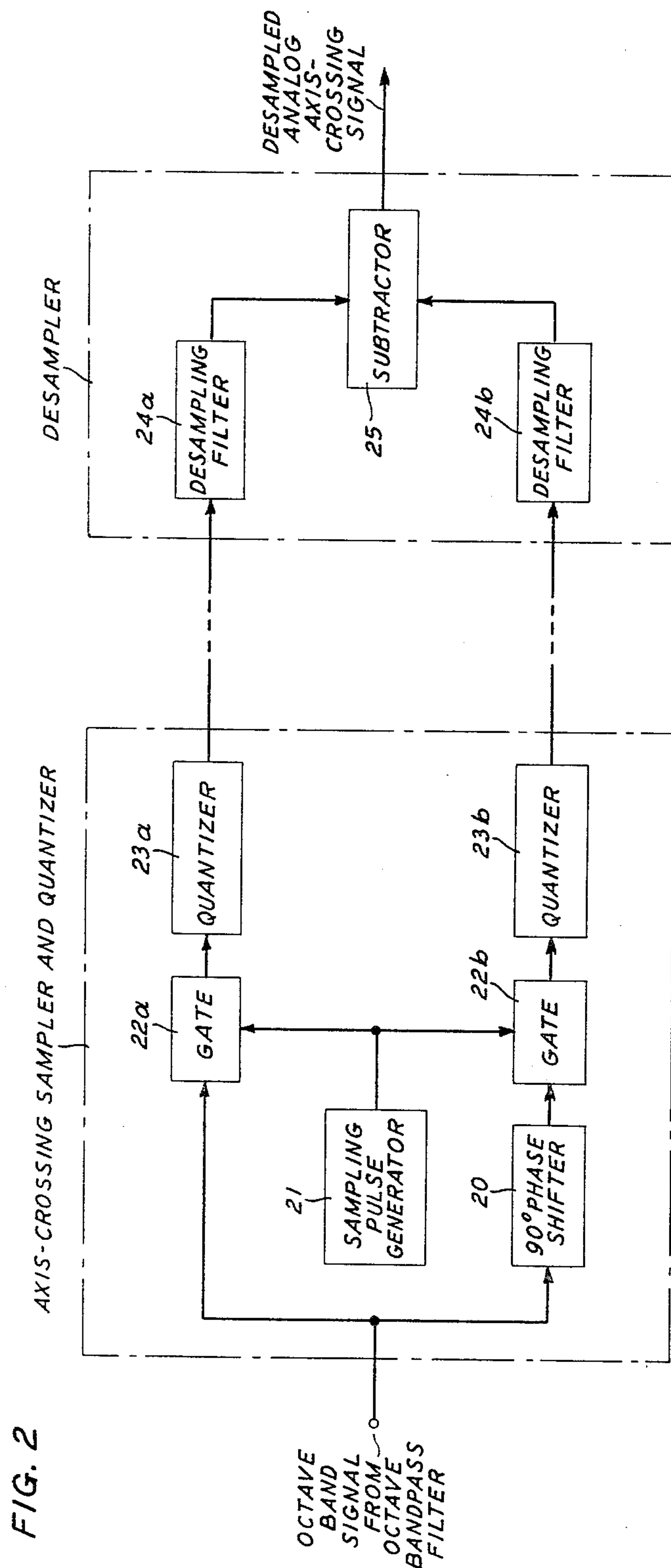
J. L. FLANAGAN

3,381,093

SPEECH CODING USING AXIS-CROSSING AND AMPLITUDE SIGNALS

Filed Aug. 4, 1965

2 Sheets-Sheet 2



1

3,381,093

SPEECH CODING USING AXIS-CROSSING AND AMPLITUDE SIGNALS

James L. Flanagan, Warren Township, Somerset County, N.J., assignor to Bell Telephone Laboratories, Incorporated, New York, N.Y., a corporation of New York
Filed Aug. 4, 1965, Ser. No. 477,152
10 Claims. (Cl. 179-15.55)

This invention relates to the transmission of human speech in coded form, and in particular to systems for transmitting human speech in coded form in order to conserve transmission channel bandwidth.

Conventional speech communication systems, for example, commercial telephone systems, typically convey human speech by transmitting an electrical facsimile of the acoustic waveform produced by a human talker. Because of the redundancy of human speech, however, facsimile transmission is a relatively inefficient way to transmit speech information, and it is well known that the information contained in a typical speech sound may be transmitted over a channel of substantially smaller information carrying capacity or frequency bandwidth than that required for facsimile transmission of the speech waveform.

In order to make more economic use of existing and proposed transmission media, a number of arrangements for compressing or reducing the amount of transmission channel capacity required for the transmission of speech information have been proposed. Such arrangements typically include a transmitter terminal at which there is derived from an incoming speech wave a group of coded signals representative of selected information-bearing characteristics of the speech wave, these coded signals collectively requiring a transmission channel of substantially smaller information carrying capacity than that required for facsimile transmission of the original speech wave. After transmission of the coded signals to a receiver station, a replica of the original speech wave is reconstructed from the coded signals.

One well-known bandwidth compression arrangement is the so-called resonance or formant vocoder, several examples of which are described in an article by E. E. David, Jr., entitled "Signal Theory in Speech Transmission," Vol. CT-3 IRE Transactions on Circuit Theory, pages 232, 239 (1956). In a typical formant vocoder, the principal selected information bearing characteristics represented in coded form are the time-varying frequency locations of selected formants or peaks in the speech amplitude spectrum and their relative amplitudes. These selected maxima correspond to vocal tract resonances, that is, they correspond to frequency regions of relatively effective transmission through a talker's vocal tract, and in general it is the maxima corresponding to the three principal vocal tract resonances which are selected for coding. In addition, conventional formant vocoders require extraction of vocal excitation information, that is, information describing whether the speech sound at a given instant is voiced or unvoiced, and if voiced, the fundamental voice frequency or pitch.

In determining the frequencies of speech formants, conventional formant vocoders divide an incoming speech wave into a number of fixed frequency subbands, each subband being selected to embrace the frequency range within which a particular formant normally occurs. From the speech frequency components lying within each subband there is derived a slowly varying control signal representative of the short-time average of the frequency at which a formant peak occurs within each subband. As pointed out in the above-mentioned David article, one way of measuring formant frequencies is to count the rate of axis crossings within each subband, since the axis-cross-

2

ing rate within each subband is generally determined by the frequency of that harmonic of the fundamental pitch frequency which has the greatest energy within each subband, and the harmonic that is closest to the formant peak within a subband usually has the greatest energy within that subband. It is usual then, for each of the slowly varying signals to represent a formant location in terms of the short-time average rate of axis crossings within a particular frequency subband.

It is recognized, however, that measurement of formant frequencies from axis-crossing rates is subject to error for a number of reasons, chief of which is the substantial overlapping of the frequency ranges within which formants normally occur. As a result, in a conventional formant vocoder a particular fixed frequency subband will contain from time to time more than one formant, and it is difficult to track the frequency of one formant in the presence of a second formant because the axis-crossing rate is influenced by the presence of more than one formant within a subband; for a description of this phenomenon, see E. Peterson, "Frequency Detection and Speech Formants," vol. 23, Journal of the Acoustical Society of America, page 668 (1951). In such a case, a control signal that specifies a formant frequency may be in error, and synthetic speech that is reconstructed from erroneous formant signals is degraded in both intelligibility and naturalness because the human hearing mechanism is relatively sensitive to formant locations.

A number of proposals have been made in an effort to improve the accuracy of formant detection; for example, see H. L. Barney Patent 2,819,341 issued Jan. 7, 1958, and M. R. Schroeder Patent 2,857,465 issued Oct. 21, 1958. However, the improvement in accuracy of formant detection is achieved at the expense of an increase in the complexity of formant locating equipment.

The present invention provides a speech coding system based primarily upon the formant coding concept, which exhibits a moderate compression ratio and which is free from the errors inherent in conventional formant extraction. In its principal form, the present invention is presented in terms of a digital implementation which requires relatively simple equipment and which provides for digital transmission of speech information at data rates substantially less than those associated with conventional digital transmission of speech. Instead of attempting to measure the frequency of formants within each of a number of subbands, the present invention transmits a signal indicative of the actual axis crossings of each of a number of contiguous frequency bands of the incoming speech wave, thereby avoiding the errors that arise from attempting to derive a slowly varying signal representative of the frequency of a specific, single speech formant. In the present invention, the frequency bands are selected to be octave bands, that is, the upper cutoff frequency of each band is twice the lower cutoff frequency of that band, in order to provide favorable discrimination against quantizing distortion which might otherwise appear in the subsequent digital processing. The axis crossings of each octave band are supplemented by an amplitude signal representative of the short-time average energy within each octave band, so that for each octave band a pair of signals is derived, an axis-crossing signal and an amplitude signal. The signals of each pair are then sampled and quantized to provide an efficient digital representation. In digital form, the pairs of axis-crossing and amplitude signals have a substantially lower bit rate than a comparable digital version of the original speech wave. After transmission over a channel having a reduced digital information carrying capacity, each pair of digital signals is desampled to an analog form, and the analog amplitude signal in each pair is used to adjust the amplitude

of the analog axis-crossing signal in order to restore natural amplitude variations to each analog axis-crossing signal. The amplitude-adjusted axis-crossing signals therefore constitute approximations of the corresponding octave bands, which when combined, form a natural sounding replica of the original speech wave.

Another important improvement over conventional formant vocoder arrangements is its elimination of the necessity for supplementing the formant signals with a so-called voiced-unvoiced pitch signal representative of the characteristics of the excitation source applied to the talker's vocal tract. The pitch signal indicates which portions of the speech wave are voiced sounds and which are unvoiced sounds, and for voiced portions, the pitch signal also indicates the pitch or fundamental frequency of the excitation that produced the sound. It has long been recognized that accurate pitch detection presents one of the most difficult problems in a workable bandwidth compression system since accurate pitch detection is essential to reconstruction of natural sounding speech.

The present invention avoids the difficulties inherent in attempting to measure pitch by transmitting the axis crossings of a number of octave bands, therein preserving sufficient fundamental pitch periodicity to provide pitch information in the reconstructed speech wave for the listener. As mentioned above and explained more fully in the Peterson article previously cited, the axis-crossing rate of an octave band is determined by the frequency of the speech frequency component with greatest energy within the octave band. Since the frequency components of voiced speech are harmonics of the fundamental pitch frequency, this means that the amplitude-adjusted axis-crossing signals which are combined at the receiver station are, in essence, harmonically related. Now it is well known that if two harmonically related waves are combined, the sum wave has an envelope that is periodic in the fundamental frequency of which the individual waves are harmonics. Hence the replica wave obtained in this invention by combining the individual-amplitude adjusted axis-crossing signals has an envelope that is essentially periodic in the fundamental pitch frequency, provided that the axis-crossing rates of not less than two of the axis-crossing signals are harmonics of the fundamental pitch frequency. It was pointed out above that the axis-crossing rate of an octave band may at times represent an average of the frequencies of two speech components when two formants approximately equal in amplitude occur in the same band, and since this average may not be a pitch harmonic, it is desirable to provide at least three and desirably four octave bands in order to ensure that at least two axis-crossing rates are always harmonics of the fundamental pitch frequency.

Another important feature of this invention is its elimination of unwanted noise during intervals when speech energy is not present. In the digital implementation of this invention, during the absence of speech, residual low-level noise is unavoidably encoded and transmitted as part of the axis-crossing signal. If permitted to be synthesized in its full amplitude, that is, as essentially infinitely clipped noise, the effect would be exceedingly annoying, and it could impair both the intelligibility and the naturalness of the reconstructed wave. However, by suitably adjusting the threshold of the amplitude signal so that zero amplitude is indicated in the absence of speech energy, the unwanted noise portions of the axis-crossing signal are eliminated in the course of adjusting the amplitude of the axis-crossing signal at the receiver; that is, in the absence of speech energy, the amplitude signal acts as a squelch signal by reducing the amplitude of the axis-crossing signal to zero.

The invention will be fully understood from the following detailed description of illustrative embodiments thereof taken in connection with the appended drawings, in which:

FIG. 1 shows apparatus for transmitting the informa-

tion content of a speech wave over a channel having a relatively low information-carrying capacity; and

FIG. 2 illustrates certain components of the apparatus of FIG. 1.

Referring first to FIG. 1, this drawing illustrates a complete speech bandwidth compression system embodying the principles of this invention. A speech sound wave is converted by transducer 10 into a corresponding electrical wave which is delivered in parallel to a plurality of bandpass filters 11-1 through 11-n, where transducer 10 may be a conventional microphone. The pass bands of filters 11-1 through 11-n are selected to divide the incoming speech wave into contiguous octave bands, or narrower bands if desired. Thus filter 11-1 may be provided with a pass band which extends from f_1 cycles per second to $2f_1$ cycles per second, filter 11-2 may be provided with a pass band that extends from $2f_1$ cycles per second to $4f_1$ cycles per second, and filter 11-n may be provided with a pass band that extends from $2^{n-1}f_1$ cycles per second to $2^n f_1$ cycles per second.

In general, the pass bands of filters 11 are selected to divide the incoming speech wave into contiguous octave bands each having a relatively high probability of containing no more than a single speech formant. However, it is important to observe at this point that it is not necessary that each octave band contain no more than one formant, since the present invention avoids the difficulties inherent in detecting the frequency locations of formants by transmitting directly and without modification the axis crossings of each octave band as an indication of formant locations. The relationship between axis crossings and formant locations is described by E. Peterson in "Frequency Detection and Speech Formants," volume 23, Journal of the Acoustical Society of America, page 668 (1951). Therefore the occurrence of more than one formant within an octave band does not constitute a potential source of error in the present invention.

By way of example, for an incoming speech wave having a frequency range extending from about 250 cycles per second to 4,000 cycles per second, $n=4$ bandpass filters 11 are sufficient. For such a speech wave, a set of octave bands may be 250 cycles per second to 500 cycles per second for filter 11-1, 500 cycles per second to 1,000 cycles per second for filter 11-2, 1,000 cycles per second to 2,000 cycles per second for filter 11-3, and 2,000 cycles per second to 4,000 cycles per second for filter 11-4. For a speech wave whose bandwidth extends over the range of frequencies from 250 to 4,000 cycles per second, which approximates the bandwidth conveyed by commercial telephone equipment, the choice of octave bands given in the preceding example is particularly desirable. First, the first formant of voiced, non-nasal sounds can span a relatively large frequency range which is on the order of two octaves. In addition, nasalized voiced sounds are often accompanied by a relatively low first formant, a nasal formant in the range around 800 to 1,000 cycles per second, and the normal second and third formants are higher in frequency. The octave bands given in this example, therefore, have a high probability of containing a single speech formant within each band. It is to be understood, however, that in general the number and widths of the octave bands are chosen to fit the specific bandwidth of an incoming speech wave and the capacity of a particular transmission channel.

Each filter is followed by a corresponding coding circuit 12-1 through 12-n, respectively. Since the coding circuits are substantially identical in construction, only circuit 12-1 is shown in detail, it being understood that the other circuits differ only in the adjustments necessary to process octave bands with different frequency ranges. Within each coding circuit the octave band from the preceding filter is applied to two parallel subpaths. As shown in coding circuit 12-1, one of these subpaths includes a rectifier 121, a low-pass filter 122, and an amplitude-sampler and quantizer 123, while the other subpath con-

tains an axis-crossing sampler and quantizer 124. By rectifying and low-pass filtering the octave band there is obtained at the output terminal of filter 122 a unidirectional signal representing the short time average amplitude of the octave band passed by filter 11-1. Element 123, which may be a conventional sampler and quantizer circuit, samples and quantizes this unidirectional signal to obtain a suitable low-bit rate representation of the short time average amplitude of the incoming octave band.

For example, if filter 122 is provided with a cutoff frequency of 20 cycles per second, then at the Nyquist rate element 123 obtains from the unidirectional signal 40 samples per second, and therefore a suitable low-bit rate representation of two binary digits or bits for each of these samples produces a digital signal indicating the short time average amplitude of the incoming octave band at a rate of 80 bits per second.

Axis-crossing sampler and quantizer 124 in the second subpath within coding circuit 12-1 is a conventional band-pass sampling and quantizing circuit that samples the octave band passed by filter 11-1 to obtain a relatively low-bit rate digital signal that indicates the axis crossings of the incoming octave band. A detailed block diagram of a suitable element 124 is shown in FIG. 2 and described below. Since axis crossings may be specified by a simple two-level code that indicates only the instantaneous polarity of the octave band, a suitable digital representation of the axis crossings may be obtained by sampling the octave band at an adequate sampling rate followed by one bit quantizing of each sample. However, it is to be understood that although one bit quantizing affords the maximum saving in bit rate, additional quantizing levels may be employed if desired, with a corresponding increase in transmission bit rate. The one bit representation thus obtained is a digital representation of the analog signal obtained by infinitely clipping speech in the manner described by J. C. R. Licklider and I. Pollack in "Effects of Differentiation, Integration, and Infinite Clipping upon the Intelligibility of Speech," Vol. 20, Journal of the Acoustical Society of America, page 42 (1948).

For example, for an octave band that extends from 250 to 500 cycles per second, an adequate sampling rate is 500 samples per second, so that with a one-bit code element 124 produces a digital axis-crossing signal at a rate of 500 bits per second. This digital axis-crossing signal together with the digital amplitude signal from element 123 comprises the pair of digital signals derived by each coding circuit 12-1 through 12-n from each incoming octave band into which the speech wave is divided.

It is observed that the transmission channel capacity required to convey the 2·n digital signals derived in accordance with the principles of this invention is significantly smaller than that required to convey a conventional digital version of the original speech wave. Thus in the example given previously, in which a speech wave was assumed to have a frequency range extending from 250 cycles per second to 4,000 cycles per second, a sampling rate of 2×3750=7500 samples is required. If a seven bit code is utilized for each sample, as suggested for speech quality comparable to that provided by analog telephone equipment, then the bit rate is 7×7500=52,500. On the other hand, in the compression arrangement provided by this invention, the sampling rate for each of the four octave bands is the sum of the short time average amplitude rate and the axis-crossing rate, that is:

Octave Band (c.p.s.)	Amplitude (b.p.s.)	Axis Crossing (b.p.s.)	Sum (b.p.s.)
250-500	80	500	580
500-1,000	80	1,000	1,080
1,000-2,000	80	2,000	2,080
2,000-4,000	80	4,000	4,080
Total			7,820

where c.p.s. denotes cycles per second and b.p.s. denotes bits per second. Hence the bit rate required for transmission of the digital amplitude and digital axis-crossing signals of this invention is substantially lower than that required for digital transmission of the original speech signal under commercial standards. In fact, in comparison with any conventional digital coding of the original speech wave which provides comparable speech quality, the present invention provides a significantly lower bit rate.

It is also noted that any one of a number of well known quantizers may be employed in elements 123 and 124 and similar elements in the other coding circuits. For example, the quantizers may be of the companded coder variety described by H. Mann, H. M. Straube, and C. P. Villars in "A Companded Coder for an Experimental PCM Terminal," Volume 41, Bell System Technical Journal, p. 173 (1962).

The n pairs of digital signals derived by each coding circuit are transmitted to a receiver station by way of suitable multiplex and transmission equipment (not shown). Because of the relatively low-bit rate of the n pairs of digital signals, a transmission medium having a limited bit rate transmission capacity may be employed. At the receiver station each digital signal is converted into a corresponding analog signal by passing each control signal through an appropriately proportioned desampler of any well-known design. Thus, the digital amplitude signal from coding circuit 12-1 is passed through desampler 131-1, for example, a conventional low pass filter, while the digital axis-crossing signal from coding circuit 12-1 is passed through a suitable band-pass desampler 132-1, which may be of the type shown in FIG. 2. Each desampler thus develops an analog version of the corresponding transmitted digital signal, with each of the analog amplitude signals having a unidirectional waveform roughly approximating the amplitude variations in the original octave band, and each of the analog axis-crossing signals having a waveform approximately uniform in amplitude which indicates only the axis crossings in the original octave band.

The output terminals of each pair of desamplers 131-1 and 132-1 through 131-n and 132-n are respectively connected to the input terminals of suitable multiplier circuits 14-1 through 14-n. Within each multiplier the analog amplitude signal is utilized to adjust the amplitude of the analog axis-crossing signal, thereby restoring a portion of the original, natural amplitude variation to each analog axis-crossing signal. Each amplitude adjusted axis-crossing signal from multipliers 14-1 through 14-n is passed to an adder 15 to be combined to form an approximate replica of the original speech wave. The replica wave may be converted into audible sound by a suitable reproducer 16, for example a conventional loudspeaker.

It is important to point out that no pitch measuring or restoring equipment is required at the transmitter or receiver stations of this invention, yet the replica wave formed by adder 15 follows the periodicity of the original speech wave. This is achieved because the axis-crossing rate conveyed by each of the axis-crossing signals is generally determined by the frequency of that one of the harmonics of the fundamental pitch frequency which has the greatest amplitude within the corresponding octave band. When the amplitude adjusted analog axis-crossing signals are added together in adder 15, the resulting wave has an envelope that is periodic in the fundamental pitch frequency because the individual signals are harmonically related, and the human ear perceives this periodicity even if there is no fundamental frequency component present in the spectrum of the wave. Moreover, even if one of the octave bands has an axis-crossing rate that does not coincide with a harmonic of the fundamental pitch frequency, for example, when two formants occur in the same octave band, it is only

necessary that two of the axis-crossing rates be harmonically related in order for adder 15 to produce a replica wave with an envelope periodic at the fundamental pitch frequency. Hence not only does the present invention avoid the errors and equipment complexities associated with locating formants but also it avoids the errors and equipment complexities associated with measuring and restoring the speech pitch characteristic.

Another advantage realized by the present invention is the suppression by the amplitude signals of unwanted noise which might be present during intervals in which no speech energy is present. Such noise, if present, is sampled and quantized to full scale in the axis-crossing sampling and quantizing circuits, and if not suppressed, would be reproduced in full amplitude in the replica wave and thereby annoy the listener. However, by setting the lowest nonzero quantizing level of element 123 at a high enough threshold to prevent the coding of most noise energy, the digital amplitude signal and its desampled analog counterpart will indicate either the presence or absence of speech energy, so that any coded noise that appears in a digital axis-crossing signal will be suppressed or squelched in the corresponding multiplier at the receiver station.

Turning now to FIG. 2, this drawing illustrates suitable band-pass sampling and quantizing equipment for deriving a digital signal representative of the axis crossings of each octave band, and a suitable desampler for converting the digital axis-crossing signal thus derived into an analog axis crossing signal. Within the sampler and quantizer, an incoming octave band is applied to two parallel subpaths, one subpath comprising a sampling gate 22a and a quantizer 23a connected in tandem, and the other subpath comprising a 90 degree phase shifter 20, a sampling gate 22b, and a quantizer 23b, all connected in series. The two sampling gates 22a and 22b are controlled by sampling pulses supplied from sampling pulse generator 21 at an appropriate rate. Each of the elements shown in the sampler and quantizer may be constructed in accordance with well-known designs.

The band-pass sampling performed by elements 20, 21, 22a and 22b is based upon the well known theory of sampling band-pass functions, such as described in S. Goldman, "Information Theory" page 75 (1953). Thus, for an octave band of frequency bandwidth W cycles per second, gate 22a samples the octave band at a sampling rate of W samples per second, while gate 22b samples the quadrature function or Hilbert transform of the octave band at W samples per second, so that the octave band is being represented by a total of 2W samples per second. This means that generator 21 supplies W sampling pulses per second simultaneously to gates 22a and 22b. In order to obtain the Hilbert transform of the octave band, the octave band is passed through a 90 degree phase shifter of well known design. This sampling technique produces a pair of digital signals having a combined spectrum that is the sum of the original octave band spectrum and its folded negative frequency image.

The samples of the octave band and the samples of its Hilbert transform are quantized to a desired number of levels in quantizers 23a and 23b, and the pair of quantized signals is appropriately coded and transmitted in the manner shown in FIG. 1. Maximum economy in transmission is achieved by two level or one bit quantizing, but additional quantizing levels may be utilized if desired.

After transmission to a receiver station, the pair of digital axis-crossing signals is decoded and then converted into a pair of analog signals by desamplers 24a and 24b, these desamplers being band-pass filters of conventional construction. Following desampling, the pair of analog signals is applied to the input terminals of subtractor 25, in which the Hilbert transform analog signal is subtracted from the other analog signal in order to cancel the folded negative frequency image from the combined spectrum of the two signals. The difference signal developed at the

output terminal of subtractor 25 thereby constitutes the desampled analog axis-crossing signal described above in connection with FIG. 1.

Although this invention has been described in terms of a speech communication system of the type shown in the appended drawing, it is to be understood that applications of the principles of this invention are not limited to the field of speech communication, but include the fields of automatic speech recognition, speech processing, and automatic message recording and reproduction. In addition, it is to be understood that the above-described arrangement is merely illustrative of the numerous arrangements which may be devised for the principles of this invention by those skilled in the art without departing from the spirit and scope of the invention.

What is claimed is:

1. A communication system for conveying the information content of a speech wave over a channel of relatively small capacity which comprises,

means for dividing said speech wave into contiguous frequency subbands,

means for transmitting to a receiver station in coded form each axis crossing of each of said subbands and the short-time average amplitude variations of each of said subbands, and at said receiver station,

means for reconstructing a replica of said speech wave from said axis crossings and said short-time average amplitude variations.

2. A speech communication system in which the information content of an incoming speech wave is conveyed in coded form from a transmitter station to a receiver station which comprises

at said transmitter station,

means for dividing said speech wave into a plurality of contiguous frequency bands,

means for deriving from each of said frequency bands a corresponding pair of first and second coded signals, wherein said first coded signal represents the short-time average amplitude of said corresponding frequency band, and said second coded signal indicates each axis crossing of said corresponding frequency band,

and at said receiver station,

a plurality of means each under the control of a corresponding one of said first coded signals for adjusting the amplitude of each of said second coded signals to form a plurality of amplitude-adjusted signals corresponding to said plurality of frequency bands, and means for combining said plurality of amplitude-adjusted signals to form a replica of said speech wave.

3. Apparatus as defined in claim 2 wherein said plurality of contiguous frequency bands comprises a plurality of contiguous octave bands, wherein the upper cutoff frequency of each octave band is twice the lower cutoff frequency of that octave band.

4. A speech communication system for conveying in coded form the information content of a speech wave which comprises

a source of an incoming speech wave,

a transmitter station including

means for dividing said speech wave into a plurality of contiguous octave bands,

means for deriving a corresponding plurality of pairs of first and second digital coded signals from said plurality of octave bands, wherein

each of said first digital coded signals represents in digital form the short-time average amplitude of said corresponding octave band, and

each of said second coded signals represents in digital form each axis crossing of said corresponding octave band.

5. Apparatus as defined in claim 4 wherein said means for dividing said speech wave into a plurality of contiguous octave bands comprises a plurality of bandpass filters having contiguous pass bands, wherein the upper

cutoff frequency of each pass band is twice the lower cutoff frequency of that pass band.

6. Apparatus as defined in claim 4 wherein said means for deriving from each of said octave bands a corresponding pair of first and second coded signals comprises a pair of parallel first and second subpaths,

wherein said first subpath includes rectifier means, low-pass filter means, and first sampler and quantizer means, all connected in series, for deriving said first coded signal, and

wherein said second subpath comprises a second sampler and quantizer means for deriving said second coded signal.

7. Apparatus as defined in claim 6 wherein said second sampler and quantizer means comprises

a sampling pulse generator for supplying sampling pulses at a predetermined rate of W pulses per second, where W is the bandwidth in cycles per second of the corresponding octave band,

means for deriving the Hilbert transform of said octave band by shifting the phase of said octave band by ninety degrees,

a pair of first and second sampling gates controlled by said sampling pulses for simultaneously sampling said octave band and its Hilbert transform at W samples per second, and

quantizing means for representing each of said octave band samples and each of said Hilbert transform samples in terms of a two level code to obtain said second digital coded signal.

8. In combination with the apparatus defined in claim 4, a receiver station for reconstructing a replica of said incoming speech wave from said plurality of pairs of first and second digital coded signals, which comprises

a plurality of corresponding pairs of first and second desampling means for converting each of said pairs of first and second digital coded signals from digital form to analog form to develop a corresponding plurality of pairs of analog coded signals,

a plurality of corresponding amplitude-adjusting means each under the control of a corresponding one of said first analog coded signals in each pair of analog coded signals for adjusting the amplitude of the corresponding second analog coded signal in each pair of analog coded signals to reconstruct a plurality of amplitude-adjusted signals each of which is a replica of a corresponding one of said octave bands, and means for combining said plurality of amplitude-

adjusted signals to synthesize a replica of said incoming speech wave.

9. Apparatus as defined in claim 8 wherein each of said second desampling means for converting a second digital coded signal into an analog coded signal, said second digital coded signal comprising quantized samples of the corresponding octave band and quantized samples of the Hilbert transform of said octave band, comprises

a first desampling filter provided with an input terminal supplied with said quantized octave band samples and an output terminal,

a second desampling filter provided with an input terminal supplied with said quantized samples of the Hilbert transform of said octave band and an output terminal,

subtracting means provided with a minuend input terminal, a subtrahend input terminal, and an output terminal,

means for connecting the output terminal of said first desampling filter to said minuend input terminal, and

means for connecting the output terminal of said second desampling filter to said subtrahend input terminal, thereby to develop said analog coded signal at said output terminal of said subtracting means.

10. The method of conveying the information content of a speech wave over a channel having a capacity substantially smaller than that required for facsimile transmission of the speech waveform, which comprises the steps of

dividing said speech wave into contiguous frequency subbands,

transmitting each axis crossing of each of said subbands to a receiver station,

transmitting the short-time average amplitude variations of each subband to said receiver station, and at said receiver station,

combining said short-time average amplitude variations with said axis crossings to form a replica of said speech wave.

References Cited

UNITED STATES PATENTS

3,102,928 9/1963 Schroeder ----- 179—1

ROBERT L. GRIFFIN, *Primary Examiner*.

W. S. FROMMER, *Assistant Examiner*.