United States Patent

[11] 3,555,191

[72] Inventor A. Michael Noll
Newark, N.J.
[21] Appl. No. 744,931
[22] Filed July 15, 1968
[45] Patented Jan. 12, 1971
[73] Assignee Bell Telephone Laboratories, Incorporated
Murray Hill, Berkeley Heights, N.J.
a corporation of New York

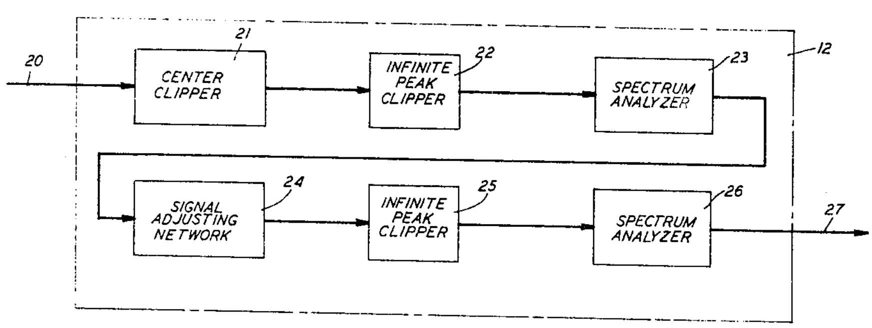
[54] **PITCH DETECTOR**
7 Claims, 9 Drawing Figs.

[52] U.S. Cl. **179/1**
[51] Int. Cl. **G10l 1/00, G10l 1/04**
[50] **Field of Search** 179/1AS, 15.55; 324/77A, 77B

[56] **References Cited**
UNITED STATES PATENTS
3,381,091 4/1968 Sondhi .... 179/1(AS)
OTHER REFERENCES
IEEE Transactions on Audio and Electroacoustics "New Methods of Pitch Extraction," Man Mohan Sondhi, Vol. AU—16, No. 2 June 1968

Primary Examiner—Kathleen H. Claffy
Assistant Examiner—Jon Bradford Leaheey
Attorneys—R. J. Guenther and William L. Keefauver

ABSTRACT: The pitch of a complex speech wave is determined by spectrum analyzing the infinitely peak-clipped log spectrum of a center-clipped and infinitely peak-clipped interval of an analogue speech wave.



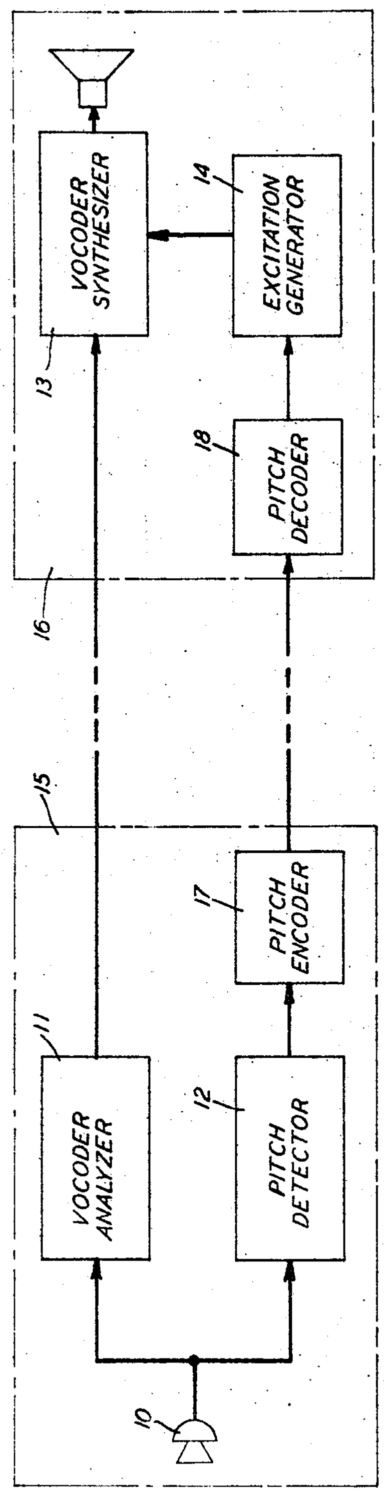


FIG. 1

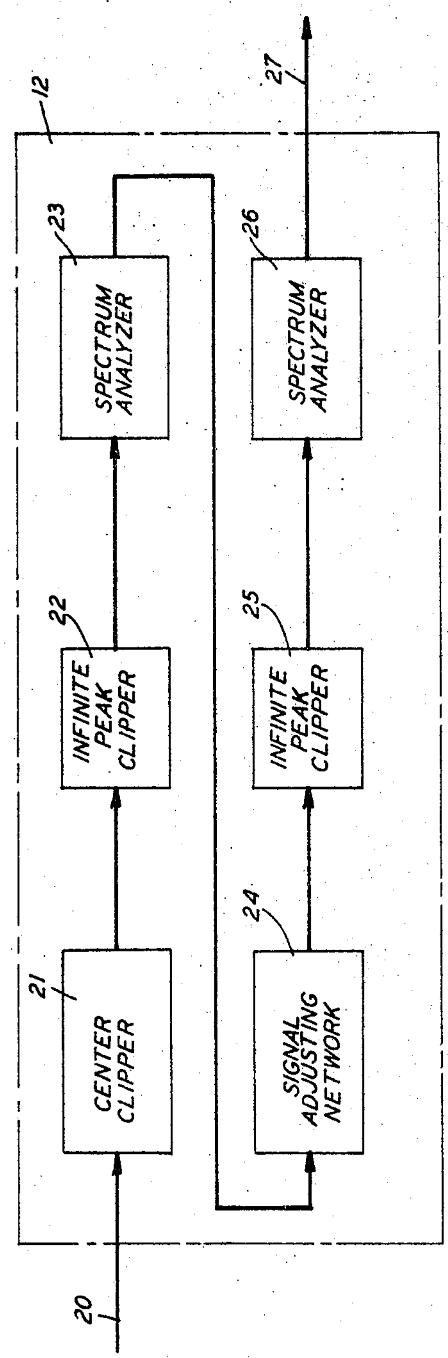
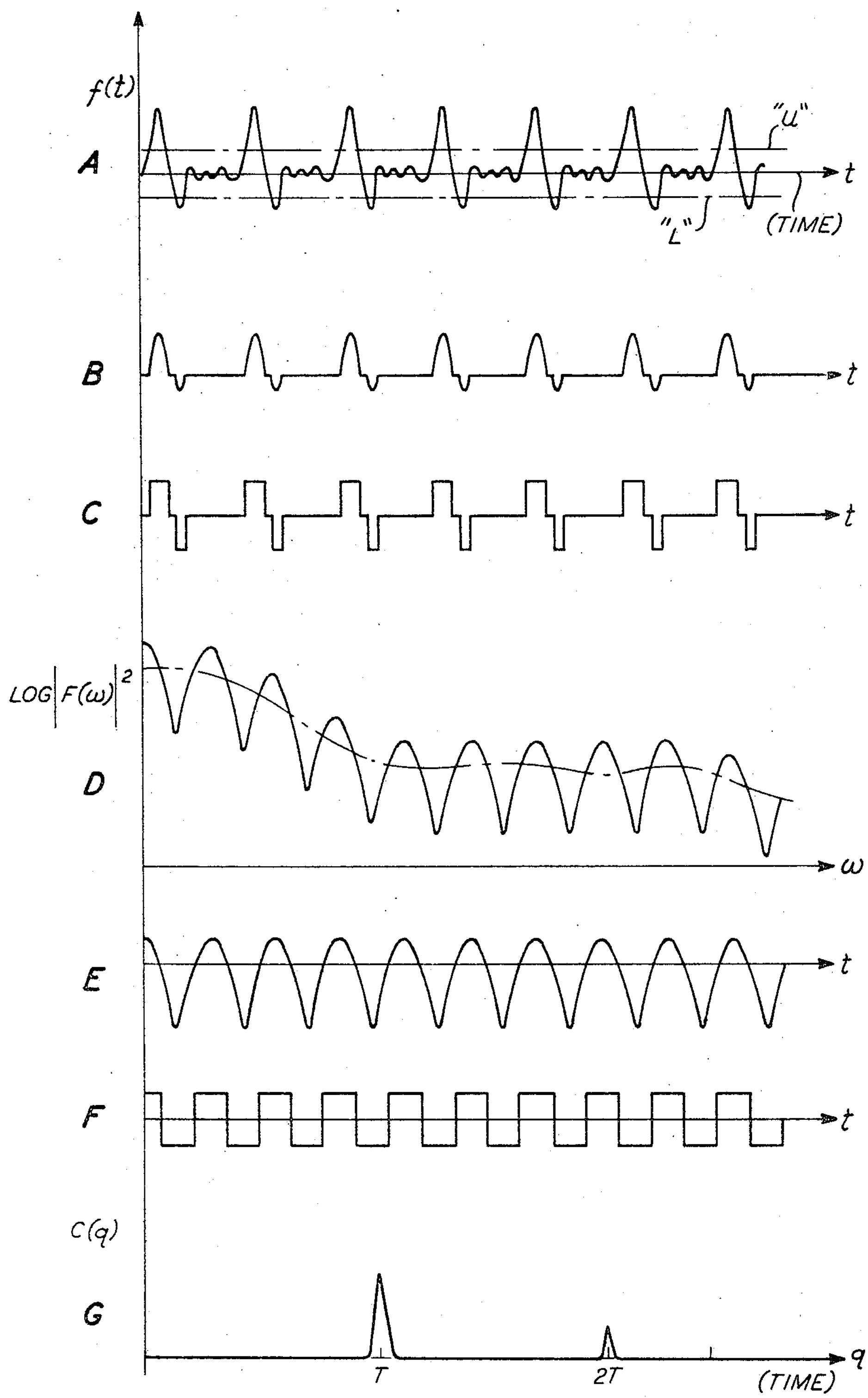


FIG. 2

INVENTOR
A. M. NOLL
BY
G. E. Hirsch Jr.
ATTORNEY

FIG. 3



PITCH DETECTOR

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to the narrow band transmission of speech, and in particular to apparatus for identifying and analyzing the pitch of complex speech waves.

In the processing of complex speech waves it is often important to determine whether a particular portion of such wave is periodic or aperiodic and, if the wave is periodic, to determine its period or pitch. For example, in communications systems of the vocoder type only selected characteristics of a complex speech wave are transmitted to a receiving station which synthesizes an artificial replica of the original speech signal. In such systems the voiced or unvoiced quality of the speech signal (voiced speech being periodic, unvoiced being aperiodic) and the pitch of the signal where voiced speech is present are among the most important characteristics transmitted. It is particularly important that pitch be determined accurately since small errors in pitch detection create a distorted and unnatural sounding output speech wave. An accurate wave analyzer for the purpose of voice pitch detection has been long sought.

2. Description of the Prior Art

In a copending application Ser. No. 420,362 filed Dec. 22, 1964 apparatus is described for determining the pitch of a complex wave such as a voiced speech wave by "cepstrum" analysis. According to the spectrum method, which is thought to be a substantial improvement over prior techniques, the square of the Fourier cosine transform of the logarithm of the power spectrum of a segment of the speech signal (defined as the "cepstrum") is computed. The resulting spectrum signal is characterized by a peak at an interval proportional to the fundamental pitch period during voiced or periodic portions of the speech signal, and by the absence of a peak during unvoiced or aperiodic portions of the signal. The theoretical considerations underlying spectral analysis and apparatus for carrying out such analysis are adequately described in the application referred to above and in an article entitled "Spectrum Pitch Determination" in the Journal of the Acoustical Society of America, Vol. 41, No. 2, pp. 293—309, Feb. 1967 .

However, as the referenced application indicates, spectrum analysis requires substantial signal processing involving repeated multiplication operations. If digital processing is employed, such multiplications require elaborate and expensive digital computing apparatus. It would thus be desirable to simplify the digital processing apparatus required to obtain the accuracy and reliability of spectrum pitch detection. This could be accomplished if the numerous digital multiplications could be eliminated and replaced with simpler additive processing.

Thus, it is an object of the present invention accurately to detect and analyze the periodicity of a complex wave with comparatively simple processing apparatus.

SUMMARY OF THE INVENTION

In attaining this and other objects and in accordance with the invention, a selected interval of a complex wave is center-clipped, thus removing a selected fraction of the amplitude of the complex wave. The center-clipped signal is infinitely peak-clipped and the logarithm of the power spectrum of the resulting signal is computed. The clipped power spectrum is treated as a time varying signal and as such is low-pass filtered to remove DC and selected low frequency components and again infinitely peak-clipped. The square of the Fourier cosine transform of the resulting signal, which is defined as the "clipstrum" of the speech signal, is then examined. The presence of peaks in this function indicates the presence of periodic waves in the input signal. Where periodic waves are present, the location of peaks in the clipstrum establish the period of such waves.

BRIEF DESCRIPTION OF THE DRAWING

The invention will be fully apprehended from the following description of an illustrative embodiment thereof taken in conjunction with the appended drawings wherein:

FIG. 1 is a block diagram illustrating a vocoder transmission system employing a pitch analyzer constructed in accordance with the invention;

FIG. 2 is a block schematic drawing of a pitch analyzer constructed in accordance with the invention and suitable for inclusion in a system of the type shown in FIG. 1;

FIG. 3A illustrates a typical speech waveform;

FIG. 3B is a center-clipped speech waveform;

FIG. 3C depicts a center-clipped and infinite peak-clipped speech waveform;

FIG. 3D shows a typical frequency spectrum of a center-clipped and infinitely peak-clipped speech waveform;

FIG. 3E is the waveform of FIG. 3D after low frequency and DC components are filtered out;

FIG. 3F is the waveform of FIG. 3E after infinite peak-clipping; and

FIG. 3G illustrates the "clipstrum" of a speech waveform.

DETAILED DESCRIPTION

The vocoder system shown in FIG. 1 can be employed in the transmission of speech or other complex signals over narrow band communications channels. The system includes a transducer, **10**, for converting an acoustic speech wave into an electrical analogue wave which occupies a frequency band approximately equivalent to the speech band. Direct transmission of such an analogue signal would normally require a transmission channel capable of accommodating the normal speech frequency band. The vocoder apparatus shown in FIG. 1 permits transmission of such a voice band signal over a narrow band communications channel by converting the analogue speech signal at the transmitting station **15** into a relatively small number of low frequency control signals which together occupy only a small frequency band but which adequately describe the speech signal. A replica of the speech signal is then synthesized at the receiving station **16** from these control signals.

The vocoder system shown in FIG. 1 includes, at the transmitting station **15**, a vocoder analyzer **11** which may for example be a conventional channel vocoder analyzer of the type shown in H. W. Dudley U.S. PAT. No. 2,151,091 issued Mar. 21, 1939 . Such analyzer derives from an input speech wave a group of narrow band control signals representing in coded form the energy within each of a number of selected frequency subbands of the speech signal. Also at the transmitting station, a pitch detector **12** to be described in detail below and a pitch encoder **17** derive and code a pitch control signal which indicates the presence and pitch of voiced speech.

The control signals generated by the vocoder analyzer and the pitch coder are transmitted over a limited bandwidth communications channel to receiving station **16**. The receiver includes a vocoder synthesizer **13** and a pitch decoder **18** in conjunction with an excitation generator **14**. The pitch decoder and excitation generator receive the coded pitch signal and generate an excitation wave with a fundamental frequency equivalent to the fundamental frequency of voiced speech in the original speech signal. The excitation signal is applied to the vocoder synthesizer which also receives the subband control signals and reconstructs a replica of the original speech wave by combining the excitation wave with the subband control signals. The vocoder synthesizer may be a conventional channel synthesizer of the type described in the above referenced Dudley patent and the pitch coder, pitch decoder and excitation generator may be of a type described in H. S. McDonald U.S. Pat. No. 3,190,142 issued Oct. 29, 1963.

It has been found that the quality of the speech signal reproduced by the vocoder apparatus described above is largely dependent on the accuracy of the pitch detector employed. Thus, in accordance with the invention, such pitch information is provided by a pitch detector of the type shown in detail in FIG. 2.

Referring to FIG. 2, speech detector 12 is designed to analyze a conventional analogue speech signal or other complex wave and to indicate the existence of voiced or periodic wave energy in that signal. When such periodic energy is present, detector 12 provides an output signal with a spike located at a time proportional to the fundamental period of the periodic wave energy in the input wave. Secondary spikes of smaller amplitude may be produced at intervals equal to multiples of the fundamental period, but these can be disregarded. When no periodic energy is present, no spike occurs. The output signal of detector 12 may thus be applied to an encoding network such as pitch encoder 17 in FIG. 1 which interprets the spike location and produces an appropriate control signal for communicating the pitch of a periodic input signal to a receiving device. This network may include apparatus for detecting peaks in a signal such as that described in a copending application Ser. No. 508,726 filed Nov. 19, 1965 by A. M. Noll now U.S. Pat. No. 3,420,955 . When the input signal is not periodic, this fact too may be communicated to the receiver.

Thus, in FIG. 2, the input signal applied to channel 20 may be a segment of a conventional analogue speech signal prepared in a manner well known in the speech processing art. Such a signal is shown in FIG. 3A. The input signal is applied to center-clipper 21, of any construction well known in the electronic arts, which has the effect of removing a center segment of the amplitude of the input wave at a selected level above and below the zero axis. The signal levels marked "U" for upper threshold and "L" for lower threshold in FIG. 3A denote one possible range of center-clipping. FIG. 3B shows the signal of FIG. 3A center-clipped between U and L. It is to be understood that the actual threshold levels U and L for center-clipper 21 may vary from signal segment to signal segment since the levels are established for each segment by applying a constant percentage to the maximum peak amplitude in the signal. The percentage of the input wave removed by center-clipper 21 may be selected in accordance with the nature of the input wave applied. It has been found that center-clipping on the order of 70 percent of the absolute maximum of the input wave in each interval is most effective in analyzing a common speech signal. Center-clipping circuits suitable for use in the network shown in FIG. 2 are described by M. M. Sondhi in U.S. Pat. No. 3,381,091 issued Apr. 30, 1968.

The center-clipped output of network 21 is applied to infinite peak-clipper 22. Peak-clipping is analogous to center-clipping except that, rather than eliminating the central amplitude portion of the input wave as center-clipping does, a peak-clipper eliminates the extreme high and low amplitude portions of the wave, leaving only the intermediate section. "Infinite" peak-clipping removes the entire wave structure except for that occurring in the immediate vicinity of the zero axis. In effect, only the zero crossing information is retained after infinite peak-clipping. One possible output of infinite peak-clipper 22 is shown in FIG. 3C. This signal has an arbitrary amplitude selected for convenience and zero crossings directly related to the zero crossings in the waveform shown in FIG. 3B. It is to be understood that other waveforms which maintain only zero crossing information could be produced by infinite peak-clipper 22. Infinite peak-clipping networks suitable for inclusion in the system shown in FIG. 2 are well known in the electronic arts.

The infinitely peak-clipped output of network 22 is applied to log spectrum analyzer 23. Analyzer 23 produces a signal of the form shown in FIG. 3D which represents the amplitude of the various frequency components of the applied signal plotted versus frequency. It is observed in FIG. 3D that the spectrum has the appearance of a waveform characterized by a fine wave structure superimposed upon a coarse wave structure. In the case where the wave applied to analyzer 12 is a speech wave, the long wavelength peaks in FIG. 3D represent remnants of the formant structure of the initial wave and other disturbances and the period of the short wavelength peaks represents the fundamental frequency of the incoming speech wave.

Log spectrum analyzer 23 may be any one of numerous spectrum analyzing devices. It may be an analogue heterodyne spectrum analyzer of the type described in the aforementioned copending application filed Dec. 22, 1964, Ser. No. 420,362 or may be similar to the analyzer described by M. R. Schroeder in U.S. Pat. No. 3,321,582 . It is to be noted that the analyzer may be analogue or a digital device. If a digital spectrum analyzer is employed, the signal processing is very much simplified by the fact that the input signal to the analyzer takes the form of a square wave.

Whatever form of spectrum analyzer is employed, the analyzer output is applied to signal adjusting network 24 wherein the DC and low frequency variations in the analyzer output signal are removed. If the signal resulting from analyzer 23 is treated as a time varying signal, this process is in effect a filtering process in the time domain. Since the output of analyzer 23 may be either in digital or analogue form, the adjusting apparatus of network 24 is selected accordingly. Digital or analogue filters suitable for performing this adjusting function are well known in the signal processing art.

The output of adjusting network 24, shown in FIG. 3E is treated as a time varying signal and is applied to infinite peak-clipper 25 which is similar in design and operation to infinite peak-clipper 22 described above. The output waveform of clipper 25 takes the form of a square wave as shown in FIG. 3F. It will be seen that the output of clipper 25, shown in FIG. 3F, is more regularly periodic than the initial speech wave shown in FIG. 3A. This regular periodicity, which is related to the fundamental period of the voiced speech elements in the initial speech signal, is detected by spectrum analyzer 26 which is similar to analyzer 23 described above. The output of analyzer 26, which appears in channel 27, contains a spike as shown in FIG. 3G at a time proportional to the fundamental period of the voiced components of the input speech wave. Smaller amplitude spikes representing higher order harmonics may also appear. If the input signal does not contain voiced or periodic speech elements, the high frequency component in the output of network 23, shown in FIG. 3D, will not exist and no spike will appear in the output signal from analyzer 26. Thus the absence of a spike in the signal output from network 26 can be taken to indicate the absence of voiced speech in the input wave.

It is to be understood that the above-described arrangements are merely illustrative of the invention. Other arrangements may be devised by those skilled in the art without departing from the spirit and scope of the invention.

I claim:

1. Apparatus comprising, in combination;
means for center-clipping a complex wave,
means supplied with signals from said center-clipping means for infinitely peak-clipping said center-clipped complex wave;
first means for developing a signal representative of the logarithm of the power spectrum of said infinitely peak-clipped complex wave;
means for removing selected low-frequency variations in said spectrum representative signals to produce an adjusted spectrum, signal;
means for infinitely peak-clipping said adjusted signal; and
second means for developing an output signal representative of the power spectrum of said infinitely peak-clipped adjusted signals.

2. Apparatus as defined in claim 1 further including means for identifying the existence of peaks in said output signal.

3. Apparatus as defined in claim 2 further including means for measuring the location on the time axis of peaks in said output signal.

4. Apparatus as defined in clam 3 wherein said first and second for means for developing spectrum signal are heterodyne spectrum analyzers.

5. Apparatus as defined in claim 3 wherein said first and second analyzer means are digitized Fourier transform analyzers.

6. Apparatus which comprises, in combination;

a source of complex wave signals;

a center-clipping network supplied with said complex wave signals;

a first infinite peak-clipping network supplied with signals from said center-clipping network;

a log spectrum analyzer network for analyzing the frequency components of signals produced by said infinite peak-clipping network;

a signal adjusting network for removing the slow variations in the output from said first spectrum analyzer;

a second infinite peak-clipping network supplied with signals from said signal adjusting network; and

a second spectrum analyzing network for analyzing the frequency components of signals from said second infinite peak-clipping network.

7. Apparatus for analyzing the periodicity of a complex wave which comprises:

means for removing a selected central portion of the amplitude of a selected interval of said complex wave;

means for generating a first square wave signal with axis crossings related to the axis crossings in the signal produced by said removing means;

means for generating a signal related to the power spectrum of said square wave signal;

A means for producing a logarithm signal proportional to the logarithm of said power spectrum signal;

means for removing the slow variations of said logarithm signal;

means for generating a second square wave signal with axis crossings related to the axis crossings of said logarithm signal; and

means for generating a signal proportional to the power spectrum of said second square wave signal.