Oct. 7, 1969 R. L. MILLER 3,471,648

VOCODER UTILIZING COMPANDING TO REDUCE BACKGROUND NOISE CAUSED BY QUANTIZING ERRORS

Filed July 28, 1966 2 Sheets-Sheet 1

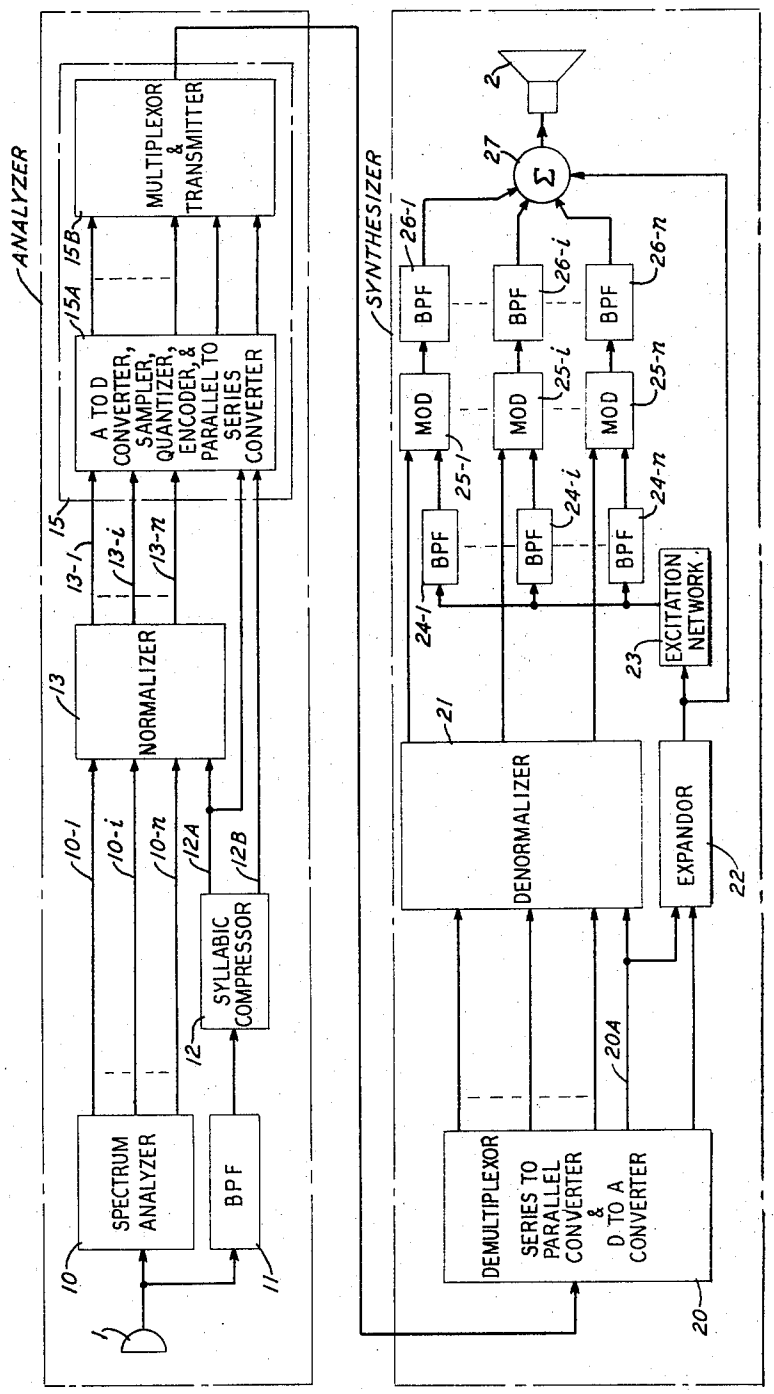

FIG. 1

INVENTOR
R. L. MILLER
BY
G. E. Hirschfeld
ATTORNEY

Oct. 7, 1969 R. L. MILLER 3,471,648
VOCODER UTILIZING COMPANDING TO REDUCE BACKGROUND NOISE CAUSED BY QUANTIZING ERRORS
Filed July 28, 1966 2 Sheets-Sheet 2

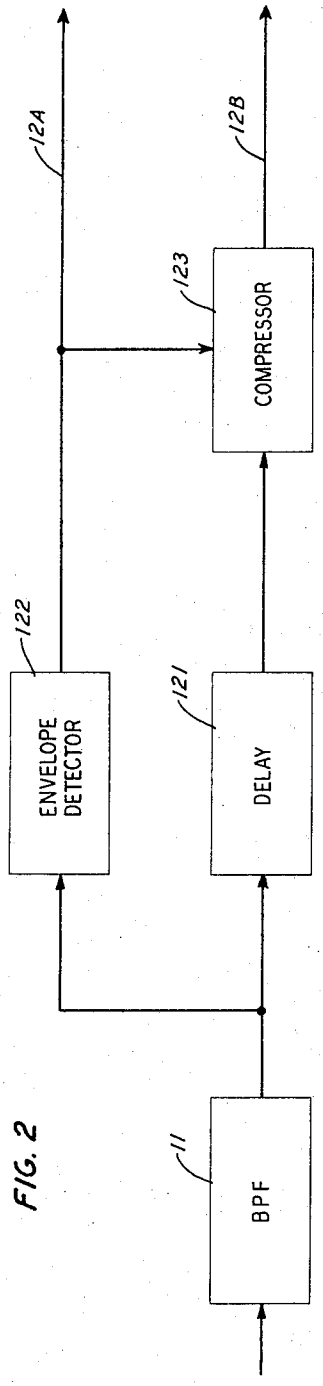

FIG. 2

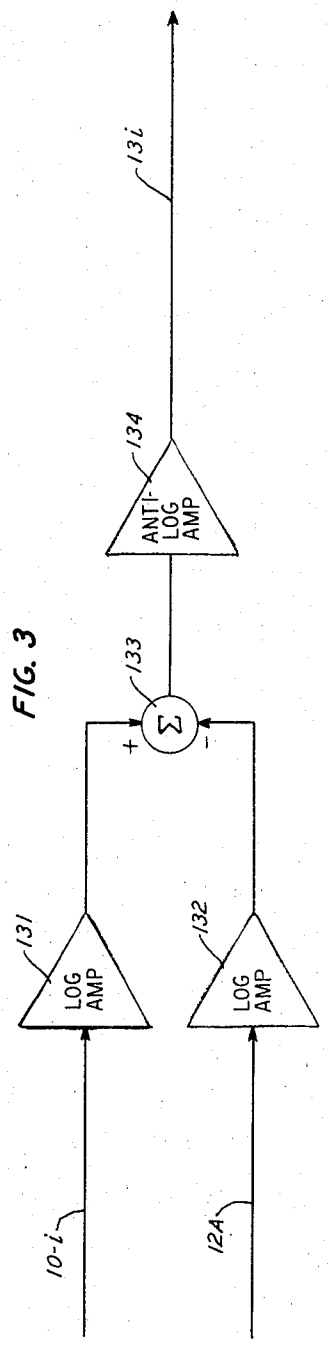

FIG. 3

United States Patent Office 3,471,648 Patented Oct. 7, 1969

3,471,648

VOCODER UTILIZING COMPANDING TO REDUCE BACKGROUND NOISE CAUSED BY QUANTIZING ERRORS

Ralph L. Miller, Chatham, N.J., assignor to Bell Telephone Laboratories, Incorporated, Murray Hill and Berkeley Heights, N.J., a corporation of New York

Filed July 28, 1966, Ser. No. 568,455
Int. Cl. H04b 1/66
U.S. Cl. 179—15.55 5 Claims

ABSTRACT OF THE DISCLOSURE

Quantizing noise is reduced in a digital voice excited vocoder by companding the baseband signal at a syllabic rate prior to transmission. Because the baseband signal is already delayed to ensure its synchronization with the low frequency vocoder control signals, no additional delay is necessary to ensure its synchronization with the syllabically varying compandor control signal. Thus, the quality of the transmitted speech signal is significantly increased over that obtained in a conventional digital voice excited vocoder and this without a commensurate increase in vocoder complexity. Further, because the compandor control signal represents essentially the amplitude of the first speech formant during voiced speech and is transmitted to the vocoder synthesizer for use in decoding, this control signal can be used as a normalizing signal at both the vocoder analyzer and synthesizer without any increase in bit rate.

This invention relates to the transmission of speech signals and in particular to the transmission of speech signals over a channel of limited bandwidth. A principal object of this invention is to improve the quality of the output speech signal from a digital voice excited vocoder.

Numerous systems have been proposed for reducing the bandwidth required to transmit speech signals. One of the best known is Dudley's channel vocoder described in Patent No. 2,151,091, issued March 21, 1939. Dudley's vocoder operates on an input speech signal to derive a low frequency signal proportional to the pitch, or fundamental frequency, of the speech signal, and a set of low frequency control signals representative of the energy in selected contiguous frequency bands of the speech. These derived signals are transmitted over a narrow bandwidth transmission channel to a synthesizer where they are used to produce a replica of the input speech signal. While in theory the speech synthesized by this and similar systems is a good replica of the input speech, in practice, the synthesized speech sounds artificial.

Various explanations have been offered for this. One is that the variations with time of the phases and amplitudes of the frequency components of the synthesized speech often differ significantly from the variations with time of the phases and amplitudes of the frequency components of the live speech. As a result, a marked decrease in the quality of the synthesized speech is often noticeable compared to the quality of live speech.

Considerable attention has been devoted to improving the quality of speech synthesized by a vocoder. The voice excited vocoder disclosed in Patent No. 3,030,450 by M. R. Schroeder achieves a noticeable improvement in speech quality by transmitting unaltered to the synthesizer a so-called baseband signal containing the pitch frequency, if present, and several harmonics thereof. A measure of the speech energy in the frequencies above the baseband signal is sent to the synthesizer in the form of low frequency control signals. At the synthesizer an excitation signal, derived from the baseband signal, is operated upon by the low frequency control signals, and the resulting

signals, together with the baseband signal, are used to form a replica of the input speech signal. The quality of the resulting speech is significantly higher than the quality of speech obtained from vocoders which do not make use of a baseband signal. This is because the synthesized speech contains the original, unaltered pitch signal, if present, and several of its harmonics.

However, when the voice excited vocoder is used in a digital transmission system, disturbing background noise is often present. This noise is caused by quantizing errors in the analog to digital encoding process and its amplitude is enhanced by the large variations in amplitude with time of most input speech signals. Heretofore, to reduce this noise, the number of bits in the binary code words used to represent samples of the low frequency control signals and the baseband signal have been increased. However, such an increase results in a commensurate increase in the bandwidth or bit rate required to transmit the speech signal.

This invention, on the other hand, reduces this quantizing noise and at the same time significantly reduces the bit rate from the rate usually associated with a digital voice excited vocoder capable of producing an equivalent high quality speech signal. This improvement in speech quality is achieved without, unexpectedly, a commensurate increase in vocoder complexity.

According to this invention, the amplitude of the baseband signal is nonlinearly altered prior to encoding to reduce or compress the range of amplitude variations which must be quantized and encoded. The altered baseband signal is transmitted to the synthesizer where it is expanded to its original form. This compression-expansion process, known as "companding," significantly improves the quality of the synthesized speech by reducing the range of baseband signal amplitudes which must be quantized into a given number of steps. Thus the quantizing noise is reduced.

A control signal, called the compandor control signal, is derived from the baseband signal prior to encoding and is used to control the amount of compression and expansion of the baseband signal. The derivation of this compandor control signal introduces a 12 to 20 millisecond delay in the compandor control signal relative to the baseband signal. This is approximately equal to the delay of the low frequency control signals relative to the baseband signal in a voice excited vocoder. Since the baseband signal already is delayed to ensure its synchronization with the low frequency control signals at the vocoder synthesizer, the baseband signal can be altered, that is, companded, without the necessity of having to provide an additional delay network to synchronize the baseband signal with the compandor control signal at the vocoder synthesizer.

Further, because during voiced speech the compandor control signal represents essentially the amplitude of that section of the speech spectrum possessing maximum energy, that is, the first formant, it is an ideal normalizing signal. Accordingly, the compandor control signal is used to normalize the low frequency vocoder control signals. The normalizing essentially eliminates the variations of speech amplitude with time leaving only the variations of speech amplitude with frequency to be encoded. No increase in bit rate attributable to the normalizing occurs because the compandor control signal must be transmitted to the vocoder synthesizer anyway to control the restaration of the altered baseband signal to its original form.

This invention may be more fully understood from the following detailed description taken in conjunction with the figures in which:

FIG. 1 shows one embodiment of a digital voice excited vocoder constructed in accordance with this invention;

FIG. 2 shows one embodiment of the syllabic speech compressor **12** shown in FIG. 1; and

FIG. 3 shows an embodiment of the normalizing circuit used in this invention.

In FIG. 1, an input speech signal detected by transducer **1** is analyzed by spectrum analyzer **10** which produces n low frequency control signals representative of the speech energy in n selected frequency bands, where n is a selected positive integer. Such spectrum analyzers are well known in the vocoder art and thus analyzer **10** will not be described in detail.

The input speech signal is simultaneously passed through bandpass filter **11** wyith cutoff frequencies selected to bracket the pitch frequency, if present. Such cutoff frequencies might be, for example, 80 and 900 cycles per second. That portion of the input speech signal passed by filter **11**, that is, the signal in the 80–900 c.p.s. passband of the filter, is known as the baseband signal.

The baseband signal is operated upon or altered in compressor **12** so that its average rectified amplitude in any given pitch period is made to approach a selected value. This reduces the range of amplitude variations of the baseband signal and thus reduces the quantizing noise associated with encoding the baseband signal into a fixed number of digital code words.

Compressor **12** is shown in more detail in FIG. 2. The baseband signal from filter **11** is sent through envelope detector **122** to generate a slowly varying direct current compandor control signal proportional to the average amplitude of the baseband signal over approximately one pitch period. Envelope detector **122**, in its simplest embodiment, consists of a rectifier and a suitable low pass filter. The baseband signal is also sent through delay **121** to ensure its synchronization with the compandor control signal from detector **122**.

The compandor control signal is used to control compressor **123** which either attenuates or amplifies the baseband signal depending on the average amplitude level of the baseband signal. When the average amplitude of the baseband signal is below a selected reference value, the compandor control signal actuates compressor **123** in such a manner that the baseband signal is amplified. Thus a low amplitude baseband signal is amplified to increase the ratio of baseband signal strength to transmission line noise. On the other hand, when the average amplitude of the baseband signal is above the selected reference value, the compandor control signal actuates compressor **123** in such a manner that the baseband signal is attenuated. The amplitude range occupied by the baseband signal is thus reduced and the baseband signal is said to be "compressed." Compressor **123** may, for example, be similar to the ones described in the August 1946 "Transactions of the American Institute of Electrical Engineers," vol. 65, page 1082, Figs. 5 and 6. Of course, other compressing devices, either active or passive, can also be utilized.

The baseband signal is restored or "expanded" to its original form in expandor **22** (FIG. 1) at the vocoder synthesizer before it is used in producing a replica of the input speech signal. As explained in the above cited article, a second device, designed to have the inverse characteristics of compressor **123**, is used for this expansion process. The above described process is called "companding," and when, as in this case, the companding occurs at a syllabic rate, it is called "syllabic companding."

The compandor control signal from envelope detector **122** (FIG. 2) represents the variations in average amplitude of the baseband signal from pitch period to pitch period and is transmitted to the vocoder synthesizer for use in restoring the companded baseband signal to its original form. The compandor control signal is derived from the baseband signal and thus essentially represents the amplitude of the first speech formant during voiced speech. Since the amplitude of the first speech formant is usually greater than the amplitudes of the other formants, the compandor control signal is an ideal signal with which to normalize the low frequency control signals derived from the input speech.

During unvoiced speech, more speech energy per unit bandwidth is often present at high frequencies than at low frequencies. The baseband signal, however, usually occupies a frequency bandwidth several times the speech bandwidths from which the low frequency vocoder control signals are derived. Thus, the baseband signal usually contains as much energy as each of the low frequency control signals even though its energy per unit bandwidth is often lower than the energy per unit bandwidth of these control signals. Hence, the amplitude of the compandor control signal is often approximately equal to the maximum amplitude of the low frequency control signals and the compandor control signal remains an acceptable normalizing signal.

Normalizing is carried out in normalizer **13** (FIG. 1) which, in one embodiment of the invention, consists of a series of dividing networks. One such network is shown in FIG. 3. In FIG. 3, a low frequency control signal from analyzer **10** in FIG. 1 is sent on lead **10–*i*** to logarithmic amplifier **131**, where i is a positive integer given by $1 \leq i \leq n$. The compandor control signal is sent to logarithmic amplifier **132** on lead **12A**. The signal from amplifier **132** is subtracted from the signal from amplifier **131** in summing network **133** and an output signal proportional to the inverse or antilogarithm of this difference signal is generated by amplifier **134**. This output signal represents the quotient of the low frequency control signal on lead **10–*i*** divided by the compandor control signal and thus is a normalized low frequency control signal. Other types of normalizing networks, such as those varying a PCM reference voltage in response to a control signal, can of course be used in this invention.

Normalizing ensures that the amplitude variations in the speech signal are limited to the variations in amplitude with frequency at any one time and do not include the potential variations in speech amplitude with time. Since speech amplitude often varies with time by several orders of magnitude, the amplitude range over which the low frequency control signals must be encoded is reduced at least an order of magnitude, if not several orders of magnitude, by normalizing. A corresponding reduction occurs in the quantizing errors associated with the encoding of these signals into digital code words.

The normalized low frequency control signals are sent to PCM transmitter **15** (FIG. 1) where they are converted from analog to digital form by means of a sampler, a quantizer, and an encoder. Transmitter **15** also includes a parallel-to-series converter, a multiplexer and a transmitting unit. Such encoding, multiplexing, and transmitting units are well known and thus transmitter **15** will not be described in detail.

The companded baseband signal which has been delayed in delay **121** (FIG. 2) to ensure synchronization at the vocoder synthesizer with the low frequency control signals and the compandor control signal, is also sent to transmitter **15** (FIG. 1) together with the compandor control signal for conversion to digital form and transmission to the synthesizer.

At the vocoder synthesizer (FIG. 1), the digital pulses representing the input speech signal are reconverted to analog form by PCM receiver **20**. Receiver **20** is likewise of well known design and thus will not be described in detail. The normalized low frequency control signals are sent to denormalizer **21** together with the compandor control signal (lead **20A**). Denormalizer **21** may, for example, consist of n modulators, each with an output signal proportional to the product of a selected normalized low frequency control signal and the compandor control signal. The companded baseband signal is similarly sent to expandor **22** together with the compandor control signal to produce a replica of the original baseband signal.

The reconstructed baseband signal is sent to excitation network **23** where it is utilized in a well known manner (see for example the above cited Schroeder patent) to generate an excitation signal suitable for use in producing a replica of the input speech. The excitation signal is passed through parallel-connected bandpass filters **24–1** through **24–*n*** to yield subsignals corresponding on a one-to-one basis to the speech subsignals from which the low frequency control signals were derived at analyzer **10**. Each of these resulting subsignals is modulated in a corresponding modulator **25–*i*** by the corresponding low frequency control signal from denormalizer **21**. Each modulator **25–*i*** produces an output signal in response to the simultaneous presence of a low frequency control signal and a filtered excitation signal. The output signal from each modulator represents that part of the input speech in a selected corresponding portion of the frequency spectrum of the input speech. The output signals from the modulators are filtered in bandpass filters **26–1** through **26–*n*** to remove undesired frequency components. Then they are summed with the reconstructed baseband signal in network **27** and converted to an acoustic signal in transducer **2** to produce a replica of the input speech signal.

The quality of the resulting speech signal is significantly above the quality expected from a digital voice excited vocoder because quantizing noise has been appreciably reduced by the normalizing of the low frequency control signals and by the syllabic companding of the baseband signal. Yet surprisingly, the complexity of the vocoder has not been significantly increased. The slight increase in bit rate necessary to transmit the compandor control signal is more than offset by the improved quality of the synthesized speech resulting from normalizing the low frequency control signals. Thus, the additional improvement in speech quality attributable to the companding process is obtained with essentially no increase in bit rate. As a result, the binary code words used to represent each sample of the baseband signal can contain, for example, only four digits while the quality of the reconstructed speech compares favorably to that obtained from a digital voice excited vocoder which encodes the baseband signal into binary code words containing five digits.

Furthermore, no additional delay networks are required to ensure synchronization of the baseband signal with the compandor control signal because the baseband signal already is delayed the required amount to ensure its synchronization with the low frequency control signals.

Other embodiments incorporating the principles of this invention will be obvious to those skilled in the vocoder arts.

What is claimed is:

1. In a digital voice excited vocoder in which a baseband signal is derived from an input speech signal, that improvement which comprises

means for deriving a control signal from said baseband signal,

means controlled by said control signal for compressing said baseband signal,

means for encoding said compressed baseband signal and said control signal for transmission to the synthesizer of said voice excited vocoder; and

at said synthesizer,

means for decoding said compressed and encoded baseband signal and said encoded control signal, and

means responsive to said decoded control signal for expanding said decoded and compressed baseband signal to its original form.

2. In a digital voice excited vocoder in which low frequency control signals and a baseband signal are derived from an input speech signal, that improvement which comprises

means for deriving a compandor control signal from said baseband signal,

compandor means responsive to said compandor control signal for compressing said baseband signal,

means for normalizing said low freqency control signals with said compandor control signal,

means for encoding said compressed baseband signal, said compandor control signal, and said normalized low frequency control signals for transmission to the synthesizer of said voice excited vocoder; and

at said synthesizer,

means for decoding said compressed and encoded baseband signal, said encoded compandor control signal and said encoded and normalized low frequency control signals,

means responsive to said decoded compandor control signal for denormalizing said normalized low frequency control signals, and

means responsive to said decoded compandor control signal for expanding said decoded and compressed baseband signal to its original form.

3. In a digital voice excited vocoder containing an analyzer in which low frequency control signals and a baseband signal are derived from an input speech signal, and a synthesizer for producing a replica of said input speech signal, that improvement which comprises:

at said analyzer,

means for altering the amplitude of said base band signal at a syllabic rate, means for deriving from said baseband signal a syllabically varying compandor control signal,

means utilizing said compandor control signal for normalizing said low frequency control signal,

means for encoding said altered baseband signal, said compandor control signal and said normalized low frequency control signals for transmitting to said synthesizer in coded form; and

at said synthesizer,

means for decoding said transmitted signals,

means for restoring said altered baseband signal to its original form, and

means responsive to said compandor control signal for denormalizing said normalized low frequency control signals.

4. In combination:

a transducer for converting a speech signal into an electrical signal,

means for deriving a baseband signal from said electrical signal,

means for deriving low frequency control signals from said electrical signal,

compandor means for altering said baseband signal at a syllabic rate and for deriving a syllabically varying compandor control signal,

means for normalizing said low frequency control signals with said compandor control signal,

means for delaying said baseband signal to ensure its synchronization with said low frequency control signals and said compandor control signal,

means for converting said normalized low frequency control signals, said compandor control signal and said altered baseband signal to digital form,

means for transmitting said normalized low frequency control signals, said compandor control signal and said altered baseband signal to a synthesizer in digital form; and

at said synthesizer,

means for converting said normalized low frequency control signals, said compandor control signal and said altered baseband signal to analog form,

means responsive to said compandor control signal for denormalizing said normalized low frequency control signals,

means controlled by said compandor control signal for restoring said altered baseband signal to its original form,

means driven by said restored baseband signal for generating an excitation signal,

modulating means responsive to said denormalized low frequency control signals and said excitation signal for generating a first set of signals representative of the signals in selected frequency bands of said speech signal,

bandpass filters corresponding on a one-to-one basis to said modulating means for removing undesired frequency components from said first set of signals,

means for combining said filtered first set of signals and said restored baseband signal, and

means for converting said combined, filtered first set of signals and said restored baseband signal into an acoustical replica of said speech signal.

5. Vocoder apparatus which comprises:

at an analyzer,

means for processing an input speech signal to derive low frequency control signals indicative of the energy in selected frequency bands of said speech signal and to derive a baseband signal,

compandor means for compressing said baseband signal at a syllabic rate and for deriving a syllabically varying compandor control signal,

means utilizing said compandor control signal for normalizing said low frequency control signals,

means for converting said compressed baseband signal, said compandor control signal, and said normalized low frequency control signals to digital form,

means for transmitting said compressed baseband signal, said compandor control signal, and said normalized

low frequency control signals in digital form to a synthesizer; and

at said synthesizer,

means for converting said transmitted signals from digital to analog form,

means responsive to said compandor control signal for denormalizing said normalized low frequency control signals,

means responsive to said compandor control signal for expanding said compressed baseband signal to its original form, and

means responsive to said denormalized low frequency control signals and said expanded baseband signal for generating a replica of said input speech signal.

References Cited

UNITED STATES PATENTS

2,953,644	9/1960	Miller	324—77
3,052,757	9/1962	Kalfaian	179—1
3,109,142	10/1963	McDonald	179—1
3,124,654	3/1964	Raisbeck	179—15.55
3,127,476	3/1964	David	179—15.55
3,139,487	6/1964	Logan	179—15.55

RALPH D. BLAKESLEE, Primary Examiner

A. B. KIMBALL, JR., Assistant Examiner

U.S. Cl. X.R.

179—1, 15; 324—77