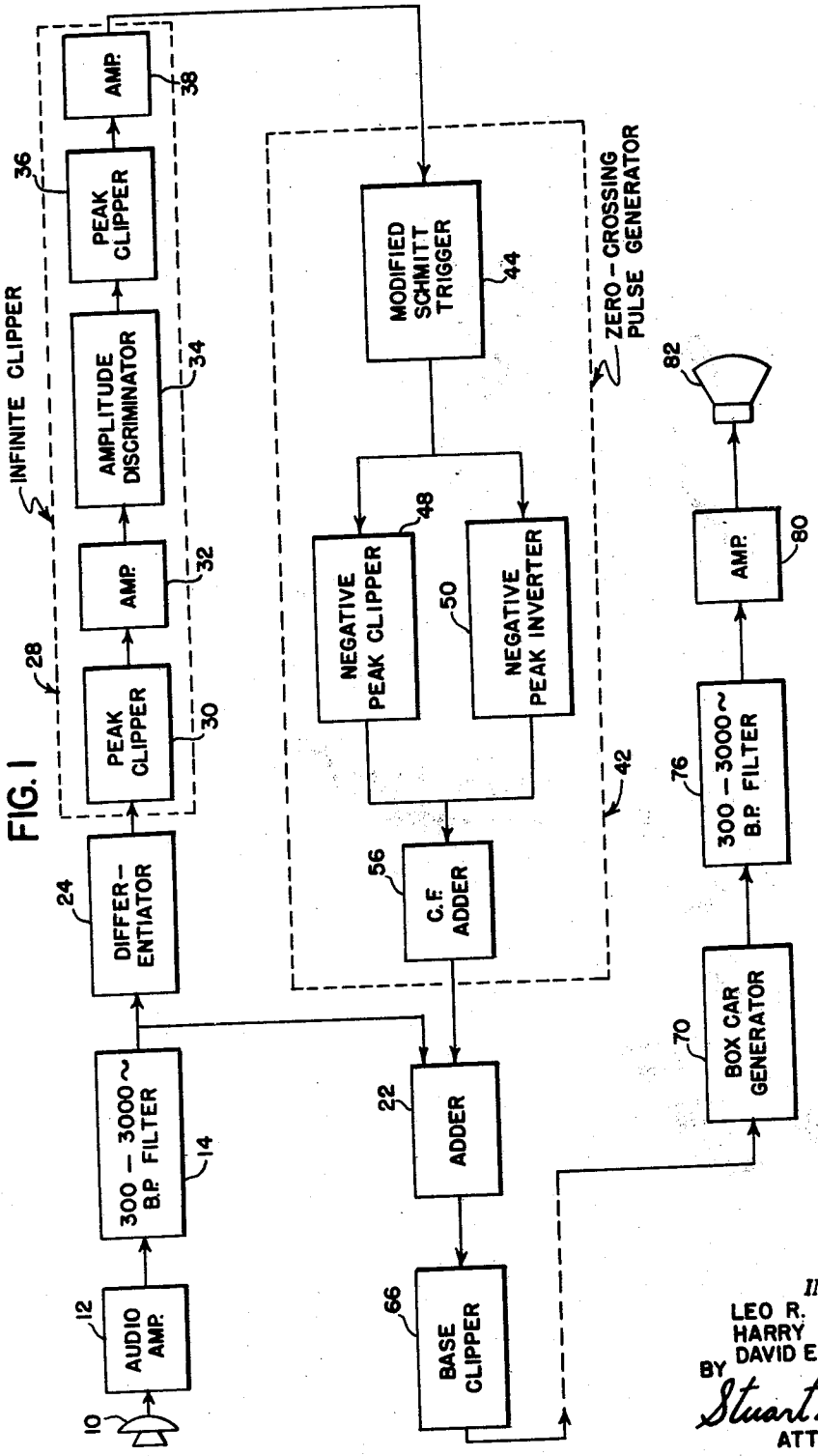
March 17, 1964 L. R. SPOGEN, JR., ETAL 3,125,723

SELECTIVE AMPLITUDE SAMPLING METHOD AND SYSTEM FOR PROCESSING SPEECH SIGNALS

Filed Dec. 5, 1960 2 Sheets-Sheet 1

FIG. 1



INVENTORS
LEO R. SPOGEN JR.
HARRY N. SHAVER
DAVID E. BAKER
BY Stuart R. Peterson
ATTORNEY

March 17, 1964 L. R. SPOGEN, JR., ETAL 3,125,723

SELECTIVE AMPLITUDE SAMPLING METHOD AND SYSTEM FOR PROCESSING SPEECH SIGNALS

Filed Dec. 5, 1960 2 Sheets-Sheet 2

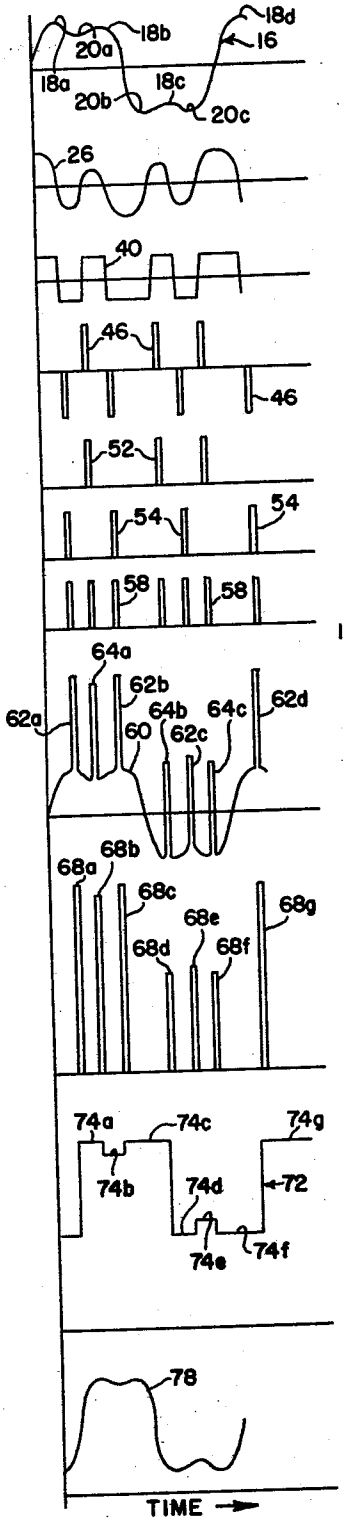
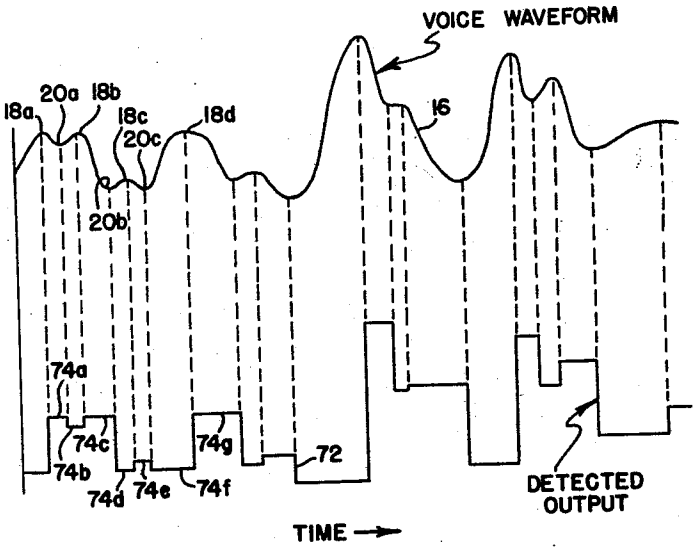


FIG. 3

FIG. 2



INVENTORS
LEO R. SPOGEN JR.
HARRY N. SHAVER
DAVID E. BAKER
BY Stuart R. Peterson
ATTORNEY

United States Patent Office 3,125,723 Patented Mar. 17, 1964

3,125,723

SELECTIVE AMPLITUDE SAMPLING METHOD AND SYSTEM FOR PROCESSING SPEECH SIGNALS

Leo R. Spogen, Jr., Harry N. Shaver, and David E. Baker, Tucson, Ariz., assignors to Arizona Research Foundation, Inc., a corporation of Arizona

Filed Dec. 5, 1960, Ser. No. 73,828
5 Claims. (Cl. 325—38)

This invention relates generally to the processing of speech waves or signals, and pertains more particularly to a method and system in which certain portions of a speech or voice signal are selected as being representative of the intelligence contained in the wave or signal.

The present invention will be better understood, it is believed, if two prior art attempts in the field of speech processing are first referred to. One of these schemes is known as Nyquist sampling where the sampling is conducted at a repetition rate at least twice as great as the highest frequency in the sampled signal. Filtering of the sample pulses can in such a situation provide a close reproduction of the sampled waveform. However, voice processing by Nyquist sampling requires a pulse repetition rate of approximately 7,000 pulses per second, assuming a band-width for the speech wave of 3,000 cycles per second.

The second noteworthy attempt concerns clipped speech work where it has been recognized that the intelligibility of a speech is contained primarily in the zero crossings of either speech or differentiated speech. Therefore, it has become possible to utilize a pulse transmission arrangement in which pulses are transmitted only at times of zero crossings. Such a technique decreases the total pulses required for the transmission of speech to approximately 2,200 pulses per second. The principal difficulty with this type of system, however, is that the quality of the output or reproduced speech is greatly degraded.

With the present speech processing system, that is employing what has been termed selective amplitude sampling, a sampling is made of the instantaneous magnitude of the voice wave which is in the form of a transduced electrical signal, the sampling being only at times of zero slope of the signal; in other words, at times of maxima or minima. While the envisaged sampling technique and detection system requires the same number (approximately 2,200) of pulses as those needed in the clipped differentiated speech procedure alluded to above, the quality obtainable with our invention compares favorably with unaltered speech. Stated somewhat differently, the present invention reduces appreciably the number of samples from that required by Nyquist sampling, yet does not seriously degrade the quality.

Accordingly, one important object of the invention is to provide a speech processing system in which the intelligibility and quality of the voice are maintained to the extent that there is only a slight degradation of these characteristics relative to the original speech wave.

Another object is to provide a voice processing system employing simple and inexpensive apparatus, it being an aim to use less complex circuitry for sampling and detection than in existing techniques for sampling rate reduction.

Another object of the invention is to provide a system of the foregoing type in which the duty cycle is approximately thirty percent of conventional voice-pulse systems, thereby providing either a decrease in the power required, an increase in channel capacity, or a decrease in the necessary bandwidth.

Yet another object of the invention is that the processed voice lends itself readily to being coded.

Other objects will be in part obvious and in part pointed out more in detail hereinafter.

The invention accordingly involves a speech processing concept which will be exemplified in the construction hereafter set forth and the scope of the application which will be indicated in the appended claims.

In the drawings:

FIGURE 1 is a block diagram of a single polarity system embodying therein the teachings of the invention;

FIGURE 2 is a graphical representation of a typical voice waveform or transduced signal and also a resulting detected output signal produced with the present system, and

FIGURE 3 illustrates a plurality of voltage waveforms or signals occurring at various processing points in the system depicted in FIGURE 1, a fragmentary portion of the transduced signal of FIGURE 2 having been utilized and from which the other signals of this figure are derived.

Referring now in detail to the drawings, the exemplary system of FIGURE 1 comprises a pickup transducer in the form of a microphone 10 which is connected to an audio amplifier 12. The amplifier signal is fed to a band-pass filter 14 where the speech is first limited to the 200–3000 c.p.s. band. It is the voltage waveform from the filter 14 that will hereinafter be referred to as the transduced signal. A typical voice waveform or transduced signal appears in FIGURE 2 and has been assigned the reference numeral 16. It will be observed that the signal 16 includes various maxima and minima, certain of the time spaced maxima having been designated at 18*a*, 18*b*, 18*c* and 18*d* with the intervening minima having been denoted as 20*a*, 20*b* and 20*c*. These same zero slope portions are also pictured in the fragment of the signal 16 appearing in FIGURES 3. It has been found that these zero slope portions contain considerable intelligence, and as the description progresses it will be seen that these portions serve as the samples utilized in the speech processing.

As can be seen from FIGURE 1, the output of the filter 14 is divided, one path of the signal 16 being directly to an adder 22 and the other path being via a processing route now to be described.

It will be discerned that the upper path or route includes therein a differentiating circuit 24 in which the transduced signal 16 is differentiated to produce a signal 26 which is shown immediately below that fragment of the signal 16 from which it is derived. It will be appreciated that the differentiated signal 26 passes through zero when the transduced signal 16 is either at a maximum or a minimum value as indicated by the reference numerals 18*a, b, c, d* and 20*a, b, c*. The differentiated signal 26 is fed to an infinite clipper designated generally by the reference numeral 28. The infinite clipper, it may be explained, provides a high grain followed by symmetrical clipping. Illustratively, the infinite clipper may include a first peak clipper 20, an amplifier 32, a center clipper or amplitude discriminator 34, a second peak clipper 36, and an amplifier 38. It will soon become apparent that low level noise passed by the differentiator 24 should be eliminated from the system and this is accomplished by the amplitude discriminator 34, the use of this center clipper or discriminator also having the desirable effect of providing a pulse count reduction as will be more readily understood later on. Stated somewhat differently, a minimum signal is required to produce certain pulses yet to be described. This minimum signal can be referred to as a threshold level, a level that may be adjusted to provide minimum degradation and quality. The output from the infinite clipper 28 bears the reference numeral 40, this signal also appearing in FIGURE 3.

Next to be referred to is a zero-crossing pulse generator

indicated generally by the numeral 42. The function of this generator is to produce a positive voltage pulse of predetermined height and width each time the signal 40 from the infinite clipper 28 passes through zero. It will be understood that these zero crossings coincide with those of the differentiated waveform 26. While the height and width of the pulses to be generated are obviously susceptible to variation, solely as a guide it might be explained that pulses of 6.0 volt magnitude having 0.6 μsec. width have been utilized and found satisfactory. The desirability of the center clipper or amplitude discriminator 34 should now be evident, for one would not wish that the zero-crossing pulse generator 42 be triggered into operation by extraneous low level signals not occurring at the previously mentioned maxima 18*a, b, c, d* and minima 20*a, b, c*.

The zero-crossing pulse generator includes in the exemplified instance a modified Schmitt trigger circuit labeled 44. By "modified," it is merely meant that the circuit differs from a conventional Schmitt trigger circuit by reason of an inductor in circuit with the free plate thereof. It is the function of the infinite clipper 28 to provide a signal 40 of sufficient magnitude and shape for operating the Schmitt trigger.

The output from the circuit 44 has been indicated by the numeral 46. Inasmuch as the signal 46 includes both positive and negative pulses, the generator 42 also has incorporated therein a negative peak clipper 48 in shunt with a negative peak inverter 50. The unit 48 simply provides a free path for all positive pulses contained in the signal 46, these being labeled 52 in FIGURE 3, and at the same time blocking passage of negative pulses via this path. The unit 50, as its name implies, inverts the negative pulses contained in the signal 46 so that they appear as positive pulses 54 (FIGURE 3). By virtue of a dual cathode follower type of adder 56 the pulses 52 and 54 are channeled over a common route, the pulse train or signal now being indicated as 58 in FIGURE 3. These are the positive voltage pulses of predetermined height and width previously alluded to when first referring to the role to be played by the zero-crossing pulse generator 42. It should be understood, though that these constant amplitude pulses occur at times of zero crossings of the differentiated wave or signal 26 and hence occur at the already alluded to various maxima and minima of the transduced signal 16.

The pulses 58 are now combined with or added to the signal 16 through the medium of the previously mentioned adder 22 to produce a composite signal 60 as also shown in FIGURE 3. It will be noted that the adding together of the voice signal 16 and the pulses 58 to produce the signal 60 results in the signal 60 having the same configuration as the original signal 16 except at times of maxima and minima when a large amplitude positive pulse or spike appears, those pulses occurring at the time of maxima being denoted by 62*a*, 62*b*, 62*c* and 62*d* and those occurring at the minima being designated by the numerals 64*a*, 64*b* and 64*c*.

By means of a base clipper 66, the wave or signal 60 is stripped of all voltage components having a magnitude less than that of maximum voice amplitude (the largest unnumbered maximum of the transduced signal 16) delivered to the adder 22. This leaves discrete positive pulses 68*a*, 68*b*, 68*c*, 68*d*, 68*e*, 68*f* and 68*g* which constitute the usable amplitude sampling of the speech being processed, these pulses 68*a–g* containing most of the intelligence originally present in the segment of the signal 16 from which said pulses have been derived. In other words, the remaining signal 68 comprises a succession of positive pulses 68*a–g* occurring at times of maxima and minima of the signal 16 and having heights which are linearly related to these maximum and minimum values.

To detect the information available in the signal 68, a boxcar circuit 70 is utilized. So called "boxcars" have been extensively used in radar applications; therefore, if

detailed information is desired, reference may be made to the Massachusetts Institute of Technology Radiation Laboratory Series, vol. 24, published by McGraw-Hill Book Company, Inc. in 1950. The boxcar circuit 70 holds or maintains the approximate amplitude of each of the pulses 68*a–g* until the occurrence of the next one. Such action is shown generally in FIGURE 2 in the form of the voltage signal 72 which has been derived from the signal 16 over a longer interval than the fragmentary portion of the signal 72 that has been presented in FIGURE 2. At any rate, it is believed evident that the boxcar 70 causes each pulse 68*a–g* in the succession of pulses to persist until the arrival of the following such pulse and so on, this being characteristic of boxcar circuitry. Thus, the pulse 68*a* will be held during a period labeled 74*a*, the pulse 68*b* over a period labeled 74*b*, the pulse 68*c* through a somewhat longer interval bearing the numeral 74*c* and so on as indicated by the various references 74*d*, 74*e*, 74*f* and 74*g*.

The collective signal 72 is fed to a band-pass filter 76, this filter corresponding to the previously mentioned filter 14. The filtered output signal has been illustrated only in FIGURE 3, being assigned the reference numeral 78. A comparison of the input and output waveforms 16 and 78 shows that the system introduces a relative phase shift, but the relative amplitudes of the frequency components are closely approximated. Thus, while the system does not reconstruct the exact waveform, it does retain the information necessary for the transmission of intelligibility and quality after appropriate amplification by an audio amplifier 80 and playout by a transducer in the form of a speaker 82.

From what has been presented, it is believed that our system and its attendant method of operation will be readily understood. Quite briefly, the transduced signal 16 is amplitude sampled at times of substantial maxima and minima. This is accomplished basically by the differentiator 24, the zero-crossing generator 42 and the adder 22. After stripping the signal 60 of that portion of the waveform below a predetermined level the resulting signal 68 is fed to the boxcar circuit 70 for detection purposes.

As many changes could be made in the above construction and many apparently widely different embodiments of the invention could be made without departing from the scope thereof, it is intended that all mattter contained in the above description or shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

It is also to be understood that the language used in the following claims is intended to cover all of the generic and specific features of the invention herein described and all statements of the scope of the invention which, as a matter of language, might be said to fall therebetween.

What is claimed:

1. A speech processing method comprising the steps of differentiating a transduced voice signal to produce a derived signal having zero crossings during the times that the transduced signal is at a maximum or minimum, said transduced signal having substantially the same wave form as the original speech generating a rectangular pulse having a certain height at each of said zero crossings, adding said rectangular pulses to said transduced signals to provide resulting signals, and maintaining the approximate value of each resulting signal until the arrival of the next resulting signal to produce a detected output signal.

2. A speech processing system comprising means for producing a transduced voice signal, said transduced signal having substantially the same wave form as the original speech means for differentiating said signal, means for generating a pulse of predetermined magnitude each time the signal from said differentiating means passes through zero, means for combining said pulses with said transduced signal, and means for detecting the signal from said combining means.

3. A speech processing system in accordance with claim 2 including means for removing a base portion of preferred magnitude from the signal produced by said combining means before detection thereof by said detecting means.

4. A speech processing system comprising a pickup transducer for producing a voice signal, a band-pass filter connected to said transducer, a differentiator connected to the output side of said filter for providing a differentiated signal having zero crossings at the times of maxima and minima of said transduced signal, an amplitude discriminator for removing those portions of said differentiated signal having a magnitude less than a desired value, a zero-crossing pulse generator for producing a pulse of a given magnitude for each zero crossing of said differentiated signal, an adder connected to the output side of said filter and to the output side of said pulse generator for adding said pulses to the filtered voice signal, a base clipper connected to said adder for stripping the output of said adder of voltages having a magnitude less than that of the voice amplitude to said adder to provide selected amplitude signals having heights linearly related to the values of said filtered transduced signal at said times of maxima and minima, a boxcar generator connected to said clipper for holding the approximate amplitude of each selected amplitude signal until the occurrence

of the next selected amplitude signal, band-pass filter connected to said boxcar generator for filtering the output of said boxcar generator, and a speaker connected to the output side of said last-mentioned filter.

5. A speech processing system comprising means for producing amplified sampled signals from a transduced voice signal at times of substantial maxima and minima which includes first means for differentiating said transduced signal to provide a differentiated signal, second means for generating constant magnitude pulses at the zero crossings of said differentiated signal, and third means for adding said pulses to said transduced signal at the time of said maxima and minima, said transduced signal having substantially the same wave form as the original speech, the system further including a boxcar circuit for maintaining the approximate value of each sampled signal until the occurrence of the next such signal.

References Cited in the file of this patent

UNITED STATES PATENTS

2,448,718	Koulicovitch	Sept. 7, 1948
2,673,893	Kalfaian	Mar. 30, 1954
2,699,464	Di Toro et al.	Jan. 11, 1955
2,962,553	Halina	Nov. 29, 1960

UNITED STATES PATENT OFFICE

CERTIFICATE OF CORRECTION

Patent No. 3,125,723 March 17, 1964

Leo R. Spogen, Jr., et al.

It is hereby certified that error appears in the above numbered patent requiring correction and that the said Letters Patent should read as corrected below.

Column 1, line 30, strike out "a"; column 2, line 24, for "200-3000" read -- 300-3000 --; line 30, for "at" read -- as --; line 34, for "FIGURES" read --FIGURE --; line 54, for "grain" read -- gain --; column 4, line 46, for "mattter" read -- matter --; line 63, for "signals" read -- signal --; column 6, line 1, before "band-pass" insert -- a --.

Signed and sealed this 28th day of July 1964.

(SEAL)

Attest:

ESTON G. JOHNSON
Attesting Officer

EDWARD J. BRENNER
Commissioner of Patents