United States Patent

[11] 3,624,302

[72] Inventor   Bishnu S. Atal
                 Murray Hill, N.J.
[21] Appl. No. 872,051
[22] Filed   Oct. 29, 1969
[45] Patented   Nov. 30, 1971
[73] Assignee   Bell Telephone Laboratories, Incorporated
                 Murray Hill, N.J.

[54] SPEECH ANALYSIS AND SYNTHESIS BY THE USE OF THE LINEAR PREDICTION OF A SPEECH WAVE
    10 Claims, 4 Drawing Figs.

[52] U.S. Cl. .... **179/1 SA**
[51] Int. Cl. **G10l 1/00**
[50] Field of Search .... 179/1 SA,
                                        15.55 R; 325/38.1

[56]   References Cited
              UNITED STATES PATENTS
2,817,707   12/1957   Weibel .... 179/1 SA
3,020,344   2/1962   Prestigiacomo .... 179/1 SA
3,158,685   11/1964   Gerstman .... 179/1 SA
3,328,525   6/1967   Kelly .... 179/1 SA

ABSTRACT: A short-time spectral analysis of a nonstationary signal, such as a speech signal, does not ordinarily yield control signal information sufficient for subsequent synthesis. However, more reliable control signals for a speech synthesizer can be obtained by making use of natural constraints, applicable to a speech wave, in the analysis procedure. For frequencies below 5 kHz., the human vocal tract can be modeled as an acoustic tube in which only plane waves propagate. Thus, for vowels and vowellike sounds, the speech output of the vocal tract at any instant of time can be assumed to be a weighted sum of its past values and the input to the vocal tract at that instant of time. In the described invention, a speech wave is represented by the output of a linear filter which simulates an acoustic tube and which is excited by a combination of a quasi-periodic pulse train and white noise. The parameters of this filter are derived from the speech wave such that the mean-squared error between the synthetic speech samples at the output of the filter and the input speech samples is minimum.
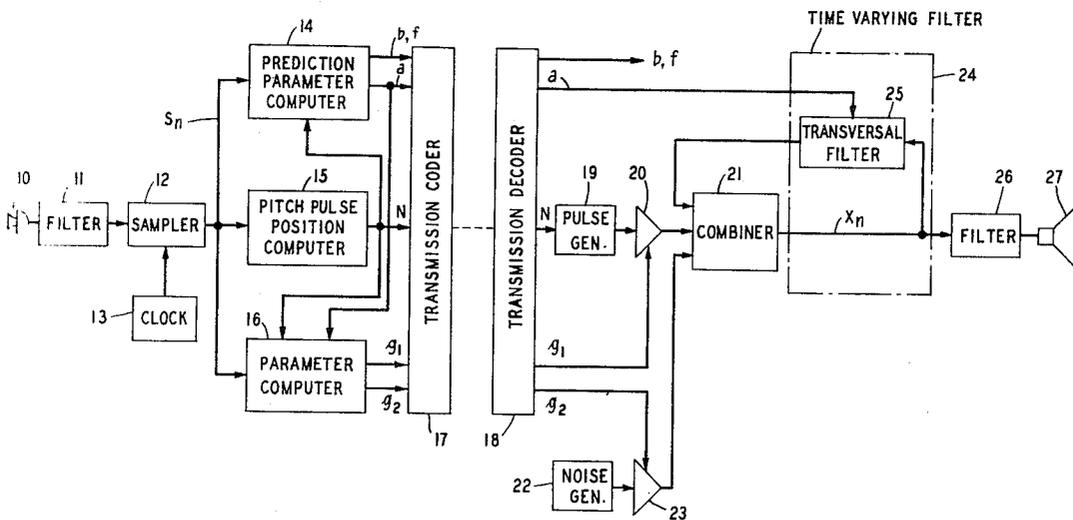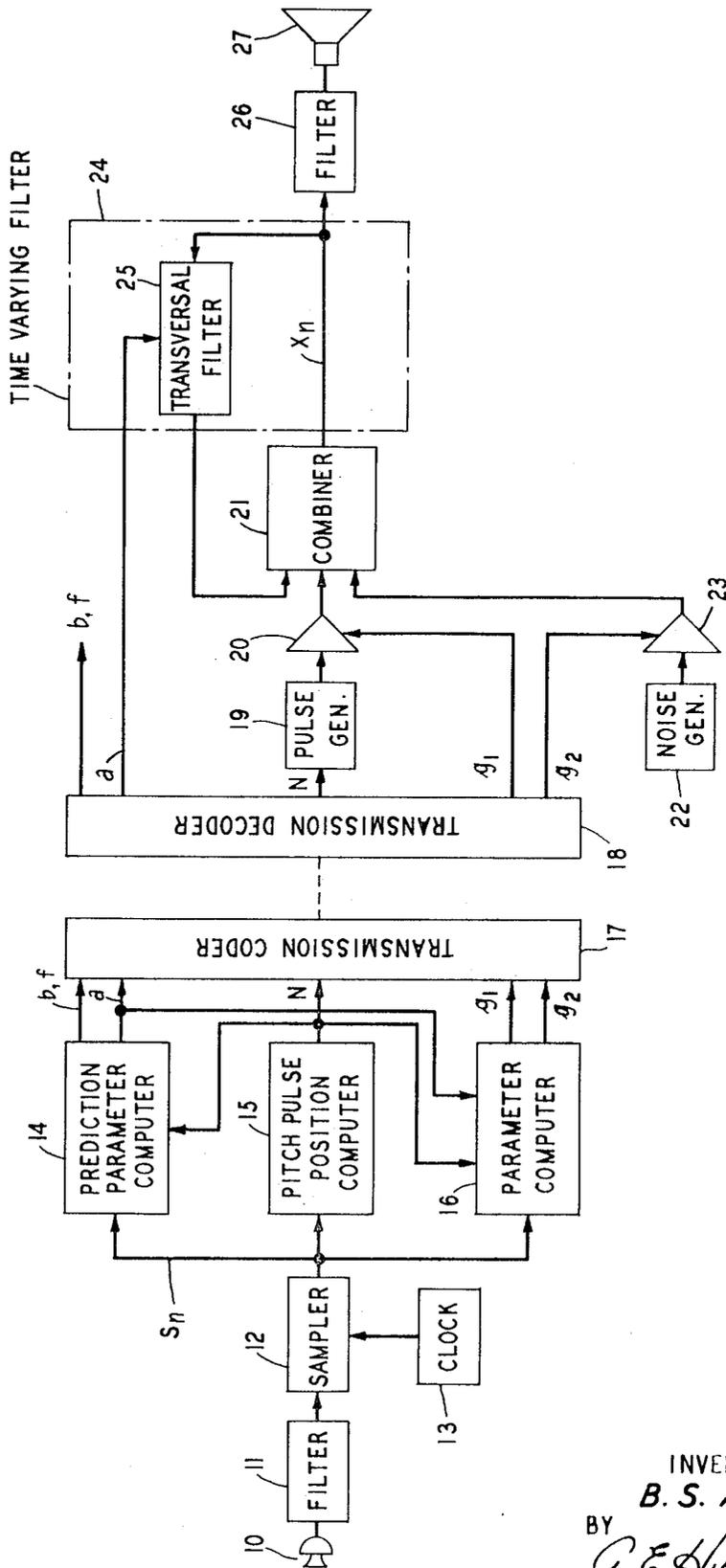
F/G. 1

INVENTOR
B. S. ATAL

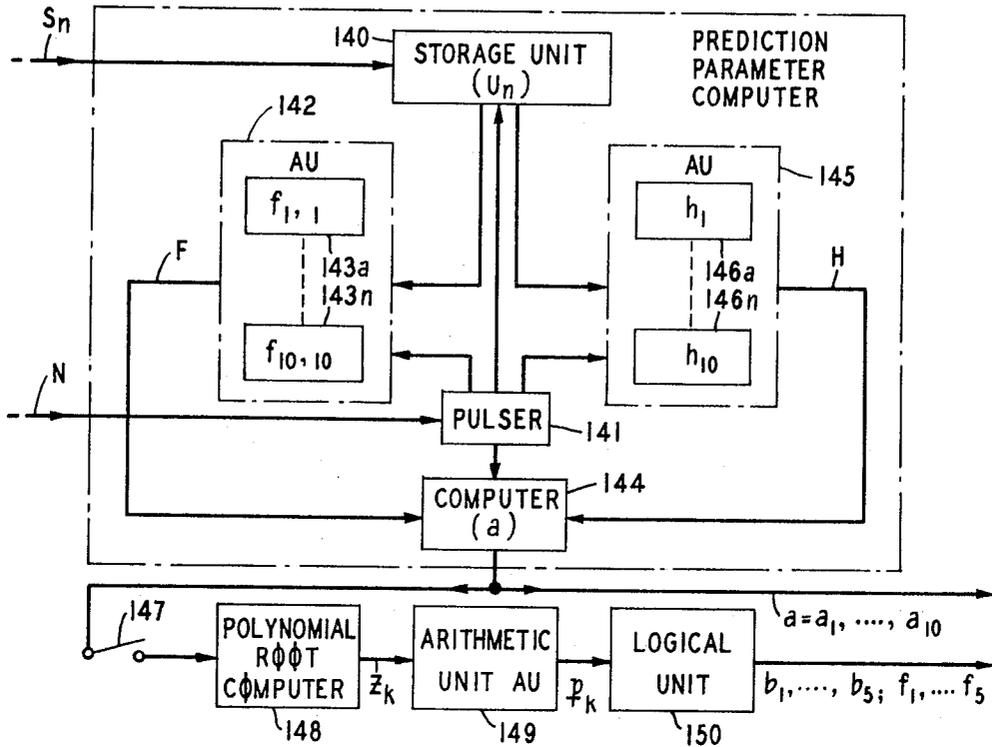BY
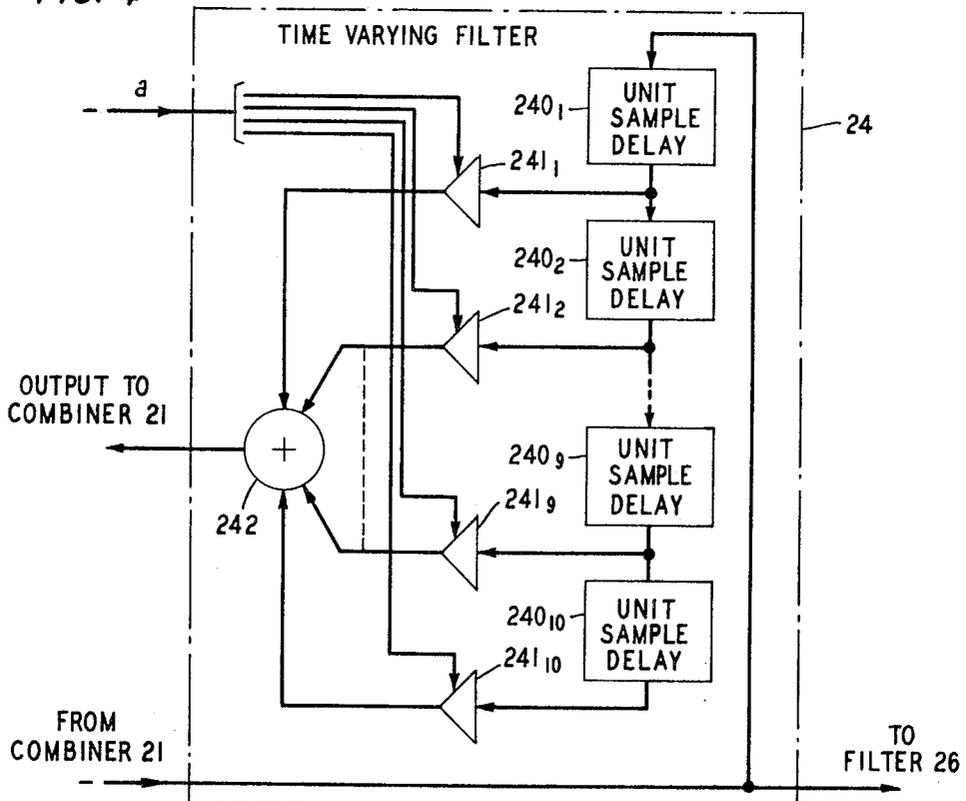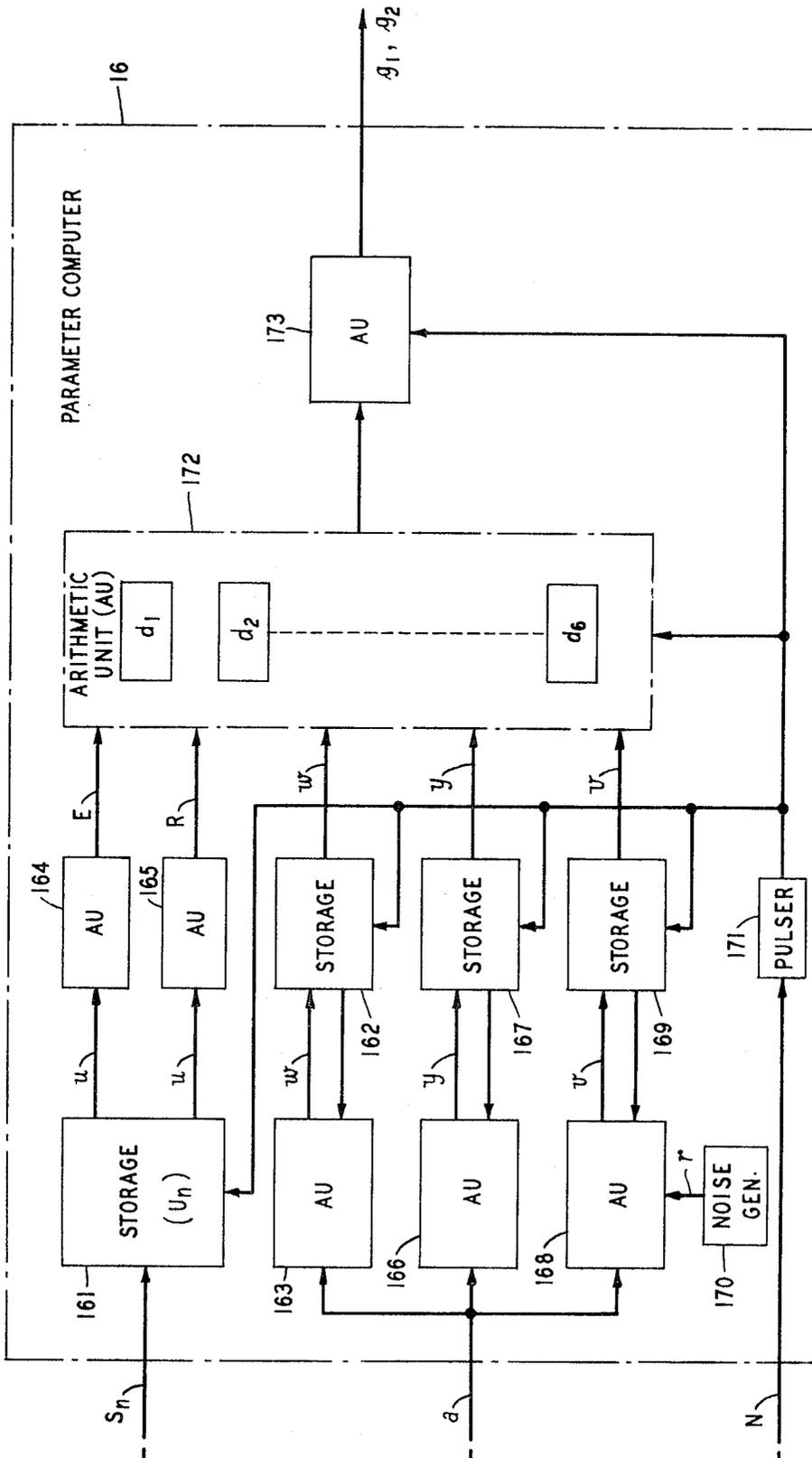*G. E. Hirsch Jr.*
ATTORNEY

FIG. 2

FIG. 4

FIG. 3

SPEECH ANALYSIS AND SYNTHESIS BY THE USE OF THE LINEAR PREDICTION OF A SPEECH WAVE

BACKGROUND OF THE INVENTION

This invention relates to the artificial production of speech or similar complex waves from control signals, and particularly to the derivation of control signals from an original speech wave that can be accommodated by storage or transmission facilities with limited channel capacity.

The principal object of the invention is to reduce, as far as possible, the channel capacity, or bit rate in the case of a digital channel, required for the storage or transmission of speech control signals without, however, a sacrifice of intelligibility or the introduction of an objectionable unnatural quality into the reconstructed speech.

1. Field of the Invention

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 speaker. Because of the redundance of human speech, however, facsimile transmission is a relatively inefficient way to transmit this information. Consequently, a number of arrangements for compressing or reducing the required channel capacity required for the transmission of speech information have been proposed. One of the best known of these arrangements is the so-called vocoder. More recently, techniques for removing inherent signal redundancy in the speech wave through the use of a linear predictor have been utilized.

2. Description of the Prior Art

Production of good quality synthetic speech is a necessary corollary to limited channel capacity transmission systems of whatever sort. However, the quality of speech obtained from priorly known synthesizers generally lacks naturalness and exhibits and undesirable quality, even when the synthesizer control signals are derived from the original speech at closely spaced intervals. There are a number of reasons for the poor quality of such synthetic speech. Consider, for example, the case of a formant synthesizer, this being a part of another typical system for the narrow band transmission of speech. Most formant analyzers attempt to isolate peaks due to various formants in the speech spectra. This is a difficult task, even for low-pitched male voices, since formants do not always show up as distinct peaks in the spectra, and the spectral peaks do not always result from the formants. Such methods usually break down completely for female speech. Further, satisfactory operation of a formant synthesizer often depends upon the correct ordering of the various formants. This, too, is difficult to achieve.

SUMMARY OF THE INVENTION

To avoid many of these problems, a different approach to speech analysis and synthesis is followed in the present invention. Speech parameter signals are continuously developed at a transmitter station using the constraint that the applied speech wave at any instant of time is a weighted sum of its past values, that is to say, speech parameter signals are developed which specify linearly predictable characteristics of an applied speech signal. To derive parameter control signals for the production of realistic synthesized speech, a suitable functional model of speech production is established and it is assumed that a close approximation to a speech wave can be produced at its output. Typically, the model includes a discrete, linear, time-varying filter which is excited by a suitable combination of a quasi-periodic pulse train (voiced excitation) and white noise (unvoiced excitation). The output of the linear filter at any sampling instant is a linear combination of past output samples and the input. In this analysis, the nth speech sample, s_n, may be expressed as:

$$s_n = \sum_{k=1}^{p} a_k s_{n-k} + \sum_{k=1}^{q} b_k x_{n-k} + x_n \qquad (1)$$

where $a_1, a_2, ..., a_p; b_1, b_2, ..., b_q$ are parameters which specify the filter at any time, and x_n is the nth input sample. For completely voiced sounds, the samples x_n represent a train of quasi-periodic pulses, whereas for completely unvoiced sounds, x_n represents the output of a white noise generator. For this model of speech production, it can be shown that in any pitch period the speech samples after the first q samples may be expressed as linear combination of the preceding p samples. The optimum linear combination, $a_1, a_2, ..., a_p$, is obtained by minimizing the mean-squared error between the actual values of the speech samples and their predicted values based on the past p samples. The values of p and q are determined by the bandwidth of the input speech signal and the length of the vocal tract. A 10th to 12th order linear predictor satisfactorily represents the speech signal band-limited to 5 kHz. with sufficient accuracy. A higher order predictor may be necessary for certain cases (some male speakers saying nasalized consonants). The parameter q is assumed to be equal to 10 in the analysis and zero in the synthesis.

Durations of individual pitch periods are determined by calculating the pitch-synchronous autocorrelation function of the third power of the input speech wave and selecting the delay for which the autocorrelation function is maximum.

A speech signal may be synthesized by a network continually adjusted by parameter signals derived in this fashion.

This invention will be more fully understood from the following detailed description taken together with the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block schematic diagram of a speech transmission system, including an analyzer and a synthesizer which illustrates the principles of the invention;

FIG. 2 is a block schematic diagram of a prediction parameter computer suitable for use in the analyzer of the speech transmission system illustrated in FIG. 1;

FIG. 3 is a block schematic diagram of a network for developing parameters, representing the relative amplitudes of voiced and unvoiced signal components, suitable for use in the analyzer of the system illustrated in FIG. 1; and

FIG. 4 is a block schematic diagram of a time-varying filter which may be used at the synthesizer of a transmission system embodying the principles of the invention.

DETAILED DESCRIPTION

A complete limited channel capacity speech transmission system which illustrates the principles of the invention is shown in FIG. 1. Speech signals, which may originate, for example, in transducer 10, are passed through low pass filter 11 which has a cutoff frequency in the neighborhood of 5 kHz. and which exhibits a 3-db. cutoff frequency in the neighborhood of 4 kHz. The resultant signal is then sampled at a frequency of approximately 10 kHz. in sampler 12. Clock 13 is employed to energize the sampler and other units in the system. Speech samples, s_n, thus derived are supplied to prediction parameter computer 14, to pitch pulse position computer 15, and to parameter computer 16.

Prediction parameter computer 14 operates on applied speech samples s_n to develop a series of parameter signals $a=a_1, a_2, ..., a_n$, for each pitch period (as indicated by signal N, from computer 15). Parameters a uniquely specify the frequencies and bandwidths of speech formants in the input signal below about 5 kHz. Parameter signals a are developed from linearly predictable characteristics of the applied speech signals delivered by sampler 12. An extensive discussion of the relation of parameter signals a to the input signal, and their development, is contained in my copending patent application, Ser. No. 753,408, filed Aug. 19, 1968. Details of the construction of a prediction parameter computer specially adapted for the practice of this invention is given hereinafter with reference to FIG. 2.

Pitch pulse position computer 15 determines the location of the glottal pulses in the applied speech wave; the difference

between the positions of successive glottal pulses specifies the duration of the pitch period. Any suitable pulse position analyzer may be employed to derive pitch period signals, N. For example, a suitable arrangement is described in *Automatic Speaker Recognition Based on Pitch Contours* by B. S. Atal, Polytechnic Institute of Brooklyn, June, 1968, pages 33–43.

Speech samples from unit 12 are also supplied to computer 16, which determines parameters g_1, and g_2. These parameters characterize the amplitudes of voiced and unvoiced signal excitation, i.e., parameter g_1 specifies the amplitude of voiced (or buzz) excitation signals, and the parameter g_2 specifies the amplitude of unvoiced (or hiss) excitation signals.

Parameter signals a, N, and g_1, g_2, thus derived uniquely determine formant frequencies and bandwidths of a speech signal, its spectrum, and the relative amplitudes of voiced and unvoiced components necessary for the synthesis of artificial speech. Since these parameters require considerably less channel capacity than the corresponding analog signal representation, they may be economically stored for future use, or transmitted to a distant station. All parameter signals may, for example, be combined for transmission, for example, by multiplexing, or the like, in transmission coder 17. At a receiver station, these signals are recovered and delivered individually through the action of transmission decoder 18. Transmission coders and decoders of any desired construction and form may be employed. Obviously, storage of the parameter signals may take place at any point in the indicated transmission arrangement; the transfer of parameter signals from one storage location to another, for example, may be considered to be a form of transmission.

At the synthesizer, voiced excitation is generated, for example, in pulse generator 19 of any desired construction, under control of pitch pulse parameter signal N. The amplitude of the voiced excitation signal is controlled continuously by parameter signal g_1 acting upon modulator 20. Typically, generator 19 produces a pulse of unit amplitude at the beginning of every pitch period. Similarly, unvoiced excitation is produced in noise generator 22. Generator 22 typically produces a sequence of random numbers uniformly distributed between +1 and −1 at sampling instants. Noise signals are controlled in amplitude by parameter signal g_2 acting on modulator 23.

The outputs of pulse generator 19 and noise generator 22, as scaled by controlled amplifiers 20 and 23, are added together with selected past signal values, available at the output of time-varying filter 24, in a combining network 21. The combined signal produced at the output of network 21 is thereupon delivered by way of low pass filter 26 to reproducer 27, for example, a loud speaker. Low pass filter 26 preferably has a cutoff frequency of about 5 kHz., its exact frequency range being commensurate with the range of filter 11 at the analyzer.

The combined signal is also delivered to the input of transversal filter 25, forming a part of filter network 24. Time-varying filter 24 serves to regenerate speech from the applied excitation and parameter signals a. Such a filter arrangement resembles the resonant filter system of the human vocal tract and typically exhibits certain natural resonances which may be tuned in accordance with formant parameter signals a. Resonant vocoder apparatus of this general form is well known in the art; a typical example is described in J. L. Kelly, Jr., U.S. Pat. No. 3,328,525, issued June 27, 1967. A transversal filter arrangement specifically adapted for use in the apparatus illustrated in FIG. 1 is described below with reference to FIG. 4.

FIG. 2 illustrates a prediction parameter computer 14 suitable for developing formant parameter signals a in accordance with the invention. For every pitch period of the applied speech wave, an array of signal values s_n from sampler 12 (FIG. 1) is transferred into storage unit 140 to replace the previous array of signal samples contained in the storage unit. Storage unit 140 thus stores an array of signal values u_{-10}, u_{-9}, ..., u_{-1}, u_0, ..., u_N, where N represents the duration of the current pitch period in samples. Every pitch period the values

u_{-10}, u_{-9}, ..., u_{-1} are replaced by values u_{N-9}, u_{N-8}, ..., u_N. Incoming samples are placed in the vacated storage locations u_0, ..., u_N. Thus, signals u_0, ..., u_N are consecutively stored as they are received in storage unit 140. Every pitch period, under the influence of timing signals from pulser 141, synchronized by signals N from pitch pulse position computer 15 to indicate the positions of glottal pulses, an array of signal values is read out of storage unit 140 and transferred to arithmetic unit (AU) 142. This unit comprises a plurality of arithmetic units 143a, ..., 143n, designated individually $f_{1,1}, f_{1,2}, ..., f_{1,10}; f_{2,2}, ..., f_{2,10}; f_{3,3}, ..., f_{10,10}$, which operate in parallel. In a typical example of practice, $n=55$, i.e., 55 arithmetic units are employed. Each individual unit serves to compute one value of f according to the following equation:

$$f_{i, j} = \sum_{n = 10}^{N-10} u_{n-i} u_{n-j} \qquad (2)$$

index i varies from 1, 10

index j varies from i, 10

Computations of $f_{i,j}$ are carried out simultaneously and output values, designated F, are periodically supplied to computer 144.

The array of signals u_n accumulated in storage unit 140 is also supplied to arithmetic unit 145 wherein an array of values is evaluated as follows:

$$h_j = \sum_{n = 10}^{N-10} u_n u_{n-j} \qquad (3)$$

index j varies from 1, 10

Arithmetic unit 145 preferably comprises an array of individual units, 146a, ..., 146m, operating in parallel to evaluate several values of h. Typically, 10 units are employed, i.e., $j=10$. The resultant array, h_1, h_2, ..., h_{10}, designated H, is delivered every pitch period to computer 144.

Computer 144 is programmed to solve the matrix equation

$$F a = -H, \qquad (4)$$

to yield values of a. Although a special purpose computer may be programmed for this evaluation, one suitable arrangement is described in copending patent application Ser. No. 753,408, filed Aug. 19, 1968.

Prediction parameters a_1, a_2, ..., a_{10}, uniquely determine the frequencies and bandwidths of all speech formants below 5 kHz. If desired, the bandwidths and frequencies of formants may be determined from values of a for use in the control of other synthesis apparatus. In accordance with the invention, this determination is made by supplying parameter values a from computer 144, by way of switch 147, to polynomial root computer 148. This unit determines the complex roots of a polynomial with real coefficients, i.e., the roots of a polynomial $f(z)$, defined as:

$$f(z) = z^{10} + a_1 z^9 + ... a_9 z + a_{10}. \qquad (5)$$

A polynomial root locater suitable for making the necessary evaluation is described in *Mathematical Methods for Digital Computers*, edited by Ralston and Wilf, John Wiley & Sons, Inc., 1967, in the section by E. R. Bareiss, at page 185. The output of the polynomial root locator 148 is 10 complex numbers (two sets of 10 real numbers) $z_1, z_2, ..., z_{10}$, which are then supplied to the arithmetic unit 149 which computes the numbers $p_1, p_2, ..., p_{10}$ in accordance with equation (6) below.

$$p_k = (1/2\pi)(1/T) \log (z_k). \qquad (6)$$

Arithmetic unit 149 is thus a device which takes the complex logarithm of numbers z_k and multiplies them with a number $(1/2\pi T)$ where $T=0.0001$ sec., the interval of sampling unit 12 (FIG. 1). The complex numbers p_k can be separated into their real and imaginary parts, b_k and f_k, respectively, as follows:

$$p_k = b_k + j f_k, \qquad (7)$$

where index k varies from 1 to 10. Logical unit 150 orders the numbers p_k such that the first number has the lowest positive imaginary part, the second number the second lowest positive imaginary part, and so on. Consequently, the numbers f_k and b_k represent the frequencies and the bandwidths of the various

5

formants of the speech signal for the pitch period under consideration. These representations may be used in any desired fashion, e.g., for controlling a formant synthesis.

Speech samples from sampling unit 12 (FIG. 1) are also supplied to parameter computer 16 which determines parameter values g_1 and g_2. These parameters denote the relative amplitudes of voiced and unvoiced signal components in the applied speech signal. The operation of computer 16 is illustrated in FIG. 3. Every pitch period, an array of signal values s_n is transferred into storage unit 161 to replace the previous array of signal samples already in storage. Storage unit 161 thus stores an array of signal values $u_{-m}, u_{-m+1}, ..., u_0, ..., u_N$, where N denotes the duration of the current pitch period in samples and m represents the largest pitch period as measured in samples. A value of $m=200$ has been found to be sufficient in most cases. Every pitch period, the values $u_{-N}, ..., u_0$, are replaced by values $u_0, ..., u_N$. Incoming samples are placed in the vacated storage locations $u_0, ..., u_N$. Arithmetic units 164 and 165 operate on array u to evaluate the values of parameters E and R in accordance with equations (8) and (9) as follows:

$$E = \sum_{n=0}^{N-1} u_n^2 \qquad (8)$$

$$R = \sum_{n=0}^{N-1} u_n u_{n-N} \Big/ \sqrt{E \sum_{n=0}^{N-1} u_{n-N}^2} \qquad (9)$$

Storage unit 162 contains an array of signal values $w_{-m}, ..., w_0, ..., w_N$. Every pitch period, the values $w_{-N}, ..., w_0$ are replaced by values $w_0, ..., w_N$. New signal values are computed in arithmetical unit 163 according to equation (10) and stored consecutively in storage locations $w_0, ..., w_N$.

$$w_n = \sum_{k=1}^{10} a_k w_{n-k} \qquad (10)$$

$$n \text{ varies from } 0, N$$

Arithmetic unit 166 computes an array of signal values $y_0, ..., y_N$, designated y, and stores them in storage unit 167 in locations designated $y_0, ..., y_N$. Storage unit 167 is equipped with 10 additional storage locations designated $y_{-10}, ..., y_{-1}$, which have a number 0 stored in them permanently. The array y is computed in arithmetic unit 166 according to the following relation:

$$y_n = \sum_{k=1}^{10} a_k y_{n-k} \text{ if } n \text{ is between 1 and N} \qquad (11)$$

$$= 1 \text{ if } n = 1$$

Arithmetic unit 168 computes another array of signal values v and stores them in storage unit 169 in locations designated $v_0, ..., v_N$. Array v is computed in accordance with (12) below.

$$v_n = \sum_{k=1}^{10} a_k v_{n-k} + r_n \qquad (12)$$

$$n \text{ varies from } 0, N$$

The array of numbers r designates the output of a white noise generator 170. Similar to storage unit 167, storage unit 169 also has 10 additional storage locations designated $v_{-10}, ..., v_{-1}$ which have the number 0 stored in them permanently.

The arrays w, y, and v, and the numbers E and R are transferred periodically, under the influence of pitch synchronized clock pulses from pulser 171, to arithmetic unit 172 which comprises six arithmetic units designated $d_1, d_2, d_3, ..., d_6$ which operate in parallel. These units of system 172 compute the numbers $d_1, ..., d_6$ in accordance with equations (13) through (18) set forth below. The index n is summed from 0 to N in each of the equations.

$$d_1 = \sum_n \hat{s}_n^2 \qquad (13)$$

$$d_2 = \sum_n y_n^2 \qquad (14)$$

6

$$d_3 = \sum_n v_n^2 \qquad (15)$$

$$d_4 = \sum_n \hat{s}_n y_n \qquad (16)$$

$$d_5 = \sum_n \hat{s}_n \hat{s}_{n-N} \qquad (17)$$

$$d_6 = \sum_n y_n \hat{s}_{n-N} \qquad (18)$$

The array of numbers $d_1, ..., d_6$, computed in the manner indicated above are delivered to arithmetic unit 173 which computes parameters g_1 and g_2 in accordance with the following relations:

$$g_1 = [-d_y + \sqrt{d_y^2 - d_2(d_1 - R^2 E)}]/d_2 \qquad (19)$$

$$0 \le d_1 \le R^2 E$$

$$g_2 = E(1 - R^2)/d_3$$

$$g_1 = 0 \qquad (20)$$

$$R^2 E \le d_1 \le E$$

$$g_2 = E(1 - R^2)/d_3$$

$$g_1 = 0 \qquad (21)$$

$$E \le d_1$$

$$g_2 = 0$$

Each of the operations indicated above is carried out sequentially every pitch period under the influence of clock signals (developed by pulser 171) synchronized with the positions of the glottal pulses from pitch pulse computer 14 (FIG. 1), as defined by signal N.

Arithmetic and storage units operative in a fashion similar to that described above are described in greater detail in the aforementioned copending application, Ser. No. 753,408.

Ordinarily, there are five resonances below the frequency of 5 kHz. in the human vocal tract. As discussed above, these resonances may be simulated by a transversal filter arrangement employing n-discrete delay elements. When $n=10$, the system can simulate $n/2$ resonances, i.e., the five resonances of the vocal tract. The synthesizer of this invention thus employs a discrete linear time-varying filter excited by a suitable combination of quasi-periodic pulses and white noise. A transversal filter arrangement is satisfactory for developing a linear combination of past output samples and the current input sample. Actual locations of resonances are determined in the transversal filter arrangement by the parameters a. Details of this form of resonance simulation is described in the above-mentioned Kelly U.S. Pat. No. 3,328,525. Transversal filter arrangements for use in speech synthesizers also have been described abundantly in the art. One suitable form is shown by way of a rudimentary block diagram in FIG. 4.

In the arrangement of FIG. 4 time-varying filter 24 (FIG. 1) includes a transversal filter network 25 composed of 10 unit delay elements $240, ..., 240_{10}$, supplying applied signals to 10 adjustable gain amplifiers $241_1, ..., 241_{10}$. Signals developed at the junctions of the several delay units thus represent past sample values of signals supplied from combiner 21 to filter 26 in the synthesizer of FIG. 1. The gains of the individual amplifiers 241 are adjusted by parameter values a to form a collection of weighted past sample values. The resultant signals are additively combined in adder network 242 and supplied to one input of combiner unit 21. As discussed above, the combined output of combiner network 21, which includes voiced and unvoiced excitation, and the combination of weighted past sample values constitute a replica of the applied speech signal. It is supplied by way of filter 26 to loud speaker 27.

Thus, in accordance with the invention an analog speech signal may be efficiently transmitted in the form of an array of numbers, viz., N; g_1, g_2; and $a_1, ..., a_{10}$. These parameters represent the necessary information concerning the speech wave in any given pitch period and are sufficient for reconstructing the speech wave. A saving of approximately 10 to 1 in transmission capacity may be achieved when using these parameters rather than the analog signal itself.

For example, a 10-kilobit signal used for representing the parameter has been found to yield excellent quality synthesized speech. A 5-kilobit signal still permits vary acceptable speech to be produced; this in contrast to the usual requirement of a 50-kilobit signal for direct coding of a speech wave.

Various other arrangements and modifications of the described arrangements will occur to those skilled in the art.

What is claimed is:

1. Speech analysis apparatus, which comprises:
   means for developing a first set of signals which specify linearly predictable characteristics of an applied speech signal,
   means for developing a second set of signals representative of the duration of individual pitch periods of said applied speech signal,
   means for developing a third set of signals representative of the energy of a speech signal and of the voicing character of speech signals within each of said pitch periods, and
   means for utilizing all of said developed signals together as a representation of said applied speech signal.

2. Speech signal analysis apparatus as defined in claim 1, wherein,
   said first set of signals which specify linearly predictable characteristics comprises a plurality of limited channel capacity parameter signals derived from past and current values of said applied speech signal for adjusting a resonant filter system, arranged to produce a replica of said applied speech signal when excited by voiced and unvoiced excitation signals.

3. Speech signal analysis apparatus, as defined in claim 1, wherein,
   said first set of signals comprises a sequence of signals $a=a_1$, ..., a_n, for each pitch period of said applied signals, which uniquely determine the frequencies and bandwidths of formants of said applied signal below approximately 5 kHz.

4. Speech signal analysis apparatus as defined in claim 3, in combination with,
   means supplied with said sequence of signals a for developing signals representative of the frequencies and bandwidths of formants of said applied speech signal during selected pitch periods.

5. Speech signal analysis apparatus as defined in claim 1, wherein,
   said first set of signals is developed by minimizing the mean-squared error between the actual values of samples of said applied speech signal and predicted values thereof based on a selected number of past sample values.

6. Speech signal apparatus, which comprises:
   at a transmitter station;
   means for developing a first set of signals which specify linearly predictable characteristics of an applied speech signal,
   means for developing a second set of signals representative of the duration of individual pitch periods of said applied speech signal,
   means for developing a third set of signals representative of the energy of a speech signal in each of said pitch periods and of the voicing character of speech signals within said pitch periods, and
   means for combining all of said developed signals for transmission to a receiver station; and
   at said receiver station;
   means responsive to received signals of said first set for developing signals representative of predicted values of a speech signal,
   means responsive to received signals of said second set for developing a sequence of pitch period pulses,
   means for generating white noise signals,
   means responsive to received signals of said third set for individually adjusting the levels of said pitch period pulses and said white noise signals, and
   means for combining said adjusted pitch period pulses, said adjusted white noise signals, and said predicted value signals to form speech signal which is a replica of said applied speech signal.

7. Speech signal apparatus as defined in claim 6, wherein,
   said means at said receiver station for developing signals representative of predicted values of said speech signal comprises,
   a transversal filter supplied with a combination of adjusted pitch period pulses, adjusted noise signals, and signals selectively representative of past values of said applied signal.

8. Synthesis apparatus for developing artificial speech from signals representative of the pitch period, voicing character, and selected predictable characteristics of an applied speech signal, which comprises:
   means responsive to received signals representative of selected predictable characteristics of an applied speech signal for developing signals representative of selected predicted values of said speech signal,
   means responsive to received signals representative of the pitch period of said applied speech signal for developing a sequence of pitch period pulses,
   means for generating white noise signals,
   means responsive to received signals representative of the voicing character of said applied speech signal for individually adjusting the levels of said pitch period pulses and said white noise signals, and
   means for combining said adjusted pitch period pulses, said adjusted white noise signals, and said predicted value signals to form speech signal which is a replica of said applied speech signal.

9. Synthesis apparatus as defined in claim 8, wherein
   said means for developing signals representative of predicted values of said speech signal comprises a transversal filter supplied with said combined replica signal and adjusted by said predictable characteristic signals.

10. Synthesis apparatus as defined in claim 8, wherein,
   said predicted value signals are selected to represent a linear combination of preceding values of said replica of said applied speech signal.

*   *   *   *   *