United States Patent

Tscheschner et al.

[15] **3,652,801**

[45] **Mar. 28, 1972**

[54] CIRCUIT ARRANGEMENT FOR SYNTHESIS OF ACOUSTIC ELEMENTS

[72] Inventors: **Walter Tscheschner**, Dresden; **Siegfried H. E. Fuchs**, Coswig, both of Germany

[73] Assignee: **Veb Elektronische Rechenmaschinen**, Karl-Marx-Stadt, Germany

[22] Filed: **Apr. 7, 1969**

[21] Appl. No.: **813,948**

[52] **U.S. Cl.** ..... **179/1 SA**

[51] **Int. Cl.** **G10l 1/00**

[58] **Field of Search** ..... 179/1 SA

[56] **References Cited**

UNITED STATES PATENTS

2,817,711 12/1957 Feldman 179/1

3,127,476 3/1964 David 179/1

Primary Examiner—Kathleen H. Claffy
Assistant Examiner—Douglas W. Olms
Attorney—Nolte and Nolte

[57] **ABSTRACT**

The artificial speech initiating electrical signal is in the form of periodic triangular vibrations or of a noise signal. The signal is selectively applied through a coordinate control matrix to input or excitation points of a conventional speech synthetizing potential field device which distributes the potentials of the signal along a number of output electrodes. Each of the electrodes is connected with a band-pass filter, respectively, and the output spectral portions are combined into the shaped spectrum of the desired articulated sound.

3 Claims, 4 Drawing Figures

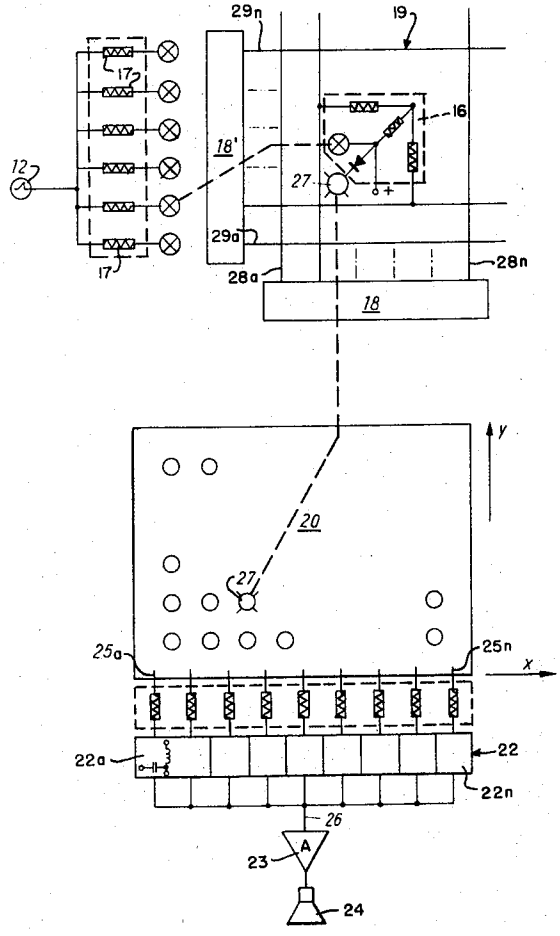
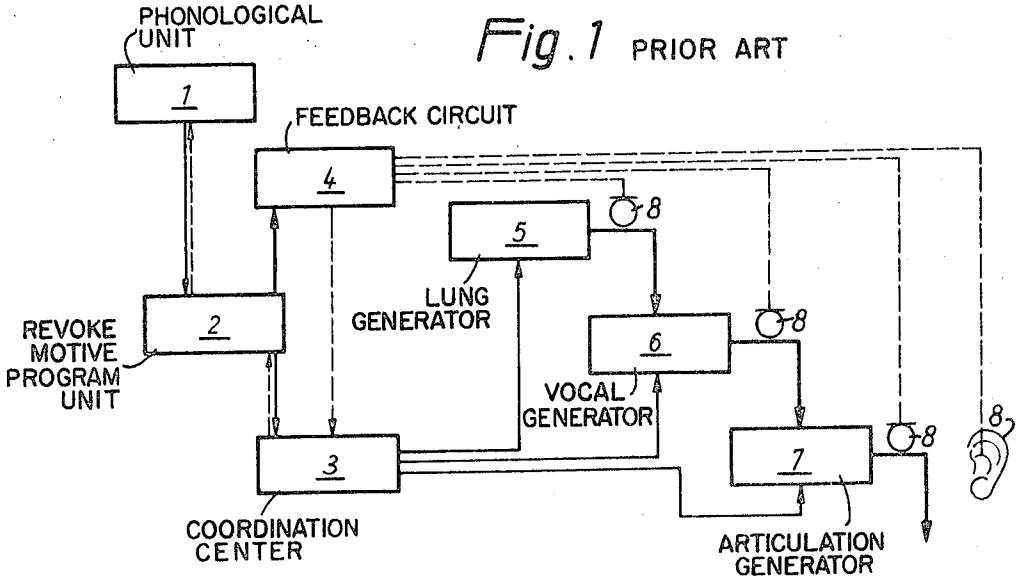


Fig. 1 PRIOR ART



PRIOR ART Fig. 2

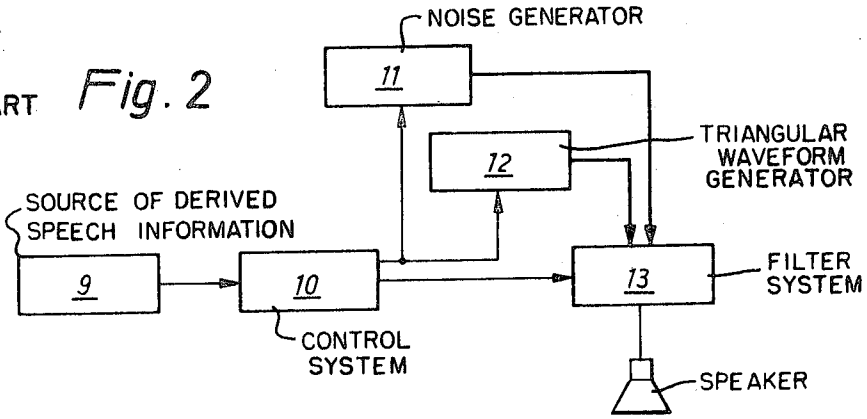
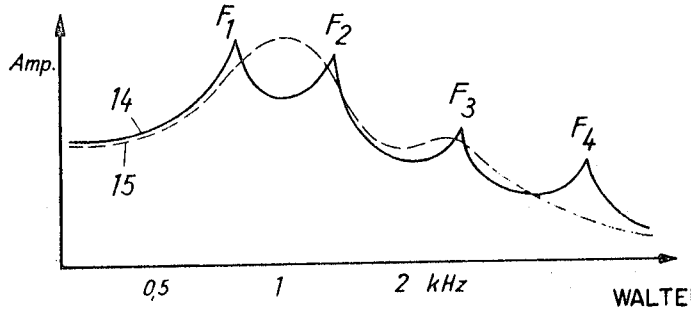


Fig. 3

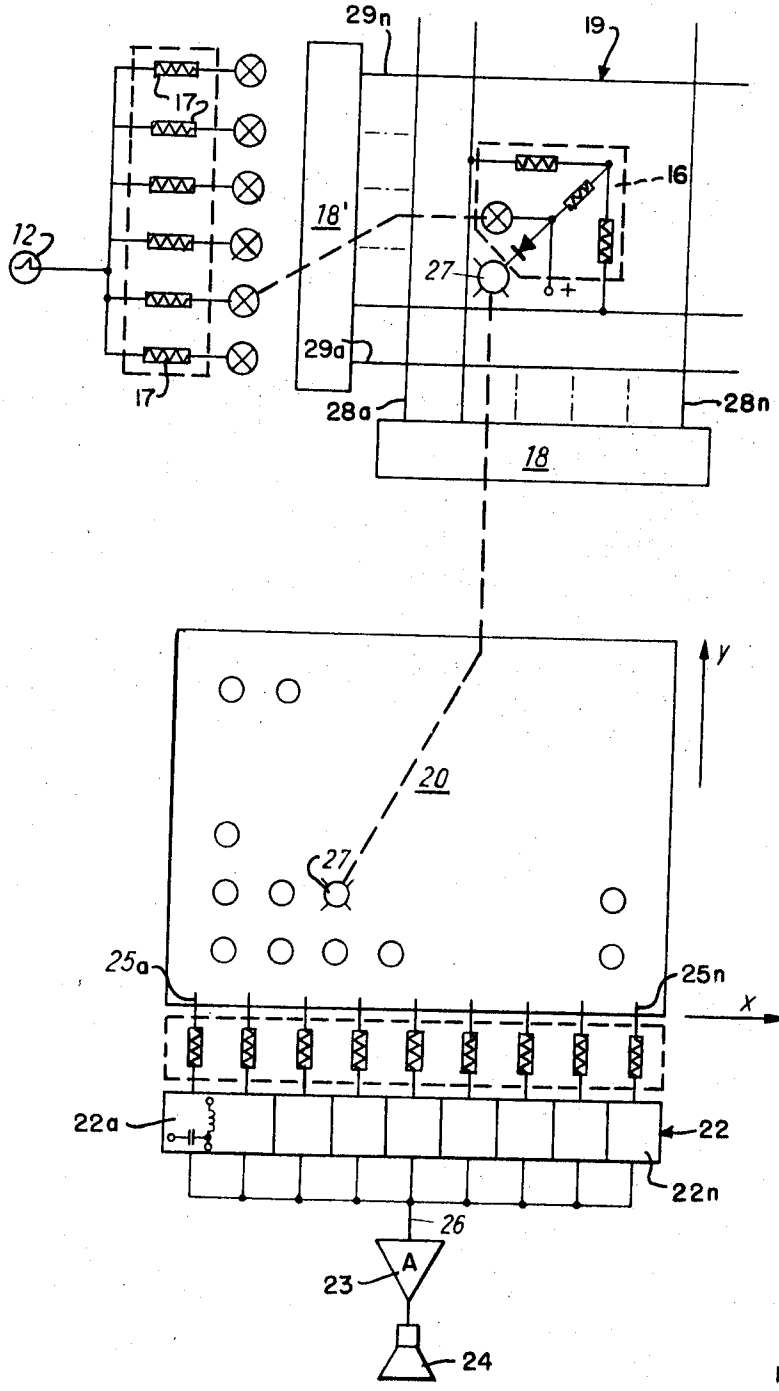


INVENTORS
WALTER TSCHESCHNER
SIEGFRIED FUCHS

BY Nolte & Nolte

ATTORNEYS

Fig. 4



INVENTORS
WALTER TSCHESCHNER
SIEGFRIED FUCHS

BY Nolte & Nolte

ATTORNEYS

CIRCUIT ARRANGEMENT FOR SYNTHESIS OF ACOUSTIC ELEMENTS

The present invention relates to a circuit arrangement for the production of artificial speech.

The essential peculiarity of organic processing of acoustic elements is that the acoustic elements which serve for human communication have a relatively high degree of accommodation to a particular organic process of conversion. As a consequence, a highly reliable identification is provided having the capability of forming numerous combinations of these elements.

Due to the very low degree of accommodation to the peculiarities of human communication which is attainable in existing technical systems, it is very difficult to make the systems for identifying and indicating the acoustic elements cooperate to a greater extent with the devices for data processing, communication compression systems, devices for rehabilitation and for the like applications.

Until now, there has been no satisfactory solution of this problem. Even the devices which are extremely costly have a very limited performance.

A substantial shortcoming of all of the known speech synthesizers is that the processes for shaping acoustic elements along the amplitude, frequency and time axes are not, due to the physical limitations of existing technical systems, controllable to the degree necessary for the articulation and especially, for the perception process. At the same time, it is very difficult to make a technical model having properties and control structures of organic partial systems and the complexity of the entire organic system.

FIG. 1 shows schematically in a block-circuit diagram an arrangement for the organic generation of acoustic elements originating, as rule, from discrete information. The information excites the independently operating generators and coordinates the cooperation of the generators according to the intended effect via process controlling systems.

Whereas in organic systems the individual physiological organs for controlling positions, such as, for example, positions of the jaw or tongue, or protrusion and the like, influence the peripheral physical system and the action thereof such as, for instance, the formant frequencies of the first up to fourth formants, the band-widths of the first to fourth formants, and the amplitudes of the formants, it is necessary in technical models to individually modify each of those parameters to attain analogous effects.

It is, therefore, indispensible in synthetizing systems operating according to a channel-vocoder principle to apply great quantity of controls. In systems having glide band-pass filters and operating on the formant-vocoder principle, the reproduction of dynamic processes is complicated as a result of the employment of the plurality of controls and also due to the physically conditioned persistency of the system. If it is attempted in such systems to adjust the physically conditioned relationships in order to simplify the control action, it will be found that the functional relations which are required for the improvement of communication cannot be accomplished.

Considerable advantages for speech forming composition of control data can be attained by means of a potential field device coupled with a channel system. A potential field device or potential distribution device as hereinafter discussed is comprised of an electrically conductive medium having a plurality of regularly arrayed spaced apart input electrodes connected to the medium, and a plurality of spaced apart output electrodes also connected to the medium. In this type of device, the potential at any given output electrode is a function of potential applied to the input electrodes as well as the geometry of the medium between the output electrode and the input electrodes to which potentials are applied. The potential field device or potential field matrix operates similarly as the so-called "electrolytic tank" in which a plurality of input electrodes or excitation points are regularly arrayed in an electrically conducting medium. Respective excitation points within the electrolytic tank or other potential distributing means are connected to a switch means for selectively applying electric

potentials thereto. The polarity of those potentials is applied via a plurality of output electrodes operatively disposed at the device to control modulators forming the frequency envelope of individual sounds. THe potential field device makes it possible to provide both the composition of control data and such physical limit conditions in the frequency and time axes which result in an excellent reproduction of speech. However, it has the disadvantage that the amplitude and time conditions thereof are subject to unfavorable limitations of physical nature and also that the unavoidable modulators due to their relatively very small operative range and unstability in time, may reduce the quality of speech.

Therefore, the object of the present invention is to provide a circuit arrangement for a speech synthesis which avoids the aforementioned disadvantages of the systems operating according to a channel or formant synthesis principle.

More particularly, the object of the present invention is to provide a circuit arrangement in which the technical parameters can be improved in such a manner that the artificial speech signals can be readily and easily produced.

In accordance with the present invention the speech stimulating signal which is applied to excitation points of a conventional potential field device, has the form of periodic triangular vibrations or of a noise signal. A n outputs of the potential field, there are coupled n band-pass filters, respectively, through which separate information pertaining to each of the sound elements are composed into acoustic spectrum of articulated and acoustically modified sounds.

By superposing a plurality of potential field devices, the amplitudes of source currents or voltages can be processed in uniformly designed channel stages and the limiting conditions can be compensated to obtain more suitable form.

The invention has the advantage that a number of very critical structural stages can be avoided, and the arrangement of the potential field can be still better adjusted to operating conditions of organic information processing systems.

The present invention will now be explained in more detail by way of an example with reference to accompanying drawings, in which:

FIG. 1 is a block-circuit diagram of an organic speech transmitting system;

FIG. 2 is a block diagram of a technical speech synthetizer;

FIG. 3 is a graph showing the envelope of the sound "a" (Curve 14) and of a simulated function of "a" Curve 15); and

FIG. 4 is a schematic diagram of an embodiment of the speech synthetizer according to the present invention.

FIG. 1 shows a block-circuit diagram of an organic speech transmitting system which usually serves as a model for technical embodiments. The speech forming process is initiated by selecting a series of elementary phonological units 1 which are mostly superimposed with super-segmental units and revoke motive programs 2 through which the realization of the speech is effected in peripheral generator systems 5, 6 and 7. The control commands for lungs generator 5, vocal generator 6 and articulation generator 7 are released by neuro-motive coordination centers 3 whereby the execution of the entire speech realization is mainly performed by neuro-motive feedback 4 originating at tactile and acoustic sensing organs.

The actual technical synthesizers have a substantially lower degree of complexity. Also, the mutual control of partial systems is relatively decreased. In the majority of technical synthesizers (as seen in FIG. 2) the time functions of spectra of natural speech are simulated by means of filter system 13 controlled according to a time function for providing a spectral envelope with variable formants. Control system 10 determines the excitation of noise signals from noise generator 11 for unvoiced sounds or consonants, or the excitation of triangular periodic functions from triangular-waveform generator 12 for voiced sounds, thus determining the position and shape of formants for an articulated sound. The input of control system 10 can be supplied with results of a continuous physical speech analysis (vocoder), with a series of vocal demonstrations (voder), with digital information (phonetic output),

with a series of sounds (from a speech machine) or with any other derived speech information **9**. There must always be provided a number of various control signals to accomplish the synthesis of speech signals in the given synthesizer system, such as a signal for the actuation of triangular-waveform generators for producing the vocal articulated sounds, a signal for the actuation of noise generators for producing the unvoiced sounds, and further signals for the adjustment of frequency of the triangular-waveform generators. Besides, there must also be provided information of the position, bandwidth and the amplitude of formants.

With reference to the evaluation of discrete singular points of envelope curve **14** of the instantaneous spectrum as seen in FIG. **3**, at least three formants are necessary, i.e., six formant control signals, eventually as many channel control signals as to enable a sound coordination according to the produced formant positions. As to the shape of function **15** defining the sound perception, at least ten channels must be controlled or, eventually, as the case may be in the controlling of a potential field device, two pole sites controlled along two coordinate axes.

FIG. **4** illustrates schematically the principle of a circuit arrangement according to this invention for accomplishing a synthesis of vocal sounds at one pole site. Although the embodiment of FIG. **4** relates to the formation of a shape of a single spectral envelope, it can be employed also for the production of a number of formant positions. The triangle-like speech excitation signal from the triangle signal generator **12** is applied through resistor network **17** to coordinate control field or matrix **19** which defines the position of actuated excitation points **27** in potential field device **20**. According to the position of actuated excitation points, a potential distribution of the triangular-waveform voltage at the *n*-number of output electrodes **25** along the X-axis of potential field device **20** is produced as a function of the position of the actuated excitation points in the potential field device **20**.

The positions of the maxima of the potential distribution of the compound triangular signal are, therefore, determined by positions of the excitation points on the x-axis and the shape thereof is defined by the position of the excitation points of the y axis. A separate filter **22***a* to **22***n* of filter set **22** is connected to each of the *n* electrodes **25**. The filters of filter set **22** always select the magnitude of the corresponding spectral portions of the broad band frequency spectrum as determined by the potential field device **20**. Consequently, the combined signal voltages at the output **26** of filter set **22** form the synthesized sound spectrum which may be amplified in amplifier **23** and reproduced by speaker **24**.

Since respective filters **22***a* to **22***n* in filter set **22** are arranged in consecutive order according to corresponding frequency groups or bands, the spectrum is in a form which is suitable for the process of acoustic resolution and evaluation.

Coordinate control matrix **19** can be designed in various manners, either electronically or by means of electromechanical coordinate switches. The preferred embodiment is based on the electronic solution, in which two counters or selectors **18** and **18'** are in operative relation to respective column and row conduits **28***n* and **29***n* arrayed in a matrix-like pattern. Respective points of intersection of the conduits are coupled to the excitation points and corelated output electrodes of the potential field device through a plurality of diode switching means **16** which actuate the electrodes in response to the action of selectors **18**.

Resistors **17** at inputs of coordinate control matrix **19** and resistors **21** between respective output electrodes of the potential field device and the filter set **22** modify in accordance with the distribution of their values the potentials at respective terminals along the X and Y coordinates to compensate technical limit conditions of the entire circuit arrangement. Furthermore, by means of resistors **17** and **21** it is possible to simulate in the potential field device the position function of the formants as predetermined by various configurations of the spectrum of an articulated sound. For a consecutive amplitude-position potential field device (not shown), resistors **17** and **21** can be combined to substitute the resistor set **17**. The application of a second amplitude-position potential field device is particularly advantageous when different modes of the speech excitation require different nonlinear amplitude-position functions. The coupling of the second potential field device to the first one is accomplished by means of coordinates control matrix similar to the matrix **19**.

What is claimed is:

1. A circuit for the production of artificial speech from acoustic elements comprising signal generator means for supplying a signal having a wide spectrum of frequencies and amplitudes, coordinate control matrix means having a plurality of input conduits and a number of intersection points, signal selector means for connecting said input conduits to said generator means, a potential field device comprised of an electrically conductive medium having a plurality of regularly arrayed spaced apart input electrodes and a plurality of spaced apart output electrodes connected thereto, whereby the potential at a given output electrode is a function of potentials applied to said input electrodes and the geometry of said medium between said output electrode and input electrodes to which potentials are applied, means coupling said input electrodes to said intersection points for providing a potential distribution of said signal along the output electrodes in response to the signals applied to said input conduit by way of said selector means, a plurality of band-pass filter means coupled to said output electrodes to select the magnitude of a spectral portion of the signal as determined by the action of said potential field device, and means for coupling the outputs of respective filter means to produce a synthesized sound spectrum.

2. The circuit according to claim **1** wherein said signal generator means is a noise signal generator.

3. The circuit according to claim **1** wherein said signal generator means is triangle-waveform signal generator.

* * * * *