



WO 96/26592

THIS SHEET HAS BEEN
REPLACED

PCT/US96/02073

**APPARATUS AND METHOD FOR CANCELING
ACOUSTIC ECHOES INCLUDING NON-LINEAR
DISTORTIONS IN LOUDSPEAKER TELEPHONES**

This application is related to Application
Serial No. 08/393,726 entitled "Apparatus and Method
for Adaptively Precompensating for Loudspeaker
Distortions" filed concurrently herewith and assigned
5 to the assignee of the present invention, the
disclosure of which is hereby incorporated in its
entirety herein by reference.

Field of the Invention

This invention relates to the field of
10 telephony, and more particularly to the suppression of
echoes in loudspeaker telephones.

Background of the Invention

A loudspeaker telephone system includes an
output transducer, such as a loudspeaker, and an input
15 transducer, such as a microphone. The loudspeaker
produces sound pressure waves in response to an input
signal received from the distant party which is
representative of a desired sound pressure wave, and
the microphone receives sound pressure waves to be
20 converted to an output signal and transmitted to the
distant party. Because the loudspeaker broadcasts
sound pressure waves into the environment around the
loudspeaker telephone, there is an acoustic path from
the loudspeaker to the microphone which may result in
25 an echo. Typically, this acoustic path includes a
plurality of channels (representing a plurality of
reflections) so that a plurality of echoes reach the
microphone at different times.

If nothing is done to compensate for this
30 acoustic path, sound pressure waves generated by the
loudspeaker will echo back through the microphone to
the distant user. In practice, this means that when

-2-

the distant party speaks, the speech will be broadcast by the loudspeaker and then transmitted back making conversation difficult. Accordingly, there have been attempts in the art to reduce these echoes.

5 For example, the reference by Park et al. entitled "Acoustic Echo Cancellation for Full-Duplex Voice Transmission on Fading Channels" discusses the implementation of an adaptive acoustic echo canceler for a hands-free cellular telephone operating on a
10 fading channel. *Proc. of International Mobile Satellite Conference*, Ottawa, Ontario, Canada, June 18-20, 1990. The adaptive lattice structure, which is particularly known for faster convergence relative to the conventional tapped delay line (TDL) structure, is
15 used in the initialization stage. After convergence, the lattice coefficients are converted into the coefficients for the TDL structure which can accommodate a larger number of taps in real-time operation due to its computational simplicity.

20 Other approaches to echo cancellation are discussed in the reference by Burnett et al. entitled "Echo Cancellation in Mobile Radio Environments", IEE Colloquium on Digitized Speech Communication via Mobile Radio, (Digest No. 139), p. 7/1-4, IEE, London, UK,
25 December 19, 1988. The conventional approach to providing echo attenuation is to use voice switched attenuators. As described therein, the half duplex channel enforced by such voice operated switches and the imperfections in the current voice activity
30 detectors may lead to unnatural conversation. Another solution is to attenuate the echo by means of an adaptive echo canceler. As also described therein, this process may be imperfect, however, because any non-linearities in the echo path may cause harmonics
35 which are inherently uncanceled. While microphones and amplifiers have more than adequate linearity

-3-

specifications, loudspeakers usually have no linearity specification at all.

Notwithstanding the above mentioned references, there continues to exist a need in the art
5 for improved loudspeaker telephone systems and methods which reduce echoes from the loudspeaker to the microphone.

Summary of the Invention

It is therefor an object of the present
10 invention to provide improved echo cancellation systems and methods for loudspeaker telephones.

It is another object of the present invention to provide improved echo cancellation systems and methods for mobile cellular radiotelephones.

15 These and other objects are provided according to the present invention by a loudspeaker telephone system comprising an echo filter, including a loudspeaker model, which generates an estimated echo signal. Most loudspeakers generate an actual sound
20 pressure wave that differs from the desired sound pressure wave represented by the input signal. This difference is due, in part, to non-linear aspects of the loudspeaker. For example, the diaphragm of a loudspeaker may have a non-linear stress-strain curve,
25 and the motion of the diaphragm may result in the delay modulation of higher frequencies by lower frequencies. Prior art echo cancelers known to applicants, however, fail to account for non-linear portions of the echo generated by the loudspeaker, meaning that these non-
30 linear portions may be transmitted back to the distant party.

Accordingly, the estimated echo signal of the present invention includes non-linear components which correspond to the non-linear portions of the echo
35 resulting from non-linear aspects of the loudspeaker. Therefore, when the estimated echo signal is subtracted



-4-

from the output signal generated by the microphone, non-linear portions of the echo generated by the loudspeaker can be reduced. This echo filter also includes an acoustic path model which generates an
5 estimate of the acoustic path from the loudspeaker to the microphone so that echo effects resulting from the acoustic path can be reduced.

In a preferred embodiment according to the present invention, an echo-canceling loudspeaker
10 telephone includes an output transducer which produces a sound pressure wave in response to an input signal which is applied to an audio input thereof. This sound pressure wave includes a desired linear component which is a linear function of the input signal, and an
15 undesired non-linear component which is a non-linear function of the input signal. This sound pressure wave is transmitted along an acoustic path, and an input transducer is positioned in the acoustic path to convert the sound pressure wave into an output signal.

20 An echo filter is responsive to the input signal and generates an estimated echo signal. The echo filter includes an output transducer model for generating an estimate of the sound pressure wave including an estimate of the linear component and an
25 estimate of the non-linear component. The echo filter also includes an acoustic path model which generates an estimate of the acoustic path from the output transducer to the input transducer. A combiner combines the estimated echo signal and the output
30 signal thereby reducing an echo portion of the output signal.

This echo canceling loudspeaker telephone system reduces echoes caused by the acoustic path from the output transducer to the input transducer which are
35 a linear function of the input signal. This system also reduces portions of the echoes generated by the non-linear aspects of the loudspeaker which are a non-

-5-

linear function of the input signal. Accordingly, this system reduces echoes in the output signal to a greater degree than would be possible without the output transducer model.

5 The echo-canceling loudspeaker telephone may also include an echo filter modifier responsive to the output signal, the input signal and the estimated echo signal for modifying the echo filter to further reduce the echo portion of the output signal. Accordingly,
10 the system can adapt its operation for maximum echo reduction in a changing environment where the acoustic path varies. The echo filter modifier may include an output transducer model modifier which modifies the estimate of the sound pressure wave including the
15 linear component and the non-linear component. The echo filter modifier may also include an acoustic path model modifier for improving the estimate of the acoustic path from the output transducer to the input transducer.

20 The acoustic path model is preferably a finite-impulse-response filter, and the echo filter is preferably a digital signal processor, allowing implementation with existing hardware. In addition, the output transducer model preferably includes means
25 for performing a transformation of the input signal. This transformation is preferably an estimate of a transfer function of the output transducer means, and the transfer function preferably includes a non-linear component thereby providing a relatively precise
30 approximation of the output transducer transfer function.

 The non-linear component of the sound pressure wave represents one of a delay modulation of the output transducer and/or a diaphragm stress-strain
35 curve of the output transducer. In a preferred embodiment, the output transducer is a loudspeaker, and the input transducer is a microphone. The output

-6-

transducer model and the acoustic path model may be connected in series between the audio input and the combiner. In addition, the acoustic path model may include a first processing block for generating an estimate of the acoustic path for the non-linear component of the sound pressure wave and a second processing block for generating an estimate of the acoustic path for the linear component of the sound pressure wave. In addition, the combiner may be a subtractor which subtracts the estimated echo signal from the output signal.

The foregoing and other objects and aspects of the present invention are explained in detail in the drawings and specification set forth below.

Brief Description of the Drawings

Figure 1 is a schematic diagram of an echo canceling loudspeaker telephone system including a loudspeaker, a microphone, and an echo filter including a model of the loudspeaker and a model of the acoustic path arranged in series, according to the present invention.

Figure 2 is a schematic diagram representing a model of the electrical characteristics of the loudspeaker shown in Figure 1.

Figure 3 is a schematic diagram of an echo canceling loudspeaker telephone system including a loudspeaker, a microphone, and an echo filter including a model of the loudspeaker in series with two separate filters used to model the acoustic path, according to the present invention.

Figure 4 is a schematic diagram of an echo canceling loudspeaker telephone system including a loudspeaker, a microphone, and an echo filter including two processing blocks to model the loudspeaker and two processing blocks to model the acoustic path, according to the present invention.

-7-

Detailed Description of the Invention

The present invention will now be described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the present invention are shown. This invention may, however, be embodied in many different forms and should not be construed as limited to the embodiment set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like numbers refer to like elements throughout.

The echo cancellation system illustrated in Figure 1 can be implemented in a loudspeaker telephone such as a hands-free loudspeaker cellular radiotelephone for use in an automobile. When implemented as a hands-free cellular telephone, speech signals received from a distant party are transmitted from a cellular base station (not shown), received by the transceiver of the cellular phone (not shown), and applied to input node 36 as input waveform $W(t)$.

As shown in Figure 1, the waveform $W(t)$ is applied in an analog format at node 36, and converted to a digital format by A-to-D converter 30 for use by the loudspeaker model 12. D-to-A converter 26 is then used to convert the waveform to an analog format. The analog signal is amplified by amplifier 27, and a sound pressure wave representative of the speech of the distant party is broadcast by an output transducer such as loudspeaker 14. Accordingly, the radiotelephone user hears sound pressure waveforms which are representative of the speech of the distant party. Alternately, the analog signal $W(t)$ at node 36 can be applied directly to amplifier 27. If the waveform $W(t)$ is applied in a digital format at node 36, then it can be applied directly to the loudspeaker model and the D-to-A converter 26.

-8-

The sound pressure wave, however, is also broadcast along the acoustic path 18 which can include multiple channels. A channel is a reflection (or echo) path from the loudspeaker to the microphone. As a
5 result, echoes of the sound pressure wave are received by an input transducer such as microphone 20. It is therefor desirable to reduce these echoes in the output signal $Z(t)$ generated by the microphone 20 so that the distant party is not confused by delayed echoes of his
10 own speech. This echo reduction is preferably achieved by using an echo filter implemented as digital signal processor ("DSP") 21 to generate an estimated echo signal, and to subtract this estimated echo signal from the microphone 20 output signal. Accordingly,
15 amplifier 37 and A-to-D converter 28 can be used to convert the output waveform $Z(t)$ to an appropriate digital format.

The echo filter is preferably implemented as a digital signal processor ("DSP") 21 which generates
20 an estimated echo signal $Z'(t)$ in response to the input waveform $W(t)$. If the input waveform is applied in an analog format, A-to-D converter 30 can be used to convert the waveform to a digital format. (If $W(t)$ is applied in a digital format, D-to-A converter 30 is not
25 needed.) The input waveform is then applied to loudspeaker model 12 within the echo filter which includes a transfer function representative of the loudspeaker. This transfer function models both linear and non-linear aspects of the loudspeaker 14. The
30 output from the loudspeaker model 12 is applied to the acoustic path model 34, implemented within the echo filter, which represents the acoustic path 18. The acoustic path model 34 is preferably implemented as an adaptive finite-impulse-response ("FIR") filter.
35 Accordingly, the estimated echo signal $Z'(t)$ from the acoustic path model 34 can closely approximate the echo from the loudspeaker received by the microphone.

-9-

Adaptive filters used in echo cancellation are discussed, for example, in U.S. Patent No. 5,237,562 to Fujii et al., entitled "Echo Path Transition Detection." Other echo cancellers including
5 adaptive echo estimation or including a finite impulse response filter are respectively discussed in U.S. Patent No. 5,131,032 to Esaki et al., entitled *Echo Canceller and Communication Apparatus Employing the Same*," and U.S. Patent No. 5,084,865 to Koike, entitled
10 *Echo Canceller Having FIR and IIR Filters for Cancelling Long Tail Echoes*." Each of the three above cited references are hereby incorporated in their entirety herein by reference.

The estimated echo signal $Z'(t)$ is combined
15 with the microphone 20 output signal $Z(t)$ by a subtractor 22, also implemented within the DSP, which subtracts the estimated echo signal from the output signal. Accordingly, only sounds generated at the radiotelephone will be transmitted to the distant
20 party. In a cellular radiotelephone application, the output waveform $Z(t)$ (minus the estimated echo signal $Z'(t)$) at output node 38 is applied to the radiotelephone transceiver (not shown) and transmitted to a remote cellular base station (not shown). If the
25 transceiver requires an analog output waveform, D-to-A converter 32 can be used to convert the signal to an analog format. If the transceiver requires a digital waveform, D-to-A converter 32 is not needed.

The echo filter also preferably includes an
30 echo filter modifier 16 which modifies the operation of the echo filter allowing further reduction of the echo portion of the output signal. The echo filter modifier monitors the various signals within the echo filter and modifies the operation of the loudspeaker model and the
35 acoustic path model to further reduce echoes in the output waveform. Accordingly, the echo filter can

-10-

modify its operation to accommodate changes in the acoustic path as well as aging of the system.

Figure 2 shows an analog model of the electrical characteristics of an output transducer such as loudspeaker 14. An electrical input signal is applied at input node A to create a current through the loudspeaker coil. The current flow is opposed by the coil resistance 40 and coil inductance 42, as well as the back EMF induced by the coil velocity in the magnetic field. By suitable choice of units and scaling in the model, the voltage at node C may be equal to the back-EMF as well as being representative of the coil velocity. The back EMF from node C is presented in opposition to the drive voltage at input A by connection to the positive input of differencing operational amplifier 44. The output of amplifier 44 is the sum of the back EMF from node C and a term proportional to the current in the coil. Amplifier 46 subtracts the back EMF to yield a voltage representing the current in the coil only, and by suitable choice of arbitrary units, this voltage also represents the force the coil exerts on the loudspeaker diaphragm by the current reacting with the magnetic field produced by the loudspeaker magnet. As will be understood by those having skill in the art, the term diaphragm is used throughout this specification in its broadest sense so as to include a planar diaphragm, a dome shaped diaphragm, or a cone shaped diaphragm.

The force causes an acceleration of the loudspeaker diaphragm to a certain velocity which is resisted by the diaphragm's mass or inertia and by air resistance encountered. Operational amplifier 48 has a feedback capacitor 50 representing the diaphragm's mass and a feedback resistor 52, which might be non-linear, representing the air resistance acting against the diaphragm. The current flow through resistor 52 opposes the accelerating force and relates to the air

-11-

pressure wave created by the diaphragm movement. Current sensor 54 generates a signal at node C' which represents this air pressure wave created by the diaphragm movement.

- 5 The pressure wave, however, emanates from a moving object, the diaphragm. When the diaphragm is instantaneously displaced to the front of the loudspeaker, it will be closer to a listener in front of the loudspeaker. Accordingly, sound waves will
- 10 reach the listener with a shorter time delay than when the diaphragm is displaced toward the rear of the loudspeaker. Diaphragm displacements occur with greatest amplitude at low frequencies giving rise to the non-linear phenomenon of delay modulation (also
- 15 known as phase modulation) of higher frequencies by lower frequencies. A signal representative of the diaphragm displacement is generated at node D by resistance 60, capacitance 62, and operational amplifier 64, which together make up integrator 65.
- 20 Thus the pressure wave signal from the diaphragm generated at node C' is subjected to delay modulation produced by delay modulator 66 according to the diaphragm displacement signal generated at node D in order to produce the net sound pressure waveform at
- 25 output node B that is transmitted to a listener.

- The diaphragm displacement signal generated at node D is also needed to model the diaphragm spring restoring force that opposes the force exerted by the coil which is represented by the coil force signal
- 30 generated by operational amplifier 46. The diaphragm spring is expected to exhibit a non-linear stress-strain curve modelled by the non-linear resistor 56. Operational amplifier 58, having non-linear resistor 56 in its feedback path, converts the displacement-related
- 35 signal generated at node D to a restoring force which adds in opposition to the coil force signal at the

-12-

input of operational amplifier 48. The resistors labeled R_0 may be equal to 1 ohm.

Thus, with appropriate choice of parameters and scalings in the above-described model of Figure 2, the sound pressure wave generated at loudspeaker output node B can be predicted from the electrical signal applied to the loudspeaker input node A.

According to a preferred embodiment of the present invention, a loudspeaker model 12 including an estimation of loudspeaker non-linearities is included in the echo cancellation path of a full-duplex echo canceler in order to improve prediction of echoes that will be received back into the microphone 20 of a loudspeaker telephone. Figure 1 shows inclusion of this model 12 into the block diagram of an echo canceler. An input speech waveform $W(t)$ is received from the telephone system, and after suitable processing is applied to input node 36 and then loudspeaker 14. In a mobile phone system, such processing can include demodulation of a digitally-modulated radio signal, error correction decoding and speech decoding using, for example, a Residually Excited Linear Prediction ("RELP") or Vector Set Excited Linear Prediction ("VSELP") speech synthesizer. The waveform $W(t)$ is the output of such processing, and may be in a digital format which is more suitable for processing by the echo canceler of the present invention. In this case, $W(t)$ can be converted by D-to-A converter 26 and amplified by amplifier 27 before being applied to the speaker 14. The loudspeaker 14 broadcasts a sound pressure wave including non-linear distortion components, into the environment, and some of this sound pressure wave reaches microphone 20 by a variety of delayed channels along acoustic path 18. The microphone output waveform $Z(t)$ is amplified by amplifier 37 and preferably sampled and A-to-D converted using A-to-D converter 28 to produce

-13-

digitized samples of waveform $Z(t)$. The original speech waveform $W(t)$ is also processed using a loudspeaker model 12 and a model 34 of the multi-path acoustic coupling from loudspeaker to microphone, in order to produce an estimate $Z'(t)$ of the microphone output signal $Z(t)$. This is subtracted by subtractor 22 to leave a residual echo signal $E(t)$ which is desired to be reduced.

If the near-end party is speaking, $E(t)$ also contains near-end speech which is transmitted to the telephone network. The signal $E(t)$, if in digital form, may be converted to analog form if necessary for onward transmission using D-to-A convertor 32. In a mobile cellular phone system, other processing such as RELP or VSELP speech coding, for example, can be used to produce a reduced bit-rate representation of the signal, error correction coding, and digital modulation on a radio frequency carrier.

An echo filter modifier 16 is used to update the coefficients of the acoustic path model 34 so as to obtain the least mean square value of the residual $E(t)$. This is done by computing the cross-correlation between the waveform $V(t)$ at the input of the linear channel model and the residual error $E(t)$, and determining by how much each coefficient shall be changed to produce better cancellation. For example, if $E(t)$ shows strong correlation with $V(t)$ delayed by 23 samples, then the FIR coefficient applied to delay tap number 23 of the filter is adjusted to remove this correlated component from the output of subtractor 22. This process will be known to those having skill in the art and will not be described further here. The inclusion of the transducer model 12 to produce the waveform $V(t)$, however, permits improved echo cancellation.

Any modification of either the loudspeaker model 12 or the acoustic path model 34 is preferably

-14-

done when only the distant party is speaking. This condition allows a comparison of the input waveform $W(t)$ to the residual echo without the interference of other sounds not generated by the loudspeaker. In one embodiment, this condition is determined by comparing the signal strength into the loudspeaker to the signal strength out of the microphone. A device that determines when the microphone signal is substantially derived from acoustic feedback is discussed, for example, in U.S. Patent No. 5,263,019 to Chu, entitled *Method and Apparatus for Estimating the Level of Acoustic Feedback Between a Loudspeaker and Microphone*, the disclosure of which is hereby incorporated in its entirety herein by reference.

The invention may also include the use of the echo filter modifier 16 to adjust the parameters of the loudspeaker model 12 in order to further improve echo cancellation by removing residual distortion components that are not modeled by the linear FIR filter of the acoustic path model 34.

The waveform $V(t)$ is calculated by the transducer model 12 from the waveform $W(t)$. These waveforms are assumed to be represented by numerical samples . . . , $W(i-1)$, $W(i)$, $W(i+1)$, . . . , and . . . , $V(i-1)$, $V(i)$, $V(i+1)$, Likewise, the internal waveforms at nodes B, C, C', D, and I, shown in Figure 2, are represented by discrete-time samples. The computation of the output waveform samples $W(i)$ proceeds using the following equations:

$$I(i) = (A(i) + Y \cdot I(i-1) - C(i-1)) / (R + Y) \quad (1)$$

$$C(i) = C(i-1) + (G(D(i-1)) - U(i-1) - I(i)) / X \quad (2)$$

$$D(i) = D(i-1) - C(i) \quad (3)$$

$$U(i) = F(C(i)) \quad (4)$$

and

$$W(i) = U(i) - 0.5(U(i+1) - U(i-1)) \cdot D(i) \cdot dT \quad (5)$$

where

-15-

Y is equal to the coil inductance L divided by the sample time spacing;

R is the coil resistance;

G is a non-linear function representing the
5 diaphragm spring stress-strain relationship;

X represents the diaphragm mass parameter 52 divided by the sample time spacing; and

F is a non-linear function representing the conversion of diaphragm velocity waveform C(i) to a
10 sound pressure waveform U(i).

Equation (5) expresses delay modulation of the pressure wave U(i) which results from the motion of the diaphragm. This calculation is done by interpolating between samples U(i) using the derivative
15 $0.5(U(i+1) - U(i-1))$, by an amount depending on the diaphragm instantaneous displacement D(i) and a scaling factor dT representing the amount of delay modulation.

For example, if the sample rate is 8k samples per second, samples . . . , (i-1), (i), (i+1), . . . are
20 125 μ S apart. In 125 μ S sound travels approximately 1.5 inches, so if the diaphragm displacement D were computed by integrator 65 in units of 1.5 inches, D=1 would signify one whole sample delay. The formula would then be changed to:

25 $W(i) = U(i-1)$ for D(i) = 1

or $W(i) = U(i+1)$ for D(i) = -1

Since D is expected to be less than 0.5 however, equation (5) is more appropriate.

In the equations discussed above, the sign of
30 the delay modulation has been arbitrarily assumed. It may be necessary to flip the sign of the delay modulation, which is one of the reasons for introducing the scaling factor dT, which can be positive or negative. The other reason to include the scaling
35 factor is to permit D(i) to be computed in units other than 1.5 inches. The units can be chosen to be suitable for computing the function G(D(i)).

-16-

It can be acceptable to assume a linear conversion of diaphragm velocity to pressure, in which case $C(i) = U(i)$. An arbitrary scaling here represents the fact that no particular units have been assumed for defining the conversion of electrical signals to sound waves. The following four equations then result:

$$I(i) = (A(i) + Y \cdot I(i-1) - C(i-1)) / (R + Y) \quad (6)$$

$$C(i) = (G(D(i-1)) + X \cdot C(i-1) - I(i)) / (1 + X) \quad (7)$$

$$D(i) = D(i-1) - C(i) \quad (8)$$

10 and

$$W(i) = C(i) - 0.5(C(i+1) - C(i-1)) \cdot D(i) \cdot dT. \quad (9)$$

There are now only two non-linear effects modelled in these equations. These non-linear effects are the delay modulation, which is represented by the addition of an amount dT times a distortion waveform which is the product of the derivative and the integral of $C(t)$; and the diaphragm stress-strain curve which is represented by a function $G(t)$.

$G(t)$ can be partitioned into a linear stress-strain curve of slope G_0 plus the non-linear remainder $G'(D) = G(D) - G_0 \cdot D$. The purpose of this is to enable equation (7) to be replaced with the following small-signal version:

$$C(i) = C(i-1) + (G_0 \cdot D(i-1) - U(i-1) - I(i)) / X. \quad (10)$$

25 This equation can then be used with equations (6) and (8) to predict small-signal behavior of the loudspeaker. The small signal and linear parameters can then be determined for the loudspeaker by measurement.

30 The determination of the coil resistance 40 and inductance 42 parameters R and Y will be understood by one having skill in the art, while the diaphragm mass and linear part of the diaphragm stress-strain curve G_0 can be determined by measuring the diaphragm's mechanical resonant frequency and Q factor when the speaker is in its intended housing.

-17-

The small-signal parameters are then fixed; and the non-linear parameters, dT representing delay modulation and G' representing the non-linear part of the stress-strain curve, may be determined by large
5 signal measurements. Delay modulation may be determined, for example, by observing with a spectrum analyzer the intermodulation produced on a two-tone test between a low frequency sine wave signal that causes large diaphragm displacements and a high
10 frequency sine wave signal that is most sensitive to phase modulation by the low-frequency diaphragm displacements.

The non-linear part of the stress-strain curve can be obtained by using a spectrum analyzer to
15 observe the harmonic distortion of a large, low-frequency, sine wave signal as a function of amplitude and finding a function G' by trial and error that explains it. The function can be represented in a numerical signal processor by a look-up table.
20 Alternatively, this curve can be directly determined by physical measurements of force or DC current required to displace the diaphragm a measured amount. The invention may include the provision of a diaphragm displacement or movement sensor for the purpose of
25 assisting in real-time determination or adaptive updating of model parameters.

In practice, a stress-strain curve G' may be assumed to be known apart from a scaling factor for a particular loudspeaker. Likewise, it may be assumed
30 that the linear model parameters resulting in particular diaphragm mechanical resonances are well known for a particular loudspeaker size and make. Small errors in small-signal parameters that effect small-signal frequency response are not of great
35 consequence as any system is assumed to have some ability to adapt linear frequency responses to

-18-

compensate. For example, a manual equalizer or tone control may be used.

In a loudspeaker telephone, the linear frequency response from the loudspeaker 14 to the microphone 20 across acoustic path 18 includes reflections from nearby objects and possible room resonances, generally referred to by the term "environment". Accordingly, this response can be modeled by an acoustic path model 34 including a complex linear FIR filter. The echo or ring-around, however, is imperfectly modeled due to the non-linear effects discussed above which are not modeled by the linear FIR filter thereby resulting in imperfect echo cancellation. Using the non-linear echo cancellation system shown in Figure 1, however, the channel from electrical input, to the loudspeaker amplifier, to the loudspeaker, across the acoustic path, to the microphone, and through the microphone amplifier is more accurately matched by the combination of loudspeaker model 12 and acoustic path model 34, thereby providing better echo cancellation. It is now described, with the aid of Figure 1 and the equations discussed above, how the non-linear transducer model parameters can be adapted in real time to continuously reduce residual uncanceled echo-distortion residuals.

Referring to equation (9), it can be seen that the output $V(t)$ of loudspeaker model 12 includes the sum of two waveforms, $C(i)$ and $B(i) = 0.5(C(i+1) - C(i-1)) * D(i)$, the latter being scaled by $-dT$. Since the FIR filter 34 discussed above is linear, its output is the sum of two waveforms obtained by filtering $C(i)$ and $B(i)$ independently, and adding the filtered waveforms with a scaling of $-dT$ for the filtered $B(i)$ output.

Since waveforms $C(t)$ and $D(t)$ are calculable from $W(t)$, $B(i)$ can be precalculated and $C(t)$ and $B(t)$ can be independently filtered to obtain samples of

-19-

waveforms $Z_1(t)$ and $Z_2(t)$ respectively. $Z'(t)$ will be equal to $Z_1(t) - dT * Z_2(t)$, and the value of dT can be computed to increase the echo cancellation. That is, dT can be computed to reduce the mean square value of the residual waveform $E(t)$, in which the sum of the squares of:

$E(i) = Z(i) - Z_1(i) + dT * Z_2(i)$ is reduced. This value of dT is given by the equation:

$$dT = -\frac{1}{N} \sum (Z(i) * (Z(i) - Z_1(i))),$$

where $Z_2(i)$ is correlated with $Z(i) - Z_1(i)$, for $i = 1$ to N samples.

The amount ALPHA of the $Z_1(t)$ waveform that is combined with BETA of the $Z_2(t)$ waveform can be jointly optimized to obtain:

$$\text{ALPHA} = \frac{de - bf}{ad - bc} ; \quad \text{BETA} = \frac{af - ce}{ad - bc}$$

where:

$$a = \sum_i [Z_1^2(i)]$$

$$d = \sum_i [Z_2^2(i)]$$

$$b = c = \sum_i [Z_1(i) * Z_2(i)]$$

$$e = \sum_i [Z_1(i) * Z(i)]$$

$$f = \sum_i [Z_2(i) * Z(i)]$$

Since changing both the amount of Z_1 and Z_2 is approximately equivalent to scaling the FIR coefficients, the values ALPHA and BETA can be implemented by an overall scaling of the FIR coefficients by multiplication with ALPHA, and by setting: $dT = -\text{BETA}/\text{ALPHA}$.

Extra degrees of freedom to reduce the residual $E(t)$ can be obtained by using separate adaptively optimized FIR filters for filtering the $B(t)$ waveform and the $C(t)$ waveform to obtain two waveforms

-20-

$Z_1(t)$ and $Z_2(t)$ in dependence on separate sets of FIR coefficients. However, there is no physical basis for expecting the coefficient sets to differ other than by a scaling $-dT$. In problems of the type considered
5 here, it is generally found that models based on physical reality are the most economic in terms of complexity. Complexity may sometimes be sacrificed, however, in favor of quick implementations through the use of existing adaptive FIR algorithms that are
10 already programmed. An echo cancellation system with separate FIR filters 34_{A1} and 34_{B1} is illustrated in Figure 3.

The transducer model 12_1 in Figure 3 produces two output waveforms, a non-delay modulated waveform
15 $C(t)$ (shown as $V_{A1}(t)$) and a distortion waveform $B(t)$ (shown as $V_{B1}(t)$). $C(t)$ and $B(t)$ pass through separate acoustic path models 34_{A1} and 34_{B1} , each comprising an FIR filter, the coefficients of which are all adapted by echo filter modifier 16_1 in order to reduce the
20 residual waveform $E_1(t)$ out of subtractor 22_1 . Those having skill in the art will be able to contrive other variations in this arrangement. For example, the two FIR filters can be separate for part of the path and then summed into a common FIR filter for the remainder
25 of the path. The foregoing embodiment of Figure 3, using coefficients ALPHA and BETA, is in fact an extreme version of this principle in which the separate parts of the filters are reduced to single taps of weight ALPHA and BETA, respectively. However, all the
30 values of ALPHA, BETA, and the coefficients of the common part of the FIR filter may be jointly optimized instead of optimizing ALPHA and BETA only while accepting the previously optimized FIR coefficients. This is a matter of design implementation and a
35 complexity trade off that can be made by persons having skill in the art while adhering to the basic principles of the non-linear echo cancellation.

-21-

An additional filter (not shown) may be included before loudspeaker models 12 and 12, of Figures 1 and 3. The purpose of this filter is to model the anti-aliasing filter contained in D-to-A converter 26.

5 This filter has some effect on the waveform driving the loudspeaker, and so in the interests of driving the loudspeaker models 12 and 12, with the same waveform that drives the real loudspeaker, a model of this filter can be included. A model of the frequency

10 response of the speaker amplifier 27 can also be included here. This filter should not need to be adapted and can be set in the factory to model the anti-aliasing filter and audio amplifier responses. An anti-aliasing filter corresponding to the signal path

15 from the microphone input 20 through the A-to-D convertor 28 can be adequately modelled by the FIR filter or filters used to model the acoustic path.

The technique described above takes care of one of the non-linear distortion mechanisms of the

20 loudspeaker. The non-linearity, caused by the diaphragm spring, can be updated in a similar way. The diaphragm spring non-linearity is modelled as a polynomial such as:

$$G_0 * D(t) + G_1 * D(t)^2 + G_2 * D(t)^3 \dots$$

25 Each of the second, third and higher order distortion waveforms from amplifier 58 of Figure 2 enters amplifier 48 and is subjected to a certain frequency response before emerging at output B. The distortion waveforms are also further distorted, but these higher

30 order effects can be neglected without significantly affecting the model.

The frequency response to which the cubic distortion term, for example, is subjected can be represented by an FIR filter having tap weights t_0 , t_1 ,

35 t_2 , t_3 , ..., etc. That is, the output distortion waveform will be:

$$t_0 * G_2 * D^3(i) + t_1 * G_2 * D^3(i-1) + t_2 * G_2 * D^3(i-2) \dots$$

-22-

This waveform then reaches the microphone through the acoustic path 18 and is also passed through the FIR filter of acoustic path model 34 to reduce the echo.

Since the FIR filter of acoustic path model 34 is linear, each of the terms in the series listed above may be filtered independently and added. Since each of the terms is in fact just a delayed version of the foregoing term, there is really only one waveform to pass through the FIR filter. This waveform is:

10 ..., $D^3(i-1)$, $D^3(i)$, $D^3(i+1)$, ...

The resulting filtered waveform is denoted by:

 ..., $Z_3(i-1)$, $Z_3(i)$, $Z_3(i+1)$, ...

This waveform is then subjected to the FIR filter defined by weights, T_0 , T_1 , T_2 , ..., and subtracted from
15 the residual echo to reduce it. The tap coefficients, T_0 , T_1 , T_2 , ..., which most effectively reduce the residual echo can be found by the same technique used to find the tap weights for the FIR filter of acoustic path model 34 to reduce the residual by subtracting
20 weighted delayed copies of the $W(t)$ waveform. The optimum taps, T_0 , T_1 , T_2 , ..., differ from the known frequency response of the model, t_0 , t_1 , t_2 , ..., only by the factor G_2 . Therefore, an updated value of the distortion coefficient G_2 can be found. In the same
25 way, any polynomial coefficient of the diaphragm spring non-linearity can be assigned an updated value which reduces non-linear echo residuals.

As discussed above with regard to Figure 1, the echo filter according to the present invention can
30 be implemented as a digital signal processor 21 including a loudspeaker model 12, an acoustic path model 34, a subtractor 22, and an echo filter modifier 16. In this embodiment, the loudspeaker model and the acoustic path model each comprise a single processing
35 block, and these processing blocks are arranged in series. The processing functions, however, can have other arrangements.

-23-

For example, the acoustic path model can be broken down into two processing blocks such as acoustic path model A 34_{A1} and acoustic path model B 34_{B1} within DSP 21₁, as shown in Figure 3. In this embodiment, the

5 loudspeaker model 12₁ has a first output $V_{A1}(t)$ which models the non-linear aspects of the loudspeaker 14 output, and a second output $V_{B1}(t)$ which models the linear aspects of the loudspeaker output. Accordingly, the acoustic path model A 34_{A1} models the acoustic path

10 for the non-linear aspects of the sound pressure wave, and the acoustic path model B 34_{B1} models the acoustic path for the linear aspects of the sound pressure wave. Each of the acoustic path models can be independently modified in order to provide a more accurate estimates

15 of each portion of the echo signal.

Accordingly, if the acoustic path 18 behaves differently with regard to the linear and non-linear aspects of the sound pressure wave, the separate acoustic path model processing blocks can accommodate

20 this difference. This arrangement allows flexibility with regard to the parameters used to model the acoustic path 18 as well as flexibility with regard to the modification of these parameters by echo filter modifier 16₁. Figure 3 shows the subtraction of the

25 non-linear echo estimate Z_{A1} and the linear echo estimate Z_{B1} by subtractor 22₁ in a single processing operation. These echo estimates may alternately be performed separately by multiple subtractors, as will be understood by those having skill in the art.

30 In yet another embodiment, the loudspeaker model is broken down into separate processing blocks 12_{A2} and 12_{B2} within DSP 21₁, as shown in Figure 4. Here loudspeaker model A 12_{A2} models the non-linear aspects of the sound pressure wave generated by the

35 loudspeaker, and loudspeaker model B 12_{B2} models the linear aspects of the sound pressure wave generated by the loudspeaker. This arrangement provides flexibility

-24-

with regard to the parameters used to characterize the sound pressure wave as well as flexibility with regard to modification of these parameters by echo filter modifier 16,.

5 In this embodiment, the acoustic path model is broken down into multiple processing blocks for modeling the acoustic path for linear and non-linear aspects of the sound pressure wave, as discussed above with regard to Figure 3. The embodiments of Figures 3
10 and 4 show that the loudspeaker model and the acoustic path model can be broken down into multiple processing blocks in order to model the linear and non-linear aspects of the echo separately. The invention, however, contemplates that these models can be broken
15 down into separate processing blocks in order to model other aspects of the echo separately. For example, separate processing blocks can be provided to model high and low frequency portions of the echo, high and low amplitude portions of the echo, etc.

20 It has thus been described how a non-linear loudspeaker model 12 of loudspeaker 14 can be obtained and included in the echo cancellation path of a full-duplex loudspeaker telephone in order to obtain better echo cancellation by also canceling non-linear
25 distortion products. The technique can be implemented using a fixed, non-adaptive model of the transducer. The parameters of these non-linearities can be determined by measurement, or the technique can include adaptive determination of the non-linear parameters of
30 the transducer using the techniques described above or variations thereof that can be derived by persons having skill in the art of adaptive signal processing.

 Accordingly, many modifications and other embodiments of the invention will come to one skilled
35 in the art having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood



-25-

that the invention is not to be limited to the specific embodiments disclosed, and that modifications are intended to be included within the scope of the appended claims.



-26-

That Which Is Claimed Is:

1. An echo canceling loudspeaker telephone comprising:

a loudspeaker for producing a sound pressure wave in response to an input signal which is applied to an audio input thereof, said sound pressure wave including a desired linear component which is a linear function of said input signal, and an undesired non-linear component which is a non-linear function of said input signal, said sound pressure wave being transmitted along an acoustic path;

a microphone positioned in said acoustic path for converting said sound pressure wave into an output signal;

an echo filter responsive to said input signal and which generates an estimated echo signal, said echo filter comprising a loudspeaker model which generates an estimate of said sound pressure wave including an estimate of said linear component and an estimate of said non-linear component, and an acoustic path model which generates an estimate of said acoustic path from said loudspeaker to said microphone; and

a subtractor for subtracting said estimated echo signal from said output signal thereby reducing an echo portion of said output signal wherein said loudspeaker model and said acoustic path model are connected in series.

2. An echo-canceling loudspeaker telephone according to Claim 1 further comprising an echo filter modifier responsive to said output signal, said input signal and said estimated echo signal for modifying said echo filter to further reduce said echo portion of said output signal.



-27-

3. An echo-canceling loudspeaker telephone according to Claim 2 wherein said echo filter modifier comprises a loudspeaker model modifier for modifying the estimate of said sound pressure wave including said
5 estimate of said linear component and said estimate of said non-linear component.

4. An echo-canceling loudspeaker telephone according to Claim 2 wherein said echo filter modifier comprises an acoustic path model modifier for improving
10 the estimate of said acoustic path from said loudspeaker to said microphone.

5. An echo-canceling loudspeaker telephone according to Claim 1 wherein said acoustic path model comprises a finite-impulse-response filter.

15 6. An echo-canceling loudspeaker telephone according to Claim 1 wherein said echo filter comprises a digital signal processor.

7. An echo-canceling loudspeaker telephone according to Claim 1 wherein said loudspeaker model
20 comprises means for performing a transformation of said input signal, said transformation being an estimate of a transfer function of said loudspeaker, said transfer function including a non-linear component.

8. An echo-canceling loudspeaker telephone
25 according to Claim 7 wherein said non-linear component represents one of a delay modulation of said loudspeaker and a diaphragm stress-strain curve of said loudspeaker.



This Sheet Has Been Replaced





-28-

9. An echo canceling loudspeaker telephone comprising:

a loudspeaker for producing a sound pressure wave in response to an input signal which is applied to
5 an audio input thereof, said sound pressure wave including a desired linear component which is a linear function of said input signal, and an undesired non-linear component which is a non-linear function of said input signal, said sound pressure wave being
10 transmitted along an acoustic path;

a microphone positioned in said acoustic path for converting said sound pressure wave into an output signal;

an echo filter responsive to said input
15 signal and which generates an estimated echo signal, said echo filter comprising a loudspeaker model which generates an estimate of said sound pressure wave including an estimate of said linear component and an estimate of said non-linear component, and an acoustic
20 path model comprising a first processing block for generating an estimate of said acoustic path for said non-linear component of said sound pressure wave and a second processing block for generating an estimate of said acoustic path for said linear component of said
25 sound pressure wave; and

a subtractor for subtracting said estimated echo signal from said output signal thereby reducing an echo portion of said sound signal.

10. An echo-canceling loudspeaker telephone
30 according to Claim 9 further comprising an echo filter modifier responsive to said output signal, said input signal and said estimated echo signal for modifying said echo filter to further reduce said echo portion of said output signal.

This Sheet Has Been Replaced





WO 96/26592

THIS SHEET HAS BEEN
REPLACED

PCT/US96/02073

-29-

11. An echo-canceling loudspeaker telephone according to Claim 10 wherein said echo filter modifier comprises an loudspeaker model modifier for modifying the estimate of said sound pressure wave including said
5 estimate of said linear component and said estimate of said non-linear component.

12. An echo-canceling loudspeaker telephone according to Claim 10 wherein said echo filter modifier comprises an acoustic path model modifier for improving
10 the estimate of said acoustic path from said loudspeaker to said microphone.

13. An echo-canceling loudspeaker telephone according to Claim 9 wherein said acoustic path model comprises a finite-impulse-response filter.

14. An echo-canceling loudspeaker telephone according to Claim 9 wherein said echo filter comprises a digital signal processor.
15

15. An echo-canceling loudspeaker telephone according to Claim 9 wherein said loudspeaker model
20 comprises means for performing a transformation of said input signal, said transformation being an estimate of a transfer function of said loudspeaker, said transfer function including a non-linear component.

16. An echo-canceling loudspeaker telephone
25 according to Claim 15 wherein said non-linear component represents one of a delay modulation of said loudspeaker and a diaphragm stress-strain curve of said loudspeaker.

This Sheet Has Been Replaced





-30-

17. An echo-canceling loudspeaker telephone comprising:

output transducer means for producing a sound pressure wave in response to an input signal which is applied to an audio input thereof, said sound pressure wave including a desired linear component which is a linear function of said input signal, and an undesired non-linear component which is a non-linear function of said input signal, said sound pressure wave being transmitted along an acoustic path;

input transducer means positioned in said acoustic path for converting said sound pressure wave into a output signal;

echo filter means responsive to said input signal for generating an estimated echo signal, said echo filter means comprising an output transducer model for generating an estimate of said sound pressure wave including an estimate of said linear component and an estimate of said non-linear component, and an acoustic path model for generating an estimate of said acoustic path from said output transducer means to said input transducer means; and

combination means for combining said estimated echo signal and said output signal thereby reducing an echo portion of said output signal.

18. An echo-canceling loudspeaker telephone according to Claim 17 further comprising an echo filter modifier responsive to said output signal, said input signal and said estimated echo signal for modifying said echo filter to further reduce said echo portion of said output signal.

19. An echo-canceling loudspeaker telephone according to Claim 18 wherein said echo filter modifier comprises an output transducer model modifier for modifying the estimate of said sound pressure wave





WO 96/26592

THIS SHEET HAS BEEN
REPLACED

PCT/US96/02073

-31-

including said estimate of said linear component and
said estimate of said non-linear component.

20. An echo-canceling loudspeaker telephone
according to Claim 18 wherein said echo filter modifier
5 comprises an acoustic path model modifier for improving
the estimate of said acoustic path from said output
transducer to said input transducer.

21. An echo-canceling loudspeaker telephone
according to Claim 17 wherein said acoustic path model
10 comprises a finite-impulse-response filter.

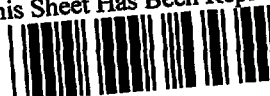
22. An echo-canceling loudspeaker telephone
according to Claim 17 wherein said echo filter means
comprises a digital signal processor.

23. An echo-canceling loudspeaker telephone
15 according to Claim 17 wherein said output transducer
model comprises means for performing a transformation
of said input signal, said transformation being an
estimate of a transfer function of said output
transducer means, said transfer function including a
20 non-linear component.

24. An echo-canceling loudspeaker telephone
according to Claim 23 wherein said non-linear component
represents one of a delay modulation of said output
transducer means and a diaphragm stress-strain curve of
25 said output transducer means.

25. An echo-canceling loudspeaker telephone
according to Claim 17 wherein said output transducer
means comprises a loudspeaker.

This Sheet Has Been Replaced





-32-

26. An echo-canceling loudspeaker telephone according to Claim 17 wherein said input transducer means comprises a microphone.

27. An echo-canceling loudspeaker telephone
5 according to Claim 17 wherein said output transducer model and said acoustic path model are connected in series between said audio input and said combination means.

28. An echo-canceling loudspeaker telephone
10 according to Claim 17 wherein said acoustic path model comprises a first processing block for generating an estimate of said acoustic path for said non-linear component of said sound pressure wave and a second
15 processing block for generating an estimate of said acoustic path for said linear component of said sound pressure wave.

29. An echo-canceling loudspeaker telephone according to Claim 17 wherein said combination means comprises a subtractor for subtracting said estimated
20 echo signal from said output signal.

30. A method for reducing echoes in a telephone system including a loudspeaker, said method comprising the steps of:

applying an input signal to said loudspeaker
25 to produce a sound pressure wave which is transmitted along an acoustic path, said sound pressure wave including a desired linear component which is a linear function of said input signal and an undesired non-linear component which is a non-linear function of said
30 input signal;

converting said sound pressure wave in said acoustic path to produce an output signal including an echo portion;





WO 96/26592

THIS SHEET HAS BEEN
REPLACED

PCT/US96/02073

-33-

generating an estimated echo signal in response to said input signal, wherein said estimated echo signal includes an estimate of distortions due to said loudspeaker, which includes linear and non-linear components, and an estimate of distortions due to said acoustic path; and

combining said output signal and said estimated echo signal to reduce said echo portion of said output signal.

10 31. A method according to Claim 30 wherein said combining step is followed by the steps of:
 determining a residual echo portion of said output signal remaining after said combining step; and
 modifying said estimates of distortions due
15 to said loudspeaker and said acoustic path in response to said residual echo portion to further reduce said echo portion of said output signal.

20 32. A method according to Claim 31 further comprising the step of comparing said input signal with said output signal to determine when said output signal substantially comprises only said echo portion, and wherein said modifying step is performed when said output signal substantially comprises only said echo portion.

25 33. A method according to Claim 32 wherein said input signal comprises an input test signal.

This Sheet Has Been Replaced