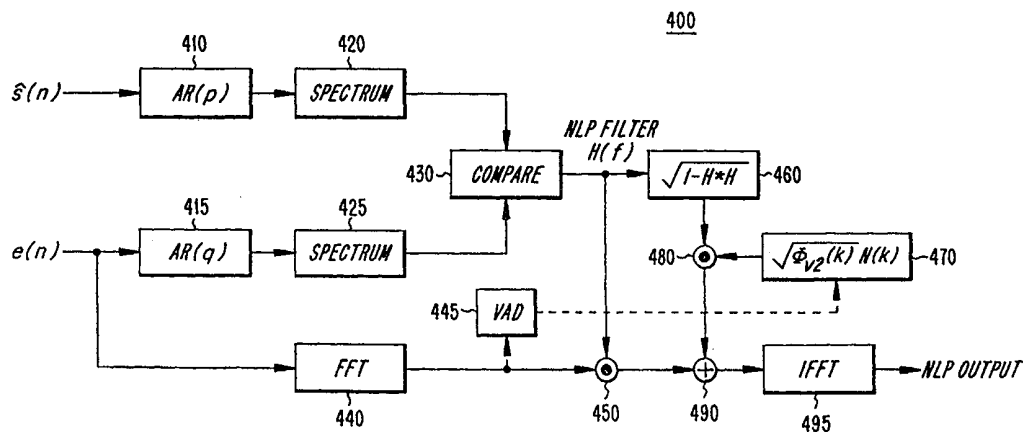




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : H04B 3/23	A1	(11) International Publication Number: WO 00/25441 (43) International Publication Date: 4 May 2000 (04.05.00)
(21) International Application Number: PCT/SE99/01909 (22) International Filing Date: 22 October 1999 (22.10.99) (30) Priority Data: 09/177,538 23 October 1998 (23.10.98) US (71) Applicant: TELEFONAKTIEBOLAGET LM ERICSSON (publ) [SE/SE]; S-126 25 Stockholm (SE). (72) Inventors: SÖRQVIST, Patrik; Stambanevägen 18, S-141 43 Huddinge (SE). ERIKSSON, Anders; Smålandsvägen 2, S-757 58 Uppsala (SE). (74) Agent: ERICSSON RADIO SYSTEMS AB; Ericsson Research / Patent Support, S-164 80 Stockholm (SE).	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	

(54) Title: METHODS AND APPARATUS FOR PROVIDING ECHO SUPPRESSION USING FREQUENCY DOMAIN NONLINEAR PROCESSING

**(57) Abstract**

Frequency domain nonlinear processing techniques are applied in the context of echo suppression. According to exemplary embodiments, the filtering characteristic of a nonlinear processor is dynamically adjusted by comparing the power spectrum of a residual echo component of a communication signal with that of the overall communication signal itself. For example, the filtering characteristic can be adjusted so that the nonlinear processor blocks only those signal frequencies in which the residual echo is, in some sense, dominant. The disclosed techniques make full duplex communication possible during the entirety of a conversation, even while residual echo is being actively suppressed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

-1-

**METHODS AND APPARATUS FOR
PROVIDING ECHO SUPPRESSION USING
FREQUENCY DOMAIN NONLINEAR PROCESSING**

5 **Field of the Invention**

The present invention relates to communications systems, and more particularly, to echo suppression in a bi-directional communications link.

10 **Background of the Invention**

In many communications systems, for example landline and wireless telephone systems, voice signals are often transmitted between two system users via a bi-directional communications link. In such systems, speech of a near-end user is typically detected by a near-end microphone at one end of the communications link and then transmitted over the link to a far-end loudspeaker for reproduction and presentation to a far-end user. Conversely, speech of the far-end user is detected by a far-end microphone and then transmitted via the communications link to a near-end loudspeaker for reproduction and presentation to the near-end user.

At either end of the communications link, loudspeaker output detected by a proximate microphone may be inadvertently transmitted back over the communications link, resulting in what may be unacceptably disruptive feedback, or echo, from a user perspective. This problem is particularly prevalent in the context of hands-free telephony, where the user's body does not effectively isolate microphone from loudspeaker. Furthermore, impedance mismatches at hybrid junctions in a landline network can result in similarly disruptive echo.

Conventionally, echo suppression has been accomplished using echo canceling circuits which employ adaptive filters to estimate and remove the unwanted echo component of a communications signal. For example, U.S. Patent

-2-

No. 5,475,731, entitled "Echo-Canceling System and Method Using Echo Estimate to Modify Error Signal" and issued December 12, 1995, describes the use of Least Mean Square (LMS) and Normalized Least Mean Square (NLMS) algorithms in updating Finite Impulse Response (FIR) adaptive filters.

5 Though such known echo cancellation techniques do remove some echo, audible residual echo will often remain even after cancellation. Such residual echo can, despite its relatively low level, be quite disturbing and should therefore be removed. Today, residual echo suppression is typically accomplished using some form of Non-Linear Processor (NLP) following an echo canceler. See, for
10 example, the above cited U.S. Patent No. 5,475,731, which describes the use of a conventional center-clipping NLP.

 Most known NLPs completely remove those parts of a communications signal containing residual echo. As a result, when both a near-end and a far-end speaker are active (i.e., during double-talk), such known NLPs either pass the
15 residual echo through or remove both the near end speech and the residual echo. Additionally, when known NLPs are used in contexts with complex background noise (e.g., when a car radio is playing while a driver is using a hands-free mobile telephone), the background sounds are also removed along with the residual echo. Though pre-computed comfort noise can be added in an attempt to fill the
20 resulting holes, the processed signal is often choppy and bothersome from a receiving user's perspective.

 Consequently, there is a need for improved methods and apparatus for removing residual echo in communications signals.

25 **Summary of the Invention**

 The present invention fulfills the above-described and other needs by providing improved frequency domain nonlinear processing techniques.

-3-

Generally, the filtering characteristic of a nonlinear processor according to the invention is dynamically adjusted by comparing the power spectrum of a residual echo component of a communication signal with that of the overall communication signal itself. More specifically, the filtering characteristic of the nonlinear processor is adjusted so that the nonlinear processor blocks only those signal frequencies in which the residual echo is dominant. Advantageously, a nonlinear processor according to the invention makes full duplex communication possible during the entirety of a conversation, even while actively reducing residual echo.

An exemplary echo suppressor according to the invention includes a processor configured to compute a power spectrum of a communications signal and to estimate a power spectrum of a residual echo component of the communications signal. The exemplary echo suppressor also includes an adaptive filter configured to process the communications signal and to thereby suppress the residual echo component. A filtering characteristic of the adaptive filter is adjusted based on the power spectrum of the communications signal and on the estimated power spectrum of the residual echo.

An exemplary method for suppressing residual echo in a communications signal according to the invention includes the steps of computing a power spectrum of the communications signal and estimating a power spectrum of the residual echo. According to the exemplary method, a filtering characteristic is adjusted based upon the computed power spectrum of the communications signal and upon the estimated power spectrum of the residual echo, and the communications signal is filtered, using the adjusted filtering characteristic, to suppress the residual echo.

According to the embodiments, the filtering characteristic can be adjusted, for example, using spectral subtraction techniques. Alternately, individual coefficients of the filtering characteristic can be set based on direct comparisons of

-4-

the power in the residual echo with the power in the overall communications signal at the frequencies corresponding to the coefficients.

For stand alone configurations (i.e., where no leading echo canceler is present), estimation of the power spectrum of the residual echo can be accomplished by estimating an attenuation factor of the echo return path and scaling power spectrum samples of the echo source signal by the estimated attenuation factor to thereby provide estimated power spectrum samples of the residual echo. When a leading adaptive echo canceler is present, estimation of the power spectrum of the residual echo can be carried out by computing a power spectrum of an echo component estimate provided by the echo canceler, and scaling the power spectrum of the echo component estimate (based on operation of the adaptive echo canceler) to thereby provide the estimate of the power spectrum of the residual echo.

Scaling of the power spectrum of the echo component estimate can include the steps of determining an echo return loss enhancement of the adaptive echo canceler, and multiplying the power spectrum of the echo component estimate by the echo return loss enhancement to provide the estimate of the power spectrum of the residual echo. Alternatively, scaling of the power spectrum of the echo component estimate can include the steps of determining the echo return loss enhancement of the adaptive echo canceler, computing an average value for the power spectrum of the echo component estimate, multiplying the average value by the echo return loss enhancement to provide a saturation value, multiplying the power spectrum of the echo component estimate by the echo return loss enhancement to provide a scaled spectrum, and clipping the scaled spectrum at the saturation value to thereby provide the estimate of the power spectrum of the residual echo.

-5-

Advantageously, the present invention also provides methods and apparatus for adding suitable comfort noise to the filtered (i.e., echo suppressed) communications signal. Generally, a power spectrum of a noise component of the communications signal is estimated, and the comfort noise is created based upon the estimated power spectrum of the noise component and on the prevailing echo-suppression filtering characteristic. According to exemplary embodiments, the comfort noise adds energy at suppressed frequencies (i.e., at frequencies where the filtering characteristic is removing energy and thereby suppressing echo) so that the total energy at those frequencies is equal that of the prevailing noise.

The above-described and other features and advantages of the invention are explained in detail hereinafter with reference to the illustrative examples shown in the accompanying drawings. Those skilled in the art will appreciate that the described embodiments are provided for purposes of illustration and understanding and that numerous equivalent embodiments are contemplated herein.

Brief Description of the Drawings

Figure 1 is a block diagram of an exemplary acoustic echo suppression system in which the teachings of the present invention can be implemented.

Figure 2 is a block diagram of an exemplary network echo suppression system in which the teachings of the present invention can be implemented.

Figure 3 provides a frequency domain comparison of a typical echo signal and a corresponding residual echo signal and thereby further provides motivation for certain aspects of the present invention.

Figure 4 is a flowchart depicting steps in an exemplary method of residual echo suppression according to the invention.

Detailed Description of the Invention

The frequency domain nonlinear processing techniques of the invention are equally applicable in the contexts of acoustic and network echo suppression.

Accordingly, Figures 1 and 2 depict, respectively, exemplary acoustic and network echo suppression systems 100, 200 in which the teachings of the present invention can be implemented. As shown, the exemplary acoustic echo suppression system 100 includes a microphone 110, a loudspeaker 120, a linear echo canceler 130, a nonlinear processor 140, a comfort noise generator 150, and first and second summing devices 160, 170. The exemplary network echo suppression system 200 includes the echo canceler 130, the nonlinear processor 140, the comfort noise generator 150, and the summing devices 160, 170 of Figure 1, as well as a hybrid junction 210 and a third summing device 220.

Those skilled in the art will appreciate that the below described functionality of the various elements of Figures 1 and 2 can be implemented, for example, using known digital signal processing components, one or more application specific integrated circuits (ASICs), or a general purpose digital computer.

In operation, the loudspeaker 120 of Figure 1 receives a far-end audio signal $x(t)$ and outputs a corresponding acoustic signal to a near-end system user (not shown). At the same time, the microphone 110 receives an acoustic signal including a near-end speech component $v_1(t)$, a near-end noise component $v_2(t)$ and an echo component $s(t)$ resulting from reflection of the loudspeaker output. The microphone 110 converts the received acoustic signal to a corresponding audio signal $y(t)$ which is coupled to an additive input of the first summing device 160. Additionally, the linear echo canceler 130 processes the far-end audio signal $x(t)$ using known adaptive filtering techniques to provide an estimate $\hat{s}(t)$ of the echo component $s(t)$ of the microphone signal $y(t)$. The resulting echo estimate $\hat{s}(t)$ is

-7-

coupled to a subtractive input of the first summing device 160 and thus subtracted from the microphone signal $y(t)$ to provide an echo-canceled, or error signal $e(t)$.

The error signal $e(t)$ is fed back to the linear echo canceler 130 for use in providing the echo estimate $\hat{s}(t)$ and is also coupled to an input of the nonlinear processor 140. As is described in detail below, the nonlinear processor 140 filters the error signal $e(t)$, based on the echo estimate $\hat{s}(t)$, to provide an echo-suppressed output signal $e_{\text{NLP}}(t)$ which is coupled to an additive input of the second summing device 170. As is also described in detail below, the comfort noise generator 150 provides a comfort noise signal $c(t)$ based on the error signal $e(t)$ and on the prevailing filtering characteristic of the nonlinear processor 140. The comfort noise signal $c(t)$ is coupled to a second additive input of the second summing device 170 and thereby added to the echo suppressed output $e_{\text{NLP}}(t)$ of the nonlinear processor 140 to provide an echo suppressed and comfort noise enhanced output signal $e_{\text{NLP}+\text{CN}}(t)$.

In the network system 200 of Figure 2, the hybrid junction 210 receives a far-end audio signal $x(t)$ and provides a corresponding network signal to a near-end system user (not shown). Additionally, impedance mismatches at the hybrid 210 cause an echo $s(t)$ of the far-end signal $x(t)$ to be additively coupled with received near-end voice and noise signals $v_1(t)$, $v_2(t)$ (shown conceptually via the third summing device 220) and passed back to the additive input of the first summing device 160. As in the system 100 of Figure 1, the linear echo canceler 130 processes the far-end audio signal $x(t)$ using known adaptive filtering techniques to provide an estimate $\hat{s}(t)$ of the echo component $s(t)$ of the near-end signal $y(t)$. The resulting echo estimate $\hat{s}(t)$ is coupled to the subtractive input of the first summing device 160 and thus subtracted from the near-end signal $y(t)$ to provide the echo-canceled, or error signal $e(t)$. The nonlinear processor 140 and the comfort noise generator 150 operate on the error signal $e(t)$ as

-8-

described with respect to Figure 1 (and as described in detail below) to provide the echo suppressed and comfort noise enhanced output signal $e_{\text{NLP+CN}}(t)$.

In both the acoustic and network systems 100, 200, the error signal $e(n)$ is equal to the sum of the near end speech $v_1(n)$, the background noise $v_2(n)$ and the echo $s(n)$, less the estimated echo $\hat{s}(n)$ (where the variable n is used to indicate time domain samples):

$$e(n) = v_1(n) + v_2(n) + s(n) - \hat{s}(n)$$

Additionally, the residual echo $\tilde{s}(n)$ is defined as the true echo $s(n)$ less the estimated echo $\hat{s}(n)$:

$$\tilde{s}(n) = s(n) - \hat{s}(n)$$

According to the invention, the filtering characteristic of the nonlinear processor 140 is dynamically computed based on the error signal $e(n)$ and the residual echo $\tilde{s}(n)$. Specifically, the spectral content of the error signal $e(n)$ is compared to an estimate of the spectral content of the residual echo $\tilde{s}(n)$, and the filtering characteristic of the nonlinear processor 140 is adjusted to pass only those frequencies for which the residual echo $\tilde{s}(n)$ is, in a sense, the dominant component of the error signal $e(n)$.

To compare the spectral content of the error signal $e(n)$ with that of the residual echo $\tilde{s}(n)$, the error signal $e(n)$ and the echo estimate $\hat{s}(n)$ are first converted to the frequency domain. For example, M -point Fast Fourier Transforms (FFTs) can be used to compute the corresponding frequency domain signals $E(k)$, $\hat{S}(k)$ directly from the time domain samples (where k represents the discrete frequencies ranging from 0 to 0.5 in steps of $2/M$). Alternatively, to

-9-

reduce the variance in the frequency domain results, auto-regressive (AR) models of order p can first be computed for the error signal e and the echo estimate $\hat{s}$ based on, for example, a set of N time domain data samples for each signal. M -point FFTs of the zero padded AR(p) sequences then provide the discrete

5 frequency domain variables $E(k)$ and $\hat{S}(k)$ as desired. In practice, suitable values for the variables p , N and M are 10, 160 and 256, respectively. Of course, the order p used for modeling the error signal e can be different than that used for modeling the echo estimate $\hat{s}$.

Once the frequency domain sequences $E(k)$ and $\hat{S}(k)$ have been computed,

10 a corresponding Power Spectral Density (PSD) can be computed for each sequence by multiplying each sequence, in a sample-wise fashion, by its complex conjugate sequence as follows:

$$\Phi_e(k) = E(k)E^*(k)$$

$$\Phi_{\hat{s}}(k) = \hat{S}(k)\hat{S}^*(k)$$

Next, the power spectral density of the residual echo, $\Phi_{\hat{r}}$, can be estimated from the power spectral density of the estimated echo, $\Phi_{\hat{s}}$. A first approximation can be obtained using the Echo Return Loss Enhancement (ERLE) provided by

20 the echo canceler. The ERLE of an echo canceler provides an indication of the level of echo suppression achieved by the echo canceler, and methods for computing the ERLE of an echo canceler are well known. Using the assumption that the spectrum for the residual echo is similar to the spectrum for the estimated echo, then the power spectral density of the residual echo can be estimated at

25 ERLE dB below the power spectral density of the estimated echo as follows:

-10-

$$\hat{\Phi}_s(k) \approx \text{ERLE} \Phi_s(k)$$

However, examination of empirical data reveals that the residual echo is actually often much higher than ERLE dB below the echo estimate, particularly for frequencies having low energy compared to the maximum energy in the spectra (for speech, usually the higher frequencies). See, for example, Figure 3, where an actual residual echo curve 320 is approximately ERLE dB (roughly 15 dB in the Figure) below an actual estimated echo curve 310 for a lower frequency range (0 to 0.25 in the Figure), and greater than ERLE dB below the estimated curve 310 for a higher frequency range (0.25 to 0.5 in the Figure).

To provide a better approximation, the ERLE measurement can be made frequency dependent. However, doing so requires a stationary distribution (when in fact the ERLE is often varying in time as well as frequency) and, in instances in which significant near-end energy exists in the higher frequency bands, an undesirable degree of near-end signal clipping could be introduced. A more practical solution can be achieved by saturating the estimate of the power spectral density of the residual echo at ERLE dB below the mean level of the power spectral density of the estimated echo, as follows:

$$\Phi_s(k) \approx \text{Max}(\text{ERLE} \Phi_s(k), \text{ERLE} \overline{\Phi_s})$$

The estimate of the power spectral density of the residual echo can then be used to derive an appropriate filtering characteristic, or transfer function $H(k)$ for the nonlinear processor. For example, the teachings of spectral subtraction (often

-11-

used in noise reduction contexts) can be applied to provide an intuitively correct echo clipping filter $H(k)$ as follows:

$$H(k) = \sqrt{\frac{\Phi_e(k) - \delta_1 \Phi_f(k)}{\Phi_e(k)}}$$

where δ_1 is the subtraction factor. Empirical studies have shown that values for δ_1 in the range of 1-4 are appropriate in the echo suppression context.

In practice, determining the filter $H(k)$ directly from the above equation can be overly complex (i.e., computing a square root using standard digital signal processing components is often undesirable), and furthermore, the level of echo attenuation may not be satisfactory due to estimation errors. A simpler and harder-clipping solution can be achieved by removing all power from frequency bins in which the residual echo is in some sense dominant, while passing all power from frequency bins in which the residual echo is not dominant (and is therefore perceptually masked already). For example, the clipping criteria for each frequency bin k can be formed as:

$$H(k) = 0 \text{ if } \Phi_e(k) < \delta_1 \hat{\Phi}_f(k), \text{ else } H(k) = 1$$

Since the above equation will provide a rapidly changing power spectral density (due to the abrupt changes between $H(k)=1$ and $H(k)=0$ over time), the filtering characteristic can be smoothed or averaged in practice to provide soft transitions. Doing so provides high performance with a minimum of distortion

-12-

(such as of musical tones). For example, a smoothened NLP filter $H_{av}(k)$ can be updated as:

$$H_{av}(k) = H(k)(1 - \alpha (1 - H_{av}(k)))$$

5

where emperical studies have shown that a suitable value for α is, e.g., 0.6.

The frequency domain version of the error signal $E(k)$ (derived by taking an M -point FFT of the time domain error signal $e(n)$) is then filtered using the smoothened NLP transfer function $H_{av}(k)$, and the result is finally transformed back to the time domain using an Inverse Fast Fourier Transform (IFFT). In other words, the echo suppressed NLP output is given by:

10

$$e_{NLP}(n) = IFFT(H_{av} \cdot E)$$

where $\cdot$ denotes elementwise multiplication of the vectors H_{av} and E . Alternately, the frequency domain vectors H_{av} and E could first be transformed to the time domain and then convolved to provide the NLP output.

15

While the NLP output given by the previous equation will effectively suppress residual echo as desired, note that the removed frequencies will also result in a varying background level from a receiving user's perspective. According to another aspect of the invention, however, this effect can be lessened by strategically replacing the removed energy. Specifically, the output signal can be smoothened by adding energy at the suppressed frequencies so that the total energy at those frequencies is equal to the prevailing background noise level.

20

For example, a Voice Activity Detector (VAD) can be used to determine whether speech (either source or echo) is present in the input to the NLP. During

-13-

speech pauses, the VAD indicates that only noise is present, and an estimate of the power spectral density of the prevailing background noise can be updated as:

$$\hat{\Phi}_{v2}(k) = \beta \hat{\Phi}_{v2}(k) + (1 - \beta) V_2(k) V_2^*(k)$$

Methods for implementing the functionality of the VAD are well known. Also, empirical studies have shown that, when updating the noise estimate, it is reasonable to assume that background noise is stationary for about one second in the context of mobile telephony, and for even longer periods in the context of a fixed telephony network.

The background noise estimated during speech pauses is then used, during periods with residual echo (either single talk or double talk), to provide a comfort noise spectrum which effectively fills the holes created by the filter $H_{av}(k)$. In other words, a comfort noise spectrum $C(k)$ can be created as:

$$C(k) = \sqrt{(1 - H(k)H^*(k)) \Phi_{v2}(k)} N(k)$$

where $N(k)$ is the Fourier transform of a white noise sequence with unit variance.

Alternatively, the comfort noise spectrum $C(k)$ can be created as:

$$C(k) = \sqrt{(1 - H(k)H^*(k)) \Phi_{v2}(k)} e^{i\varphi(k)}$$

-14-

where the phase $\varphi(k)$ is randomly distributed between $-\pi$ and π for k in the interval 1 to $M/2-1$, and $\varphi(0) = \varphi(M/2) = 0$. For frequencies above the Nyquist frequency $M/2$, the phase is given by:

$$\varphi(k) = -\varphi(M-k) \quad k \in \left[\frac{M}{2} + 1, M-1 \right]$$

5

Since the phase in $C(k)$ is conjugate symmetric, a real time domain sequence $c(n)$ is guaranteed. The resulting comfort noise is then added to the echo suppressed output signal to obtain the echo suppressed and comfort noise enhanced output signal as follows:

10

$$e_{\text{NLP+CN}}(n) = e_{\text{NLP}}(n) + c(n)$$

Alternatively, the echo suppressed and comfort noise enhanced output signal can first be computed in the frequency domain and then transformed to the time domain as follows:

15

$$\begin{aligned} E_{\text{NLP+CN}}(k) &= H_{\text{av}}(k) \cdot E(k) + C(k) \\ e_{\text{NLP+CN}} &= \text{IFFT}(E_{\text{NLP+CN}}) \end{aligned}$$

20

In sum, the present invention provides frequency domain, nonlinear processing techniques for, e.g., echo suppression. Generally, a nonlinear processor constructed in accordance with the present invention removes energy at frequencies where, e.g., residual echo in a communications signal is dominating (e.g., where the residual echo power constitutes more than a certain percentage of

-15-

the overall power in the communications signal). The present invention further teaches methods for replacing the removed energy to thereby smoothen the resulting signal, e.g., to improve a receiving user's perception of the echo suppressed signal. Advantageously, a nonlinear processor according to the invention can suppress residual echo during the entirety of a conversation. Those skilled in the art will appreciate that the above described algorithms of the invention can be applied in either sample-wise or block-wise fashion. Those skilled in the art will also appreciate that computational complexity can be reduced, and speed advantages can be gained, by grouping frequency bins when carrying out the above described procedures.

A particularly useful algorithm 400 according to the invention is depicted by way of a flowchart in Figure 4. At step 410, an auto-regressive model of order p is computed from the time domain estimate (e.g., using the well known Levinson-Durbin algorithm), and at step 420, the resulting time domain model is converted to the frequency domain (e.g., by way of an FFT). Similarly, at steps 415 and 425, a time domain error signal is modeled (using an auto-regressive model of order q) and converted to the frequency domain. The resulting frequency domain sequences are then compared at step 430 to provide a filter transfer function H for a nonlinear processor. For example, each tap of the NLP filter H can be set to 1 or 0 depending upon whether, for the corresponding tap frequency, the power of the estimated echo is greater than a predefined fraction of the power of the error signal.

At step 440, an FFT of the error signal is computed, and at step 450 the resulting frequency domain error sequence is filtered (i.e., element-wise multiplied) by the NLP filter H to provide an echo suppressed output signal in the frequency domain. Meanwhile, the filter function H is scaled at step 460, and randomized comfort noise having a magnitude equal to a prevailing level of error

-16-

signal background noise is generated at step 470. As described above, the comfort noise generation is based on estimates of the prevailing background noise, the noise estimates being updated only when no speech is present in the error signal (shown via the VAD at step 445 in the figure). The results of steps 460 and 470 are combined at step 480 to provide comfort noise specifically tailored to appropriately replace energy removed from the error signal by the NLP filter H. At step 490, the resulting comfort noise is element-wise added to the echo suppressed output signal to provide an echo suppressed and comfort noise enhanced output signal in the frequency domain. Finally, at step 495, an IFFT is used to convert the echo suppressed and comfort noise enhanced output signal to the time domain.

Those skilled in the art will appreciate that the present invention is not limited to the specific exemplary embodiments which have been described herein for purposes of illustration and that numerous alternative embodiments are also contemplated. For example, although the disclosed NLP embodiments have been described with reference to systems including a leading echo canceler, these same NLP embodiments can also be used as stand alone devices to provide solutions superior to those provided by conventional on/off switching. The derivation syntax provided above with respect to a leading echo canceler and a trailing NLP can be readily applied to a stand alone NLP by everywhere replacing the echo estimate $\hat{s}$ with the echo causing signal $x(t)$. In such case, the residual echo can be estimated at some fixed or dynamically adjusted level below the echo-causing (e.g., loudspeaker) signal, based on the expected or measured echo return path. The scope of the invention is therefore defined by the claims appended hereto, rather than the foregoing description, and all equivalents which are consistent with the meaning of the claims are intended to be embraced therein.

-17-

We Claim:

1. A method for suppressing residual echo in a communications signal, comprising the steps of:

5 computing a power spectrum of the communications signal;
 estimating a power spectrum of the residual echo;
 comparing the computed power spectrum of the communications signal and the estimated power spectrum of the residual echo;
 adjusting a filtering characteristic based upon a result of said step of
10 comparing; and
 filtering the communications signal, using the adjusted filtering characteristic, to suppress the residual echo.

2. The method of claim 1, wherein said step of computing a power spectrum of the communications signal comprises the steps of:

15 converting a number of time domain samples of the communications signal to a corresponding number of complex frequency domain samples of the communications signal; and
 multiplying each complex frequency domain sample of the
20 communications signal by its complex conjugate to provide power spectrum samples of the communications signal.

3. The method of claim 2, wherein said step of converting a number of time domain samples of the communications signal to a corresponding number of complex frequency domain samples of the communications signal includes the
25 step of computing a Fast Fourier Transform of the number of time domain samples of the communications signal.

-18-

4. The method of claim 3, wherein said step of converting a number of time domain samples of the communications signal to a corresponding number of complex frequency domain samples of the communications signal also includes the step of computing an auto-regressive model for the time domain samples prior to computing the Fast Fourier Transform.

5. The method of claim 1, wherein the residual echo results from a source signal being transmitted through an echo return path, and wherein said step of estimating a power spectrum of the residual echo comprises the steps of:

converting a number of time domain samples of the source signal to a corresponding number of complex frequency domain samples of the source signal;

multiplying each complex frequency domain sample of the source signal by its complex conjugate to provide power spectrum samples of the source signal;

estimating an attenuation factor of the echo return path; and
scaling the power spectrum samples of the source signal by the estimated attenuation factor to provide estimated power spectrum samples of the residual echo.

6. The method of claim 5, wherein the estimate of the attenuation factor of the echo return path is static.

7. The method of claim 5, wherein the attenuation factor of the echo return path is estimated dynamically.

-19-

8. The method of claim 5, wherein said step of converting a number of time domain samples of the source signal to a corresponding number of complex frequency domain samples of the source signal includes the step of computing a Fast Fourier Transform of the number of time domain samples of the source signal.

9. The method of claim 8, wherein said step of converting a number of time domain samples of the source signal to a corresponding number of complex frequency domain samples of the source signal also includes the step of computing an auto-regressive model for the time domain samples prior to computing the Fast Fourier Transform.

10. The method of claim 1, wherein the communications signal is output by an adaptive echo canceler based on an input signal and based on an estimate of an echo component of the input signal, and wherein said step of estimating a power spectrum of the residual echo comprises the steps of:
 computing a power spectrum of the echo component estimate; and
 scaling the power spectrum of the echo component estimate, based on operation of the adaptive echo canceler, to provide the estimate of the power spectrum of the residual echo.

11. The method of claim 10, wherein said step of computing a power spectrum of the echo component estimate comprises the steps of:
 converting a number of time domain samples of the echo component estimate to a corresponding number of complex frequency domain samples of the echo component estimate; and

-20-

multiplying each complex frequency domain sample of the echo component estimate by its complex conjugate to provide power spectrum samples of the echo component estimate.

5 12. The method of claim 11, wherein said step of converting a number of time domain samples of the echo component estimate to a corresponding number of complex frequency domain samples of the echo component estimate includes the step of computing a Fast Fourier Transform of the number of time domain samples of the echo component estimate.

10

 13. The method of claim 12, wherein said step of converting a number of time domain samples of the echo component estimate to a corresponding number of complex frequency domain samples of the echo component estimate also includes the step of computing an auto-regressive model for the time domain samples prior to computing the Fast Fourier Transform.

15

 14. The method of claim 10, wherein said step of scaling the power spectrum of the echo component estimate comprises the steps of:

 determining an echo return loss enhancement of the adaptive echo canceler; and

20

 multiplying the power spectrum of the echo component estimate by the echo return loss enhancement to provide the estimate of the power spectrum of the residual echo.

25

-21-

15. The method of claim 10, wherein said step of scaling the power spectrum of the echo component estimate comprises the steps of:

determining an echo return loss enhancement of the adaptive echo canceler;

5 computing an average value for the power spectrum of the echo component estimate;

multiplying the average value by the echo return loss enhancement to provide a saturation value;

10 multiplying the power spectrum of the echo component estimate by the echo return loss enhancement to provide a scaled spectrum; and

clipping the scaled spectrum at the saturation value to provide the estimate of the power spectrum of the residual echo.

16. The method of claim 10, wherein the power spectrum of the echo component estimate is given by $\Phi(k)$, k a finite positive integer, wherein an echo return loss enhancement of the adaptive echo canceler is given by ERLE, and wherein the estimate of the power spectrum of the residual echo is computed as:

$$\text{ERLE} \Phi_s(k) .$$

17. The method of claim 16, wherein the echo return loss enhancement can vary as a function of frequency and is given by $\text{ERLE}(k)$.

-22-

18. The method of claim 10, wherein the power spectrum of the echo component estimate is given by $\Phi(k)$, k a finite positive integer, wherein an echo return loss enhancement of the adaptive echo canceler is given by ERLE, and wherein the estimate of the power spectrum of the residual echo is computed as:

$$\text{Max}(\text{ERLE}\Phi_s(k), \text{ERLE}\overline{\Phi_s}) .$$

5 19. The method of claim 18, wherein the echo return loss enhancement can vary as a function of frequency and is given by $\text{ERLE}(k)$.

10 20. The method of claim 1, wherein the filtering characteristic is given by a number of coefficients $H(k)$, k a finite positive integer, wherein the computed power spectrum of the communications signal is given by $\Phi_e(k)$, wherein the estimated power spectrum of the residual echo is given by $\Phi_s(k)$, where δ_1 is a subtraction factor, and wherein the filtering characteristic is computed as:

$$H(k) = \sqrt{\frac{\Phi_e(k) - \delta_1 \Phi_s(k)}{\Phi_e(k)}} .$$

15 21. The method of claim 1, wherein the filtering characteristic is given by a number of coefficients $H(k)$, k a finite positive integer, wherein the computed power spectrum of the communications signal is given by $\Phi_e(k)$, wherein the estimated power spectrum of the residual echo is given by $\Phi_s(k)$, wherein δ_1 is a subtraction factor, and wherein the filtering characteristic is computed as:

$$H(k) = 0 \text{ if } \Phi_e(k) < \delta_1 \Phi_s(k) , \text{ else } H(k) = 1 .$$

-23-

22. The method of claim 21, wherein the filtering characteristic coefficients $H(k)$ are averaged over time.

23. The method of claim 1, further comprising the step of adding comfort noise to the filtered communications signal.

24. The method of claim 23, wherein said step of adding comfort noise to the filtered communications signal comprises the steps of:

estimating a power spectrum of a noise component of the communications signal; and

creating the comfort noise based upon the estimated power spectrum of the noise component of the communications signal.

25. The method of claim 24, wherein the filtering characteristic is given by a number of coefficients $H(k)$, k a finite positive integer, wherein the estimated power spectrum of the noise component of the communications signal is given by $\Phi_{v_2}(k)$, wherein $N(k)$ represents a Fourier transform of a white noise sequence of unit variance, and wherein a spectrum $C(k)$ of the comfort noise is computed as:

$$C(k) = \sqrt{(1 - H(k)H^*(k))\Phi_{v_2}(k)} N(k) .$$

26. An echo suppressor, comprising:
a processor configured to compute a power spectrum of a communications signal and to estimate a power spectrum of a residual echo component of the communications signal; and

-24-

an adaptive filter configured to process the communications signal and to thereby suppress the residual echo component,

wherein a filtering characteristic of the adaptive filter is adjusted by comparing the power spectrum of the communications signal and the estimated power spectrum of the residual echo.

27. A device for suppressing residual echo in a communications signal, comprising:

means for computing a power spectrum of the communications signal;

means for estimating a power spectrum of the residual echo;

means for comparing the computed power spectrum of the communications signal and the estimated power spectrum of the residual echo;

means for adjusting a filtering characteristic based upon a result of said step of comparing; and

means for filtering the communications signal, using the adjusted filtering characteristic, to suppress the residual echo.

28. A method for suppressing residual echo in a communications signal, comprising the steps of:

computing a power spectrum of the communications signal;

estimating a power spectrum of the residual echo;

adjusting a filtering characteristic based upon the computed power spectrum of the communications signal and upon the estimated power spectrum of the residual echo; and

filtering the communications signal, using the adjusted filtering characteristic, to suppress the residual echo,

-25-

wherein, in at least one of said computing and estimating steps, a power spectrum is derived from an auto-regressive model of a time-domain signal.

29. A method according to claim 28, wherein said step of computing a power spectrum of the communications signal comprises the steps of:

computing an auto-regressive model of a number of time domain samples of the communications signal;

computing a Fast Fourier Transform of the auto-regressive model to provide a number of complex frequency domain samples of the communications signal; and

multiplying each complex frequency domain sample of the communications signal by its complex conjugate to provide power spectrum samples of the communications signal.

30. A method according to claim 28, wherein the residual echo results from a source signal being transmitted through an echo return path, and wherein said step of estimating a power spectrum of the residual echo comprises the steps of:

computing an auto-regressive model of a number of time domain samples of the source signal;

computing a Fast Fourier Transform of the auto-regressive model to provide a number of complex frequency domain samples of the source signal;

multiplying each complex frequency domain sample of the source signal by its complex conjugate to provide power spectrum samples of the source signal;

estimating an attenuation factor of the echo return path; and

-26-

scaling the power spectrum samples of the source signal by the estimated attenuation factor to provide estimated power spectrum samples of the residual echo.

5 31. A method according to claim 28, wherein the communications signal is output by an adaptive echo canceler based on an input signal and based on an estimate of an echo component of the input signal, and wherein said step of estimating a power spectrum of the residual echo comprises the steps of:

10 computing an auto-regressive model of a number of time domain samples of the echo component estimate;

 computing a Fast Fourier Transform of the auto-regressive model to provide a number of complex frequency domain samples of the echo component estimate;

15 multiplying each complex frequency domain sample of the echo component estimate by its complex conjugate to provide power spectrum samples of the echo component estimate; and

 scaling the power spectrum samples of the echo component estimate, based on operation of the adaptive echo canceler, to provide the estimate of the power spectrum of the residual echo.

20

Fig. 1

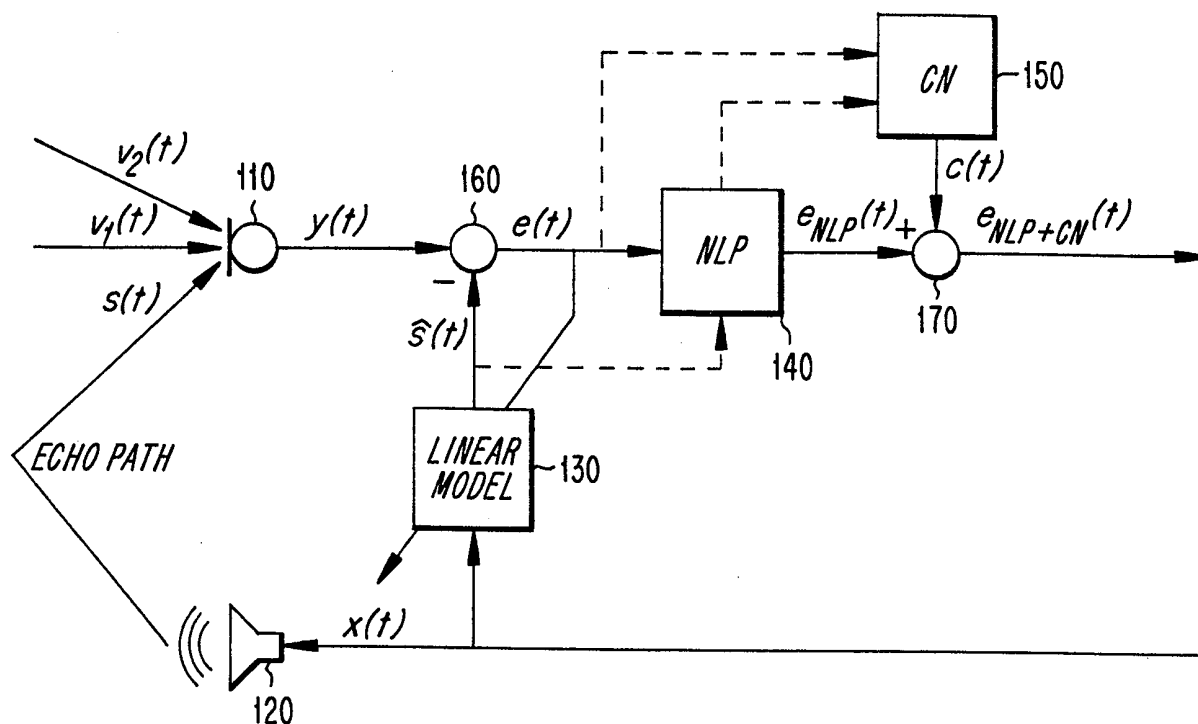


Fig. 2

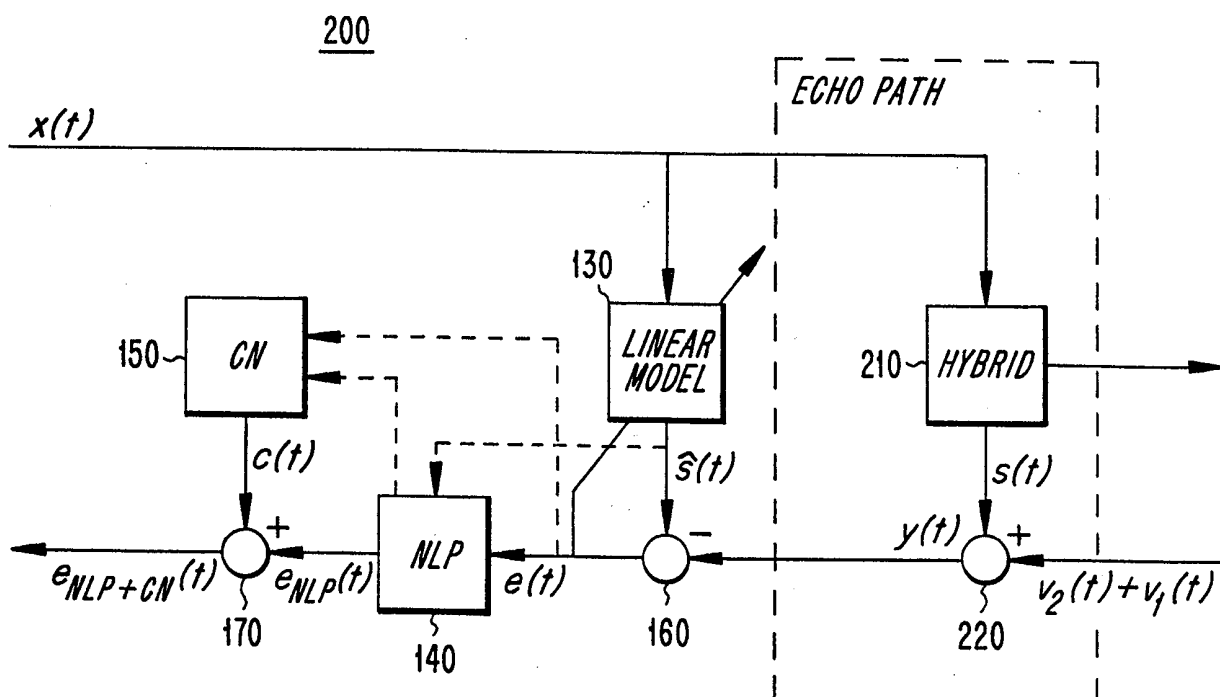


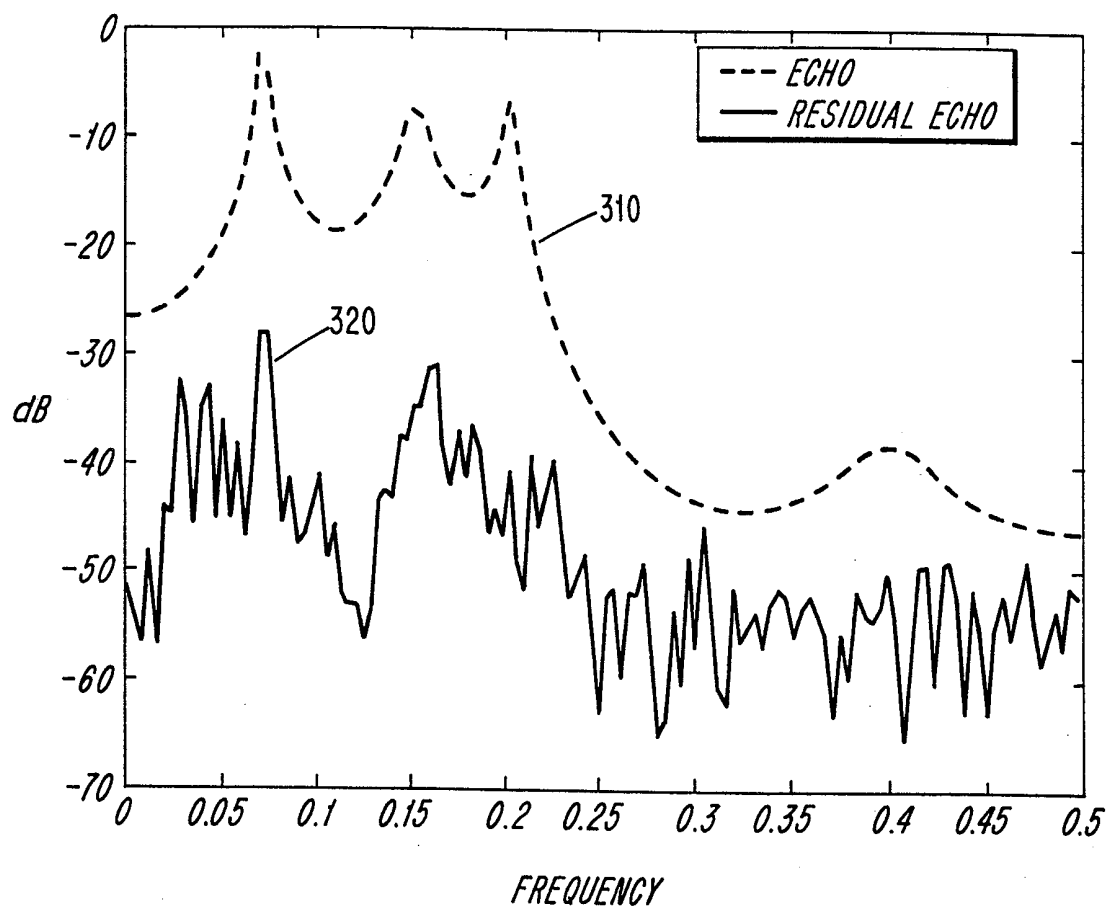
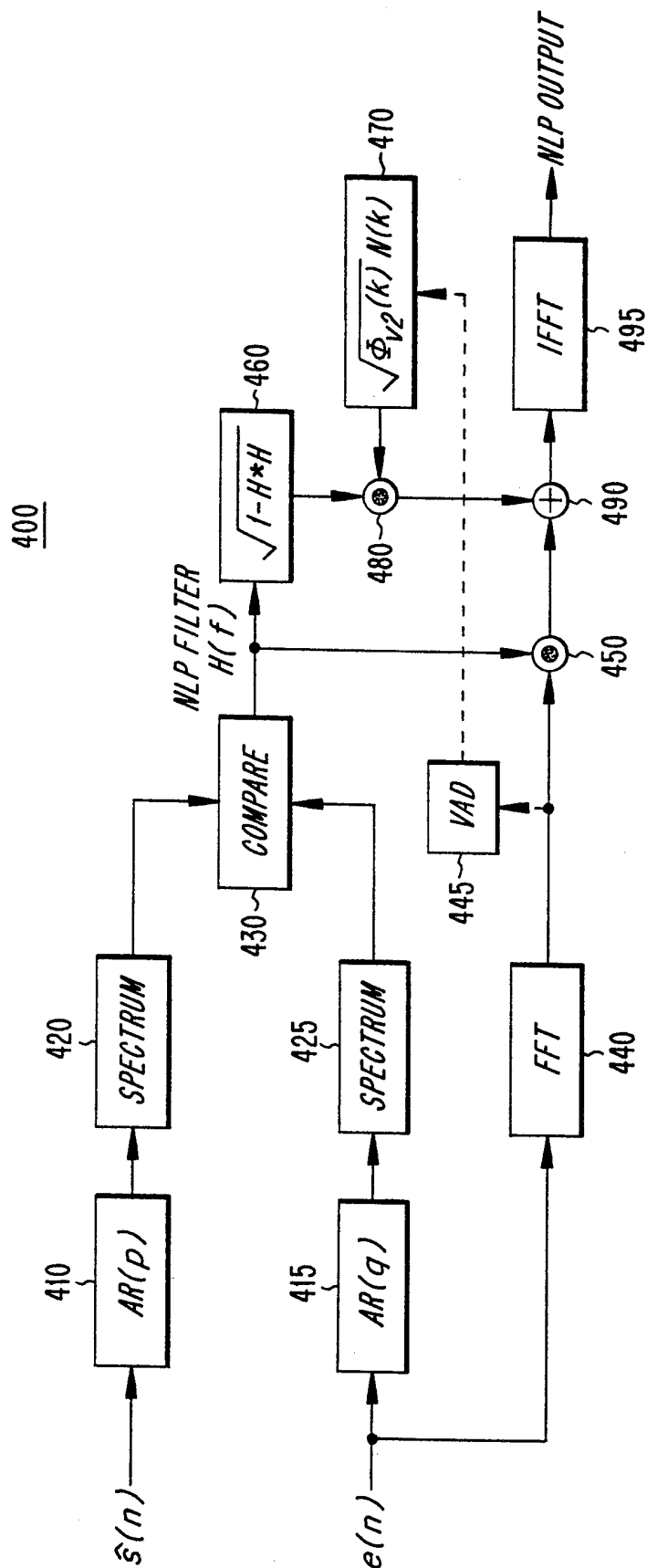
Fig. 3

Fig. 4



INTERNATIONAL SEARCH REPORT

International application No. —

PCT/SE 99/01909

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04B 3/23

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H04B, H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GUSTAFSSON S ET AL: "Combined acoustic echo control and noise reduction for hands-free telephony" SIGNAL PROCESSING EUROPEAN JOURNAL DEVOTED TO THE METHODS AND APPLICATIONS OF SIGNAL PROCESSING, vol. 64, no. 1, 1 January 1998, page 21-32 XP004108821, see the whole document	1-3,5-8, 10-12,14,16, 17,26,27
Y	--	23,24

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

23 March 2000

Date of mailing of the international search report

29 -03- 2000

Name and mailing address of the ISA/

Swedish Patent Office

Box 5055, S-102 42 STOCKHOLM

Facsimile No. +46 8 666 02 86

Authorized officer

Fredrik Blomqvist/AE

Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No. —

PCT/SE 99/01909

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5 561 668 A (GENTER ROLAND E) 1 October 1996 abstract; figures 1,2 column 1, line 5 - line 10 column 2, line 7 - line 64 column 6, line 66 - column 7, line 48	23,24
A	---	1,10,26,27
A	US 5 587 998 A (WYNN WOODSON D ET AL) 24 December 1996 abstract; figures 4-6 column 3, line 1 - line 42 column 6, line 8 - line 67 column 9, line 26 - line 48	1-3,5-7,10, 21-23,26,27
A	US 5 646 991 A (SIHGILBERT C) 8 July 1997 abstract; figure 5 column 2, line 19 - line 27 column 10, line 16 - line 46	1,23-25
A	WO 9833311 A1 (MOTOROLA INC.), 30 July 1998 (30.07.98), page 4, line 13 - page 6, line 3, figures 2,3,4,5,8, abstract	1,26,27

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No. —

PCT/SE 99/01909

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9833311 A1	30/07/98	CA 2247429 A	30/07/98
		CN 1214171 A	14/04/99
		EP 0895688 A	10/02/99
		IL 124578 D	00/00/00
