(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 848 243 A1

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
24.10.2007 Bulletin 2007/43

(51) Int Cl.:
H04R 3/02 $^{(2006.01)}$ ***H04M 9/08*** $^{(2006.01)}$

(21) Application number: **06008006.6**

(22) Date of filing: **18.04.2006**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR MK YU

(71) Applicant: **Harman Becker Automotive Systems GmbH**
76307 Karlsbad (DE)

(72) Inventors:
- **Buck, Markus, Dr.**
  88400 - Biberach (DE)
- **Haulick, Tim**
  89143 - Blaubeuren (DE)
- **Rössler, Martin**
  89077 - Ulm (DE)
- **Schmidt, Gerhard**
  89081 - Ulm (DE)

(74) Representative: **Grünecker, Kinkeldey, Stockmair & Schwanhäusser Anwaltssozietät**
Maximilianstrasse 58
80538 München (DE)

(54) Multi-channel echo compensation system and method

(57) The invention is directed to a multi-channel echo compensation system, comprising two loudspeaker input channels, each loudspeaker input channel being connected to a loudspeaker for providing a loudspeaker input signal to be emanated by the loudspeaker, a microphone output channel being connected to at least one microphone for receiving a microphone output signal from the at least one microphone, wherein each microphone is configured to acquire a signal emanating from the loudspeakers, a compensation channel for each loudspeaker input channel, each compensation channel connecting a respective loudspeaker input channel and the microphone output channel, an adaptive compensation filter for each compensation channel, wherein each adaptive compensation filter is configured to filter a signal on the respective compensation channel such that a compensation output signal is provided to compensate a microphone output signal for a signal emanating from the loudspeakers, a pre-processing means for pre-processing loudspeaker input signals on the compensation channels, the pre-processing means being configured to determine a correlation value of the loudspeaker input signals for the two loudspeakers according to a pre-determined criterion and to de-activate one of the adaptive compensation filters if the determined correlation value passes a pre-determined threshold.

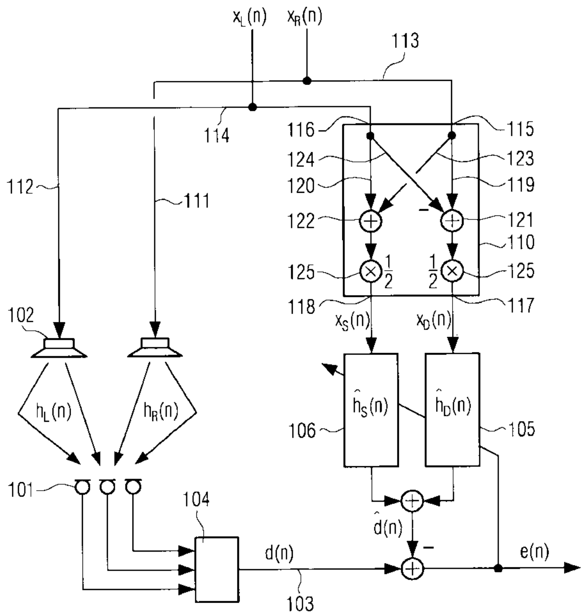


FIG. 1

EP 1 848 243 A1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The invention is directed to a multi-channel echo compensation system and method, in particular, to compensate for echoes that are present in a microphone signal and result from a multi-channel source.

[0002] The presence of echoes in a microphone signal is a problem occurring in different kinds of communication systems. A typical example is a hands-free telephony system in a vehicular cabin. Such a telephony system usually comprises one or more microphones to acquire speech signals from a speaker. However, some loudspeakers are mounted in the vehicular cabin as well. Via these loudspeakers, signals, from a radio or a CD player, for example, are output. These signals are also acquired by the microphones of the hands-free telephony system resulting in a distortion of the microphone signal.

[0003] In order to compensate for these unwanted signals, adaptive filters are used to provide a compensating signal corresponding to the unwanted signals contained in the microphone signal. For this, the adaptive filters use the input signals of the loudspeakers to determine a compensation signal that will be subtracted from the microphone signal. The structure of a conventional echo compensation system is illustrated in Fig. 6.

[0004] In the system illustrated in this figure, three microphones 601 are provided to acquire, first of all, speech signals from a speaker. However, these microphones also acquire audio signals coming from loudspeakers 602. On the microphone output channel 603 connected to the microphones, a beam-forming means 604 is provided to combine the microphone output signals in a suitable way.

[0005] The signals $\boldsymbol{h}_L(n)$ and $\boldsymbol{h}_R(n)$ acquired by the microphone array are imitated by adaptive compensation filters 605 and 606. These adaptive filters use the loudspeaker input signals $x_L(n)$ and $x_R(n)$ (which can possibly be amplified in an amplifier 607) to provide a compensation signal $\hat{d}(n)$ which comes as close as possible to the loudspeaker signal acquired by the microphones (which is denoted by $d(n)$). By subtracting the compensation signal from the microphone signal, the portion in the microphone signal stemming from the loudspeaker is removed and the resulting signal $e(n)$ becomes minimal.

[0006] When compensating signals stemming from a car radio 608, there is the problem that the absolute value squared of the coherence

$$C(\Omega) = \left| \frac{S_{x_L x_R}(\Omega)}{\sqrt{S_{x_L x_L}(\Omega) S_{x_R x_R}(\Omega)}} \right|^2$$

can vary a lot. In particular, in the case of music signals, the coherence is very small, whereas in the case of news or interviews, the signals $x_L(n)$ and $x_R(n)$ are (almost) linearly dependent and the coherence approximately equals 1. In the above equation, $S_{x_L x_R}(\Omega)$, $S_{x_L x_L}(\Omega)$ and $S_{x_R x_R}(\Omega)$ denote the cross-power spectral density and the auto-power spectral densities, respectively, of the signals $x_L(n)$ and $x_R(n)$. If the coherence is very high, i.e. if the two signals are almost linearly dependent, the cost functions used in multi-channel adaptation algorithms do not have a unique solution. This has the consequence that, for example in the case of an interview, after a change of the speaker, the filters have to be balanced anew such that echoes occur again for a short time.

[0007] Other examples in which echoes are distorting a microphone signal are a hands-free system that is provided to control devices such as a car radio or a system for passenger communication in a vehicular cabin.

[0008] During the past years, many efforts have been made to solve the above problem. According to one proposal, a non-linear pre-processing is performed on the radio signals which is indicated in Fig. 6 by a corresponding pre-processing means 609. A possible non-linear pre-processing may consist of the addition of a half-wave rectifier (see for example, J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, S. L. Gay, Advances in Network and Acoustic Echo Cancellation, Springer Verlag, Berlin, 2001). Another possibility is a time variant pre-filtering as disclosed in A. Sugiyama, Y. Joncour, A. Hirano, A Stereo Echo Canceller with Correct Echo-Path Identification Based on an Input-Sliding Technique, IEEE Transactions on Signal Processing, Vol. 49, Nr. 1, Pages 2577-2587, 2001.

[0009] It is a drawback of these methods that the signals on their path from the signal source (i.e. the car radio) to the loudspeakers are modified resulting in audible artefacts.

[0010] In view of the above-mentioned drawback of the prior art, it is the problem underlying the present invention to provide a multi-channel echo compensation system with improved echo compensation. This problem is solved by a system according to claim 1 and a method according to claim 17.

[0011] Accordingly, the invention provides a multi-channel echo compensation system, comprising:

two loudspeaker input channels, each loudspeaker input channel being connected to a loudspeaker for providing a loudspeaker input signal to be emanated by the loudspeaker,

a microphone output channel being connected to at least one microphone for receiving a microphone output signal from the at least one microphone, wherein each microphone is configured to acquire a signal emanating from the loudspeakers,

a compensation channel for each loudspeaker input channel, each compensation channel connecting a respective loudspeaker input channel and the microphone output channel,

an adaptive compensation filter for each compensation channel, wherein each adaptive compensation filter is configured to filter a signal on the respective compensation channel such that a compensation output signal is provided to compensate a microphone output signal for a signal emanating from the loudspeakers,

a pre-processing means for pre-processing loudspeaker input signals on the compensation channels, the pre-processing means being configured to determine a correlation value of the loudspeaker input signals for the two loudspeakers according to a pre-determined criterion and to de-activate one of the adaptive compensation filters if the determined correlation value passes a pre-determined threshold.

[0012] By de-activating one of the adaptive compensation filters, the above-mentioned problem of a non-unique solution of the adaptation process is overcome.

[0013] The pre-determined criterion to determine a correlation value can have different forms. For example, the correlation value can be determined as the coherence value. In this case, for example, the threshold can be selected to be 0.97. If the coherence value is greater than or equal to this value, the signals on the different input channels are considered to be highly correlated so that the adaptive filtering to provide a compensation signal can be performed with only one filter.

[0014] The multi-channel echo compensation system according to the invention has the further advantage that the required computing power can be reduced considerably in the case of highly correlated signals on different input channels. In particular, in the case of exactly two loudspeakers, i.e. a stereo system, the computing power required to perform the adaptive filtering and the update of the filters can be reduced by 50%.

[0015] The above multi-channel echo compensation system can be concerned with the case of exactly two loudspeakers or more than two loudspeakers. In the latter case, a corresponding number of loudspeaker input channels, compensation channels, and adaptive compensation filters is provided. Then, if a correlation value is determined for two of the loudspeaker input channels which passes the pre-determined threshold, all but one of the adaptive compensation filters can be de-activated.

[0016] The pre-processing means can be configured to provide a linear combination, particularly a difference and/or a sum, of signals of the two loudspeaker input channels to at least one of the adaptive compensation filters.

[0017] Linear combinations of signals, particularly differences of signals, allow to detect correlations between the signals in a very simple way.

[0018] In particular, the pre-processing means can be configured to provide a sum of the signals of the two loudspeaker input channels to a first of the adaptive compensation filters and a difference of the signals of the two second loudspeaker input channels to a second of the adaptive compensation filters.

[0019] In this way, if the signals of the first and the second loudspeaker input channel are highly correlated (for example, if the signal to be output by the loudspeaker is a mono-signal), the difference signal is very small or even zero. Therefore, it can be easily determined whether the two input signals are correlated or not.

[0020] In this case, in particular, the pre-processing means can be configured to de-activate the second of the adaptive compensation filters if the determined correlation value passes the pre-determined threshold.

[0021] The pre-processing means can be configured to determine a correlation value according to a pre-determined criterion based on the signal power of a signal, in particular, representing the difference and/or the sum of the signals of the two loudspeaker input channels.

[0022] If the input signals are highly correlated, the difference of these signals has almost vanishing power. In particular, the signal power can be determined as the norm of a corresponding signal vector; thus, a very good indication of the correlation between the two signals is obtained.

[0023] In the above multi-channel echo compensation system, the correlation value can be determined recursively. This reduces the required computing power to determine the correlation value. For example, if the signal vectors have a length N, the squared norm of this vector at time n equals the squared norm of the vector at time n-1 plus the n-th value squared of the signal vector minus the (n-N)-th value squared of the signal vector.

[0024] Each adaptive compensation filter can be configured such that no adaptation of the adaptive compensation filter is performed if the adaptive compensation filter is de-activated. Such a configuration of the adaptive compensation filters further reduces the required computing power and increases the stability of the filters.

[0025] In the above-described multi-channel echo compensation systems, the pre-processing means can comprise:

two inputs, each input being connected to a respective loudspeaker input channel,

two outputs, each output being connected to a respective adaptive compensation filter,

a first signal path connecting the first input and the first output and a second signal path connecting the second input and the second output,

a subtracting means on the first signal path and a summing means on the second signal path,

a third signal path connecting the first input and the summing means such that a signal on the first signal path and a signal on the second signal path are summed to obtain a summed signal,

and a fourth signal path connecting the second input and the subtracting means such that a signal on the first signal path and a signal on the second signal path are subtracted from each other to obtain a subtracted signal.

[0026] This is a very advantageous implementation of a multi-channel echo compensation system. In particular, if the signals on the two loudspeaker input channels are highly correlated, the signal power of the subtracted signal will be very low. On the other hand, in this case, the summed signal will have a signal power which essentially corresponds to twice the signal power of the respective signals in the loudspeaker input channels. If the signal power of the subtracted signal is below a pre-determined threshold, this is a criterion to determine that the correlation is very high. Then, the adaptive filter being connected to the first output of the pre-processing means can be de-activated. The compensation of the microphone output signal can be based on the summed signal only which is input into the corresponding adaptive compensation filter.

[0027] The pre-processing means can comprise a multiplying means on the first and/or the second signal path to multiply a signal on the first and/or on the second signal path by a weighting factor, in particular, to multiply the subtracted signal and the summed signal by a common weighting factor.

[0028] In this way, one can take into account that the signals after summation or subtraction are a superposition of different signals. In particular, if the signals on a first and a second loudspeaker input channel are highly correlated and summed in the pre-processing means, the corresponding summed signal corresponds to the signal on one of the loudspeaker input channels, but with twice the signal power. In that case, multiplying the summed signal by a factor of 0.5 leads to a resulting signal which essentially corresponds to the signal of one of the loudspeaker input channels.

[0029] The pre-processing means can comprise a first adaptive pre-processing filter on the third signal path, the adaptation of which is based on the difference between the signals on the first and on the third signal path, and a second adaptive pre-processing filter on the fourth signal path, the adaptation of which is based on the difference between the signals on the second and on the fourth signal path.

[0030] When playing back an interview on a stereo channel system, for example, the two speakers are often placed on different sides in an acoustic way. In other words, one speaker is more present on a first channel, whereas the other speaker is more present on a second channel. In such a case, the speech signals on both loudspeaker input channels are still highly correlated, however, their difference would be non-zero. Providing adaptive pre-processing filters on the third and fourth signal paths is useful in overcoming this problem.

[0031] The adaptive compensation filters and the adaptive pre-processing filters can be configured such that the adaptation of the pre-processing filters is performed slower than the adaptation of the compensation filters.

[0032] Such a configuration increases the stability of the system. In particular, the increments in the adaptation process of the pre-processing filters can be smaller than the increments in the adaptation process of the compensation filters.

[0033] The adaptation increments of the compensation filters can be larger than the adaptation increments of the pre-processing filters. In this way, the adaptation of the pre-processing filters will be slower than that of the compensation filters.

[0034] The pre-processing means can comprise a delay means on the first and on the second signal path, respectively, each delay means being configured to delay the signals on the first and on the second signal path before arriving at the summing means and subtracting means, respectively.

[0035] In this way, one overcomes the problem that the adaptive pre-processing filters do not necessarily converge towards a causal optimal solution.

[0036] In particular, the delay of the delay means can be selected such that it corresponds to about half of the length of a corresponding adaptive pre-processing filter. Then, about half of this adaptive pre-processing filter reproduces non-causal parts.

[0037] The previously described multi-channel echo compensation systems can comprise a summing means which is provided between the adaptive compensation filters and the microphone output channel and is configured to sum the signals emanating from the adaptive compensation filters.

[0038] In this way, an advantageous compensation signal for subtraction from the microphone output signal is provided.

[0039] The at least one microphone of the above multi-channel echo compensation systems can be configured as an array of at least two microphones being connected to a beam-forming means on the microphone output channel.

[0040] Such a configuration improves the signal to noise ratio in the microphone output channel. In particular, the beam-forming means yields a suitable directivity of the microphone array.

[0041] The invention also provides a method for compensating echoes in a multi-channel system, the multi-channel system comprising two loudspeakers, each loudspeaker being connected to a loudspeaker input channel for providing a loudspeaker input signal to be emanated by the loudspeaker, at least one microphone for acquiring a signal emanating from the loudspeakers, the at least one microphone being connected to a microphone output channel, a compensation channel for each loudspeaker input channel, each compensation channel connecting a respective loudspeaker input channel and the microphone output channel, an adaptive compensation filter for each compensation channel, wherein each adaptive compensation filter is configured to filter a signal on the respective compensation channel such that a compensation output signal is provided to compensate a microphone output signal for a signal emanating from the loudspeakers, the method comprising the steps of:

receiving loudspeaker input signals, wherein each loudspeaker input signal is received on a compensation channel,

pre-processing the loudspeaker input signals on the compensation channels, the pre-processing step comprising determining a correlation value of two loudspeaker signals for the two loudspeakers according to a pre-determined criterion and de-activating one of the adaptive compensation filters if the determined correlation value passes a pre-determined threshold.

[0042] The features and advantages described above in the context of the multi-channel echo compensation system apply to the case of the compensation method as well.

[0043] The pre-processing step can comprise providing a linear combination, particularly a difference and/or a sum, of signals of the two loudspeaker input channels to at least one of the adaptive compensation filters.

[0044] The pre-processing step can comprise providing a sum of the signals of the first and the second loudspeaker input channel to a first of the adaptive compensation filters and a difference of the signals of the first and second loudspeaker input channels to a second of the adaptive compensation filters.

[0045] The pre-processing step can comprise de-activating the second of the adaptive compensation filters if the determined correlation value passes the pre-determined threshold.

[0046] The pre-processing step can comprise determining a correlation value according to a pre-determined criterion based on the signal power of a signal, in particular, representing the difference and/or the sum of the signals of the two loudspeaker input channels.

[0047] In the above described methods, the correlation value can be determined recursively.

[0048] The pre-processing step can comprise summing the loudspeaker input signals on the two compensation input signals to obtain a summed signal and subtracting the loudspeaker input signals on the two compensation input signals to obtain a subtracted signal.

[0049] The pre-processing step can comprise multiplying the summed signal and the subtracted signal by a weighting factor, in particular, by a common weighting factor.

[0050] In the above methods, the pre-processing step can comprise:

adaptively filtering a loudspeaker signal of a first loudspeaker input channel before being added to a loudspeaker signal of a second loudspeaker channel, wherein the adaptation is based on the difference between the loudspeaker signals of the first and the second loudspeaker input channel, and

adaptively filtering a loudspeaker signal of a second loudspeaker input channel before being subtracted from a loudspeaker signal of a second loudspeaker channel, wherein the adaptation is based on the difference between the loudspeaker signals of the first and the second loudspeaker input channel.

[0051] In particular, the adaptively filtering of the pre-processing step can be performed slower than the adaptation of the adaptive compensation filters. For example, the adaptation increments of the compensation filters can be chosen to be larger than the adaptation increments of the adaptively filtering of the pre-processing step.

[0052] The pre-processing step can comprise:

delaying a loudspeaker signal of a second loudspeaker input channel before being added to an adaptively filtered loudspeaker signal of a first loudspeaker input channel,

delaying a loudspeaker signal of a first loudspeaker input channel before subtracting an adaptively filtered loudspeaker signal of a first loudspeaker input channel therefrom.

[0053] In particular, the delay can be selected such that it corresponds to about half of the length of a corresponding filter for adaptively filtering in the pre-processing step.
[0054] The above described methods can comprise the step of summing the signals emanating from the adaptive compensation filters.
[0055] The invention also provides a computer program product comprising one or more computer readable media having computer-executable instructions for performing the steps of the above described methods when run on a computer.
[0056] Further features and advantages of the invention will be described in the following with reference to the figures.

Fig. 1 illustrates the structure of a multi-channel echo compensation system;

Fig. 2 illustrates an example of the variation in time of a radio signal on two loudspeaker input channels;

Fig. 3 illustrates the variation in time of the sum and the difference of a radio signal;

Fig. 4 illustrates an example of the short time power spectrum of different signals;

Fig. 5 illustrates the structure of another multi-channel echo compensation system;

Fig. 6 illustrates the structure of a prior art multi-channel echo compensation system.

[0057] Fig. 1 illustrates schematically an embodiment of a multi-channel echo compensation system according to the present invention. In this example, a stereo signal source such as a car radio (not shown) outputs radio signals for a left loudspeaker channel $x_L(n)$ and for a right loudspeaker channel $x_R(n)$. These radio signals are emanated by two loudspeakers 102 and are acquired by a microphone array consisting of three microphones 101. The transmission from the loudspeakers to the microphones, for example, in a vehicular cabin, can be described by finite impulse responses:

$$\boldsymbol{h}_L(n) = \left[h_{L,0}(n), h_{L,1}(n), \ldots, h_{L,L-1}(n)\right]^T$$

and

$$\boldsymbol{h}_R(n) = \left[h_{R,0}(n), h_{R,1}(n), \ldots, h_{R,L-1}(n)\right]^T.$$

[0058] The variable n indicates the time dependence of the coefficients.
[0059] The signals acquired by the microphones 101 are output to a microphone output channel 103, on which a processing means 104 is provided. Such a processing means can perform a linear time invariant processing, for example, as it is done by a beam-former or a high-pass filter. The microphone output signal (in the example after the processing means 104) is denoted by *d(n).*
[0060] The microphones are mainly used to acquire speech signals from a speaker, for example, for a hands-free telephony system or for a hands-free control system of specific devices. In view of this, it is desirable to reduce radio signal components in the microphone output signal *d(n).* To reduce these components, two adaptive compensation filters 105 and 106 are provided, the impulse responses of which are given by

$$\hat{\boldsymbol{h}}_L(n) = \left[\hat{h}_{L,0}(n), \hat{h}_{L,1}(n), \ldots, \hat{h}_{L,N-1}(n)\right]^T$$

and

$$\hat{\boldsymbol{h}}_R(n) = \left[\hat{h}_{R,0}(n), \hat{h}_{R,1}(n), \ldots, \hat{h}_{R,N-1}(n)\right]^T.$$

[0061] In general, the order of the adaptive filters N is smaller than the order of the impulse responses. As an example, 300 to 500 coefficients at a sampling rate of 11 kHz can be used for the adaptive filters. The adaptive filters are to provide a compensation signal that is to be subtracted from the microphone output signal *d(n)*. This results in a signal *e(n)* which is sometimes called error signal and which has the form:

$$e(n) = d(n) - \sum_{i=0}^{N-1} \hat{h}_{L,i}(n) x_L(n-i) - \sum_{i=0}^{N-1} \hat{h}_{R,i}(n) x_R(n-i).$$

[0062] This output signal is used to adapt the adaptive compensation filters. The adaptation of the filters is to be performed in such a way that the estimated impulse response $\hat{h}_{L,i}(n)$ and $\hat{h}_{R,i}(n)$ are as close as possible to the real impulse responses $h_{L,i}(n)$ and $h_{R,i}(n)$ and that a high number of coefficients is estimated. The adaptation of the filters can be performed via the NLMS algorithm

[0063] In the example of Fig. 1, each loudspeaker input channel 111 and 112 is connected to the microphone output channel 103 by a corresponding compensation channel 113 and 114 such that each compensation channel receives a respective loudspeaker input signal. The pre-processing means 110 comprises two inputs 115 and 116, each input being connected to a respective loudspeaker input channel via a compensation channel. The pre-processing means 110 further comprises two outputs 117 and 118, these outputs being connected to the respective adaptive compensation filters 105 and 106.

[0064] Within the pre-processing means 110, there is a first signal path 119 connecting the first input 115 and the first output 117; furthermore, there is a second signal path connecting the second input 116 and the second output 118. On the first signal path 119, a subtracting means 121, and on the second signal path 120, a summing means 122, are provided.

[0065] A third signal path 123 connects the first input 115 and the summing means 122, whereas a fourth signal path 124 connects the second input 116 and the subtracting means 121. On the first and the second signal path, the summing and the subtracting means are followed by a multiplying means 125 to multiply the summed and the subtracted signals by a common weighting factor of 0.5. As a result, the following linear combinations of the loudspeaker input channels are emanating from the pre-processing means at the outputs 117 and 118:

$$x_s(n) = \frac{1}{2}\left[x_L(n) + x_R(n)\right],$$

$$x_D(n) = \frac{1}{2}\left[x_L(n) - x_R(n)\right]$$

[0066] In conventional signal processors, weighting with a factor of 0.5 can be realized by shifting the result in the accumulation register by one bit.

[0067] Fig. 2 shows the variation in time of a typical stereo radio signal. The upper graph corresponds to the left loudspeaker input channel and the lower graph to the right loudspeaker input channel. In the left part of each graph, the power spectrum for the playback of news is depicted; the right part corresponds to the playback of classical music.

[0068] Fig. 3 illustrates the signals $x_S(n)$ and $x_D(n)$ corresponding to the signals of Fig. 2. The upper graph shows $x_S(n)$ weighted by a factor of 0.5 and the lower graph shows $x_D(n)$ weighted by a factor of 0.5. As a signal corresponding to the playback of news is usually a mono-signal, the difference signal $x_D(n)$ vanishes or at least almost vanishes during this time period. In the case of classical music, the summed signals and the subtracted signals differ only slightly from the original signals.

[0069] Returning to Fig. 1, as soon as the difference signal $x_D(n)$ has a power spectrum below a pre-determined threshold, the corresponding adaptive compensation filter 105 is de-activated. Then, no adaptation of this adaptive compensation filter is performed so that the computing power required to adapt the filters is halved. However, although only one of the adaptive compensation filters is activated, the resulting echo reduction is unaltered to a large degree.

[0070] This is illustrated in Fig. 3. In this figure, the variation of a microphone output signal $d(n)$ in time without additional echo compensation is shown. Furthermore, an output signal $e(n)$, for which a conventional echo compensation as illustrated in Fig. 6 was performed, is depicted as a comparative example. Furthermore, the graph also shows the short time power of an output signal $e(n)$ for which an echo compensation according to Fig. 1 was performed. However, 4 dB were added to the values of this last curve for better illustration. In other words, without the additional 4 dB, the curves of the output signals with the conventional and the novel echo compensation method would be almost indistinguishable.

[0071] It is to be pointed out that, although the resulting echo compensation with the novel method is as good as the conventional method, the required computing power is significantly reduced. As can be seen, after some tuning, the radio signal can be reduced by approximately 30 dB.

[0072] In Fig. 1, the pre-processing means 110 and the adaptive filters 105 and 106 have been depicted as separate elements. Then, control of the adaptive filters by the pre-processing means, in particular, de-activation of one of the adaptive filters, has to be performed via a corresponding control connection (not shown). However, it is to be understood that checking whether a correlation value is below a pre-determined threshold can also be performed at the input of an adaptive filter; in this case, the pre-processing means would comprise part of the adaptive filter as well.

[0073] Determining a correlation value according to a pre-determined criterion can be performed by determining the squared norm, i.e. in the signal power, of the signal vector at the output of the pre-processing means or the input of the adaptive compensation filter. This can be done in a recursive way:

$$\|\boldsymbol{x}_S(n)\|^2 = \|\boldsymbol{x}_S(n-1)\|^2 + x_S^2(n) - x_S^2(n-N),$$

$$\|\boldsymbol{x}_D(n)\|^2 = \|\boldsymbol{x}_D(n-1)\|^2 + x_D^2(n) - x_D^2(n-N).$$

[0074] If the norm determined in this way falls below a pre-determined threshold, a corresponding release variable $a_S(n)$ and $a_D(n)$ can be set to zero:

$$a_S(n) = \begin{cases} 1, \; if \; \|\boldsymbol{x}_S(n)\|^2 > P_0, \\ 0, \; else, \end{cases}$$

$$a_D(n) = \begin{cases} 1, \; if \; \|\boldsymbol{x}_D(n)\|^2 > P_0, \\ 0, \; else. \end{cases}$$

[0075] The pre-determined threshold, for example, can be set to 0.03. Then, the determination of the output signal $e(n)$ after subtraction of the compensation signal is:

$$e(n) = d(n) - a_S(n)\sum_{i=0}^{N-1}\hat{h}_{S,i}(n)x_S(n-i) - a_D(n)\sum_{i=0}^{N-1}\hat{h}_{D,i}(n)x_D(n-i).$$

[0076] In this equation, the sums (corresponding to a convolution) are to be determined only if the corresponding release variable is non-zero. Correspondingly, the adaptation of the adaptive compensation filters is to be performed only under these circumstances:

$$\hat{h}_{S,i}(n+1)=\begin{cases}\hat{h}_{S,i}(n)+\mu\dfrac{x_S(n-i)e(n)}{\|\boldsymbol{x}_S(n)\|^2+\|\boldsymbol{x}_D(n)\|^2}, & if\ a_S(n)=1\\ \hat{h}_{S,i}(n), & else\end{cases}$$

$$\hat{h}_{D,i}(n+1)=\begin{cases}\hat{h}_{D,i}(n)+\mu\dfrac{x_D(n-i)e(n)}{\|\boldsymbol{x}_S(n)\|^2+\|\boldsymbol{x}_D(n)\|^2}, & if\ a_S(n)=1\\ \hat{h}_{D,i}(n), & else.\end{cases}$$

[0077] It is to be pointed out that the conventional echo compensation method according to Fig. 6 has the following convergence behavior for the adaptive compensation filters:

$$\hat{\boldsymbol{h}}_L(n)\Big|_{E\{e^2(n)\}\to\min}=\boldsymbol{h}_L(n),$$

$$\hat{\boldsymbol{h}}_R(n)\Big|_{E\{e^2(n)\}\to\min}=\boldsymbol{h}_R(n).$$

under the condition that the signals are not completely correlated. The novel echo compensation method according to Fig. 1 has the following convergence behavior, even if the input signals are fully correlated:

$$\hat{\boldsymbol{h}}_S(n)\Big|_{E\{e^2(n)\}\to\min}=\boldsymbol{h}_L(n)+\boldsymbol{h}_R(n),$$

$$\hat{\boldsymbol{h}}_D(n)\Big|_{E\{e^2(n)\}\to\min}=\boldsymbol{h}_L(n)-\boldsymbol{h}_R(n).$$

[0078] Thus, there is no non-uniqueness with the novel method.

[0079] An extension of the multi-channel echo compensation system illustrated in Fig. 1 is depicted schematically in Fig. 5. In this figure, the structure of the corresponding pre-processing means is shown.

[0080] This embodiment is particularly useful if, in an interview, one of the speakers is (acoustically) placed to the left and the other to the right. This can be done by varying the amplification of the right and the left channel, by inserting delay time elements, or by a combination of both. In addition, further filters for amending the tone can be used.

[0081] In such a case, a modified pre-processing means as shown in Fig. 5 is advantageous. In this embodiment, the output signals of the pre-processing means have the form:

$$x_D(n)=\frac{1}{2}\left[x_R(n-N_V)-\sum_{i=0}^{N_G}x_L(n-i)\hat{g}_{D,i}(n)\right],$$

$$x_S(n) = \frac{1}{2}\left[x_L(n-N_V) + \sum_{i=0}^{N_G} x_R(n-i)\hat{g}_{S,i}(n)\right].$$

[0082] Here, additional adaptive pre-processing filters 526 and 527 are provided along the third signal path 523 and the fourth signal path 524, respectively. Furthermore, on the first signal path 519 and the second signal path 520, a delay means 528 and 529, respectively, is provided. These delay means are provided before the summing and the subtracting means in the direction of the signal flow.

[0083] The adaptation of the pre-processing filters 526 and 527 can be performed via the NLMS algorithm, as in the case of the compensation filters. First of all, the two error signals for the adaptation are determined as:

$$e_D(n) = 2x_D(n)$$

and

$$e_S(n) = x_L(n-N_V) - \sum_{i=0}^{N_G} x_R(n-i)\hat{g}_{S,i}(n).$$

[0084] Then, the filter adaptation is performed according to

$$\hat{g}_{D,i}(n+1) = \hat{g}_{D,i}(n) + \mu_G \frac{x_L(n-i)e_D(n)}{\sum_{p=0}^{N_G} x_L^2(n-p)},$$

$$\hat{g}_{S,i}(n+1) = \hat{g}_{S,i}(n) + \mu_G \frac{x_R(n-i)e_S(n)}{\sum_{p=0}^{N_G} x_R^2(n-p)}.$$

[0085] Preferably, the adaptation of the pre-processing filters 526 and 527 is performed slower than the adaptation of the compensation filters 105 and 106. This can be achieved, for example, by choosing smaller increments in the case of the adaptation of the pre-processing filters:

$$0 \le \mu_G << \mu \le 1.$$

[0086] As the filters 526 and 527 do not necessarily converge towards causal optimal solutions, the delay elements 528 and 529 can be configured in such a way that the delay times of N_V cycles are selected that about half of the corresponding filter reproduce non-causal parts:

$$N_V \approx \frac{N_G}{2}.$$

[0087] In the case of a mono-signal ($x_L(n) = x_R(n)$), both filters converge to the optimal solutions with the transfer functions

$$\hat{G}_S(e^{j\Omega})\Big|_{E\{e_S^2(n)\}\to\min} = e^{-j\Omega N_V},$$

$$\hat{G}_D(e^{j\Omega})\Big|_{E\{e_D^2(n)\}\to\min} = e^{-j\Omega N_V}.$$

[0088] Then, the signals output by the pre-processing means have the form

$$x_S(n) = \frac{1}{2}\left[x_R(n-N_V) + x_L(n-N_V)\right],$$

$$x_D(n) = \frac{1}{2}\left[x_R(n-N_V) - x_L(n-N_V)\right]$$

[0089] Thus, except for a delay by N_V cycles, the adaptive pre-processing of Fig. 5 corresponds to the fixed pre-processing shown in Fig. 1.

[0090] In the above examples, the multi-channel echo compensation system was described for the case of a two-channel, i.e. stereo, system. However, an extension to more than two channels is also possible. In this case, for example, the pre-processing means can be configured to de-activate all but one of the adaptive filters if a correlation value passes a pre-determined threshold.

[0091] Further modifications and variations of the present invention will be apparent to those skilled in the art in view of this description. Accordingly, the description is to be construed as illustrative only and is for the purpose of teaching those skilled in the art the general manner of carrying out the present invention. It is to be understood that the forms of the invention shown and described herein are to be taken as the presently preferred embodiments.

Claims

1. Multi-channel echo compensation system, comprising:

two loudspeaker input channels, each loudspeaker input channel being connected to a loudspeaker for providing a loudspeaker input signal to be emanated by the loudspeaker,
a microphone output channel being connected to at least one microphone for receiving a microphone output signal from the at least one microphone, wherein each microphone is configured to acquire a signal emanating from the loudspeakers,
a compensation channel for each loudspeaker input channel, each compensation channel connecting a respective loudspeaker input channel and the microphone output channel,
an adaptive compensation filter for each compensation channel, wherein each adaptive compensation filter is configured to filter a signal on the respective compensation channel such that a compensation output signal is provided to compensate a microphone output signal for a signal emanating from the loudspeakers,
a pre-processing means for pre-processing loudspeaker input signals on the compensation channels, the pre-processing means being configured to determine a correlation value of the loudspeaker input signals for the

two loudspeakers according to a pre-determined criterion and to de-activate one of the adaptive compensation filters if the determined correlation value passes a pre-determined threshold.

2. Multi-channel echo compensation system according to claim 1, wherein the pre-processing means is configured to provide a linear combination, particularly a difference and/or a sum, of signals of the two loudspeaker input channels to at least one of the adaptive compensation filters.

3. Multi-channel echo compensation system according to claim 1 or 2, wherein the pre-processing means is configured to provide a sum of the signals of the two loudspeaker input channels to a first of the adaptive compensation filters and a difference of the signals of the two loudspeaker input channels to a second of the adaptive compensation filters.

4. Multi-channel echo compensation system according to claim 3, wherein the pre-processing means is configured to de-activate the second of the adaptive compensation filters if the determined correlation value passes the pre-determined threshold.

5. Multi-channel echo compensation system according to one of the preceding claims, wherein the pre-processing means is configured to determine a correlation value according to a pre-determined criterion based on the signal power of a signal, in particular, representing the difference and/or the sum of the signals of the two loudspeaker input channels.

6. Multi-channel echo compensation system according to one of the preceding claims, wherein the correlation value is determined recursively.

7. Multi-channel echo compensation system according to one of the preceding claims, wherein each adaptive compensation filter is configured such that no adaptation of the adaptive compensation filter is performed if the adaptive compensation filter is de-activated.

8. Multi-channel echo compensation system according to one of the preceding claims, wherein the pre-processing means comprises:

two inputs, each input being connected to a respective loudspeaker input channel,
two outputs, each output being connected to a respective adaptive compensation filter,
a first signal path connecting the first input and the first output and a second signal path connecting the second input and the second output,
a subtracting means on the first signal path and a summing means on the second signal path,
a third signal path connecting the first input and the summing means such that a signal on the first signal path and a signal on the second signal path are summed to obtain a summed signal,
and a fourth signal path connecting the second input and the subtracting means such that a signal on the first signal path and a signal on the second signal path are subtracted from each other to obtain a subtracted signal.

9. Multi-channel echo compensation system according to claim 8, wherein the pre-processing means comprises a multiplying means on the first and/or the second signal path to multiply a signal on the first and/or on the second signal path by a weighting factor, in particular, to multiply the subtracted signal and the summed signal by a common weighting factor.

10. Multi-channel echo compensation system according to claim 8 or 9, wherein the pre-processing means comprises:

a first adaptive pre-processing filter on the third signal path, the adaptation of which is based on the difference between the signals on the first and on the third signal path,
a second adaptive pre-processing filter on the fourth signal path, the adaptation of which is based on the difference between the signals on the second and on the fourth signal path.

11. Multi-channel echo compensation system according to claim 10, wherein the adaptive compensation filters and the adaptive pre-processing filters are configured such that the adaptation of the pre-processing filters is performed slower than the adaptation of the compensation filters.

12. Multi-channel echo compensation system according to claim 11, wherein the adaptation increments of the compensation filters are larger than the adaptation increments of the pre-processing filters.

13. Multi-channel echo compensation system according to one of the claims 10 - 12, wherein the pre-processing means comprises a delay means on the first and on the second signal path, respectively, each delay means being configured to delay the signals on the first and second signal path before arriving at the summing means and subtracting means, respectively.

14. Multi-channel echo compensation system according to claim 13, wherein the delay of the delay means is selected such that it corresponds to about half of the length of a corresponding adaptive pre-processing filter.

15. Multi-channel echo compensation system according to one of the preceding claims, comprising a summing means which is provided between the adaptive compensation filters and the microphone output channel and is configured to sum the signals emanating from the adaptive compensation filters.

16. Multi-channel echo compensation system according to one of the preceding claims, wherein the at least one microphone is configured as an array of at least two microphones being connected to a beam-forming means on the microphone output channel.

17. Method for compensating echoes in a multi-channel system, the multi-channel system comprising two loudspeakers, each loudspeaker being connected to a loudspeaker input channel for providing a loudspeaker input signal to be emanated by the loudspeaker, at least one microphone for acquiring a signal emanating from the loudspeakers, the at least one microphone being connected to a microphone output channel, a compensation channel for each loudspeaker input channel, each compensation channel connecting a respective loudspeaker input channel and the microphone output channel, an adaptive compensation filter for each compensation channel, wherein each adaptive compensation filter is configured to filter a signal on the respective compensation channel such that a compensation output signal is provided to compensate a microphone output signal for a signal emanating from the loudspeakers, the method comprising the steps of:

    receiving loudspeaker input signals, wherein each loudspeaker input signal is received on a compensation channel,
    pre-processing the loudspeaker input signals on the compensation channels, the pre-processing step comprising determining a correlation value of two loudspeaker signals for the two loudspeakers according to a pre-determined criterion and de-activating one of the adaptive compensation filters if the determined correlation value passes a pre-determined threshold.

18. Method according to claim 17, wherein the pre-processing step comprises providing a linear combination, particularly a difference and/or a sum, of signals of the two loudspeaker input channels to at least one of the adaptive compensation filters.

19. Method according to claim 17 or 18, wherein the pre-processing step comprises providing a sum of the signals of the first and the second loudspeaker input channel to a first of the adaptive compensation filters and a difference of the signals of the first and second loudspeaker input channels to a second of the adaptive compensation filters.

20. Method according to claim 19, wherein the pre-processing step comprises de-activating the second of the adaptive compensation filters if the determined correlation value passes the pre-determined threshold.

21. Method according to one of the claims 17 - 20, wherein the pre-processing step comprises determining a correlation value according to a pre-determined criterion based on the signal power of a signal, in particular, representing the difference and/or the sum of the signals of the two loudspeaker input channels.

22. Method according to one of the claims 17 - 21, wherein the correlation value is determined recursively.

23. Method according to one of the claims 17 - 22, wherein the pre-processing step comprises summing the loudspeaker input signals on the two compensation input signals to obtain a summed signal and subtracting the loudspeaker input signals on the two compensation input signals to obtain a subtracted signal.

24. Method according to claim 23, wherein the pre-processing step comprises multiplying the summed signal and the subtracted signal by a weighting factor, in particular, by a common weighting factor.

25. Method according to claim 23 or 24, wherein the pre-processing step comprises:

adaptively filtering a loudspeaker signal of a first loudspeaker input channel before being added to a loudspeaker signal of a second loudspeaker channel, wherein the adaptation is based on the difference between the loudspeaker signals of the first and the second loudspeaker input channel, and
adaptively filtering a loudspeaker signal of a second loudspeaker input channel before being subtracted from a loudspeaker signal of a second loudspeaker channel, wherein the adaptation is based on the difference between the loudspeaker signals of the first and the second loudspeaker input channel.

26. Method according to claim 26 wherein the adaptively filtering of the pre-processing step is performed slower than the adaptation of the adaptive compensation filters.

27. Method according to claim 16, wherein the adaptation increments of the compensation filters are chosen to be larger than the adaptation increments of the adaptively filtering of the pre-processing step.

28. Method according to one of the claims 25 - 27, wherein the pre-processing step comprises

delaying a loudspeaker signal of a second loudspeaker input channel before being added to an adaptively filtered loudspeaker signal of a first loudspeaker input channel,
delaying a loudspeaker signal of a first loudspeaker input channel before subtracting an adaptively filtered loudspeaker signal of a first loudspeaker input channel therefrom.

29. Method according to claim 28, wherein the delay is selected such that it corresponds to about half of the length of a corresponding filter for adaptively filtering in the pre-processing step.

30. Method according to one of the claims 17 - 29, comprising the step of summing the signals emanating from the adaptive compensation filters.

31. Computer program product comprising one or more computer readable media having computer-executable instructions for performing the steps of the method according to one of the claims 17 - 30 when run on a computer.

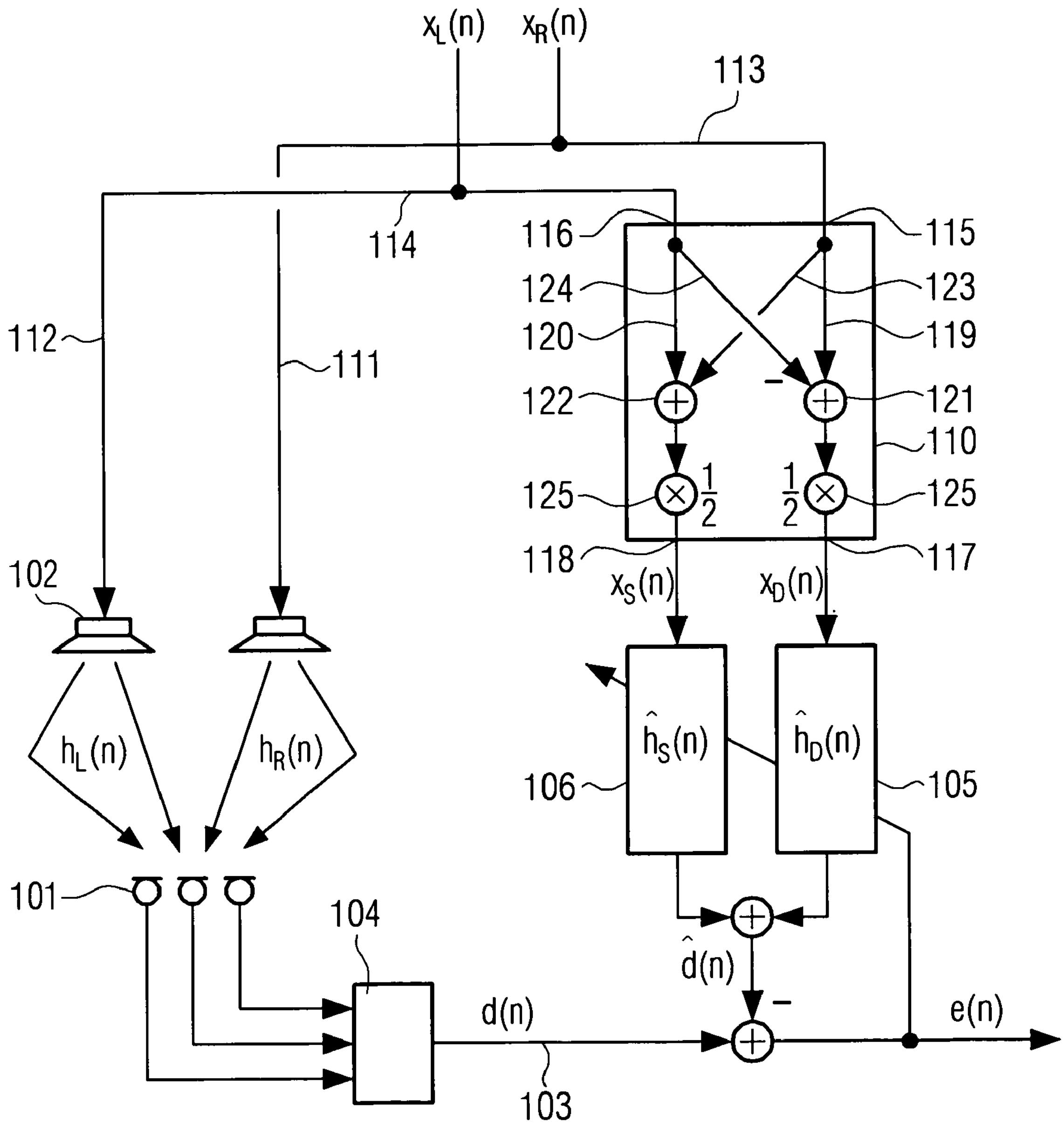

$x_L(n)$
$x_R(n)$
113
114
116
115
124
123
120
119
112
111
122
121
110
125
125
$\frac{1}{2}$
$\frac{1}{2}$
118
117
$x_S(n)$
$x_D(n)$
102
$h_L(n)$
$h_R(n)$
$\hat{h}_S(n)$
$\hat{h}_D(n)$
106
105
101
104
$\hat{d}(n)$
d(n)
e(n)
103

FIG. 1

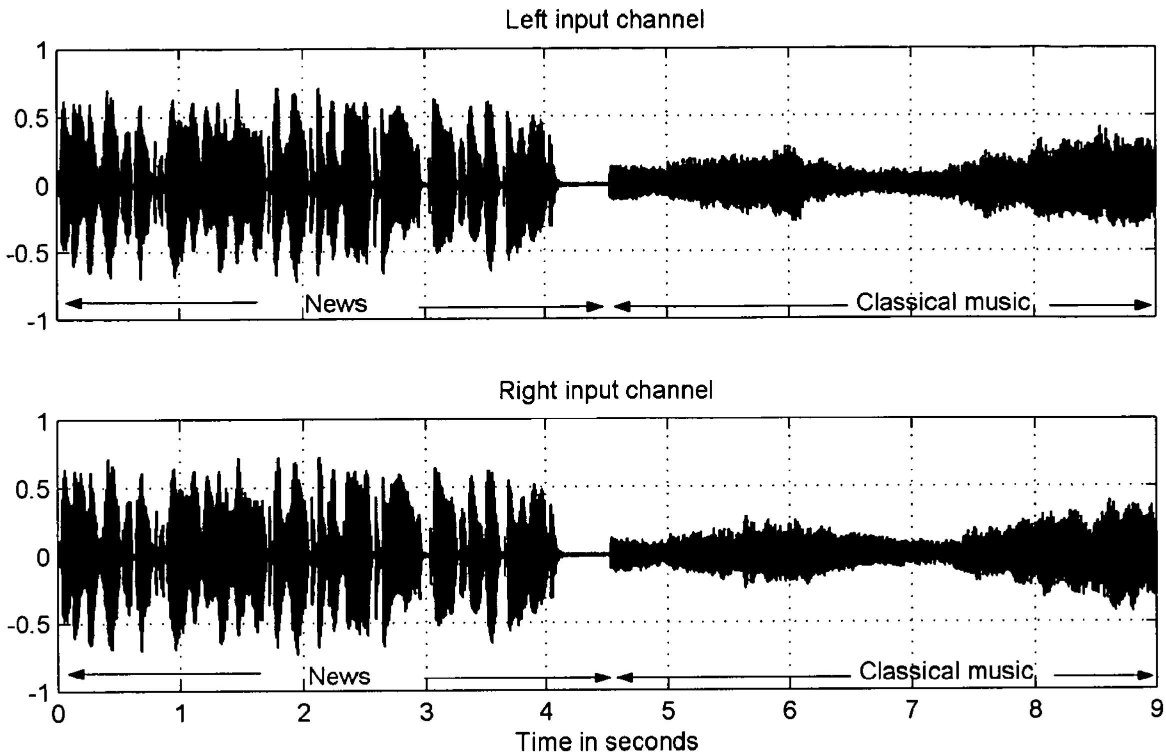

Left input channel
1
0.5
0
-0.5
-1
News
Classical music
Right input channel
1
0.5
0
-0.5
-1
News
Classical music
0
1
2
3
4
5
6
7
8
9
Time in seconds

Fig. 2

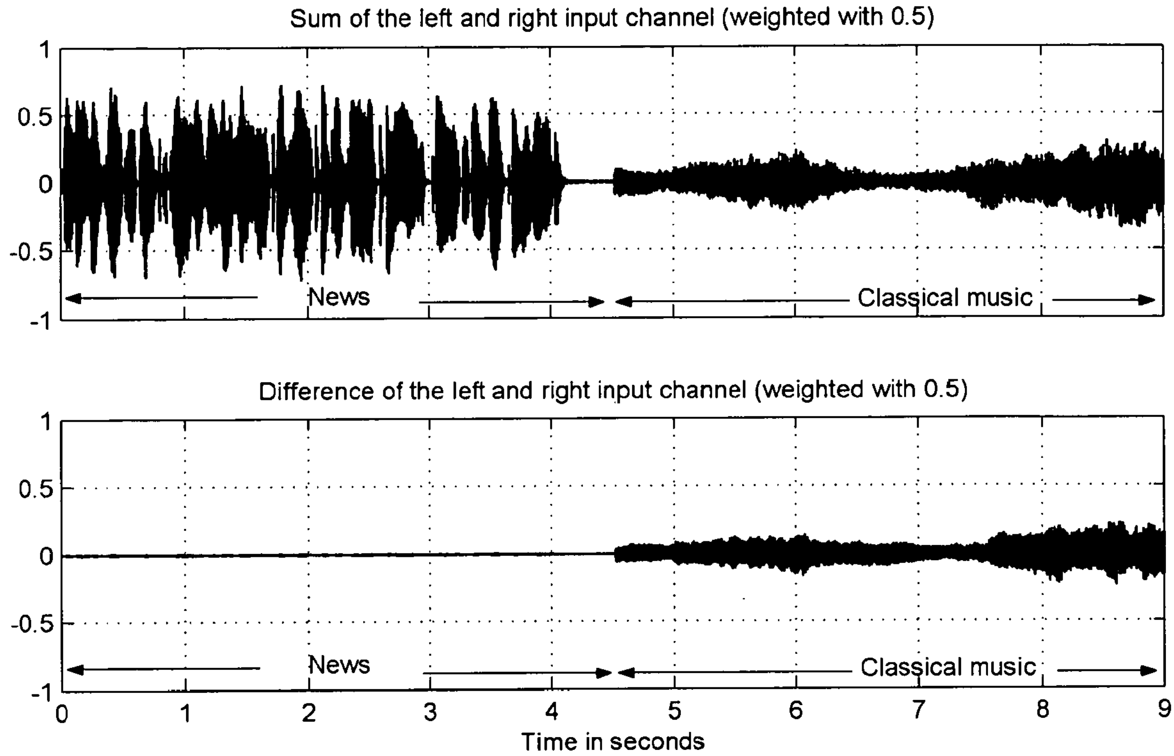

Sum of the left and right input channel (weighted with 0.5)
1
0.5
0
-0.5
-1
News
Classical music
Difference of the left and right input channel (weighted with 0.5)
1
0.5
0
-0.5
-1
News
Classical music
0
1
2
3
4
5
6
7
8
9
Time in seconds

Fig. 3

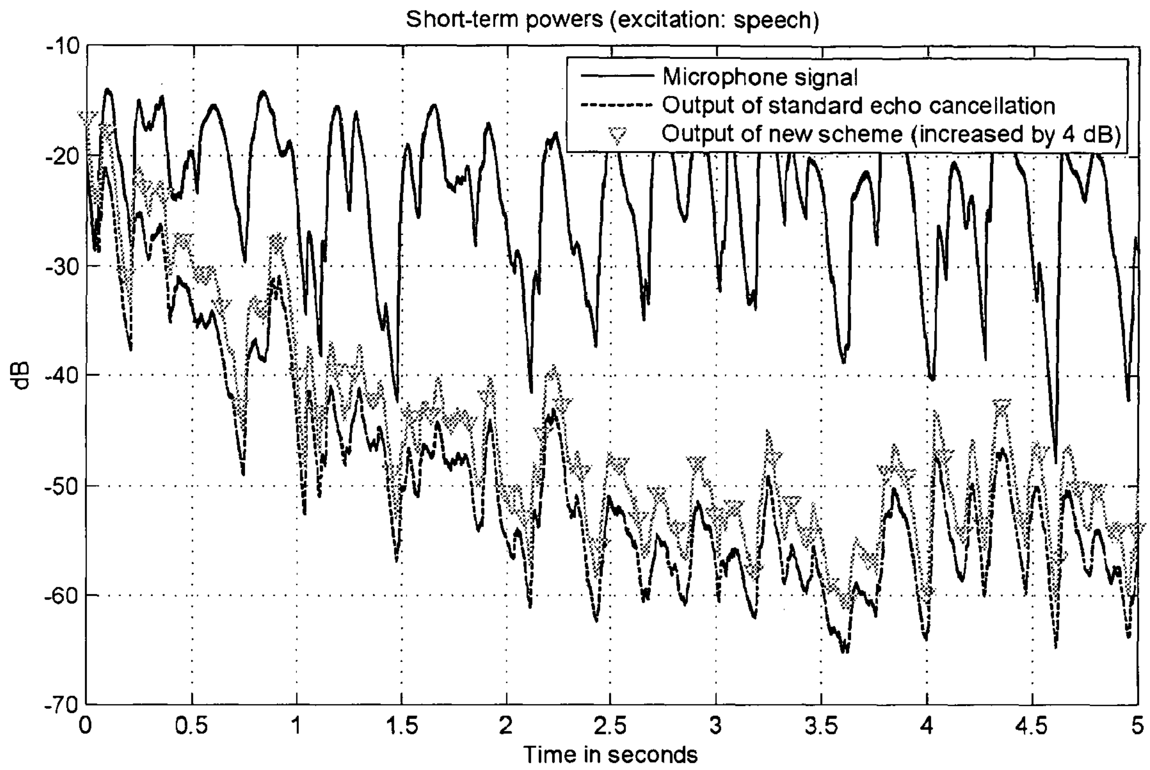
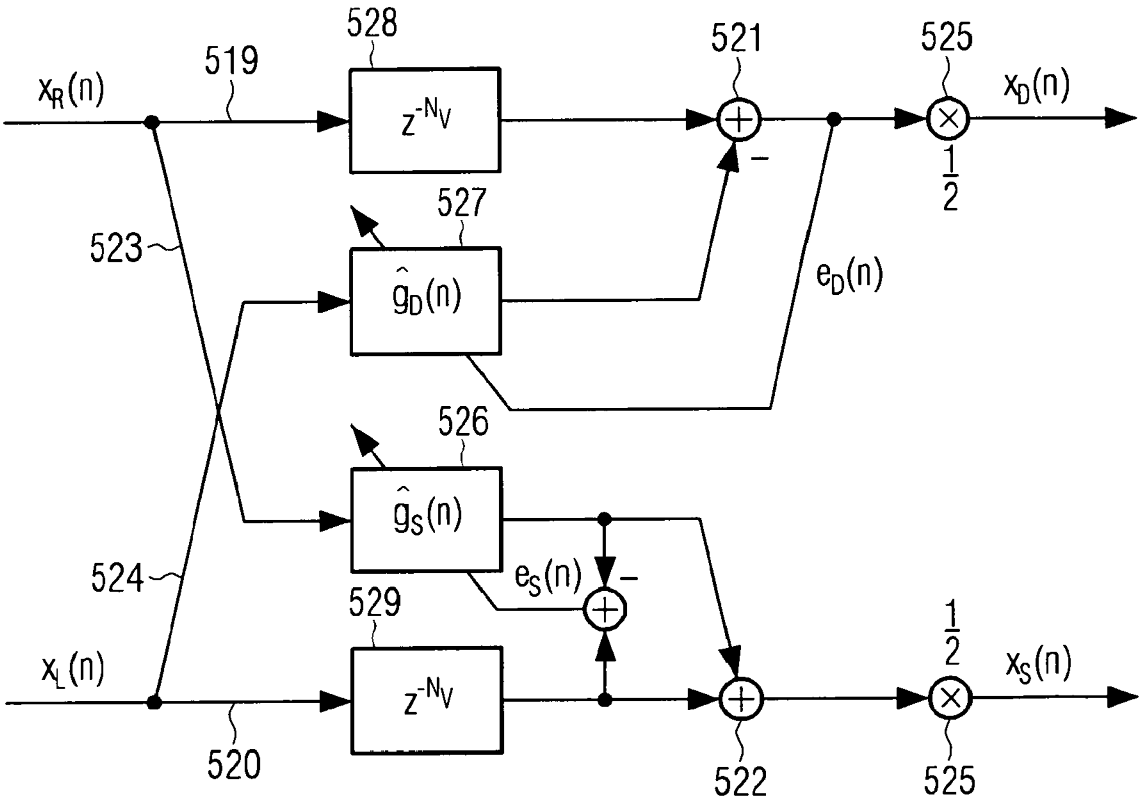

Short-term powers (excitation: speech)
Microphone signal
Output of standard echo cancellation
Output of new scheme (increased by 4 dB)
dB
-10
-20
-30
-40
-50
-60
-70
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time in seconds

Fig. 4



$x_R(n)$
519
528
z^{-N_V}
521
525
$x_D(n)$
−
$\frac{1}{2}$
527
523
$\hat{g}_D(n)$
$e_D(n)$
526
$\hat{g}_S(n)$
524
$e_S(n)$
−
529
$\frac{1}{2}$
$x_S(n)$
$x_L(n)$
z^{-N_V}
520
522
525

FIG. 5

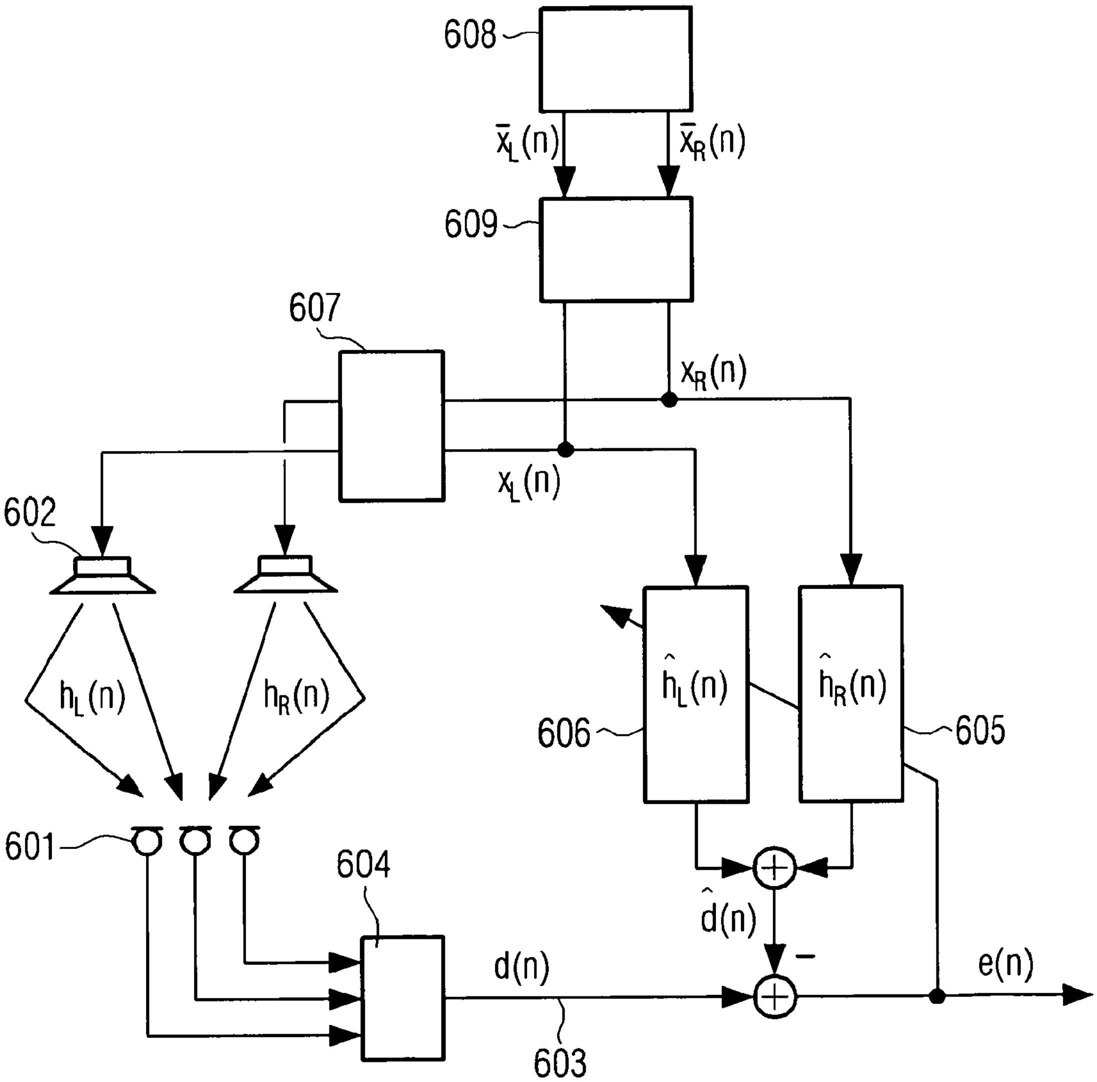


FIG. 6
(prior art)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 06 00 8006

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	US 6 738 480 B1 (BERTHAULT FREDERIC [FR] ET AL) 18 May 2004 (2004-05-18) * figures 1-3 * * column 1, line 7 - line 55 * * column 4, line 4 - line 14 * * column 4, line 24 - line 28 * * column 4, line 52 - column 5, line 6 * * column 9, line 14 - line 60 * * column 10, line 24 - line 30 * -----	1-31	INV. H04R3/02 H04M9/08
A	JACOB BENESTY ET AL: "A Hybrid Mono/Stereo Acoustic Echo Canceler" September 1998 (1998-09), IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US , XP011054323 ISSN: 1063-6676 * the whole document * -----	1-31	
A	US 2005/213747 A1 (POPOVICH STEVEN [US] ET AL) 29 September 2005 (2005-09-29) * the whole document * -----	1-31	TECHNICAL FIELDS SEARCHED (IPC)
A	EP 1 406 397 A (YAMAHA CORP [JP]) 7 April 2004 (2004-04-07) * the whole document * -----	1-31	H04M H04R H04B

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	13 November 2006	Moscu, Viorel

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

1

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 06 00 8006

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

13-11-2006

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6738480 B1	18-05-2004	CA 2307999 A1	12-11-2000
		EP 1052838 A1	15-11-2000
		FR 2793629 A1	17-11-2000
US 2005213747 A1	29-09-2005	NONE	
EP 1406397 A	07-04-2004	CN 1541456 A	27-10-2004
		WO 03007500 A1	23-01-2003
		JP 3506138 B2	15-03-2004
		JP 2003102085 A	04-04-2003
		US 2004174991 A1	09-09-2004

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **J. BENESTY ; T. GÄNSLER ; D. R. MORGAN ; M. M. SONDHI ; S. L. GAY.** Advances in Network and Acoustic Echo Cancellation. Springer Verlag, 2001 **[0008]**
- **A. SUGIYAMA ; Y. JONCOUR ; A. HIRANO.** A Stereo Echo Canceller with Correct Echo-Path Identification Based on an Input-Sliding Technique. *IEEE Transactions on Signal Processing,* 2001, vol. 49 (1), 2577-2587 **[0008]**