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

(11) EP 2 490 218 A1

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
22.08.2012 Bulletin 2012/34

(51) Int Cl.:
G10L 21/02 *(2006.01)* ***H04M 9/08*** *(2006.01)*

(21) Application number: **11155047.1**

(22) Date of filing: **18.02.2011**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(71) Applicant: **Svox AG**
8048 Zürich (CH)

(72) Inventors:
- **Wolf, Arthur**
  89073, Ulm (DE)
- **Iser, Bernd**
  89077, Ulm (DE)
- **Hannon, Patrick**
  89077, Ulm (DE)
- **Krini, Mohamed**
  89075, Ulm (DE)

(74) Representative: **Stocker, Kurt**
Büchel, von Révy & Partner
Zedernpark
Bronschhoferstrasse 31
9500 Wil (CH)

(54) **Method for interference suppression**

(57) In a method for interference suppression for a communication system as an indoor communication system or a handsfree telephony system or an automatic speech recognition system, at least one loudspeaker and at least one microphone are provided, particularly in a vehicle. A test signal is sent for determining system characteristics on the base of the test signal received by a microphone. The interference may be an ICC-system feedback component or the feedback signal of the audio component. The system characteristics detected by the test signal are used to determine interference model parameters in the form of frequency dependent coupling factors and frequency dependent decay factors and frequency dependent delay factors. Then, an estimated interference signal is formed on the base of the sub-band coupling factor times the magnitude square of a sub-band delayed loudspeaker signal plus a last interference signal level times a sub-band decay. The estimated interference signal is used to suppress the interference signal accordingly.

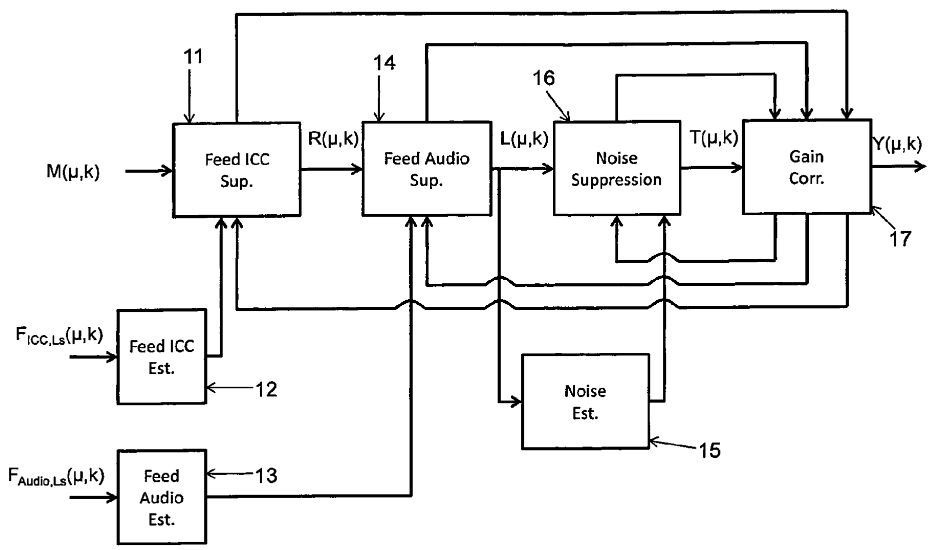

Fig. 3b

EP 2 490 218 A1

Description

Background of the invention

[0001] In limousines and vans communication between passengers in the front and in the rear may be difficult - especially if the car is driven at medium or high speed, resulting in a large background noise level. Furthermore, driver and front passenger speak towards the windshield. Thus, they are hardly intelligible for those sitting behind them. To improve the speech intelligibility the passengers start speaking louder and lean or turn towards their communication partners. For longer conversations this is usually tiring and uncomfortable. A way to improve the speech intelligibility within a passenger compartment is to use an incar communication system, often shortly called ICC. These systems record the speech of the speaking passengers by means of microphones and improve the communication by playing the recorded signals via those loudspeakers located close to the listening passengers. Examples for such ICCs can be found in E. Lleida, E. Masgrau, A. Ortega: Acoustic echo and noise reduction for car cabin communication, Proc. EUROSPEECH '01, 3, 1585-1588, Aalborg, Denmark, 2001 or in T. Haulick, G. Schmidt, Signal processing for in-car communication systems, Signal Processing, page 1307 - 1326, Juni 2006. However, problems with intelligibility are similar with handsfree telephony systems or with automatic speech recognition systems.

[0002] Indoor or incar communication systems, to refer only to one of above mentioned systems, operate in a closed electro-acoustic loop, and the speech signal is disturbed by background noise, and by interfering signals, such as by audio playback and ICC-system feedback. The microphone in the respective system picks up at least a portion of the loudspeaker signal. If this portion is not sufficiently small, sustained oscillations appear - which can be heard as howling or whistling. Cancellation of such ICC-system feedback turns out to be extremely difficult, since the adaptation of the filter is disturbed by the strong correlation between the feedback signal and the local speech signal. Thus, the above article of T. Haulick and G. Schmidt mentions that feedback suppression with the methods of echo cancellation is rather difficult, and similarly E. Hänsler, G. Schmidt: Topics in acoustic echo and noise control, Springer, page 549-598, 2006 comes to the same conclusion.

[0003] In addition to feedback signal, an indoor or car sound player can also be active while the ICC is working, and is also coupling into the microphone. Typical audio systems support stereo or even multi channel (e.g. dolby digital) audio playback. Due to the strong correlation between the single audio channels, the echo cancellation of multi channel audio output signal is also very challenging.

[0004] In addition, as the microphone also picks up the background noise inside a driving car, the noise components should be attenuated by noise suppression, as has already been mentioned in the above article of E. Lleida, E. Masgrau and A. Ortega.

[0005] As has already been mentioned above, other speech signal processing systems are faced with very similar problems. The microphone of a handsfree system also picks up the background noise and the played back audio signal. These interferences should be suppressed, before transmitting the signal to the remote subscriber. The microphone of an automatic speech recognition system also picks up the background noise and the played back audio signal. Noise and these interferences should be suppressed by adequate signal processing to improve the speech recognition results.

[0006] An attempt to realize feedback suppression has been disclosed in EP1718103B1. This solution works with an echo compensator filter which, as the present inventors found out, does not lead to satisfying results. Since background noise, and interferences, such as music signal and system feedback, should to be suppressed in a disturbed microphone signal, it is difficult due to these signal properties to calculate adaptive filters for cancelling interferences, such as a music and feedback signal. However, suppression of interferences has an influence to the signal output level of the processed speech and the residual interferences.

Summary of the invention

[0007] Therefore, it is an object of the present invention, starting from known features according to the introductory clause of claim 1, to improve the intelligibility in the above mentioned systems and, above all, to suppress interferences in a satisfying manner. This object is achieved by the features of the characterizing clause of claim 1. The method according to the invention ensures that the suppression of interferences remains constant or varies only very slowly in order to not annoy the passengers by fluctuating background noise and speech signal level.

[0008] An indoor communication system or a handsfree telephony system or an automatic speech recognition system, comprising at least one loudspeaker, at least one microphone and a signal processing system, particularly in a vehicle, wherein the microphone is recording a signal comprising communication information and interferences, the signal processing system is processing the microphone signal and providing a loudspeaker signal and the loudspeaker is emitting a sound signal corresponding to the loudspeaker signal. The method for interference suppression in the communication system includes the step of the signal processing system estimating an interference signal by an energy decay model with frequency dependent coupling factors, frequency dependent decay factors and frequency dependent delay factors, wherein an estimated interference signal includes at least one product of the respective coupling factor times a respective part of a loudspeaker signal delayed by the respective delay factor plus at least one product of an estimated interference signal at an earlier moment times the respective decay factor. The estimated interference signal is used for generating an interference suppressed loudspeaker signal.

[0009] There are different interferences for example feedback of the communication system, feedback of the audio system and background noise. For at least one interference signal an estimation is made.

[0010] The signal quality can be enhanced by applying stepwise independently or combined at least two suppressing steps out of communication system feedback suppression, audio system feedback suppression and noise suppression, wherein for all the suppression steps corresponding estimations of interference signals are used, and the suppression modules can be arranged in any order, but are preferably arranged in the following order: communication system feedback, audio system feedback and noise suppression.

[0011] In a preferred embodiment the estimated interference signal is calculated according to the formula

$$S_{ff,Mic}(\mu,k) = P_{LsMic}(\mu)\ S_{ff,Ls}(\mu,k-D(\mu)) \ + \ S_{ff,Mic}(\mu,k-1)\ e^{-\varphi(\mu)},$$

wherein $S_{ff,Mic}(\mu,k)$ is the estimated interference signal for example the feedback or audio level at the microphone Mic, for a predetermined frequency sub-band μ at time k; $P_{LsMic}(\mu)$ is the coupling factor of the sub-band μ between the loudspeaker Ls and the microphone Mic; $S_{ff,Ls}(\mu,k-D(\mu))$ is the interference signal for example the feedback or audio signal level at the loudspeaker Ls for a predetermined frequency sub-band at time k-D(μ), D(μ) is the delay factor of the sub-band μ, $S_{ff,Mic}(\mu,k-1)$ is the interference signal level for example the feedback signal level or the audio signal level at the microphone Mic for a predetermined frequency sub-band at time k-1 and $e^{-\varphi(\mu)}$ is an exponential decay factor of the sub-band μ.

[0012] In preferred embodiments the estimated interference signal is a communication system feedback signal or an audio-system feedback, wherein the audio signal is a part or all of the loudspeaker signal at loudspeaker Ls. If there is more than one communication system or a system with multiple communication directions then the interference of a further communication system or a further communication direction can be treated like an additional interference and be suppressed in the above described manner.

[0013] A test signal can be used to determine interference model parameters. A test signal is sent to the respective loudspeaker for determining system characteristics on the base of the test signal received by the respective microphone, wherein the system characteristics detected by the application of the test signal are used to determine the frequency dependent coupling factors, the frequency dependent decay factors and frequency dependent delay factors. In some applications, it may be sufficient to do it once. Considering, particularly for a car, that persons may leave the room, may open a window or the like, the model parameters may vary over time.

[0014] The interference model parameters can be deduced from at least one microphone signal or by automatically detecting and interpreting decaying signal slopes, wherein after deducting the parameters to be applied are updated and the parameter deduction occurs preferably when there is no local speech.

[0015] The interference model parameters can be deduced from the coefficients of an echo compensator preferably by placing an echo compensator parallel to the signal processing path, wherein the coefficients of the echo compensator are corresponding to the room impulse response and the parameters of the echo compensator should only be updated when there is no local speech detected, e.g. only at the decaying slopes of the microphone signal or only at the feedback of the audio signal.

[0016] In a preferred embodiment an interference suppressed loudspeaker signal is generated by Spectral Subtraction, preferably by the application of a Wiener-filter using the estimated interference signal.

[0017] An overestimation $\gamma(\mu,k)$ of the estimated interference signal of the sub-band μ can be adjusted to an estimated noise to speech signal ratio SNR(μ,k) of the sub-band μ deduced by a noise suppression module. A maximum attenuation $\beta(\mu,k)$ of the sub-band μ is adjusted to a gain correction factor V(μ,k) of the sub-band μ.

[0018] A communication system equalizer $H_{Eq}(\mu,k)$ can be adjusted according to the energy decay model frequency dependent coupling factors $P_{LsMic}(\mu,k)$.

[0019] A communication system gain can be adjusted according to the estimated interference signal level, wherein this gain is dependent on the background noise level and on the level of the communication system feedback and audio system feedback, for a high communication system feedback or audio system feedback signal level the system gain should be reduced.

[0020] Furthermore, the invention relates to a software product according to claim 13 and 14 and to a system according to claim 15.

Brief description of the drawings

[0021] The invention will be better understood by the following detailed description with reference to the following drawings in which:

Fig. 1 sketches the structure of a simple ICC system aimed to support front-to-rear conversations with one microphone and one loudspeaker;

Fig. 2 is an overview of the system;

Fig. 3a shows the sub-band energy decay curve from the test impulse response which may be used for estimating parameters for this sub-band energy decay model;

Fig. 3b shows a block diagram of a circuit for carrying out a preferred embodiment of the method according to the invention;

Fig. 4a illustrates how an echo compensator can be used for decay model parameter estimation; and

Fig. 4b is another block diagram of the ICC-system together with noise dependent gain control (NDGC) and equalizer (Eq).

Detailed description of the drawings

[0022] In a room 1, such as a car cabin, there are a driver 2a and a passenger 2b behind. In front of the driver 2a is a microphone 3 so that the driver's speech is better intelligible for the passenger 2b to whom a loudspeaker 4 is assigned. This is a simplified example, because, of course, a (further) microphone might be near the passenger 2b to be better understood by the driver 2a to whom a loudspeaker of the type of loudspeaker 4 may also be assigned.

[0023] Clearly, the microphone 3 will not only take the speech signal s(n) of the driver 2a, but also the noise b(n), and noise suppression in the line from the microphone 3 to the loudspeaker 4 is known per se. However, in addition, there are interferences received by the microphone, such as the audio signal $f_{Audio,Mic}(n)$ from a further loudspeaker 5, by means of which the driver wants, for example, to become informed about street conditions or obtains a navigation aid from an audio-source 6. Thus, the output m(n) of the microphone 3, which is composed of the voice signal s(n) of the driver 2a, the noise b(n), the audio-signal $f_{Audio,Mic}(n)$ and the ICC-system feedback signal $f_{ICC,Mic}(n)$ coming from loudspeaker 4, the two latter signals forming interferences. If one feeds at least one of the interferences, such as $f_{Audio,Ls}(n)$, into the ICC-system 7 for suppression, speech of the driver 2a is better intelligible, because the signal to the loudspeaker 4 is enhanced (vide also Fig. 2). For this invention the signals in time domain with the time index n are defined as lower case characters and signals in sub-band domain with the sub-band index μ and the frame index k are defined as upper case characters. In which for signals played back inside the acoustic room, e.g. vehicle interior, the index Mic is used for the signal at the microphone and the index Ls is used for the signal at the loudspeaker. The index Mic has a different value for each microphone, where it runs for example from 0 to the maximum number of microphones minus 1. The index Ls has a different value for each loudspeaker, where it runs for example from 0 to the maximum number of loudspeakers minus 1 For this invention the upper case character S with an index, i.e. $S_{bb}(\mu,k)$ is used for power signals, while other capital letters, e.g. $B(\mu,k)$, are used for complex sub-band signals. The power signal can be approximated as the square of the sub-band signal, e.g. $S_{bb}(\mu,k) = |B(\mu,k|.^2$

[0024] Now in detail, the sub-band microphone signal $M(\mu,k)$ of a given sub-band μ at a given time k consists of the local speech signal $S(\mu,k)$, the background noise $B(\mu,k)$, the feedback of the ICC-system output $F_{ICC,Mic}(\mu,k)$ and the feedback of the audio system $F_{Audio,Mic}(\mu,k)$, in which μ is the sub-band index and k is the time frame index. The invention provides a method of interference suppression using a mathematical model on the base of sound energy decay inside a room. In addition to the energy decay there is also a delay effect.

[0025] The signal processing system is estimating an interference signal by an energy decay model with frequency dependent coupling factors, frequency dependent decay factors and frequency dependent delay factors, wherein an estimated interference signal includes at least one product of the respective coupling factor times a respective part of a loudspeaker signal delayed by the respective delay factor plus at least one product of an estimated interference signal at an earlier moment times the respective decay factor, and that the estimated interference signal is used for generating an interference suppressed loudspeaker signal. Frequency dependent coupling factors, frequency dependent decay factors and frequency dependent delay factors can be calculated preferably from a room impulse response.

[0026] The estimated interference signal is a system feedback signal, which can be calculated according to the formula

$$S_{ff,Mic}(\mu,k) = P_{Ls,Mic}(\mu)\ S_{ff,Ls}(\mu,k-D(\mu)) + S_{ff,Mic}(\mu,k-1)\ e^{-\varphi(\mu)},$$

wherein $S_{ff,Mic}$,(μ,k) is the estimated feedback level at the microphone Mic, for a predetermined frequency sub-band μ at time k; $P_{Ls,Mic}(\mu)$ is the coupling factor of the sub-band μ between the loudspeaker Ls and the microphone Mic; $S_{ff,Ls}$ $(\mu,k-D(\mu))$ is the feedback signal level at the loudspeaker Ls for a predetermined frequency sub-band at time k-D(μ), D (μ) is the delay factor of the sub-band μ, $S_{ff,Mic}$,$(\mu,k-1)$ is the feedback signal level at the microphone Mic for a predetermined frequency sub-band at time k-1 and $e^{-\varphi(\mu)}$ is an exponential decay factor of the sub-band μ.

[0027] The interference model parameters are deduced form at least one microphone signal or from automatically detecting and interpreting decaying signal slopes, wherein after deducting the parameters to be applied are updated and the parameter deduction occurs preferably when there is no speech, no audio signal or no ICC-system output.

[0028] The coefficients of an echo compensator can be used for the adaption of the interference model parameters preferably by placing an echo compensator parallel to the signal processing path, wherein the coefficients of the echo compensator are corresponding to the room impulse response and the parameters of the echo compensator should only be updated when there is no local speech detected, e.g. only at the decaying slopes of the microphone signal or only at the feedback of the audio signal.

[0029] By preparing at least two different interference models on the basis of different parameters for different occupancy or different environment conditions and by detecting the actual occupancy of the vehicle or environment condition it is possible to select the interference model in accordance with the actual occupancy or environment condition detected.

[0030] Noise components, as is known, can be suppressed by a Wiener-filter in sub-band domain. Signal processing can be applied in sub-band or also in a melband domain to take the psychoacoustics into the account or to reduce the algorithmic complexity. The difficulty is the estimation of the ICC-system feedback and the audio signal feedback at the microphone. However, the ICC-system feedback and the audio signal feedback are known at the loudspeaker or can be supplied as a reference channel from the output of the ICC and the audio system (Fig. 2). The ICC-system feedback and the audio signal feedback at the microphone 3 can be calculated now as a convolution of the feedback signal at the loudspeaker 4 and the room impulse response between the loudspeaker 4 and the microphone $h_{LsMic}(n)$:

$$f_{ICC,Mic}(n)=f_{ICC,Ls}(n)*h_{LsMic}(n),$$

and

$$f_{Audio,Mic}(n)=f_{Audio,Ls}(n)*h_{LsMic}(n).$$

[0031] Because it is very difficult to update or estimate the loudspeaker-microphone impulse response during operating time the method according to the invention uses a model for the energy decay of the room impulse response. The energy decay of the room impulse response is modelled as the outcome of a non-stationary random process.

$$E\{h_{LsMic}^2(n)\} = 0 \qquad \text{for } n< 0$$

$$E\{h_{LsMic}^2(n)\} = \sigma^2 e^{-gn} \qquad \text{for } n\geq 0.$$

[0032] In which the energy decay is modeled as

$$g=6*\ln 10/(T60*f_S),$$

with the reverberation time T_{60}, the sampling frequency f_s. and the scaling factor for the signal energy σ^2.

[0033] Similar to the time domain description, a constant decay of the energy in the sub-band domain is assumed as

$$G_{LsMic}(\mu,k) = 0 \qquad \text{for } k< D(\mu)$$

$$G_{LsMic}(\mu,k) = P_{LsMic}(\mu)e^{-\varphi(\mu,k-D)} \quad \text{for } k\geq D(\mu).$$

[0034] In which $P_{LsMic}(\mu)$ is the coupling factor of the sub-band μ between the loudspeaker and the microphone and the energy decay $\varphi(\mu)$ is modelled as

$$\varphi(\mu) = 6*\ln10*N/(T60(\mu)*f_S),$$

with the sub-band reverberation time $T_{60}(\mu)$, the sampling frequency f_s, the frame shift N, the frame index k and the delay parameter $D(\mu)$.

[0035] The parameters for this sub-band energy decay model $G_{LsMic}(\mu,k)$ can be estimated from the sub-band energy decay curve from the impulse response as shown in Fig. 3a. Dependent on the use case, there are maybe different loudspeakers used for playing back the audio and the ICC system output. Therefore different model parameters from different impulse responses are used for ICC-system and audio system feedback suppression.

[0036] Very similar energy decay models can be used e.g. for dereverberation of a microphone signal. This has been disclosed in US2009/0117948 and also by E. Habets in "Multichannel speech dereverberation based on a statistical model of late reverberation," in ICASSP, 2005. In the case of the present invention, the model is used for estimation of the ICC-system output and the audio signal at the microphone, i.e. for interference estimation. In the sub-band domain the estimation of the feedback signal at the loudspeaker 4 can be estimated with the recursive notation:

$$S_{ff,ICC,Mic}(\mu,k) = P_{LsMic}(\mu)\ |F_{ICC,Ls}(\mu,k-D(\mu))|^2 + S_{ff,ICC,Mic}(\mu,k-1)\ e^{-\varphi(\mu)},$$

$$S_{ff,Audio,Mic}(\mu,k) = P_{LsMic}(\mu)\ |F_{Audio,Ls}(\mu,k-D(\mu))|^2 + S_{ff,Audio,Mic}(\mu,k-1)\ e^{-\varphi(\mu)}.$$

[0037] Where $S_{ff,ICC,Mic}(\mu,k)$ is the estimated feedback of the ICC-system output at the microphone and $S_{ff,Audio,Mic}(\mu,k)$ is the estimated feedback of the audio signal at the microphone. It should be noted that the sign S with an index, i.e. $S_{xx}(\mu,k)$ is used in this specification for power, while other capital letters, such as $M(\mu,k)$ or $B(\mu,k)$, are used for complex sub-band signals. The power signal can be approximated as the square of the sub-band signal, e.g. $S_{bb}(\mu,k) = |B(\mu,k|^2$.

[0038] After the estimation of the interfering signals these components can be suppressed. Therefore Spectral Subtraction, e.g. Wiener-filter in the sub-band domain can be used. Here the attenuation of the filter coefficients is constrained to a maximum attenuation (spectral floor) $\beta(\mu)$.

$$0 \leq \beta(\mu) \leq 1, \text{ conventionally } 0.1 \leq \beta(\mu) \leq 0.3.$$

[0039] For reduction of the artefacts caused by the interference suppression, also called musical noise, an overestimation factor $\gamma(\mu)$ will be used. Conventionally the overestimation factor $\gamma(\mu)$ is a fixed value, with e.g. $1 \leq \gamma(\mu)$ 3. Because these artefacts are masked by the residual noise primarily caused by the noise suppression the improved solution contains a $SNR(k,\mu)$ (signal to noise ration) dependent overestimation factor $\gamma(k, \mu)$. Where the $SNR(k,\mu)$ is defined as

$$SNR(k,\mu) = |T(\mu,k)|^2/\ S_{bb}(\mu,k).$$

[0040] Where $T(\mu,k)$ is the feedback and noise suppressed signal and approximates the clean local speech signal

$S_{ss}(\mu,k)$ power and $S_{bb}(\mu,k)$ is the estimated nose signal power.

[0041] This adaptation of the SNR(k,μ) dependent overestimation factor $\gamma(k,\mu)$ depends on a characteristic which maps the SNR(k,μ) to the overestimation factor $\gamma(k,\mu)$. Some sample characteristics are depicted in Fig. 5. Common parameters are $\gamma_{min}(\mu) = 1$; $\gamma_{max}(\mu) = 3$; $SNR_{min}(\mu) = 5$ dB; $SNR_{max}(\mu) = 15$ dB.

$$\gamma(k,\mu) = \gamma_{min}(\mu) \qquad \text{for } SNR(k,\mu) \leq SNR_{min}(\mu)$$

$$\gamma(k,\mu) = \gamma_{max}(\mu) \qquad \text{for } SNR(k,\mu) \geq SNR_{max}(\mu)$$

$$\gamma(k,\mu) = SNR(k,\mu) \cdot (\gamma_{max}(\mu) - \gamma_{min}(\mu)) / (SNR_{max}(\mu) - SNR_{min}(\mu)) \qquad \text{else}$$

[0042] The overestimation factor $\gamma(k,\mu)$ can also be defined and determined for every processing step as $\gamma_{F,ICC}(k,\mu)$, $\gamma_{F,Adio}(k,\mu)$ and $\gamma_B(k,\mu)$.

[0043] It is beneficial to suppress the feedback and the audio signal before estimating the noise power and suppress the background noise. Due to the assumption that the power of the single signal components is uncorrelated, the power of the microphone signal can be defined as:

$$S_{mm}(\mu,k) = S_{ss}(\mu,k) + S_{ff,ICC,Mic}(\mu,k) + S_{ff,Audio,Mic}(\mu,k) + S_{bb}(\mu,k)$$

[0044] Where all components are known or can be estimated as shown before, beside the power of the clean speech signal $S_{ss}(\mu,k)$ which is unknown.

[0045] In the illustrated embodiment, it is started with the ICC-system feedback suppression, but other arrangements are also possible. The coefficients of the feedback suppression filter can be calculated, e.g.

$$H_{F,ICC}(\mu,k) = \min(\beta_{F,ICC},\ 1 - \gamma_{F,ICC}(k,\mu) \cdot S_{ff,ICC,Mic}(\mu,k) / |M(\mu,k)|^2).$$

[0046] This filter can be applied to the disturbed microphone signal.

$$R(\mu,k) = M(\mu,k)\ H_{F,ICC}(\mu,k).$$

to obtain the ICC-system feedback suppressed signal $R(\mu,k)$.

[0047] Now this ICC-system feedback suppressed signal $R(\mu,k)$ can advantageously be used for suppressing the audio system feedback.

$$H_{F,Audio}(\mu,k) = \min(\beta_{F,Audio},\ 1 - \gamma_{F,Audio}(k,\mu) \cdot S_{ff,Audio,Mic}(\mu,k) / |R(\mu,k)|^2).$$

[0048] This filter can be applied to the processed microphone signal $R(\mu,k)$

$$L(\mu,k) = R(\mu,k)\ H_{F,Audio}(\mu,k)$$

to obtain the ICC-system and audio system feedback suppressed signal $L(\mu,k)$.

[0049] Now the ICC-system and audio system feedback suppressed signal $L(\mu,k)$ can be used for noise signal suppression. The power of the background noise $S_{bb}(\mu,k)$ is a stationary process and can be estimated from noisy speech signal, e.g. in speech pauses as:

$$S_{bb}(\mu,k) = |L(\mu,k)|^2 \text{ while speech pause,}$$

$$S_{bb}(\mu,k) = S_{bb}(\mu,k-1) \text{ while speech activity.}$$

[0050] With the estimated noise power the filter coefficients of the noise suppression filter can be calculated as:

$$H_B(\mu,k) = \min(\beta_B,\ 1 - \gamma_B(k,\mu) \cdot S_{bb}(\mu,k) \,/\, |L(\mu,k)|^2).$$

[0051] This filter can be applied to the noisy already ICC-system and audio system feedback suppressed signal $L(\mu,k)$

$$T(\mu,k) = L(\mu,k)\ H_B(\mu,k).$$

to obtain the feedback and noise suppressed signal $T(\mu,k)$.

[0052] The suppression of these interferences has an influence to the signal output level of the processed speech and the residual interferences. Therefore the signal output level needs to be adjusted due to the signal suppression to keep the long term output and residual interference level constant. This amplification factor can be calculated for every sub-band $V(\mu,k)$ or as scalar fullband parameter $v(k)$. One possible implementation for the update of the amplification factor is to calculate the update terms for every filter.

$$v_{ICC}(k) = \alpha\ v_{ICC}(k-1)\ + (1 - \alpha)\ N_{sub} \,/ \sum_{l=0}^{Nsub-1} H_{ICC}(l,k)$$

[0053] Where the mean value of the filter coefficients is used for the update of the with smoothing parameter α smoothed amplification correction factor. Therefore $0 \le \alpha \le 0.1$ and Nsub is the number of the sub-bands. The parameter $V_{Audio}(k)$ and $V_B(k)$ are calculated the same way.

[0054] Now the gain correction factors are combined to get the final gain correction factor

$$v(k) = (v_{ICC}(k) \cdot v_{Audio}(k) \cdot v_B(k)).$$

[0055] The calculated amplification factor can be applied to the processed signal to correct the long term power level difference caused by the signal processing

$$Y(\mu,k) = v(k)\ T(\mu,k).$$

[0056] Where $\gamma(\mu,k)$ is the output of the signal enhancement.

[0057] Certainly the amplification of the output signal changes the level of the residual interferences of the processed signal. To correct the power level of the residual interference signal the amplification factor can be used for adjusting the spectral floor of the filter calculation:

$$\beta_{ICC}(\mu,k) = \beta_{ICC,Start}(\mu) / v_{ICC}(k).$$

[0058] Where $\beta_{start}(\mu)$ is the initial value for the spectral floor. The parameter $\beta_{Audio}(\mu,k)$ and $\beta_B(\mu,k)$ are updated the same way.

[0059] The described method enables to enhance the interfered signal by a very robust and efficient way with a circuit schematically shown in Fig. 3b. The configuration shown in Fig. 3b depends on the actual system setup. Feedback suppression, audio suppression and noise suppression is applied stepwise, where for all these suppression steps corresponding estimations of interference signals are used. In Fig. 3b the suppression modules are arranged in the following order: Feedback, Audio and Noise. Rearrangements of the used modules are possible and may in some cases be necessary. It is possible to perform every processing step independently like shown before. There the modules can also be rearranged and/or combined.

[0060] Combination and application of the filter coefficients of different modules can be described as follows

$$H(\mu,k) = \text{minimum}(H_{F,ICC}(\mu,k),\ H_{F,Audio}(\mu,k),\ H_B(\mu,k)\ ).$$

[0061] Where $H(\mu,k)$ is the combined interference suppression filter coefficients dependent on the single components ICC-system feedback suppression filter coefficients $H_{F,ICC}(\mu,k)$, audio system feedback suppression filter coefficients $H_{F,Audio}(\mu,k)$ and noise suppression filter coefficients $H_B(\mu,k)$.

[0062] In that case the enhanced signal can be calculated now as

$$T(\mu,k) = H(\mu,k)\ M(\mu,k).$$

[0063] According to Fig. 3b, the microphone signal $M(\mu,k)$ is transformed by a ICC-system feedback suppression step 11 to a feedback reduced signal $R(\mu,k)$. In order to ensure that the suppression in this step 11 works precisely, there is an ICC-system feedback estimation step 12 preposed. In the step 12 the ICC-system feedback signal at the loudspeaker $F_{ICC,Ls}(\mu,k)$ and determines the interference signal level $S_{ff,ICC,Mic}(\mu,k)$ by applying the room energy decay model parameters on the base of the sub-band coupling factor $P_{LsMic}(\mu)$ times the magnitude square of a sub-band delayed loudspeaker signal $F_{ICC,Ls}(\mu,k\text{-}D(\mu))$ plus a last interference signal level $S_{ff,ICC,Mic}(\mu,k\text{-}1)$ times a sub-band decay factor $e^{-\varphi(\mu)}$. The output of module 12, the estimated interference signal level $S_{ff,ICC,Mic}(\mu,k)$, is delivered to module 11 and is used there to suppress the interference signal accordingly.

[0064] The same applies to the feedback of the audio system and feeding the estimated signal level $S_{ff,Audio,Mic}(\mu,k)$ of the stage 13 to an audio system feedback suppression stage 14.

[0065] The output of module 14 is now freed from feedback interference components and can, therefore, better be used for noise estimation in module 15, to feed a noise suppression stage 16. Since the enhanced signal has lost power, it is useful, to correct the signal level by a gain control module 17 which forms a power level corrected signal $\gamma(\mu,k)$ and is in connection with modules 11, 14 and 16. Therefore the gain control stage 17 analyzes the filter coefficients of the modules 11, 14, 16 and returns the for adjusted spectral floor factors back to the modules 11, 14, 16.

Update of the decay model parameter:

[0066] In general, from the three energy decay model parameter, the delay $D(\mu)$ and the energy decay $e^{-\varphi(\mu)}$ are related to the used hardware and the room characteristics e.g. the reverberation time T_{60}. The changes of these parameters are slow and small. The coupling factor $P_{LsMic}(\mu)$ depends on the actual position of the passengers inside the car and is changing faster. In the majority of cases it is sufficient only to adapt this parameter during signal processing.

[0067] As described before the room energy decay parameters can be estimated from the impulse response respectively the sub-band impulse response. This impulse response can be measured, before calculating the signal processing and the estimated model parameters $D(\mu)$, $P_{LsMic}(\mu)$ and $e^{-\varphi(\mu)}$, see also Fig 3a. With these parameters, signal processing can be applied.

[0068] Due to the changes of the impulse response, caused by changes of the car occupancy and environment conditions, e.g. open window or door, it is suitable to repeat the impulse response measurements for different car occupancies and environment conditions, e.g. to have different decay models for different occupancies and environment conditions. The occupancy or environment conditions of the car can be detected, e.g. with seat sensors or window sensor, and the signal processing can switch to the actual predefined decay model.

[0069] Another possibility to estimate the room energy decay model parameters, is the use of an echo compensator 18 (Fig. 4a) which is placed parallel to the signal processing path 7. The output of the echo compensator 18 is not used for feedback compensation, but only for updating the echo compensator. Due to the correspondence of the coefficients of the echo compensator to the room impulse response (as is known from EP-A-2151983), the estimated coefficients can be used in a very similar way to estimate or to update the decay model parameter during the signal processing. The parameters of the echo compensator should only be updated when there is no local speech detected, e.g. only at the decaying slopes of the microphone signal or only at the feedback of the audio signal.

[0070] A further possibility to estimate the room energy decay model parameters, is to use of frequency/phase shift methods or other decorrelation methods like nonlinearities at the system output or additional noise signal. The decay model parameter can be easily updated from this decorrelated loudspeaker signal. This method can be used together to the parallel echo compensator 18 to support and accelerate the adaptation of the echo compensator 18.

[0071] Still another possibility to estimate the room energy decay model parameters, is to automatically detect and interpret the decaying signal slopes. In this case, the energy decay of the slope needs to be monitored. The fastest decay appears when there is no additional excitation signal, e.g. no local speech, no audio signal or no ICC-system output. After detecting these moments the room energy decay model parameters can be updated by the estimated sub-band decay and the sub-band transfer at the beginning of the slope.

[0072] Another possibility is to update the decay model parameter from the calculated cross correlation between the loudspeaker signal $F_{ICC,Ls}(\mu,k)$ and the microphone signal $M(\mu,k)$ to estimate the room energy decay model parameters.

Additional use cases for this decay model:

[0073] The described model for the energy decay can also be used for adjusting the coefficients of the equalizer which can also be a part of the ICC system. Therefore the sub-band coupling parameter $P_{LsMic}(\mu,k)$ can be used to set up the sub-band equalizer 19 (vide Fig 4b) to improve the stability ICC-system gain in term of maximum ICC-system gain, due to the correlation between the room impulse response $h_{LsMic}(n)$ and the sub-band coupling parameter $P_{LsMic}(\mu,k)$. The sub-band attenuation of the equalizer $H_{Eq}(\mu,k)$ can be determined directly from the sub-band coupling parameter $P_{LsMic}(\mu,k)$

$$H_{Eq}(\mu,k) = 1 / P_{LsMic}(\mu,k).$$

[0074] The estimation of the interference components can also be used to set up the ICC system gain. Of course, this gain is dependent on the background noise level $S_{bb}(\mu,k)$. But it is also dependent on the level of the feedback and audio signal. Because for a high feedback the ICC system produces many artefacts the system gain should be reduced. For very high audio signal level, the passengers prefer to listen to the audio system. The audio level at the microphone $S_{ff,Audio,Mic}(\mu,k)$ can be estimated with the described method. This signal correlates to the sound level inside the car. In relation to the ratio between the estimated audio signal and the processed signal the system gain can be reduced or the system can be deactivated in order to not disturb the passengers, while listening music. For very high ICC-system feedback level $S_{ff,ICC,Mic}(\mu,k)$ at the microphone the processed signal contains many artefacts caused by the signal processing. For reduction of this artefacts the ICC-system gain should be reduced to reduce the level of the feedback signal $S_{ff,ICC,Mic}(\mu,k)$ or switch off the ICC-system while the ICC-system is working under inconvenient or not acceptable conditions.

[0075] The communication system gain can be adjusted according to the estimated interference signal level, wherein this gain is dependent on the background noise level and on the level of the communication system feedback and audio system signal, for a high communication system feedback or for a high audio signal level the system gain should be reduced.

Claims

1. Method for interference suppression for an communication system as an indoor communication system or a hands-free telephony system or an automatic speech recognition system, comprising at least one loudspeaker, at least one microphone and a signal processing system, particularly in a vehicle, wherein the microphone is recording a signal comprising communication information and interferences, the signal processing system is processing the microphone signal and providing a loudspeaker signal and the loudspeaker is emitting a sound signal corresponding to the loudspeaker signal, **characterized in that** the signal processing system is estimating an interference signal by an energy decay model with frequency dependent coupling factors, frequency dependent decay factors and frequency dependent delay factors, wherein an estimated interference signal includes at least one product of the respective coupling factor times a respective part of a loudspeaker signal delayed by the respective delay factor

plus at least one product of an estimated interference signal at an earlier moment times the respective decay factor, and that the estimated interference signal is used for generating an interference suppressed loudspeaker signal.

2. Method according to claim 1, **characterized in that** the estimated interference signal is calculated according to the formula

$$S_{ff,Mic}(\mu,k) = P_{LsMic}(\mu)\ S_{ff,Ls}(\mu,k\text{-}D(\mu)) + S_{ff,Mic}(\mu,k\text{-}1)\ e^{-\varphi(\mu)},$$

wherein $S_{ff,Mic}(\mu,k)$ is the estimated interference signal level at the microphone Mic, for a predetermined frequency sub-band μ at time k; $P_{LsMic}(\mu)$ is the coupling factor of the sub-band μ between the loudspeaker Ls and the microphone Mic; $S_{ff,Ls}(\mu,k\text{-}D(\mu))$ is the interference signal level at the loudspeaker Ls for a predetermined frequency sub-band at time k-$D(\mu)$, $D(\mu)$ is the delay factor of the sub-band μ, $S_{ff,Mic}(\mu,k\text{-}1)$ is the interference signal level at the microphone Mic for a predetermined frequency sub-band at time k-1 and $e^{-\varphi(\mu)}$ is an exponential decay factor of the sub-band μ.

3. Method according to claim 1 or 2, **characterized in that** the estimated interference signal is a communication system feedback signal or an audio-system feedback, wherein the audio signal is a part or all of the loudspeaker signal at loudspeaker Ls.

4. Method according to claim 1, wherein a test signal is sent to the respective loudspeaker for determining system characteristics on the base of the test signal received by the respective microphone, wherein the system characteristics detected by the application of the test signal are used to determine the frequency dependent coupling factors, the frequency dependent decay factors and frequency dependent delay factors.

5. Method according to any of the preceding claims, **characterized in that** generating an interference suppressed loudspeaker signal is performed by Spectral Subtraction, preferably by the application of a Wiener-filter using the estimated interference signal.

6. Method according to any of the preceding claims, **characterized in that** an overestimation $\gamma(\mu,k)$ of the estimated interference signal of the sub-band μ is adjusted to an estimated noise to speech signal ratio $SNR(\mu,k)$ of the sub-band μ deduced by a noise suppression module and that a maximum attenuation $\beta(\mu,k)$ of the sub-band μ is adjusted to a gain correction factor $V(\mu,k)$ of the sub-band μ.

7. Method according to any of the preceding claims, further **characterized by** applying stepwise independently or combined at least two suppressing steps out of communication system feedback suppression, audio system feedback suppression and noise suppression, wherein for all the suppression steps corresponding estimations of interference signals are used, and the suppression modules can be arranged in any order, but are preferably arranged in the following order: communication system feedback, audio system feedback and noise.

8. Method according to any of the preceding claims, **characterized in that** the interference model parameters are deduced from at least one microphone signal or by automatically detecting and interpreting decaying signal slopes, wherein after deducting the parameters to be applied are updated and the parameter deduction occurs preferably when there is no local speech.

9. Method according to any of the preceding claims, **characterized by** the further steps of preparing at least two different interference models on the basis of different parameters for different occupancy or different environment conditions; detecting the actual occupancy of the vehicle or environment condition; and selecting the interference models in accordance with the actual occupancy or environment condition detected.

10. Method according to any of the preceding claims, **characterized in that** the coefficients of an echo compensator are used for the adaption of the interference model parameters preferably by placing an echo compensator parallel to the signal processing path, wherein the coefficients of the echo compensator are corresponding to the room impulse response and the parameters of the echo compensator should only be updated when there is no local speech detected, e.g. only at the decaying slopes of the microphone signal or only at the feedback of the audio signal.

11. Method, **characterized by** adjusting a communication system equalizer $H_{Eq}(\mu,k)$ according to the energy decay

model frequency dependent coupling factors $P_{LsMic}(\mu,k)$.

12. Method, **characterized by** adjusting the communication system gain according to the estimated interference signal level, wherein this gain is dependent on the background noise level and on the level of the communication system feedback and audio system feedback, for a high communication system feedback or audio system feedback signal level the system gain should be reduced.

13. Software product for determining interference model parameters in the form of frequency dependent coupling factors and frequency dependent decay values that an estimated interference signal is formed on the base of the coupling factor times a loudspeaker signal plus a microphone signal at an earlier moment times a decay factor depending on the decay value.

14. Software product for carrying out the method according to any of the preceding claims.

15. Communication system as an indoor communication system or handsfree telephony system or automatic speech recognition system, comprising at least one loudspeaker and at least one microphone, as well as a signal treatment device, which carries out a method according to any of claims 1 to 12.

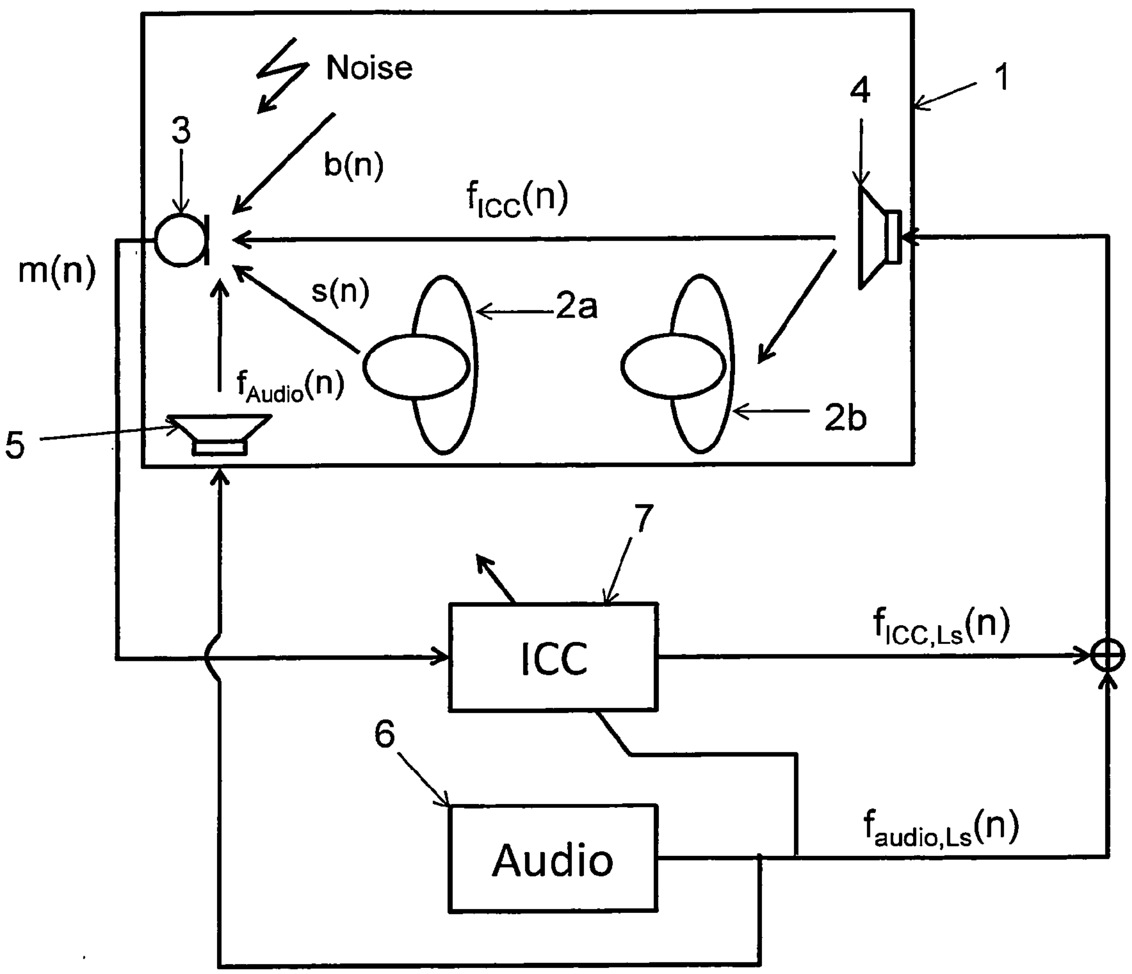
Noise
1
4
3
b(n)
$f_{ICC}(n)$
m(n)
s(n)
2a
$f_{Audio}(n)$
2b
5
7
ICC
$f_{ICC,Ls}(n)$
6
Audio
$f_{audio,Ls}(n)$

Fig. 1

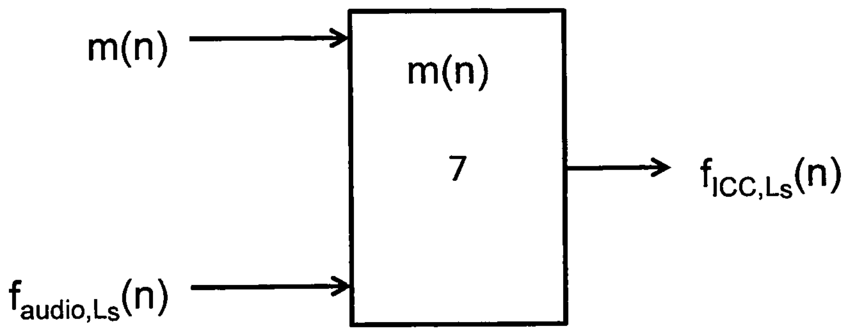
m(n)
m(n)
7
$f_{ICC,Ls}(n)$
$f_{audio,Ls}(n)$

Fig. 2

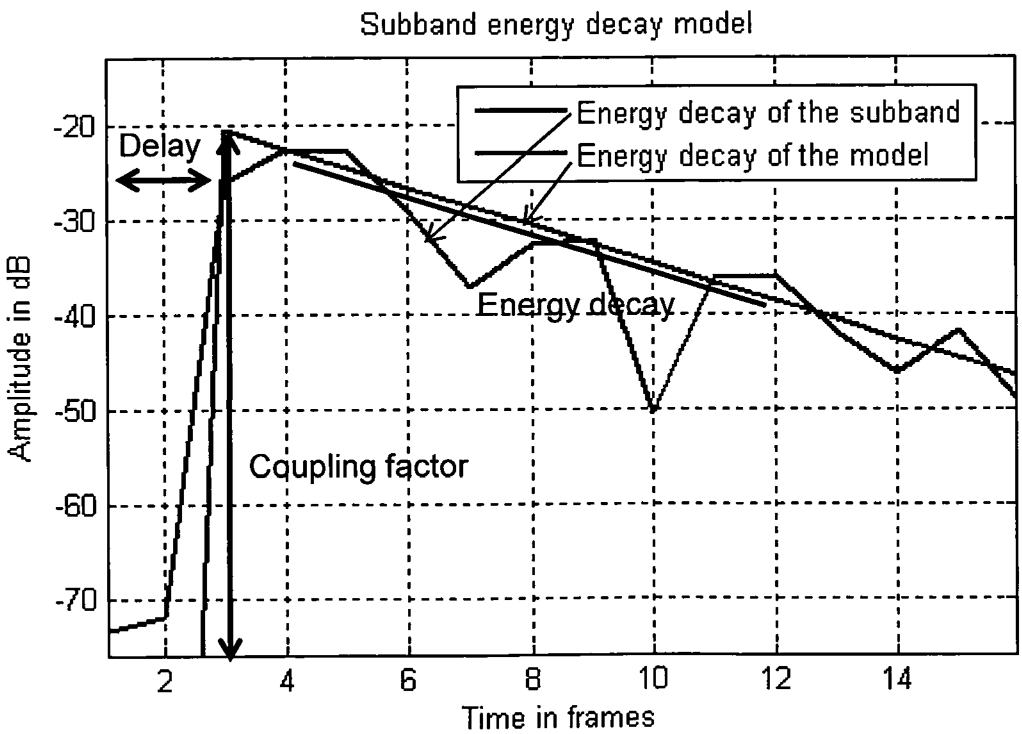
Subband energy decay model
Energy decay of the subband
Energy decay of the model
Delay
Energy decay
Coupling factor
Amplitude in dB
Time in frames
-20
-30
-40
-50
-60
-70
2
4
6
8
10
12
14

Fig. 3a

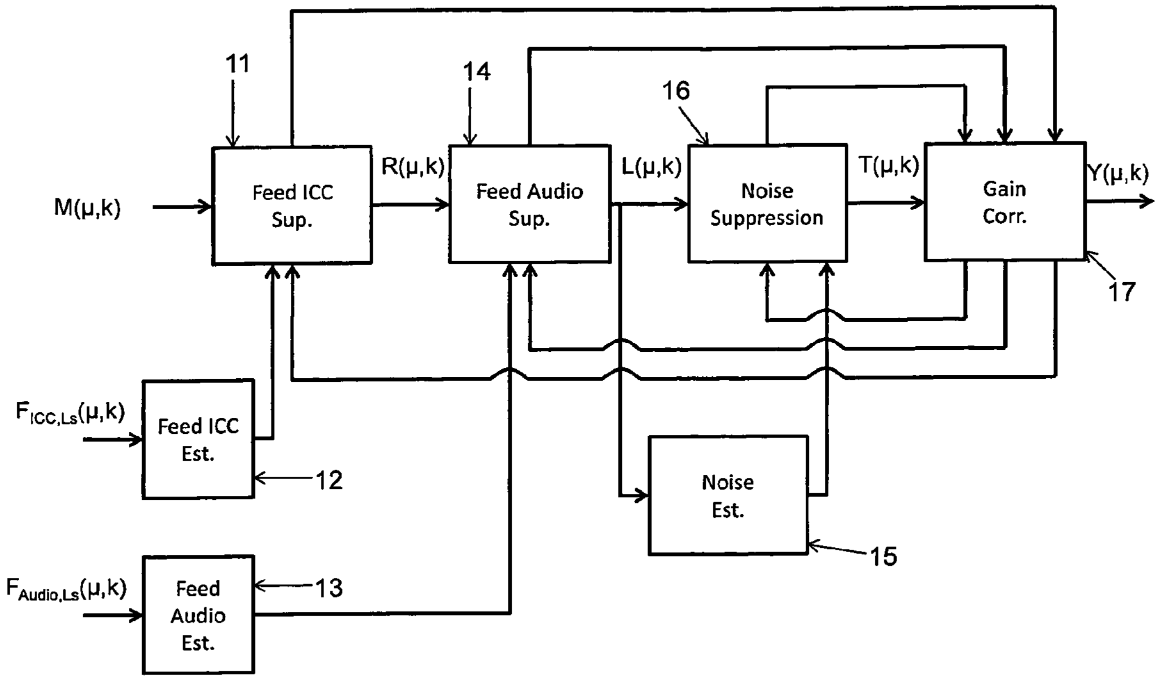
11
14
16
M(μ,k)
Feed ICC
Sup.
R(μ,k)
Feed Audio
Sup.
L(μ,k)
Noise
Suppression
T(μ,k)
Gain
Corr.
Y(μ,k)
17
F_ICC,Ls(μ,k)
Feed ICC
Est.
12
Noise
Est.
15
F_Audio,Ls(μ,k)
Feed
Audio
Est.
13

Fig. 3b

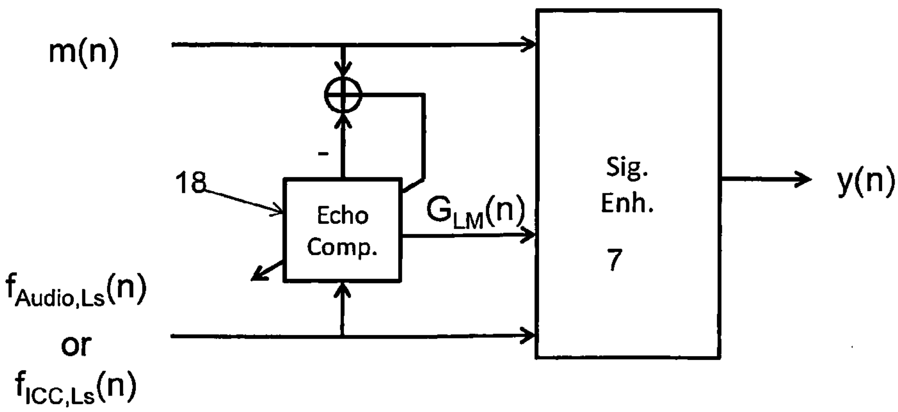
m(n)
-
18
Echo
Comp.
G_LM(n)
Sig.
Enh.
7
y(n)
f_Audio,Ls(n)
or
f_ICC,Ls(n)

Fig. 4a

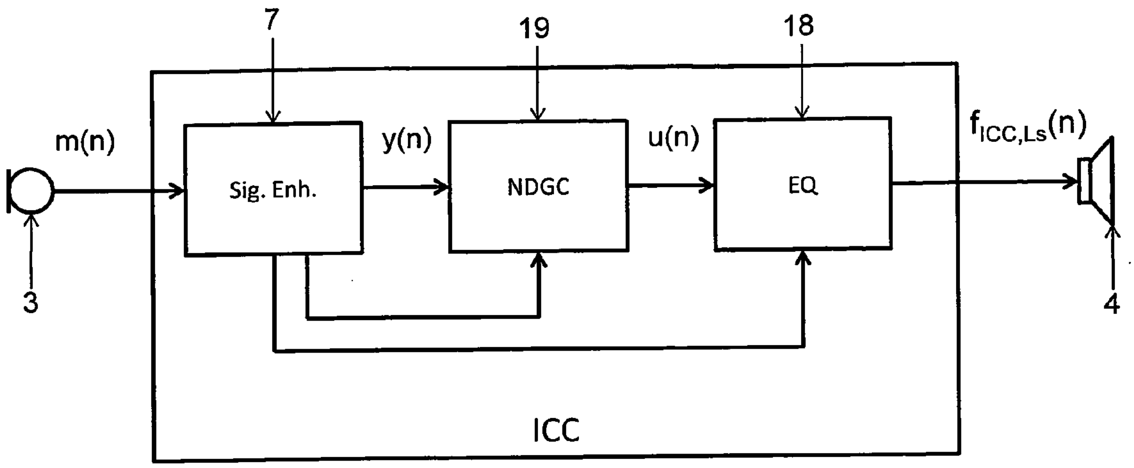
7
19
18
m(n)
Sig. Enh.
y(n)
NDGC
u(n)
EQ
$f_{ICC,Ls}(n)$
3
4
ICC

Fig. 4b

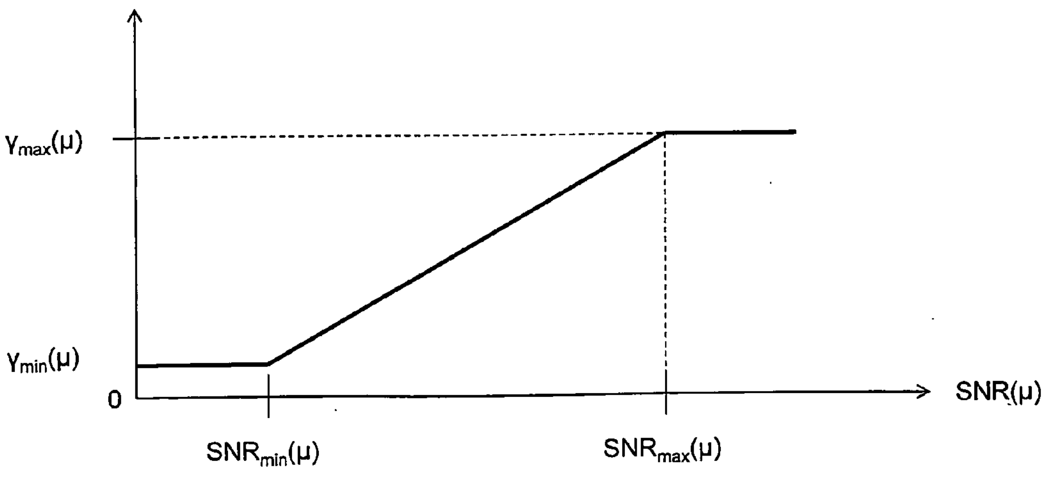
$\gamma_{max}(\mu)$
$\gamma_{min}(\mu)$
0
SNR(μ)
$SNR_{min}(\mu)$
$SNR_{max}(\mu)$

Fig. 5

Europäisches Patentamt
European Patent Office
Office européen des brevets

PARTIAL EUROPEAN SEARCH REPORT

under Rule 62a and/or 63 of the European Patent Convention. This report shall be considered, for the purposes of subsequent proceedings, as the European search report

Application Number
EP 11 15 5047

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
Y,D	SCHMIDT G ET AL: "Signal processing for in-car communication systems", SIGNAL PROCESSING, ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, NL, vol. 86, no. 6, 1 June 2006 (2006-06-01), pages 1307-1326, XP024997680, ISSN: 0165-1684, DOI: 10.1016/J.SIGPRO.2005.07.040 [retrieved on 2006-06-01] * section 1; section 3.5; figure 1 * -----	1-10,14, 15	INV. G10L21/02 H04M9/08
Y,D	US 2009/117948 A1 (BUCK MARKUS [DE] ET AL) 7 May 2009 (2009-05-07) * abstract * * paragraph [0020] * * paragraph [0023] - paragraph [0033] * * paragraph [0038] - paragraph [0061] * -----	1-10,14, 15	
	-/--		

TECHNICAL FIELDS SEARCHED (IPC)

G10L
H04M

INCOMPLETE SEARCH

The Search Division considers that the present application, or one or more of its claims, does/do not comply with the EPC so that only a partial search (R.62a, 63) has been carried out.

Claims searched completely :

Claims searched incompletely :

Claims not searched :

Reason for the limitation of the search:

see sheet C

1

Place of search	Date of completion of the search	Examiner
The Hague	27 September 2011	De Meuleneire, M

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

EPO FORM 1503 03.82 (P04E07)

Europäisches Patentamt
European Patent Office
Office européen des brevets

PARTIAL EUROPEAN SEARCH REPORT

Application Number
EP 11 15 5047

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	WO 2008/122930 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; DERKX RENE M M [NL]; JANSE CORNEL) 16 October 2008 (2008-10-16) * page 4, line 22 - line 28 * * page 5, line 6 - line 16 * -----	1,9,14,15	
A	EP 1 885 154 A1 (HARMAN BECKER AUTOMOTIVE SYS [DE]) 6 February 2008 (2008-02-06) * paragraph [0046] - paragraph [0050] * -----	1,14,15	
			TECHNICAL FIELDS SEARCHED (IPC)

1

EPO FORM 1503 03.82 (P04C10)

Europäisches Patentamt
European Patent Office
Office européen des brevets

INCOMPLETE SEARCH SHEET C

Application Number
EP 11 15 5047

Claim(s) completely searchable:
1-10, 14, 15

Claim(s) not searched:
11-13

Reason for the limitation of the search:

Claims 1, 11, 12 and 13-14 have been drafted as separate independent claims.
Under Article 84 in combination with Rule 43(2) EPC, an application may contain more than one independent claim in a particular category only if the subject-matter claimed falls within one or more of the exceptional situations set out in paragraph (a), (b) or (c) of Rule 43(2) EPC. However, this is not the case in the present application.
Moreover, independent claims 11 and 12 are not clear (Article 84 EPC). In claim 11, the parameters PLsMic are not defined. In claims 11 and 12, the communication system is not defined. In claim 12, it is not clear what a communication system gain is. Therefore, it is not possible to carry out a meaningful search regarding the state of the art on the basis of independent claims 11 and 12 (Rule 63(1)).
The applicant was therefore invited to file a statement indicating the claims complying with Rule 43(2) on the basis of which the search is to be carried out within the time limit indicated in the present communication (Rules 62a(1) and 63(1) EPC).
In his letter dated 25.08.2011, the applicant requested to carry out the search on the basis of the independent claims 1, 14 and 15.

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO.

EP 11 15 5047

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.

27-09-2011

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009117948 A1	07-05-2009	EP 2058804 A1	13-05-2009
WO 2008122930 A1	16-10-2008	NONE	
EP 1885154 A1	06-02-2008	US 2008292108 A1	27-11-2008

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.

Patent documents cited in the description

- EP 1718103 B1 **[0006]**
- US 20090117948 A **[0036]**
- EP 2151983 A **[0069]**

Non-patent literature cited in the description

- **E. LLEIDA ; E. MASGRAU ; A. ORTEGA.** Acoustic echo and noise reduction for car cabin communication. *Proc. EUROSPEECH '01,* 2001, vol. 3, 1585-1588 **[0001]**
- **T. HAULICK ; G. SCHMIDT.** Signal processing for in-car communication systems. *Signal Processing,* June 2006, 1307-1326 **[0001]**
- **E. HÄNSLER ; G. SCHMIDT.** Topics in acoustic echo and noise control. Springer, 2006, 549-598 **[0002]**
- **E. HABETS.** Multichannel speech dereverberation based on a statistical model of late reverberation. *ICASSP,* 2005 **[0036]**