

(19)



(11)

EP 2 363 853 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:

07.09.2011 Bulletin 2011/36

(51) Int Cl.:

G10L 21/02 (2006.01)(21) Application number: **10450036.8**(22) Date of filing: **04.03.2010**

(84) Designated Contracting States:

**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 SE SI SK SM TR**

Designated Extension States:

AL BA ME RS

- **Austria Wirtschaftsservice Gesellschaft mbH
1030 Wien (AT)**

(72) Inventor: **Weruaga, Luis
1090 Wien (AT)**(74) Representative: **Weiser, Andreas
Patentanwalt
Kopfgasse 7
1130 Wien (AT)**

(71) Applicants:

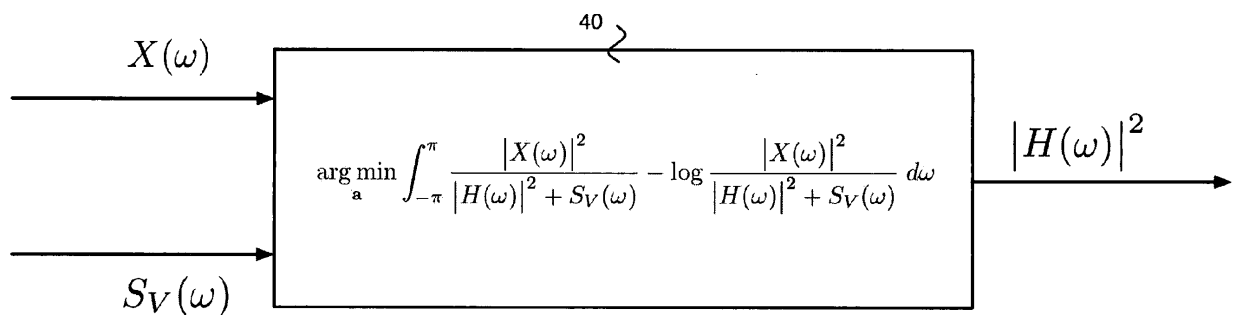
- **Österreichische Akademie der Wissenschaften
1010 Wien (AT)**

(54) **A method for estimating the clean spectrum of a signal**

(57) The invention proposes a method for estimating the clean spectrum of a signal degraded by additive noise, in particular a speech signal, by determining the coefficients of a predictive model of said clean spectrum, comprising:
computing the spectrum of said signal;
estimating the power spectrum of said noise; and
determining said coefficients by minimizing the cost function

$$\int_{2\pi} \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_V(\omega)} - \log \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_V(\omega)} d\omega$$

with respect to said coefficients, with

 $X(\omega)$ being the spectrum of said signal, $S_V(\omega)$ being the power spectrum of said noise, and $H(\omega)$ being the transfer function of said model based on said coefficients.**Fig. 2****EP 2 363 853 A1**

Description

[0001] The present invention relates to a method for estimating the clean spectrum of a signal degraded by additive noise, in particular a speech signal, by determining the coefficients of a predictive model of said clean spectrum. The invention further relates to a method for enhancing a signal based on this clean spectrum estimation.

[0002] Restoration of single-channel digital audio recordings degraded by additive noise is a technical problem that currently arouses large interest from scientific and commercial points of view. The enhancement of speech by digital signal processing means improves the quality and intelligibility of voice communication for a wide fan of applications, such as mobile telephony, hearing aids, teleconference systems, dictation systems, voice coders and automatic speech recognition systems.

[0003] Among different solutions proposed for the enhancement of noisy speech, restoration of short-time speech spectrum has been extensively studied, see e.g. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator", IEEE Trans. Acoust., Speech, Signal Processing, Vol. 32, No. 6, pp. 1109-1121, 1984; B. Sim, Y. Tong, J. Chang, and C. Tan, "A parametric formulation of the generalized spectral subtraction method", IEEE Transactions on Speech and Audio Processing, Vol. 6, No. 4, pp. 328-337, Jul. 1998; P. J. Wolfe and S. J. Godsill, "Efficient alternatives to the Ephraim and Malah suppression rule for audio signal enhancement", EURASIP J. Applied Signal Processing, Vol. 2003, No. 10, pp. 1043-1051, 2003; and P. C. Loizou, "Speech enhancement: Theory and practice", CRC Press, 2007. This approach is based on estimation of the short-time spectral amplitude of the clean speech from an estimate of the signal-to-noise ratio (SNR) at each frequency. In other cases, the clean speech is assumed to follow a parametric model, such as an autoregressive model (AR); upon the estimation of that model, an enhancement filter, such as the Wiener filter, is employed to enhance the noisy signal, see J. H. L. Hansen and M. A. Clements, "Constrained iterative speech enhancement with application to speech, recognition", IEEE Trans. Signal Processing, Vol. 39, No. 4, pp. 795-805, Apr. 1991. In all cases, an accurate estimation of the power spectrum of the noise is required. This can be accomplished by several techniques, such as minimum statistics, tracking of the spectral floor, or by detecting silences in the speech activity (P. C. Loizou, l.c.).

[0004] The biggest technical challenge in this problem is thus to obtain the a priori signal-to-noise ratio at each frequency. Since the noise is assumed to be available with state-of-the art techniques, the previous challenge is equivalent to the estimation of the clean speech spectrum from the available noisy spectrum. This problem has coped the efforts of many researchers in the last twenty five years: a decision-directed method (Y. Ephraim and D. Malah, loc.cit.), subspace methods (P. C. Loizou, loc.cit.), iterative Wiener filter (J. H. L. Hansen and M. A. Clements, loc.cit.), or Kalman filters (see e.g. E. Zarevei, S. Vaseghi, and Q. Yan, "Speech enhancement using Kalman filters for restoration of short-time DFT trajectories," IEEE Workshop Automatic Speech Recognition and Understanding, 2005, pp. 313-318.) are some of the most popular techniques thereto. From the previous techniques, the iterative Wiener filter is particularly interesting because it aims to estimate the clean speech spectrum only from the current noisy spectrum, combining iteratively Wiener filtering (IWF) with autoregressive analysis. The problem of that technique is its tendency to generate high resonance peaks, which introduces an unpleasant distortion in the enhanced speech. Further attempts for stabilizing the IWF have been made in T. V. Sreenivas and P. Kirnapure, "Codebook constrained Wiener filtering for speech enhancement," IEEE Trans. Speech, Audio Processing, Vol. 4, No. 5, pp. 383-389, Sep. 1996, but the performance of this technique and its variants is still clearly insufficient.

[0005] One possible solution to the problem of estimating parameters of a predictive clean speech model has been disclosed in the earlier application WO 2008/109904 A1 of the same applicant. This prior solution fails to estimate the clean speech spectrum in cases where the signal-to-noise (SNR) ratio of the signal is low. Likewise, the mentioned IWF method is not appropriate in such applications.

[0006] It is therefore an object of the invention to provide a method for estimating the clean spectrum of a noise-corrupted signal with improved accuracy.

[0007] This object is achieved by means of a method for estimating the clean spectrum of a signal degraded by additive noise, in particular a speech signal, by determining the coefficients of a predictive model of said clean spectrum, comprising:

computing the spectrum of said signal;
estimating the power spectrum of said noise; and
determining said coefficients by minimizing the cost function

$$\int_{2\pi} \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} - \log \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} d\omega$$

with respect to said coefficients, with
 $X(\omega)$ being the spectrum of said signal,
 $S_v(\omega)$ being the power spectrum of said noise, and
 $H(\omega)$ being the transfer function of said model based on said coefficients.

[0008] In the present disclosure the term "minimizing" is intended to comprise both, making the cost function minimal as well as making the cost function at least a sufficiently low value, i.e. a value within a given or acceptable tolerance interval from that minimum.

[0009] From the field of bioacoustics it is known that the biological hearing sense responds to the logarithm of the sound intensity. The invention is based on the insight that this bio-acoustic principle of logarithmic sense can be introduced into a novel cost function as stated above which takes into account the actual signal-to-noise ratio in each portion of the signal spectrum. Loosely speaking the proposed cost function fits the model to the data for those regions with high SNR, and - as will be detailed later on - in low-SNR areas the fitting process is driven by the mentioned good fitting performance taking place on adjacent high-SNR areas. The inventive method thus leads to an interpolation effect from high-SNR to low-SNR spectral regions.

[0010] According to a preferred embodiment of the invention said cost function is minimized by solving the equation

$$\int_{-\pi}^{\pi} M(\omega) E(\omega) A(\omega) e^{j\omega l} d\omega = 0$$

with $A(\omega)$ being the predictive model based on its coefficients a_m according to

$$A(\omega) = \sum_{m=0}^M a_m e^{-j\omega m} ,$$

$E(\omega)$ being the prediction error according to

$$E(\omega) = |X(\omega)|^2 - |H(\omega)|^2 - S_v(\omega) ,$$

and $M(\omega)$ being a spectral mask defined as

$$M(\omega) = \left(\frac{SNR(\omega)}{SNR(\omega) + 1} \right)^2 \quad \text{with} \quad SNR(\omega) = \frac{|H(\omega)|^2}{S_v(\omega)} .$$

[0011] In particular said equation can be solved by holding $E(\omega)$ and $M(\omega)$ constant, solving the remaining linear problem, using the solution to re-evaluate the previous constant terms, and proceeding further iteratively.

[0012] In general, the method of the invention is suited for any predictive model known in the art. Preferably, a parametric all-pole filter model, an autoregressive coefficients filter (ARC) model, a reflection coefficients filter (RC) model, and/or a line spectral frequencies (LSF) model is used.

[0013] In a second aspect of the invention a method for enhancing a digital signal, in particular a speech signal, with increased quality is provided. The inventive method comprises the further steps of

[0014] calculating a spectral signal-to-noise ratio on the basis of the clean spectrum and the noise spectrum, and

[0015] using the spectral signal-to-noise ratio to enhance the signal.

[0016] Preferably, the signal is enhanced by means of a Wiener filter, a MMSE-based enhancement, or variants thereof, using said spectral signal-to-noise ratio.

[0017] Further details and advantages of the invention will become apparent from the appended claims and the

following detailed description of a preferred embodiment under reference to the enclosed drawings in which

Fig. 1 shows in block diagram form an apparatus for enhancing a digital speech signal, the blocks concurrently illustrating the steps of the method of the invention, and

Fig. 2 shows the function of the clean speech estimation block and step of Fig. 1 in detail.

[0018] As a first basis of the present invention, the inventor has found out analytically that the IWF method is equivalent to a method that results from the following minimization problem

$$\arg \min_a \int_{2\pi} \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} d\omega \quad (1)$$

where ω is frequency, $X(\omega)$ is the Fourier transform of a short-time segment of the input noisy signal, $S_v(\omega)$ is the estimate of the noise power spectral density, and $H(\omega)$ is the transfer function of the autoregressive model which relates to the clean speech spectrum. The said transfer function of the autoregressive model is equal to

$$H(\omega) = \frac{1}{\sum_{m=0}^M a_m e^{-j\omega m}} \quad (2)$$

where $a = (a_0, a_1, \dots, a_M)$ are the autoregressive coefficients, and M is the autoregressive model order.

[0019] As a second basis of the invention, the inventor has found out that a functional built on the ratio between the samples and the model, such as in (1), does not possess the desirable property of frequency selectivity while such a property would be desirable when not all spectral samples are available: In case of the spectrum of the noisy signal $X(\omega)$, the spectral samples at which the a priori SNR is low or very low do not represent a trustful reference for the estimation of the autoregressive model.

[0020] To this end, the method of the present invention for estimating the clean speech spectrum is related to the minimization of the maximum likelihood (ML) of the ratio between the input noisy spectrum $X(\omega)$ and the model of clean speech corrupted by additive noise. Assuming that $X(\omega)$ is modelled by a Gaussian distribution, said maximum likelihood estimation turns out

$$\arg \min_a \int_{2\pi} \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} - \log \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} d\omega \quad (3)$$

where the clean speech follows the autoregressive model defined in (2), a is the vector containing the autoregressive coefficients, and $S_v(\omega)$ is the power spectral density of the noise which is available a priori.

[0021] By computing the gradient of the functional (3) with respect to the autoregressive coefficients a , one gets to the solution of this problem, given by the following equation

$$\int_{-\pi}^{\pi} M(\omega) E(\omega) A(\omega) e^{j\omega l} d\omega = 0 \quad (4)$$

where $A(\omega)$ is the linear prediction error filter, defined in terms of the autoregressive coefficients as

$$A(\omega) = \sum_{m=0}^M a_m e^{-j\omega m} \quad (5)$$

$E(\omega)$ is the prediction error of the model according to

$$E(\omega) = |X(\omega)|^2 - |H(\omega)|^2 - S_v(\omega) \quad (6)$$

and $M(\omega)$ is a so-called spectral mask defined as

$$M(\omega) = \left(\frac{SNR(\omega)}{SNR(\omega) + 1} \right)^2 \quad (7)$$

[0022] Here, the spectral mask is defined in terms of the a-priori signal-to-noise ratio for each frequency ("spectral" signal-to-noise ratio), $SNR(\omega)$. The a-priori (spectral) SNR is defined as the ratio between the clean speech power spectrum and the noise power spectrum,

$$SNR(\omega) = \frac{|H(\omega)|^2}{S_v(\omega)} \quad (8)$$

[0023] Since equation (4) is nonlinear with respect to the autoregressive coefficients, its solution must and can be obtained by means of an iterative procedure, in which at each iteration a positive-definite Toeplitz linear system must be solved. Several techniques are available to solve Toeplitz systems, such as the well-known Levinson algorithm. One skilled in the art will immediately recognize that this choice does not affect the essence of the present invention. It is important to mention that the spectral mask (7) weights the importance of the spectral error between the noisy samples and the model of clean speech plus additive noise. This weight at each frequency depends on the respective signal-to-noise ratio. Thus, if the SNR is high at a given frequency, the spectral mask is close to 1 at that frequency, and the information at that frequency is valuable in the estimation. On the contrary, if the SNR is low, the spectral mask tends to zero, which implies that the relevance of the information at the frequency is low.

[0024] The spectral mask, the signal-to-noise ratio, and therewith the clean speech model are estimated in an iterative fashion. The final solution is obtained either after several iterations or when successive partial solutions do not differ from each other substantially.

[0025] One iterative approach to solve equation (4) will be discussed in detail. This approach is based on considering $E(\omega)$ and $M(\omega)$ constant, and solving the remaining linear problem; this partial solution is used to re-evaluate the previous constant terms, and proceeding further iteratively. Thus, the linear residue filter $A(\omega)$ is obtained with the following iterative algorithm

$$S_{X,\omega}^{(\kappa)} = \frac{1}{|A_{\omega}^{(\kappa)}|^2} \quad (8a)$$

$$\xi_{\omega}^{(\kappa)} = \frac{S_{X,\omega}^{(\kappa)}}{S_{V,\omega}} \quad (8b)$$

$$M_{\omega}^{(\kappa)} = \left(\frac{\xi_{\omega}^{(\kappa)}}{\xi_{\omega}^{(\kappa)} + 1} \right)^2 \quad (8c)$$

$$h_l^{(\kappa)} = \int_{\langle M^{(\kappa)} \rangle} \frac{e^{j\omega l}}{A_{\omega}^{(\kappa)*}} d\omega \quad (8d)$$

$$g_l^{(\kappa)} = \int_{\langle M^{(\kappa)} \rangle} S_{V,\omega} A_{\omega}^{(\kappa)} e^{j\omega l} d\omega \quad (8e)$$

$$\int_{\langle M^{(\kappa)} \rangle} |X_{\omega}|^2 A^{(\kappa+1)}(e^{j\omega}) e^{j\omega l} d\omega = h_l^{(\kappa)} + g_l^{(\kappa)} \quad (8f)$$

for $\ell = 0, 1, \dots, M$, where subindex K denotes iteration, and superscript $*$ complex conjugate. The noise-subtracted power spectrum can be used as initial seed, i.e.,

$$S_{X,\omega}^1 = \left| |X_{\omega}|^2 - S_{V,\omega} \right|_{\epsilon} \quad (9)$$

[0026] The notation in the integrals refers to

$$\int_{\langle M^{(\kappa)} \rangle} \cdot \equiv \int_{-\pi}^{\pi} M_{\omega}^{(\kappa)} \quad (10)$$

where $M_{\omega}^{(\kappa)}$ is the spectral weight (mask) $M(\omega)$ at the κ iteration. Since the spectral weight is present in all terms of the inverse problem (8f), its effect is that of weighting the relevance of the spectral samples. The magnitude of the weight depends on the local SNR ξ_{ω} , such that in areas with high SNR $\gg 1$) the spectral weight tends to one, while in low-SNR areas ($\xi_{\omega} \leq 1$) it tends to zero. Note as comparison that in the noiseless case the spectral weight turns one for all frequencies, this meaning that the noiseless case need not require spectral selectivity.

[0027] Finally, the step (8f) is a linear inverse problem involving a positive-semidefinite symmetric Toeplitz system. Thus, it can be efficiently solved with the Levinson algorithm or any other algorithm to solve Toeplitz systems.

[0028] Fig. 1 shows in a simplified fashion the processing-block diagram of a speech enhancement front-end (apparatus 100) that uses the method of the present invention. Fig. 2 shows the function of the clean speech estimation step (block

40) of Fig. 1 in detail.

[0029] Block 10 performs the usual segmentation of the input digital signal into segments.

[0030] Block 20 performs the spectral transformation of said segment. Said spectral transformation corresponds to the "Discrete Fourier Transform", "Discrete Sinus Transform" and/or to the "Fan-Chirp Transform", among other popular choices.

[0031] Block 30 carries out the estimation of the power spectrum of the noise according to known ad-hoc techniques. It is assumed that this block has memory facilities in such a way that the spectrum of the previous segments are stored therein. Therefore, if required, the estimation of the noise power spectrum can be performed by statistical methods over spectral data stretching within a reasonably long time span.

[0032] Block 40 carries out the estimation of the clean speech model from the spectrum of the segment and the estimation of the noise power spectrum. The estimation of the clean speech model is based on the numerical implementation of the minimization problem (3), which represents the core method of the present invention.

[0033] Block 50 computes numerically the signal-to-noise ratio for each frequency (spectral signal-to-noise ratio) from the estimated clean speech model and noise model.

[0034] Block 60 enhances the spectrum of the input signal by means of state-of-art techniques that require the signal-to-noise ratio for each frequency. Among these techniques, we can cite the Wiener filter and its variants, e.g. the root-square of the Wiener filter, and the minimum-mean-square-error (MMSE)-based enhancement (see *Y. Ephraim and D. Malah*, loc.cit.) and its variants, e.g. the log-MMSE, et cet. (see *P. J. Wolfe and S. J. Godsill*, loc.cit.).

[0035] Block 70 performs the inverse spectral transformation to block 20. The output of block 70 is the enhanced segment of the audio signal.

[0036] Although all processor blocks of the apparatus 100 operate with time-discrete and frequency-discrete samples, for the sake of clarity the mathematical description of the invention has been given in continuous frequency. One skilled in the art will immediately recognize that this choice does not affect the essence of the present invention.

Claims

1. A method for estimating the clean spectrum of a signal degraded by additive noise, in particular a speech signal, by determining the coefficients of a predictive model of said clean spectrum, comprising:

computing the spectrum of said signal;
estimating the power spectrum of said noise; and
determining said coefficients by minimizing the cost function

$$\int_{2\pi} \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} - \log \frac{|X(\omega)|^2}{|H(\omega)|^2 + S_v(\omega)} d\omega$$

with respect to said coefficients, with
 $X(\omega)$ being the spectrum of said signal,
 $S_v(\omega)$ being the power spectrum of said noise, and
 $H(\omega)$ being the transfer function of said model based on said coefficients.

2. The method of claim 1, wherein said cost function is minimized by solving the equation

$$\int_{-\pi}^{\pi} M(\omega)E(\omega)A(\omega)e^{j\omega l} d\omega = 0$$

with $A(\omega)$ being the predictive model based on its coefficients a_m according to

$$A(\omega) = \sum_{m=0}^M a_m e^{-j\omega m} ,$$

$E(\omega)$ being the prediction error according to

$$E(\omega) = |X(\omega)|^2 - |H(\omega)|^2 - S_v(\omega) ,$$

and $M(\omega)$ being a spectral mask defined as

$$M(\omega) = \left(\frac{SNR(\omega)}{SNR(\omega) + 1} \right)^2 \text{ with } SNR(\omega) = \frac{|H(\omega)|^2}{S_v(\omega)} .$$

3. The method of claim 2, wherein said equation is solved by holding $E(\omega)$ and $M(\omega)$ constant, solving the remaining linear problem, using the solution to re-evaluate the previous constant terms, and proceeding further iteratively.
4. The method of any of the claims 1 to 3, wherein said predictive model is a parametric all-pole filter model, an autoregressive coefficients filter (ARC) model, a reflection coefficients filter (RC) model, and/or a line spectral frequencies (LSF) model.
5. The method of any of the claims 1 to 4, further for enhancing the signal, comprising the further steps of calculating a spectral signal-to-noise ratio on the basis of the clean spectrum and the noise spectrum, and using the spectral signal-to-noise ratio to enhance the signal.
6. The method of claim 5, wherein the signal is enhanced by means of a Wiener filter, a MMSE-based enhancement, or variants thereof, using said spectral signal-to-noise ratio.

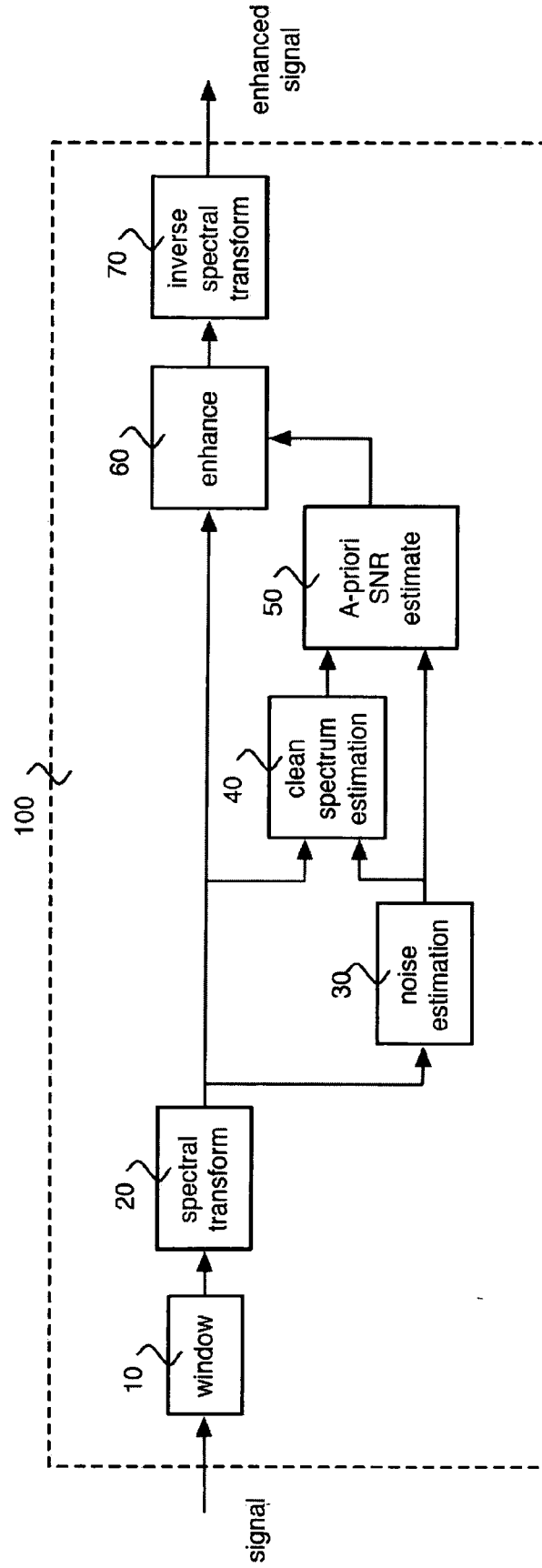
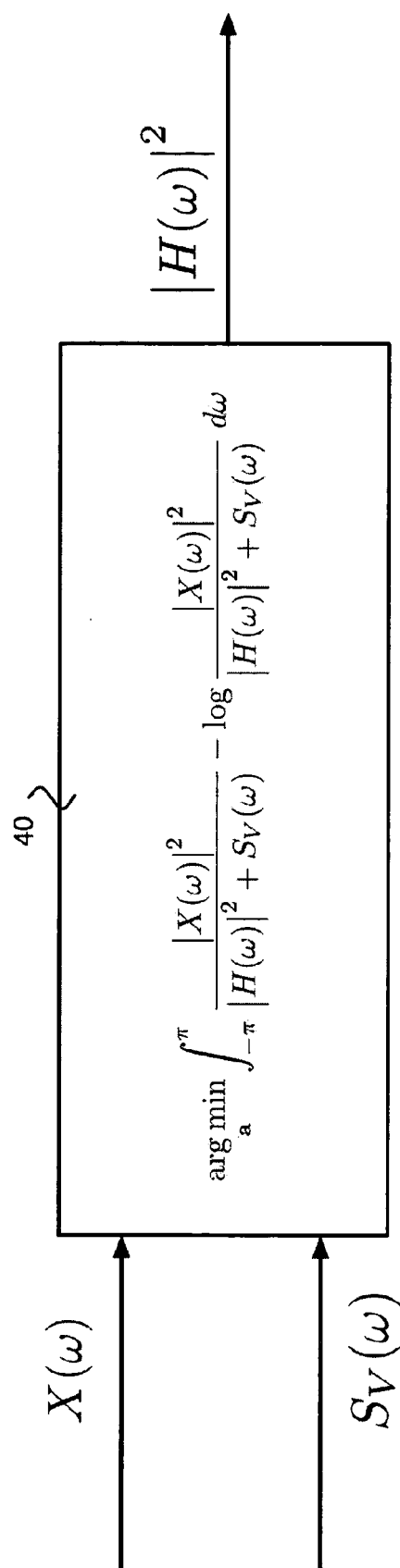


Fig. 1

**Fig. 2**



EUROPEAN SEARCH REPORT

Application Number
EP 10 45 0036

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,D	EP 1 970 893 A1 (OESTERREICHISCHE AKADEMIE DER [AT]; AUSTRIA WIRTSCHAFTSSERV GMBH [AT]) 17 September 2008 (2008-09-17) * paragraph [0022] - paragraph [0025] * -----	1-6	INV. G10L21/02
A	WO 2008/031124 A1 (UNIV GRAZ TECH [AT]; FORSCHUNGSHOLDING TU GRAZ GMBH [AT]; RANK ERHARD) 20 March 2008 (2008-03-20) * the whole document * -----	1-6	
A	K. Funaki: "Speech enhancement based on iterative Wiener filter using complex speech analysis" Proc. EUSIPCO 2008 29 August 2008 (2008-08-29), pages 1-5, XP002593133 Lausanne, Switzerland Retrieved from the Internet: URL: http://www.eurasip.org/Proceedings/Eusipco/Eusipco2008/papers/1569105040.pdf [retrieved on 2010-07-22] * the whole document * -----	1-6	TECHNICAL FIELDS SEARCHED (IPC) G10L
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 22 July 2010	Examiner Burchett, Stefanie
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
EPO FORM 1503 03.82 (P04C01)

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

EP 10 45 0036

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.

22-07-2010

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 1970893	A1	17-09-2008	WO 2008109904 A1	18-09-2008

WO 2008031124	A1	20-03-2008	AT 504164 A1	15-03-2008
			EP 2080197 A1	22-07-2009
			US 2010049507 A1	25-02-2010

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

- WO 2008109904 A1 [0005]

Non-patent literature cited in the description

- **Y. EPHRAIM ; D. MALAH.** Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Trans. Acoust., Speech, Signal Processing*, 1984, vol. 32 (6), 1109-1121 [0003]
- **B. SIM ; Y. TONG ; J. CHANG ; C. TAN.** A parametric formulation of the generalized spectral subtraction method. *IEEE Transactions on Speech and Audio Processing*, July 1998, vol. 6 (4), 328-337 [0003]
- **P. J. WOLFE ; S. J. GODSILL.** Efficient alternatives to the Ephraim and Malah suppression rule for audio signal enhancement. *EURASIP J. Applied Signal Processing*, 2003, vol. 2003 (10), 1043-1051 [0003]
- **P. C. LOIZOU.** Speech enhancement: Theory and practice. CRC Press, 2007 [0003]
- **J. H. L. HANSEN ; M. A. CLEMENTS.** Constrained iterative speech enhancement with application to speech, recognition. *IEEE Trans. Signal Processing*, April 1991, vol. 39 (4), 795-805 [0003]
- **E. ZA-VAREHEI ; S. VASEGHI ; Q. YAN.** Speech enhancement using Kalman filters for restoration of short-time DFT trajectories. *IEEE Workshop Automatic Speech Recognition and Understanding*, 2005, 313-318 [0004]
- **T. V. SREENIVAS ; P. KIRNAPURE.** Codebook constrained Wiener filtering for speech enhancement. *IEEE Trans. Speech, Audio Processing*, September 1996, vol. 4 (5), 383-389 [0004]