



(11) **EP 2 515 300 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention
de la délivrance du brevet:
14.05.2014 Bulletin 2014/20

(51) Int Cl.:
G10L 21/0208 ^(2013.01) **G10L 25/18** ^(2013.01)

(21) Numéro de dépôt: **12163679.9**

(22) Date de dépôt: **11.04.2012**

(54) **Procédé et système de réduction du bruit**

Verfahren und System für die Geräuschunterdrückung

Method and system for noise reduction

(84) Etats contractants désignés:
**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**

(30) Priorité: **22.04.2011 FR 1153507**

(43) Date de publication de la demande:
24.10.2012 Bulletin 2012/43

(73) Titulaire: **ADEUNIS RF
38920 Crolles (FR)**

(72) Inventeurs:
• **Saguin, Pascal
38830 Saint Pierre d'Allevard (FR)**
• **Maouche, Karim
38000 GRENOBLE (FR)**

(74) Mandataire: **Chevalier, Renaud Philippe et al
Cabinet Germain & Maureau
BP 6153
69466 Lyon Cedex 06 (FR)**

(56) Documents cités:
US-B1- 7 016 507

- **KAMATH S ET AL: "Multi-band spectral subtraction method for enhancing speech corrupted by colored noise", 2002 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS (CAT. NO. 02CH37334) IEEE PISCATAWAY, NJ, USA, vol. 4, 2002, XP002649825, ISBN: 0-7803-7402-9**
- **BING-YIN XIA ET AL: "A modified spectral subtraction method for speech enhancement based on masking property of human auditory system", WIRELESS COMMUNICATIONS&SIGNAL PROCESSING, 2009. WCSP 2009. INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 13 novembre 2009 (2009-11-13), pages 1-5, XP031594664, ISBN: 978-1-4244-4856-2**
- **COHEN I ET AL: "Speech enhancement for non-stationary noise environments", SIGNAL PROCESSING, ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, NL, vol. 81, no. 11, 1 novembre 2001 (2001-11-01), pages 2403-2418, XP004308517, ISSN: 0165-1684, DOI: DOI:10.1016/S0165-1684(01)00128-1**
- **LUO JUN ET AL: "Spectral Subtraction Speech Enhancement Technology Based on Fast Noise Estimation", INFORMATION ENGINEERING AND COMPUTER SCIENCE, 2009. ICIECS 2009. INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 19 décembre 2009 (2009-12-19), pages 1-3, XP031587156, ISBN: 978-1-4244-4994-1**

Il est rappelé que: Dans un délai de neuf mois à compter de la publication de la mention de la délivrance du brevet européen au Bulletin européen des brevets, toute personne peut faire opposition à ce brevet auprès de l'Office européen des brevets, conformément au règlement d'exécution. L'opposition n'est réputée formée qu'après le paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

EP 2 515 300 B1

Description

[0001] La présente invention se rapporte à un procédé de réduction du bruit et à un système de réduction du bruit associé.

[0002] Elle se rapporte plus particulièrement à un procédé et à un système de réduction du bruit sur un signal acoustique bruité $y(t)$ issu d'un microphone opérant dans un milieu bruité.

[0003] La présente invention trouve une application particulière dans les systèmes de communication audio full-duplex, sans fils et mono-capteur, autrement dit mono-microphone, qui permettent d'établir une communication audio entre plusieurs utilisateurs, de manière autonome (c'est-à-dire sans raccordement à une base de transmission ou à un réseau) et qui soit simple d'utilisation (c'est-à-dire ne nécessitant aucune intervention d'un technicien pour établir la communication).

[0004] De tels systèmes de communication sont généralement employés dans un environnement bruité, comme par exemple un milieu marin ou une salle de spectacle, voire extrêmement bruité, comme par exemple un chantier de travaux publics ou une salle ou stade accueillant un événement sportif.

[0005] Ainsi, il est nécessaire, voire indispensable, de prévoir un procédé et/ou système de réduction du bruit (ou de rehaussement de parole) dans la chaîne de traitement du signal afin d'améliorer la qualité audio pour que la conversation soit audible entre les utilisateurs des systèmes de communication, autrement dit que la communication soit compréhensible.

[0006] Il est connu d'estimer la composante de bruit contenue dans le signal bruité par un algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon une méthode de moyennage récursif des minima contrôlés dite « MCRA » pour « Minima Controlled Recursive Averaging ».

[0007] Cette méthode d'estimation du bruit dite « MCRA » est connue de la littérature scientifique, notamment des articles suivants :

- « Speech enhancement for non-stationary noise environments », par I. Cohen et B. Berdugo, Signal Processing, 2001, vol.81, pp. 2403-2418 ;
- « Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement », par I. Cohen et B. Berdugo, IEEE Signal Processing Letters, Janvier 2002, vol. 9, No.1, pp. 12-15 ;

[0008] « A Modified Spectral Subtraction Method For Speech Enhancement Based on Masking Property of Human Auditory System », par X. Bing-Yin et al, Wireless Communications & Signals Processing, International Conference on IEEE, 13 novembre 2009, pp. 1-5, XP031594664.

[0009] La méthode d'estimation du bruit dite « MCRA » est ainsi particulièrement bien adaptée dans les environnements où le signal de bruit est fortement non-stationnaire et évolue relativement rapidement dans le temps.

[0010] L'état de la technique peut également être illustré par l'enseignement de l'article « Multi-band Spectral Subtraction for Enhancing Speech Corrupted by Colored Noise », 2002 IEEE International Conference on Acoustics, Speech And Signal Processing. Proceedings (Car. NO.02CH37334), vol.4, 2002. Cet article divulgue une méthode de soustraction spectrale multi-bandes.

[0011] Un autre état de la technique publié dans « A Multi-Band Spectral Subtraction Method for Enhancing Speech Corrupted by Colored Noise », par S.D. Kamath et P.C. Loizou, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2002, Vol.4, décrit une méthode de soustraction spectrale à décomposition multi-bandes tenant compte des effets de bruit coloré à différentes fréquences.

[0012] La présente invention a pour but de proposer un procédé selon la revendication 1 et un système selon la revendication 10 de réduction du bruit adapté à des environnements fortement bruités et où la composante de bruit est fortement non-stationnaire et évolue relativement rapidement dans le temps.

[0013] A titre d'exemple, un procédé de réduction du bruit sur un signal acoustique bruité $y(t)$ issu d'un microphone, de préférence unique, opérant dans un milieu bruité, comporte les étapes successives suivantes :

a) conversion du signal acoustique bruité $y(t)$ dans le domaine temporel en un signal bruité $Y_{k,l}$ dans le domaine fréquentiel, par découpage temporel du signal acoustique bruité $y(t)$ en signaux échantillonnés y_l dans des trames temporelles l successives, fenêtrage des signaux échantillonnés y_l par application d'une fenêtre de pondération, et application d'une transformée de Fourier discrète, avec extraction du module carré $|Y_{k,l}|^2$, et éventuellement de la phase $\theta_{k,l}$, du signal bruité $Y_{k,l}$;

b) estimation d'une composante de bruit $\hat{D}_{k,l}$ contenue dans le signal bruité $Y_{k,l}$ à partir du module carré $|Y_{k,l}|^2$, par un algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon une méthode de moyennage récursif des minima contrôlés dite « MCRA » ;

ledit procédé étant remarquable en ce qu'il comprend en outre, après l'étape b), les étapes successives suivantes :

c) découpage de la bande fréquentielle en plusieurs sous-bandes fréquentielles $SB_i=[e_i, b_i]$, suivie d'une décompo-

sition multi-bandes du module carré $|Y_{k,l}|^2$ et de la composante de bruit $\hat{D}_{k,l}$, consistant à décomposer le module carré $|Y_{k,l}|^2$ et la composante de bruit $\hat{D}_{k,l}$ en respectivement plusieurs modules carré de sous-bande $|Y_{k,l,i}|^2$ et plusieurs composantes de bruit de sous-bande $\hat{D}_{k,l,i}$ propres à chacune des sous-bandes SB_i ;

d) estimation, pour chacune des sous-bandes SB_i , du module carré $|\hat{X}_{k,l,i}|^2$ d'une composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i d'un signal débruité $\hat{X}_{k,l}$, par un algorithme de soustraction spectrale multi-bandes dit « SSMB » à partir des modules carrés de sous-bande $|Y_{k,l,i}|^2$ et des composantes de bruit de sous-bande $\hat{D}_{k,l,i}$;

e) détermination d'un signal débruité de sortie $SD_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$ issus de l'étape d), et éventuellement des phases $\theta_{k,l}$ extraites lors de l'étape a) ;

f) conversion du signal débruité de sortie $SD_{k,l}$ en un signal vocal débruité de sortie $sd(t)$ dans le domaine temporel, par une étape f.1) de calcul d'un signal échantillonné de sortie Sd_l propre à chaque trame temporelle l par application d'une transformée de Fourier inverse du signal débruité de sortie $SD_{k,l}$, suivie d'une étape f.2) de reconstruction temporelle du signal vocal débruité de sortie $sd(t)$ à partir des signaux échantillonnés de sortie Sd_l .

[0014] Ainsi, cet exemple propose de mettre en oeuvre un algorithme de soustraction spectrale multi-bandes « SSMB » qui consiste à partager la bande spectrale entière en sous bandes et à adapter dans chaque sous bande le calcul de soustraction entre le signal bruité $Y_{k,l}$ et la composante de bruit $\hat{D}_{k,l}$ pour extraire un signal débruité de sortie $SD_{k,l}$; appliquer l'opération de soustraction spectrale dans chaque sous-bande permettant d'améliorer la sensibilité du procédé de réduction du bruit.

[0015] Les simulations avec ce procédé qui combine les algorithmes « MCRA » et « SSMB » démontrent une nette amélioration de la qualité audio du signal bruité traité. Pour un large éventail de bruits, le résultat s'avère largement supérieur aux procédés classiques.

[0016] Lors de l'étape a), l'extraction de la phase $\theta_{k,l}$, du signal bruité $Y_{k,l}$ est optionnelle. Il est en effet envisageable, lors de l'étape e), de déterminer le signal débruité de sortie $SD_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$ et du signal bruité $Y_{k,l}$ sans avoir à calculer en préalable sa phase.

[0017] Selon une caractéristique, l'algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon la méthode de moyennage récursif des minima contrôlés dite « MCRA » lors de l'étape b) met en oeuvre les phases de calcul suivantes :

b.1) calcul d'une composante bruitée filtrée $S_{k,l}$ répondant à l'équation :

$$S_{k,l} = \alpha_s S_{k,l-1} + (1 - \alpha_s) |Y_{k,l}|^2$$

où α_s est une constante prédéterminée caractéristique d'un filtre passe-bas ;

b.2) calcul d'une densité de probabilité de présence de parole $\tilde{p}_{k,l}$ par la mise en oeuvre du calcul progressif suivant :

(i) calcul d'une composante minimale spectrale $S_{min,k,l}$ avec :

- si $\text{rem}(k,1) = 0$
alors $S_{min,k,l} = \min(S_{min,k,l-1}; S_{k,l})$ et $S_{tmp,k,l} = S_{k,l}$
- si $\text{rem}(k,1) \neq 0$
alors $S_{min,k,l} = \min(S_{tmp,k,l-1}; S_{k,l})$ et
 $S_{tmp,k,l} = \min(S_{tmp,k,l-1}; S_{k,l})$
où $\text{rem}(k,l)$ est le reste de la division entière de k par l , puis

(ii) calcul d'un rapport spectral $Sr_{k,l}$ répondant à l'équation :

$$Sr_{k,l} = \frac{S_{k,l}}{S_{min,k,l}}$$

(iii) calcul d'une variable indicatrice $I_{k,l}$ avec :

- si $Sr_{k,l} > \delta_{TH}$ alors $I_{k,l} = 1$
- si $Sr_{k,l} \leq \delta_{TH}$ alors $I_{k,l} = 0$

où δ_{TH} est un paramètre prédéterminé de seuil fixe de détection de parole ;

(iv) calcul de la densité de probabilité de présence de parole $\tilde{p}_{k,l}$ avec :

$$\tilde{p}_{k,l} = \alpha_p \tilde{p}_{k,l-1} + (1 - \alpha_p) I_{k,l}$$

où α_p est une constante prédéterminée ;

b.3) calcul d'un coefficient $\tilde{\alpha}_{k,l}$ répondant à l'équation suivante :

$$\tilde{\alpha}_{k,l} = \alpha + (1 - \alpha) \tilde{p}_{k,l} ,$$

où α est une constante prédéterminée ;

b.4) calcul de la composante de bruit $\hat{D}_{k,l}$ répondant à l'équation suivante :

$$\hat{D}_{k,l} = \tilde{\alpha}_{k,l} \hat{D}_{k,l-1} + (1 - \tilde{\alpha}_{k,l}) |Y_{k,l}|^2 .$$

[0018] La méthode dite « MCRA » utilise le principe suivant : sur une fenêtre à horizon fini et pour une fréquence donnée, le minimum de la densité spectrale de puissance (DSP) du signal bruité $Y_{k,l}$ correspond à la valeur de la densité spectrale de puissance (DSP) de la composante de bruit $\hat{D}_{k,l}$. Ainsi, une estimation des différentes valeurs minimales du spectre sur une fenêtre glissante permet d'obtenir une estimation de la densité spectrale de puissance (DSP) de la composante de bruit $\hat{D}_{k,l}$.

[0019] Selon une autre caractéristique, l'algorithme de soustraction spectrale multi-bande dit « SSMB » de l'étape d) met en oeuvre les phases de calcul suivantes, pour chacune des sous-bandes SB_i :

d.1) calcul d'un rapport signal à bruit $SNR_{k,l,i}$ propre à chaque sous-bande SB_i répondant à l'équation :

$$SNR_{k,l,i} = 10 \cdot \log_{10} \left(\frac{\sum_{k=ei}^{bi} |Y_{k,l,i}|^2}{\sum_{k=ei}^{bi} |\hat{D}_{k,l,i}|} \right)$$

d.2) calcul du module carré $|\hat{X}_{k,l,i}|^2$ de la composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i , selon l'équation :

$$|\hat{X}_{k,l,i}|^2 = \begin{cases} |Y_{k,l,i}|^2 - \alpha_i \delta_i |\hat{D}_{k,l,i}| & \text{si } |Y_{k,l,i}|^2 > \alpha_i \delta_i |\hat{D}_{k,l,i}| \\ \beta |Y_{k,l,i}|^2 & \text{si } |Y_{k,l,i}|^2 \leq \alpha_i \delta_i |\hat{D}_{k,l,i}| \end{cases}$$

où - δ_i est un paramètre variable en fonction de la sous-bande SB_i correspondante, prenant des valeurs distinctes d'une sous-bande à l'autre ;

- α_i est un paramètre variable qui dépend de la valeur du rapport signal à bruit $SNR_{k,l,i}$ calculée dans la sous-bande SB_i correspondante ; et
- β est une constante.

[0020] Un tel algorithme est particulièrement avantageux car il prend en compte la valeur du rapport signal à bruit $SNR_{k,l,i}$ propre à chaque sous-bande SB_i pour effectuer la soustraction spectrale dans chaque sous-bande.

[0021] Comme la composante de bruit n'est pas distribué de façon uniforme le long des fréquences, la mise en oeuvre de cet algorithme de soustraction spectrale multi-bande « SSMB » permet de faire des traitements avec des paramètres

différents selon la sous-bande où l'on se trouve. De ce fait, la composante de bruit pourra être réduite de manière plus significative dans une sous-bande si elle y est plus prépondérante, contrairement à un algorithme de soustraction pleine bande qui réduira le bruit de façon égale sur toutes les sous-bandes du spectre et sera donc moins précis et donc moins efficace.

5 **[0022]** Dans une réalisation particulière, les paramètres α_i répondent aux équations suivantes :

$$10 \quad \alpha_i = \begin{cases} \alpha_{c1} & \text{si } \text{SNR}_{k,i} < \text{SNR}_1 \\ \alpha_{c2} + \alpha_{c3} \text{SNR}_{k,i} & \text{si } \text{SNR}_1 \leq \text{SNR}_{k,i} \leq \text{SNR}_2 \\ \alpha_{c4} & \text{si } \text{SNR}_{k,i} > \text{SNR}_2 \end{cases}$$

où α_{c1} , α_{c2} , α_{c3} et α_{c4} sont des constantes prédéterminées, et SNR_1 et SNR_2 sont des seuils prédéterminés.

15 **[0023]** Les valeurs α_{c1} , α_{c2} , α_{c3} et α_{c4} et SNR_1 et SNR_2 sont choisies de façon expérimentale, notamment par simulation numérique. A titre d'exemple non limitatif, les valeurs suivantes ont donné de bons résultats en simulation numérique :

$$\text{SNR}_1 = -5 \text{ dB}, \text{SNR}_2 = 20 \text{ dB} ;$$

20 et

$$\alpha_{c1} = 5, \alpha_{c2} = 4, \alpha_{c3} = -0.15, \alpha_{c4} = 1$$

25 **[0024]** Selon le procédé exemplatif, l'étape e) consiste à déterminer le signal débruité $\hat{X}_{k,i}$ à partir des modules carré $|\hat{X}_{k,i}|^2$ des composantes débruitées de sous-bande $\hat{X}_{k,i}$, et éventuellement des phases $\theta_{k,i}$ extraites lors de l'étape a), de sorte que le signal débruité de sortie $\text{SD}_{k,i}$ corresponde au signal débruité $\hat{X}_{k,i}$, soit $\text{SD}_{k,i} = \hat{X}_{k,i}$.

30 **[0025]** Ainsi, on considère que le signal débruité de sortie $\text{SD}_{k,i}$ corresponde au signal débruité $\hat{X}_{k,i}$ dont les modules carré $|\hat{X}_{k,i}|^2$ sont issus directement de l'étape d) mettant en oeuvre l'algorithme de soustraction spectrale multi-bandes « SSMB ».

[0026] Selon l'invention, en variante de l'exemple, l'étape e) consiste à :

- déterminer, pour chacune des sous-bandes SB_i , le module carré $|\bar{X}_{k,i}|^2$ d'une composante débruitée combinée

35 de sous-bande $\bar{X}_{k,i}$ propre à chaque sous-bande SB_i d'un signal débruité combiné $\bar{X}_{k,i}$, répondant à l'équation

correspondante : $|\bar{X}_{k,i}|^2 = |\bar{X}_{k,i}|^2 + \gamma |Y_{k,i}|^2$ où γ est un coefficient d'amplification prédéterminé, préférentiellement compris entre 0,01 et 0,1 ;

- déterminer un signal débruité combiné $\bar{X}_{k,i}$ à partir des modules carrés $|\bar{X}_{k,i}|^2$ des composantes débruitées

combinées de sous-bande $\bar{X}_{k,i}$, et éventuellement des phases $\theta_{k,i}$ extraites lors de l'étape a), de sorte que le

45 signal débruité de sortie $\text{SD}_{k,i}$ corresponde au signal débruité combiné $\bar{X}_{k,i}$, soit $\text{SD}_{k,i} = \bar{X}_{k,i}$.

50 **[0027]** Le signal traité, c'est-à-dire le signal débruité $\hat{X}_{k,i}$, peut souffrir au niveau de la qualité et de l'intelligibilité d'un problème de distorsion. Cette distorsion, dont l'origine est généralement le traitement de réduction du bruit, dépend des paramètres des algorithmes de réduction du bruit (algorithmes « MCRA » et « SSMB ») mais aussi du niveau et du type de composante de bruit à réduire.

55 **[0028]** Ainsi, dans le cadre de l'invention, on procède, après l'étape d) mettant en oeuvre l'algorithme de soustraction spectrale multi-bandes « SSMB », à une étape de réinjection du signal bruité $Y_{k,i}$ issu du microphone dans le signal débruité $\hat{X}_{k,i}$; cette réinjection étant contrôlée par le coefficient d'amplification (autrement appelé paramètre d'atténuation) choisi très faible, de l'ordre de quelques pourcents. Le signal réinjecté correspond au signal $\gamma |Y_{k,i}|^2$ et il ne présente pas de distorsion, car il n'a pas fait l'objet d'un traitement par les algorithmes « MCRA » et « SSMB ».

[0029] Ce principe de réinjection d'une partie très faible du signal bruité $Y_{k,i}$ issu du microphone permet de remédier

au moins en partie à ce problème de distorsion.

[0030] Dans une première réalisation, l'étape f.2) consiste à reconstruire le signal vocal débruité de sortie $sd(t)$ uniquement à partir des signaux de sortie Sd_i issus de l'étape f.1), ces signaux de sortie Sd_i correspondant aux transformées de Fourier inverse du signal débruité de sortie $SD_{k,i}$ propres à chaque trame temporelle i .

[0031] Ainsi, dans cette première réalisation, le signal vocal débruité de sortie $sd(t)$ est reconstruit à partir des seuls signaux de sortie Sd_i issus de l'étape f.1).

[0032] Dans une seconde réalisation, en variante de la première réalisation, l'étape f.2) consiste à, pour chaque trame temporelle i :

g) calculer un rapport moyen signal à bruit r_i propre à la trame temporelle i à partir du module carré $|Y_{k,i}|^2$ et de la composante de bruit $\hat{D}_{k,i}$;

h) comparer le rapport moyen signal à bruit r_i avec un seuil Ψ_{TH} prédéterminé ;

i) reconstruire le signal vocal débruité de sortie $sd(t)$ à partir en considérant que :

- si le rapport moyen signal à bruit r_i est inférieur audit seuil Ψ_{TH} pour la trame temporelle i , alors le signal considéré avant reconstruction temporelle pour cette trame temporelle i correspond au signal de sortie Sd_i issu de l'étape f.1) ;

- si le rapport moyen signal à bruit r_i est supérieur audit seuil Ψ_{TH} pour la trame temporelle i , alors le signal considéré avant reconstruction temporelle pour cette trame temporelle i correspond au signal échantillonné y_i issu de l'étape de découpage de l'étape a).

[0033] Ainsi, dans cette seconde réalisation, on met en oeuvre :

- lors de l'étape g), une détection de bruit par calcul d'un rapport moyen signal à bruit r_i ;
- lors de l'étape h), une comparaison de ce rapport moyen signal à bruit r_i avec un seuil Ψ_{TH} pour établir si le bruit est présent ($r_i < \Psi_{TH}$) ou si le bruit est absent ou du moins est extrêmement faible ($r_i > \Psi_{TH}$) ; et

- lors de l'étape i), on fait en sorte, pour chaque trame i , que :

- si le bruit est présent ($r_i < \Psi_{TH}$), alors on prend en compte le signal traité numériquement, c'est-à-dire le signal de sortie Sd_i pour la reconstruction temporelle ;

- si le bruit est absent ou extrêmement faible ($r_i > \Psi_{TH}$), alors on ne prend pas en compte le signal de sortie Sd_i mais on prend en compte directement le signal échantillonné y_i pour la reconstruction temporelle, ce qui revient à ignorer les traitements de réduction du bruit (MCRA, SSMB) pour cette trame i , avec l'avantage d'éviter des distorsions inutiles lorsque le niveau de bruit est tel qu'un traitement de réduction du bruit n'est pas nécessaire.

[0034] Cette seconde réalisation permet en quelque sorte la désactivation des étapes b), c) et d) de réduction du bruit lorsque le bruit n'est pas présent. De ce fait, les distorsions qui peuvent être apportées par le traitement de réduction du bruit, dans cette situation d'absence ou de quasi-absence de bruit, seront éliminées.

[0035] Pour rendre le procédé plus robuste, il serait avantageux que la même décision (présence ou non de bruit) soit prise sur une succession de trames temporelles i .

[0036] Dans un mode de réalisation particulier, l'étape g) met en oeuvre l'algorithme de calcul suivant, pour chaque trame temporelle i :

g.1) calcul d'une composante de bruit moyenne $\bar{D}_i$ à partir de la composante de bruit $\hat{D}_{k,i}$ estimée lors de l'étape b) et répondant à l'équation :

$$\bar{D}_i = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,i}$$

où M est une constante prédéterminée, de préférence égale à N ou à $[1+N/2]$, N étant le nombre de points de la transformée de Fourier ;

g.2) calcul d'un module carré moyen $|Y_{k,i}|^2$ du signal bruité $Y_{k,i}$ répondant à l'équation :

$$\overline{|Y_{k,l}|^2} = \frac{1}{M} \sum_{k=0}^{M-1} |Y_{k,l}|^2$$

5 g.3) calcul d'une composante filtrée P_l du module carré moyen $|Y_{k,l}|^2$ répondant à l'équation :

$$P_l = \lambda P_{l-1} + (1 - \lambda) \overline{|Y_{k,l}|^2}$$

10

où - λ est une constante prédéterminée caractéristique d'un filtre passe-bas, de préférence compris entre 0,80 et 0,99 ;

15

- $P_0 = \bar{D}_0 = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,0}$ pour initialiser l'algorithme.

g.4) calcul du rapport moyen signal à bruit r_l répondant à l'équation :

20

- si $\bar{D}_l > 0$ alors $r_l = \frac{P_l}{\bar{D}_l}$,
- si $\bar{D}_l \leq 0$ alors $r_l = 0$.

[0037] Avantageusement, les étapes a) et f) de conversion mettent en oeuvre une méthode de recouvrement et addition dite « OLA », avec :

25

- pour l'étape a), un découpage du signal acoustique bruité $y(t)$ en trames temporelles avec un recouvrement entre les trames temporelles successives ;
- pour l'étape f.2), la reconstruction du signal vocal débruité de sortie $sd(t)$ est réalisée par les additions successives des parties en recouvrement des signaux de deux trames temporelles successives.

30

[0038] Cette méthode de recouvrement et addition dite « OLA » pour « OverLap and Add method » est une méthode classique de reconstruction temporelle qui utilise des fenêtres de pondération juxtaposées (c'est-à-dire des fenêtres qui se superposent ou qui se recouvrent partiellement), puis qui additionne les signaux de sortie en tenant compte du recouvrement des trames temporelles.

35

[0039] Un autre exemple se rapporte également à un système de réduction du bruit sur un signal acoustique bruité $y(t)$ issu d'un microphone opérant dans un milieu bruité, comportant :

40

- une unité de conversion du signal acoustique bruité $y(t)$ dans le domaine temporel en un signal bruité $Y_{k,l}$ dans le domaine fréquentiel, comportant :
 - un module de découpage du signal acoustique bruité $y(t)$ en signaux échantillonnés y_l dans des trames temporelles l successives ;
 - en sortie du module de découpage, un module de fenêtrage des signaux échantillonnés y_l par application d'une fenêtre de pondération ;
 - en sortie du module de fenêtrage, un module de calcul d'une transformée de Fourier discrète qui délivre en sortie le signal bruité $Y_{k,l}$;
- une unité de traitement numérique dans le domaine fréquentiel comportant, en sortie de l'unité de conversion :

50

- un premier module d'extraction du module carré $|Y_{k,l}|^2$ du signal bruité $Y_{k,l}$; et éventuellement un deuxième module d'extraction de la phase $\theta_{k,l}$ du signal bruité $Y_{k,l}$;
- en sortie du premier module d'extraction, un module d'estimation, dit « MCRA », d'une composante de bruit $\bar{D}_{k,l}$ contenue dans le signal bruité $Y_{k,l}$ à partir du module carré $|Y_{k,l}|^2$ issu du premier module d'extraction, par un algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon une méthode de moyennage récursif des minima contrôlés dite « MCRA » ;

55

- en sortie du premier module d'extraction, un module de découpage de la bande fréquentielle en plusieurs sous-bandes fréquentielles $SB_i=[e_i, b_i]$;
- en sortie du module d'estimation « MCRA » et du module de découpage de la bande fréquentielle, un module d'estimation, dit « SSMB », du module carré $|\hat{X}_{k,l,i}|^2$ d'une composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i d'un signal débruité $\hat{X}_{k,l}$, par un algorithme de soustraction spectrale multi-bandes à partir de modules carrés de sous-bande $|Y_{k,l,i}|^2$ et de composantes de bruit de sous-bande $\hat{D}_{k,l,i}$;
- en sortie du module d'estimation « SSMB », et éventuellement du deuxième module d'extraction, un module de détermination d'un signal débruité de sortie $SD_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$, et éventuellement des phases $\theta_{k,l}$;
- une unité de conversion dans le domaine temporel comportant, en sortie de l'unité de traitement numérique :
 - un module de calcul d'un signal de sortie Sd_l propre à chaque trame temporelle l par application d'une transformée de Fourier inverse du signal débruité de sortie $SD_{k,l}$; et
 - un module de reconstruction d'un signal vocal débruité de sortie $sd(t)$ dans le domaine temporel à partir desdits signaux de sortie Sd_l .

[0040] Selon cet exemple, le module de détermination du signal débruité de sortie $SD_{k,l}$ comprend :

- un sous-module racine-carré pour calculer le module $|\hat{X}_{k,l,i}|$ des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$, et
- un sous-module de recombinaison des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$ pour obtenir le signal débruité $\hat{X}_{k,l}$ à partir des modules $|\hat{X}_{k,l,i}|$, et éventuellement des phases $\theta_{k,l}$, de sorte que le signal débruité de sortie $SD_{k,l}$ corresponde au signal débruité $\hat{X}_{k,l}$, soit $SD_{k,l} = \hat{X}_{k,l}$

[0041] Selon l'invention, en variante de cet exemple, le module de détermination d'un signal débruité de sortie $SD_{k,l}$ comprend :

- en sortie du premier module d'extraction, un sous-module d'amplification selon un coefficient d'amplification γ , préférentiellement compris entre 0,01 et 0,1, afin de délivrer un signal amplifié $\gamma |Y_{k,l}|^2$;
- en sortie du module d'estimation « SSMB », un sous-module additionneur propre à additionner le signal amplifié $\gamma |Y_{k,l}|^2$ et les modules carré $|\hat{X}_{k,l,i}|^2$, afin de délivrer en sortie les module carré $|\bar{X}_{k,l,i}|^2$ de composantes débruitées combinées de sous-bande $\bar{X}_{k,l,i}$ propres à chaque sous-bande SB_i d'un signal débruité combiné $\bar{X}_{k,l}$, répondant à l'équation correspondante :

$$|\bar{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2 ;$$

- un sous-module racine-carré pour calculer le module $|\bar{X}_{k,l,i}|$ des composantes débruitées combinées de sous-bande $\bar{X}_{k,l,i}$; et
- un sous-module de recombinaison des composantes combinées de sous-bande $\bar{X}_{k,l,i}$ pour obtenir le signal débruité combiné $\bar{X}_{k,l}$, à partir des modules $|\bar{X}_{k,l,i}|$, et éventuellement des phases $\theta_{k,l}$, de sorte que le signal débruité de sortie $SD_{k,l}$ corresponde au signal débruité combiné $\bar{X}_{k,l}$, soit $SD_{k,l} = \bar{X}_{k,l}$.

[0042] Dans cette seconde possibilité, le système met en oeuvre une étape de réinjection d'une partie très faible du signal bruité $Y_{k,l}$ issu du microphone dans le signal débruité $\hat{X}_{k,l}$ pour remédier au moins en partie aux problèmes de distorsion induits par les modules d'estimation « MCRA » et « SSMB ».

[0043] Dans une réalisation avantageuse, l'unité de traitement numérique comprend en outre, en sortie du module d'estimation « MCRA », un module de détection du bruit comprenant

- un module de calcul d'un rapport moyen signal à bruit r_l propre à chaque trame temporelle l à partir du module carré $|Y_{k,l}|^2$ et de la composante de bruit $\hat{D}_{k,l}$;
- un module de comparaison du rapport moyen signal à bruit r_l propre à chaque trame temporelle l avec un seuil Ψ_{TH} prédéterminé ;
- un module de contrôle du module de reconstruction du signal vocal débruité de sortie $sd(t)$ qui est conçu pour que :
 - si le rapport moyen signal à bruit r_l est inférieur audit seuil Ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction pour cette trame temporelle l correspond au signal de sortie Sd_l issu du module de calcul dudit signal de sortie Sd_l ;
 - si le rapport moyen signal à bruit r_l est supérieur audit seuil Ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction pour cette trame temporelle l correspond au signal échantillonné y_l issu du module de découpage du signal acoustique bruité $y(t)$.

[0044] D'autres caractéristiques et avantages de la présente invention apparaîtront à la lecture de la description détaillée ci-après, d'un exemple de mise en oeuvre non limitatif, faite en référence aux figures annexées dans lesquelles :

- la figure 1 est une vue schématique d'un premier système de réduction du bruit exemplatif;
- la figure 2 est une vue schématique d'un système de réduction du bruit conforme à l'invention ;
- la figure 3 est une vue schématique d'une variante du second système de la figure 2.

[0045] La description d'un système 1 de réduction du bruit sur un signal acoustique bruité $y(t)$ issu d'un microphone unique opérant dans un milieu bruité, et du procédé de réduction du bruit associé, est faite en référence aux figures 1 à 3.

[0046] Dans les trois modes de réalisation, le système 1 comprend une unité de conversion 2 du signal acoustique bruité $y(t)$ dans le domaine temporel en un signal bruité $Y_{k,l}$ dans le domaine fréquentiel.

[0047] Cette unité de conversion 2 comporte un module de découpage 21 du signal acoustique bruité $y(t)$ en signaux échantillonnés y_l dans des trames temporelles l successives.

[0048] Dans ce module de découpage 21, le signal acoustique bruité $y(t)$ est découpé en trames de 240 échantillons, ce qui à une fréquence d'échantillonnage de 8 kHz correspond à des trames temporelles de 30 millisecondes.

[0049] Il est également envisageable d'avoir un découpage en trames de 256 échantillons, ce qui à une fréquence d'échantillonnage de 8 kHz correspond à des trames temporelles de 32 millisecondes.

[0050] Dans ce module de découpage 21, les trames temporelles successives se recouvrent ou se chevauchent. Par exemple, les trames temporelles successives se chevauchent sur 120 échantillons, ce qui correspond à cinquante pourcent (50 %) de recouvrement.

[0051] Ce recouvrement des trames temporelles est destiné à permettre la mise en oeuvre d'une méthode de recouvrement et addition dite « OLA », qui permet le découpage temporel initial du signal acoustique bruité $y(t)$ puis la restitution finale en sortie du système 1 dans le domaine temporel.

[0052] Cette unité de conversion 2 comporte, en sortie du module de découpage 21, un module de fenêtrage 22 des signaux échantillonnés y_l par application d'une fenêtre de pondération, notamment du type fenêtre de Hanning ou fenêtre de Hamming, afin de délivrer en sortie des signaux échantillonnés pondérés $\{y_l\}$.

[0053] Ainsi, les trames temporelles sont alors apodisées avec une fenêtre de pondération, avant d'appliquer une transformée de Fourier, afin de minimiser les effets de bords dus au découpage-recouvrement effectué par le module de découpage 21.

[0054] Cette unité de conversion 2 comporte, en sortie du module de fenêtrage 22, un module de calcul 23 d'une transformée de Fourier discrète qui délivre en sortie le signal bruité $Y_{k,l}$.

[0055] D'un point de vue mathématique, on note :

$$y(t) = x(t) + d(t) ,$$

avec $x(t)$ le signal de parole utile et $d(t)$ la composante de bruit.

[0056] Le module de découpage 21 reçoit en entrée le signal acoustique bruité $y(t)$ et délivre en sortie le signal échantillonné y_l , où l est l'indice temporel (ou indice de la trame temporelle).

[0057] Le module de découpage 21 reçoit en entrée le signal échantillonné y_l et délivre en sortie le signal échantillonné pondéré $\{y_l\}$, avec :

$$\{y_l\} = \omega_l \cdot y_l$$

5 avec ω_l , le signal représentatif de la fenêtre de pondération.

[0058] Le module de calcul 23 reçoit en entrée le signal échantillonné pondéré $\{y_l\}$ et délivre en sortie le signal bruité $Y_{k,l}$ qui correspond à la transformée de Fourier Discrète de $\{y(l)\}$, où k représente l'indice de fréquence, soit :

$$10 \quad Y_{k,l} = \text{DFT}(\{y_l\}) = r_{k,l} + j i_{k,l}$$

[0059] Le calcul de la transformée de Fourier Discrète (DFT pour « Discrete Fourier Transform ») est par exemple réalisé par un calcul de transformée de Fourier rapide (FFT pour « Fast Fourier Transform ») avec une taille N qui peut être égale à 256 (N correspond au nombre de points de la transformée de Fourier).

[0060] Le système 1 comprend, en sortie de l'unité de conversion 2, une unité de traitement numérique 3 dans le domaine fréquentiel qui réalise le traitement de débruitage ou de rehaussement de la parole sur le signal bruité $Y_{k,l}$.

[0061] Cette unité de traitement numérique 3 comprend un premier module d'extraction 31 du module carré $|Y_{k,l}|^2$ du signal bruité $Y_{k,l}$.

20 **[0062]** D'un point de vue mathématique, le premier module d'extraction 31 réalise le calcul suivant :

$$|Y_{k,l}|^2 = r_{k,l}^2 + i_{k,l}^2.$$

25 **[0063]** Dans les modes de réalisation des figures 1 et 2, cette unité de traitement numérique 3 comprend un deuxième module d'extraction 32 de la phase $\theta_{k,l}$ du signal bruité $Y_{k,l}$. Comme détaillé ultérieurement, il est également envisageable de se passer de ce deuxième module d'extraction 32.

[0064] D'un point de vue mathématique, le deuxième module d'extraction 32 réalise le calcul suivant :

30

$$\cos \theta_{k,l} = \frac{r_{k,l}}{|Y_{k,l}|}, \text{ et } \sin \theta_{k,l} = \frac{i_{k,l}}{|Y_{k,l}|}$$

35

avec

$$|Y_{k,l}| = \sqrt{r_{k,l}^2 + i_{k,l}^2}$$

40

[0065] L'unité de traitement numérique 3 comprend, en sortie du premier module d'extraction 31, un module d'estimation 33, dit « MCRA », d'une composante de bruit $\hat{D}_{k,l}$ contenue dans le signal bruité $Y_{k,l}$ à partir du module carré $|Y_{k,l}|^2$ issu du premier module d'extraction 31, par un algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon une méthode de moyennage récursif des minima contrôlés dite « MCRA ».

45

[0066] D'un point de vue mathématique, l'algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon la méthode de moyennage récursif des minima contrôlés dite « MCRA » met en oeuvre les phases de calcul suivantes :

50

b.1) calcul d'une composante bruitée filtrée $S_{k,l}$ répondant à l'équation :

$$S_{k,l} = \alpha_s S_{k,l-1} + (1 - \alpha_s) |Y_{k,l}|^2$$

55

où α_s est une constante prédéterminée caractéristique d'un filtre passe-bas ;

b.2) calcul d'une densité de probabilité de présence de parole $\tilde{p}_{k,l}$ par la mise en oeuvre du calcul progressif suivant :

(i) calcul d'une composante minimale spectrale $S_{min_{k,l}}$ avec :

- si $rem(k,1) = 0$

alors $S_{min_{k,l}} = \min (S_{min_{k,l-1}} ; S_{k,l})$ et $Stmp_{k,l} = S_{k,l}$

- si $rem(k,l) \neq 0$

alors $S_{min_{k,l}} = \min (Stmp_{k,l-1} ; S_{k,l})$ et

$Stmp_{k,l} = \min (Stmp_{k,l-1} ; S_{k,l})$

où $rem(k,l)$ est le reste de la division entière de k par l , puis

(ii) calcul d'un rapport spectral $Sr_{k,l}$ répondant à l'équation :

$$Sr_{k,l} = \frac{S_{k,l}}{S_{min_{k,l}}}$$

(iii) calcul d'une variable indicatrice $I_{k,l}$ avec :

- si $Sr_{k,l} > \delta_{TH}$ alors $I_{k,l} = 1$

- si $Sr_{k,l} \leq \delta_{TH}$ alors $I_{k,l} = 0$

où δ_{TH} est un paramètre prédéterminé de seuil fixe de détection de parole ;

(iv) calcul de la densité de probabilité de présence de parole $\tilde{p}_{k,l}$ avec :

$$\tilde{p}_{k,l} = \alpha_p \tilde{p}_{k,l-1} + (1 - \alpha_p) I_{k,l}$$

où α_p est une constante prédéterminée ;

b.3) calcul d'un coefficient $\tilde{\alpha}_{k,l}$ répondant à l'équation suivante :

$$\tilde{\alpha}_{k,l} = \alpha + (1 - \alpha) \tilde{p}_{k,l} ,$$

où α est une constante prédéterminée ;

b.4) calcul de la composante de bruit $\hat{D}_{k,l}$ répondant à l'équation suivante :

$$\hat{D}_{k,l} = \tilde{\alpha}_{k,l} \hat{D}_{k,l-1} + (1 - \tilde{\alpha}_{k,l}) |Y_{k,l}|^2.$$

[0067] Le module d'estimation « MCRA » 33 délivre ainsi en sortie la composante de bruit $\hat{D}_{k,l}$.

[0068] L'unité de traitement numérique 3 comprend, en sortie du premier module d'extraction 31, un module de découpage 34 de la bande fréquentielle en plusieurs sous-bandes fréquentielles $SB_i = [e_i, b_i]$.

[0069] La bande fréquentielle peut par exemple être divisée en trois sous-bandes fréquentielles, à savoir SB_1 pour $f_i < 1000$ Hz, SB_2 pour $1000 \text{ Hz} \leq f_i \leq 2000$ Hz et enfin SB_3 pour $f_i > 2000$ Hz, où f_i est la fréquence de sous-bande.

[0070] L'unité de traitement numérique 3 comprend, en sortie du module d'estimation « MCRA » 33 et du module de découpage 34 de la bande fréquentielle, un module d'estimation 35, dit « SSMB », du module carré $|\hat{X}_{k,l,i}|^2$ d'une composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i d'un signal débruité $\hat{X}_{k,l}$, par un algorithme de soustraction spectrale multi-bandes à partir de modules carrés de sous-bande $|Y_{k,l,i}|^2$ et de composantes de bruit de sous-bande $\hat{D}_{k,l,i}$.

[0071] Le principe de l'algorithme de soustraction spectrale multi-bandes (SSMB) procède d'une généralisation de l'algorithme de soustraction spectrale qui consiste à soustraire de la densité spectrale de puissance du signal bruité $Y_{k,l}$ issu du microphone une portion de la densité spectrale de puissance de la composante de bruit estimé par la méthode

« MCRA ». L'algorithme de soustraction spectrale multi-bandes applique le même principe en découpant l'espace spectral en plusieurs sous-bandes SB_i fréquentielles et, ensuite, dans chaque sous-bande SB_i , l'opération de soustraction spectrale est appliquée, comme suit :

$$|\hat{X}_{k,l,i}|^2 = |Y_{k,l,i}|^2 - \mu_i |\hat{D}_{k,l,i}|^2$$

avec μ_i un coefficient prédéterminé.

[0072] Cet algorithme de soustraction spectrale multi-bandes est combiné avec la méthode de recouvrement et addition dite « OLA ».

[0073] Il est bien entendu envisageable d'affiner la relation de soustraction spectrale multi-bandes donnée ci-dessus, comme décrit ci-après.

[0074] D'un point de vue mathématique, l'algorithme de soustraction spectrale multi-bande « SSMB » mis en oeuvre par le module d'estimation « SSMB » 35, réalise les phases de calcul suivantes, pour chacune des sous-bandes SB_i :

d.1) calcul d'un rapport signal à bruit $SNR_{k,l,i}$ propre à chaque sous-bande SB_i répondant à l'équation :

$$SNR_{k,l,i} = 10 \cdot \log_{10} \left(\frac{\sum_{k=ei}^{bi} |Y_{k,l,i}|^2}{\sum_{k=ei}^{bi} |\hat{D}_{k,l,i}|^2} \right)$$

d.2) calcul du module carré $|\hat{X}_{k,l,i}|^2$ de la composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i , selon l'équation :

$$|\hat{X}_{k,l,i}|^2 = \begin{cases} |Y_{k,l,i}|^2 - \alpha_i \delta_i |\hat{D}_{k,l,i}|^2 & \text{si } |Y_{k,l,i}|^2 > \alpha_i \delta_i |\hat{D}_{k,l,i}|^2 \\ \beta |Y_{k,l,i}|^2 & \text{si } |Y_{k,l,i}|^2 \leq \alpha_i \delta_i |\hat{D}_{k,l,i}|^2 \end{cases}$$

où - δ_i est un paramètre variable en fonction de la sous-bande SB_i correspondante, prenant des valeurs distinctes d'une sous-bande à l'autre ;

- α_i est un paramètre variable qui dépend de la valeur du rapport signal à bruit $SNR_{k,l,i}$ calculée dans la sous-bande SB_i correspondante ; et
- β est une constante.

[0075] Il est envisageable de comparer la valeur du rapport signal à bruit $SNR_{k,l,i}$ avec deux seuils SNR_1 et SNR_2 pour établir la valeur des paramètres α_i . Ainsi, les paramètres α_i répondent aux équations suivantes :

$$\alpha_i = \begin{cases} \alpha_{c1} & \text{si } SNR_{k,l,i} < SNR_1 \\ \alpha_{c2} + \alpha_{c3} SNR_{k,l,i} & \text{si } SNR_1 \leq SNR_{k,l,i} \leq SNR_2 \\ \alpha_{c4} & \text{si } SNR_{k,l,i} > SNR_2 \end{cases}$$

où α_{c1} , α_{c2} , α_{c3} et α_{c4} sont des constantes prédéterminées, et SNR_1 et SNR_2 sont les seuils prédéterminés.

[0076] Par exemple, il est envisageable d'avoir les relations suivantes pour les paramètres α_i :

$$\alpha_i = \begin{cases} 5 & \text{si } SNR_{k,l,i} < -5 \\ 4 - \frac{3}{20} SNR_{k,l,i} & \text{si } -5 \leq SNR_{k,l,i} \leq 20 \\ 1 & \text{si } SNR_{k,l,i} > 20 \end{cases}$$

[0077] Concernant les paramètres variables δ_i , il envisageable d'avoir les relations suivantes, dans le cas de la division en trois sous-bandes fréquentielles décrite ci-dessus :

$$\delta_i = \begin{cases} 1 & \text{si } f_i < 1000 \text{ (première sous – bande SB}_1\text{)} \\ 2.75 & \text{si } 1000 \leq f_i \leq 2000 \text{ (deuxième sous – bande SB}_2\text{)} \\ 1.75 & \text{si } f_i > 2000 \text{ (troisième sous – bande SB}_3\text{)} \end{cases}$$

[0078] Concernant la constante β , il est envisageable d'avoir des valeurs de l'ordre de 0,002 ou 0,0015, soit par exemple $\beta = 0.002$ ou $\beta = 0.0015$.

[0079] L'unité de traitement numérique 3 comprend également, en sortie du module d'estimation « SSMB » 35 et du deuxième module d'extraction 32, un module de détermination 36 d'un signal débruité de sortie $SD_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$ et des phases $\theta_{k,l}$.

[0080] Dans le premier mode de réalisation illustré sur la figure 1, le module de détermination 36 du signal débruité de sortie $SD_{k,l}$ comprend :

- un sous-module racine-carré 361 pour calculer le module $|\hat{X}_{k,l,i}|$ des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$, et
- un sous-module de recombinaison 362 des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$ pour obtenir le signal débruité $\hat{X}_{k,l}$ à partir des modules $|\hat{X}_{k,l,i}|$ et des phases $\theta_{k,l}$, de sorte que le signal débruité de sortie $SD_{k,l}$ corresponde au signal débruité $\hat{X}_{k,l}$, soit $SD_{k,l} = \hat{X}_{k,l}$.

[0081] D'un point de vue mathématique, on a :

- le sous-module racine-carré 361 réalise le calcul : $|\hat{X}_{k,l}| = \sqrt{|\hat{X}_{k,l}|^2}$; et
- le sous-module de recombinaison 362 effectue la réinjection de la phase, comme suit : $SD_{k,l} = \hat{X}_{k,l} = |\hat{X}_{k,l}| \cos \theta_{k,l} + j |\hat{X}_{k,l}| \sin \theta_{k,l}$.

[0082] Dans le second mode de réalisation illustré sur la figure 2, le module de détermination 36 du signal débruité de sortie $SD_{k,l}$ comprend :

- en sortie du premier module d'extraction 31, un sous-module d'amplification 363 selon un coefficient d'amplification γ , préférentiellement compris entre 0,01 et 0,1, afin de délivrer un signal amplifié $\gamma |Y_{k,l}|^2$;
- en sortie du module d'estimation « SSMB » 35, un sous-module additionneur 364 propre à additionner le signal amplifié $\gamma |Y_{k,l}|^2$ et les modules carré $|\hat{X}_{k,l,i}|^2$, afin de délivrer en sortie les module carré $|\hat{X}_{k,l,i}|^2$ de composantes débruitées combinées de sous-bande $\hat{X}_{k,l,i}$ propres à chaque sous-bande SB_i d'un signal débruité combiné $\hat{X}_{k,l}$, répondant à l'équation correspondante :

$$|\hat{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2 ;$$

- un sous-module racine-carré 361 pour calculer le module $|\hat{X}_{k,l,i}|$ des composantes débruitées combinées de sous-bande $\hat{X}_{k,l,i}$; et
- un sous-module de recombinaison 362 des composantes combinées de sous-bande $\hat{X}_{k,l,i}$ pour obtenir le signal débruité combiné $\hat{X}_{k,l}$, à partir des modules $|\hat{X}_{k,l,i}|$ et des phases $\theta_{k,l}$, de sorte que le signal débruité de sortie $SD_{k,l}$ corresponde au signal débruité combiné $\hat{X}_{k,l}$, soit $SD_{k,l} = \hat{X}_{k,l}$.

[0083] D'un point de vue mathématique, on a :

- le sous-module d'amplification 363 délivre en sortie le signal amplifié

$$\gamma |Y_{k,l}|^2 ;$$

- 5 - le sous-module additionneur 364 réalise le calcul :

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2$$

- 10 - le sous-module racine-carré 361 réalise le calcul : $|\tilde{X}_{k,l,i}| = \sqrt{|\tilde{X}_{k,l,i}|^2}$; et
- le sous-module de recombinaison 362 effectue la réinjection de la phase, comme suit : $SD_{k,l} = \hat{X}_{k,l} = |\hat{X}_{k,l}| \cos \theta_{k,l} + j |\hat{X}_{k,l}| \sin \theta_{k,l}$

15 **[0084]** Ainsi, dans ce second mode de réalisation, le système 1 met en oeuvre une étape de réinjection d'une partie très faible du signal bruité $Y_{k,l}$, issu du microphone, dans le signal débruité $\hat{X}_{k,l}$; le signal réinjecté correspondant au signal amplifié $\gamma |Y_{k,l}|^2$. De cette manière, on remédie au moins en partie aux problèmes de distorsion introduits par les modules d'estimation « MCRA » 33 et « SSMB » 35.

20 **[0085]** Le système 1 comprend, en sortie de l'unité de traitement numérique 3, une unité de conversion 4 dans le domaine temporel.

[0086] Cette unité de conversion 4 comporte un module de calcul 41 d'un signal de sortie Sd_l propre à chaque trame temporelle l par application d'une transformée de Fourier inverse du signal débruité de sortie $SD_{k,l}$.

25 **[0087]** D'un point de vue mathématique, ce module de calcul 41 met en oeuvre le calcul de la transformée de Fourier rapide inverse (IFFT pour « Inverse Fast Fourier Transform ») avec une taille N égale à 256 (N correspondant pour rappel au nombre de points de la transformée de Fourier), avec la relation suivante :

$$Sd_l = \text{IDFT}(SD_{k,l})$$

30 où IDFT correspond à la fonction de transformée de Fourier discrète inverse, qui peut être du type transformée de Fourier rapide inverse (IFFT).

35 **[0088]** En outre, du fait de la symétrie de l'amplitude de la transformée de Fourier (DFT ou FFT) des signaux réels, le traitement de réduction du bruit et de reconstruction selon la méthode « OLA » se fera uniquement sur les premiers $[N/2 + 1]$ premiers points d'échantillonnage, soit sur les 129 premiers points d'échantillonnage pour N égal à 256, sachant que l'on a la relation suivante de symétrie hermitienne :

$$SD_{-k,l} = SD_{k,l}^*$$

40 **[0089]** Dans le cas du premier mode de réalisation illustré sur la figure 1, on a la relation suivante : $Sd_l = \hat{x}_l = \text{IFFT}(\hat{X}_{k,l})$.

[0090] Dans le cas du second mode de réalisation illustré sur la figure 2, on a la relation suivante : $Sd_l = \hat{x}_l = \text{IFFT}(\hat{X}_{k,l})$.

45 **[0091]** Cette unité de conversion 4 comporte, en sortie du module de calcul 41, un module de reconstruction 42 d'un signal vocal débruité de sortie $sd(t)$ dans le domaine temporel à partir des signaux de sortie Sd_l .

[0092] Le signal temporel est restitué selon la méthode de recouvrement et addition « OLA », la reconstruction du signal vocal débruité de sortie $sd(t)$ étant réalisée par les additions successives des parties en recouvrement des signaux de deux trames temporelles successives, selon le principe :

$$\{\widetilde{sd}_l\} = \{Sd_{l-1}\} + \{Sd_l\},$$

55 qui se traduit dans le premier mode de réalisation par : $\{\widetilde{sd}_l\} = \{\hat{x}_{l-1}\} + \{\hat{x}_l\}$, et qui se traduit dans le second mode de réalisation par : $\{\widetilde{sd}_l\} = \{\tilde{x}_{l-1}\} + \{\tilde{x}_l\}$.

[0093] A chaque fois qu'une trame temporelle du signal débruité de sortie $SD_{k,l}$ est délivrée dans le domaine fréquentiel et que sa transformée de Fourier inverse Sd_l est calculée, les premiers $N/2$ points d'échantillonnage seront additionnés avec les derniers $N/2$ points d'échantillonnage de la trame traitée précédente. Les derniers $N/2$ points d'échantillonnage de la trame traitée en cours seront quand à eux stockés en mémoire pour être à leur tour utilisés lors du traitement de la trame suivante.

[0094] Autrement dit, après traitement dans le domaine spectral, la trame temporelle du signal débruité de sortie $SD_{k,l}$ passe par le module de calcul 41 de la transformée de Fourier inverse, puis sa première moitié ($N/2$ premiers points d'échantillonnage) est additionnée avec la seconde moitié ($N/2$ derniers points d'échantillonnage) sauvegardée de la trame précédente, tandis que sa seconde moitié ($N/2$ premiers points d'échantillonnage) est sauvegardée pour le prochain bloc.

[0095] Avec N égal à 256, on a $N/2$ qui est égal à 128, étant rappelé que le taux de recouvrement de deux trames successives est de cinquante pourcent (50 %).

[0096] De manière optionnelle et avantageuse, l'unité de traitement numérique 3 comprend en outre, en sortie du module d'estimation « MCRA » 33, un module de détection du bruit 37 qui pilote le module de reconstruction 42 selon le principe suivant :

- mise en oeuvre d'un calcul d'un rapport moyen signal à bruit r_l pour effectuer une détection de bruit ;
- comparaison de ce rapport moyen signal à bruit r_l avec un seuil ψ_{TH} pour établir si le bruit est présent ($r_l < \psi_{TH}$) ou si le bruit est absent ou du moins extrêmement faible ($r_l > \psi_{TH}$) ; et
- pilotage du module de reconstruction 42 selon les règles suivantes :
- si le bruit est présent ($r_l < \psi_{TH}$), alors on prend en compte le signal traité numériquement, c'est-à-dire le signal de sortie Sd_l pour la reconstruction temporelle ;
- si le bruit est absent ou extrêmement faible ($r_l > \psi_{TH}$), alors on ne prend pas en compte le signal de sortie Sd_l mais on prend en compte directement le signal échantillonné y_l pour la reconstruction temporelle, ce qui revient à ignorer les traitements de réduction du bruit (« MCRA », « SSMB ») pour cette trame l , avec l'avantage d'éviter des distorsions inutiles lorsque le niveau de bruit est tel qu'un traitement de réduction du bruit n'est pas nécessaire.

[0097] De ce fait, dans une situation d'absence ou quasi absence de bruit, les distorsions qui peuvent être apportées par le traitement de réduction du bruit seront éliminées.

[0098] D'un point de vue structurel, le module de détection du bruit 37 comprend :

- un module de calcul d'un rapport moyen signal à bruit r_l propre à chaque trame temporelle 1 à partir du module carré $|Y_{k,l}|^2$ et de la composante de bruit $D_{k,l}$;
- un module de comparaison du rapport moyen signal à bruit r_l propre à chaque trame temporelle l avec un seuil ψ_{TH} prédéterminé ;
- un module de contrôle du module de reconstruction 42 du signal vocal débruité de sortie $sd(t)$ qui est conçu pour que :
 - si le rapport moyen signal à bruit r_l est inférieur audit seuil ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction pour cette trame temporelle l correspond au signal de sortie Sd_l issu du module de calcul dudit signal de sortie Sd_l ;
 - si le rapport moyen signal à bruit r_l est supérieur audit seuil ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction pour cette trame temporelle l correspond au signal échantillonné y_l issu du module de découpage du signal acoustique bruité $y(t)$.

[0099] D'un point de vue mathématique, le module de détection du bruit 37 met en oeuvre l'algorithme de calcul suivant, pour chaque trame temporelle l :

g.1) calcul d'une composante de bruit moyenne $\bar{D}_l$ à partir de la composante de bruit $\hat{D}_{k,l}$ estimée par le module d'estimation « MCRA » 33 et répondant à l'équation :

$$\bar{D}_l = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,l}$$

où M est une constante prédéterminée égale à $N/2$, N étant pour rappel le nombre de points de la transformée de Fourier ;

g.2) calcul d'un module carré moyen $|Y_{k,l}|^2$ du signal bruité $Y_{k,l}$ répondant à l'équation :

$$\overline{|Y_{k,l}|^2} = \frac{1}{M} \sum_{k=0}^{M-1} |Y_{k,l}|^2$$

g.3) calcul d'une composante filtrée P_l du module carré moyen $|Y_{k,l}|^2$ répondant à l'équation :

$$P_l = \lambda P_{l-1} + (1 - \lambda) \overline{|Y_{k,l}|^2}$$

où - λ est une constante prédéterminée caractéristique d'un filtre passe-bas, de préférence compris entre 0,80 et 0,99 ;

- $P_0 = \bar{D}_0 = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,0}$ pour initialiser l'algorithme.

g.4) calcul du rapport moyen signal à bruit r_l répondant à l'équation :

- si $\bar{D}_l > 0$ alors $r_l = \frac{P_l}{\bar{D}_l}$,

- si $\bar{D}_l < 0$ alors $r_l = 0$.

[0100] Bien entendu l'exemple de mise en oeuvre évoqué ci-dessus ne présente aucun caractère limitatif et d'autres améliorations et détails peuvent être apportés au système de réduction selon l'invention, sans pour autant sortir du cadre de l'invention.

[0101] Ainsi, il est envisageable de se passer de l'extraction de la phase de la phase $\theta_{k,l}$ du signal bruité $Y_{k,l}$ comme illustré sur la figure 3 où le système 1 ne comporte pas le deuxième module d'extraction 32 de la phase $\theta_{k,l}$ du signal bruité $Y_{k,l}$. Dans ce cas, le sous-module de recombinaison 362 des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$ effectue le calcul du signal débruité $\hat{X}_{k,l}$ à partir des modules $|X_{k,l,i}|$ et du signal bruité $Y_{k,l}$. En effet, la réinjection de la phase peut être effectuée directement à partir du signal bruité $Y_{k,l}$ qui intègre intrinsèquement cette phase.

[0102] A cet effet, le sous-module de recombinaison 362 réalise le calcul suivant :

$$\hat{X}_{k,l} = |\hat{X}_{k,l}| \frac{r_{k,l}}{|Y_{k,l}|} + j |\hat{X}_{k,l}| \frac{i_{k,l}}{|Y_{k,l}|} = G_{k,l} r_{k,l} + j G_{k,l} i_{k,l},$$

soit $\hat{X}_{k,l} = G_{k,l} Y_{k,l}$.

[0103] Avec $G_{k,l}$ le gain de l'algorithme de réduction de bruit. Avec ce calcul, il n'est donc plus nécessaire de calculer, stocker et réinjecter la phase.

[0104] Le système 1 de la figure 3 est une variante du système 1 de la figure 2 sans le deuxième module d'extraction 32, mais il est bien entendu également envisageable de prévoir la suppression du deuxième module d'extraction 32 dans le système 1 de la figure 1.

Revendications

1. Procédé de réduction du bruit sur un signal acoustique bruité $y(t)$ issu d'un microphone opérant dans un milieu bruité, comportant les étapes successives suivantes :

a) conversion du signal acoustique bruité $y(t)$ dans le domaine temporel en un signal bruité $Y_{k,l}$ dans le domaine fréquentiel, par découpage temporel du signal acoustique bruité $y(t)$ en signaux échantillonnés y_l dans des trames temporelles l successives, fenêtrage des signaux échantillonnés y_l par application d'une fenêtre de pondération, et application d'une transformée de Fourier discrète, avec extraction du module carré $|Y_{k,l}|^2$, et éventuellement de la phase $\theta_{k,l}$ du signal bruité $Y_{k,l}$;

b) estimation d'une composante de bruit $\hat{D}_{k,l}$ contenue dans le signal bruité $Y_{k,l}$ à partir du module carré $|Y_{k,l}|^2$, par un algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon une méthode de moyennage récursif des minima contrôlés dite « MCRA » ;

caractérisé en ce qu'il comprend en outre, après l'étape b), les étapes successives suivantes :

c) découpage de la bande fréquentielle en plusieurs sous-bandes fréquentielles $SB_i = [e_i, b_i]$, suivie d'une décomposition multi-bandes du module carré $|Y_{k,l}|^2$ et de la composante de bruit $\hat{D}_{k,l}$, consistant à décomposer le module carré $|Y_{k,l}|^2$ et la composante de bruit $\hat{D}_{k,l}$ en respectivement plusieurs modules carré de sous-bande $|Y_{k,l,i}|^2$ et plusieurs composantes de bruit de sous-bande $\hat{D}_{k,l,i}$ propres à chacune des sous-bandes SB_i ;

d) estimation, pour chacune des sous-bandes SB_i , du module carré $|\hat{X}_{k,l,i}|^2$ d'une composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i , d'un signal débruité $X_{k,l}$, par un algorithme de soustraction spectrale multi-bandes dit « SSMB » à partir des modules carrés de sous-bande $|Y_{k,l,i}|^2$ et des composantes de bruit de sous-bande $\hat{D}_{k,l,i}$;

e) détermination d'un signal débruité de sortie $SD_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$ issus de l'étape d), et éventuellement des phases $\theta_{k,l}$ extraites lors de l'étape a) ;

f) conversion du signal débruité de sortie $SD_{k,l}$ en un signal vocal débruité de sortie $sd(t)$ dans le domaine temporel, par une étape f.1) de calcul d'un signal échantillonné de sortie Sd_l propre à chaque trame temporelle l par application d'une transformée de Fourier inverse du signal débruité de sortie $SD_{k,l}$, suivie d'une étape f.2) de reconstruction temporelle du signal vocal débruité de sortie $sd(t)$ à partir des signaux échantillonnés de sortie Sd_l ;

et dans lequel l'étape e) consiste à :

- déterminer, pour chacune des sous-bandes SB_i , le module carré $|\tilde{X}_{k,l,i}|^2$ d'une composante débruitée

combinée de sous-bande $\tilde{X}_{k,l,i}$ propre à chaque sous-bande SB_i d'un signal débruité combiné $\tilde{X}_{k,l}$, répondant à l'équation correspondante :

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2$$

où γ est un coefficient d'amplification prédéterminé, préférentiellement compris entre 0,01 et 0,1 ;

- déterminer un signal débruité combiné $\tilde{X}_{k,l}$ à partir des modules carrés $|\tilde{X}_{k,l,i}|^2$ des composantes débruitées combinées de sous-bande $\tilde{X}_{k,l,i}$, et éventuellement des phases $\theta_{k,l}$ extraites lors de l'étape

a), de sorte que le signal débruité de sortie $SD_{k,l}$ corresponde au signal débruité combiné $\tilde{X}_{k,l}$, soit

$$SD_{k,l} = \tilde{X}_{k,l}.$$

2. Procédé selon la revendication 1, dans lequel l'algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon la méthode de moyennage récursif des minimas contrôlés dite « MCRA » lors de l'étape b) met en oeuvre les phases de calcul suivantes :

b.1) calcul d'une composante bruitée filtrée $S_{k,l}$ répondant à l'équation :

$$S_{k,l} = \alpha_s S_{k,l-1} + (1 - \alpha_s) |Y_{k,l}|^2$$

où α_s est une constante prédéterminée caractéristique d'un filtre passe-bas ;

b.2) calcul d'une densité de probabilité de présence de parole $\tilde{p}_{k,l}$ par la mise en oeuvre du calcul progressif suivant :

(i) calcul d'une composante minimale spectrale $S_{min,k,l}$ avec :

- si $\text{rem}(k,l) = 0$
alors $S_{min,k,l} = \min(S_{min,k,l-1}; S_{k,l})$ et $\text{Stmp}_{k,l} = S_{k,l}$
- si $\text{rem}(k,l) \neq 0$
alors $S_{min,k,l} = \min(\text{Stmp}_{k,l-1}; S_{k,l})$ et

$$\text{Stmp}_{k,l} = \min (\text{Stmp}_{k,l-1} ; S_{k,l})$$

où $\text{rem}(k,l)$ est le reste de la division entière de k par l , puis

(ii) calcul d'un rapport spectral $\text{Sr}_{k,l}$ répondant à l'équation :

$$\text{Sr}_{k,l} = \frac{S_{k,l}}{S_{\min_{k,l}}}$$

(iii) calcul d'une variable indicatrice $I_{k,l}$ avec :

- si $\text{Sr}_{k,l} > \delta_{TH}$ alors $I_{k,l} = 1$

- si $\text{Sr}_{k,l} \leq \delta_{TH}$ alors $I_{k,l} = 0$

où δ_{TH} est un paramètre prédéterminé de seuil fixe de détection de parole ;

(iv) calcul de la densité de probabilité de présence de parole $\tilde{p}_{k,l}$ avec :

$$\tilde{p}_{k,l} = \alpha_p \tilde{p}_{k,l-1} + (1 - \alpha_p) I_{k,l}$$

où α_p est une constante prédéterminée ;

b.3) calcul d'un coefficient $\alpha_{p,l}$ répondant à l'équation suivante :

$$\tilde{\alpha}_{k,l} = \alpha + (1 - \alpha) \tilde{p}_{k,l} ,$$

où α est une constante prédéterminée ;

b.4) calcul de la composante de bruit $\hat{D}_{k,l}$ répondant à l'équation suivante :

$$\hat{D}_{k,l} = \tilde{\alpha}_{k,l} \hat{D}_{k,l-1} + (1 - \tilde{\alpha}_{k,l}) |Y_{k,l}|^2.$$

3. Procédé selon l'une quelconque des revendications 1 et 2, dans lequel l'algorithme de soustraction spectrale multi-bande dit « SSMB » de l'étape d) met en oeuvre les phases de calcul suivantes, pour chacune des sous-bandes SB_i :

d.1) calcul d'un rapport signal à bruit $\text{SNR}_{k,l,i}$ propre à chaque sous-bande SB_i répondant à l'équation :

$$\text{SNR}_{k,l,i} = 10. \log_{10} \left(\frac{\sum_{k=ei}^{bi} |Y_{k,l,i}|^2}{\sum_{k=ei}^{bi} |\hat{D}_{k,l,i}|} \right)$$

d.2) calcul du module carré $|\hat{X}_{k,l,i}|^2$ de la composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i , selon l'équation :

$$|\hat{X}_{k,l,i}|^2 = \begin{cases} |Y_{k,l,i}|^2 - \alpha_i \delta_i |\hat{D}_{k,l,i}| & \text{si } |Y_{k,l,i}|^2 > \alpha_i \delta_i |\hat{D}_{k,l,i}| \\ \beta |Y_{k,l,i}|^2 & \text{si } |Y_{k,l,i}|^2 \leq \alpha_i \delta_i |\hat{D}_{k,l,i}| \end{cases}$$

où δ_i est un paramètre variable en fonction de la sous-bande SB_i correspondante, prenant des valeurs distinctes d'une sous-bande à l'autre ;

- α_i est un paramètre variable qui dépend de la valeur du rapport signal à bruit $SNR_{k,l,i}$ calculée dans la sous-bande SB_i correspondante ; et
- β est une constante.

4. Procédé selon la revendication 3, dans lequel les paramètres α_i répondent aux équations suivantes :

$$\alpha_i = \begin{cases} \alpha_{c1} & \text{si } SNR_{k,l,i} < SNR_1 \\ \alpha_{c2} + \alpha_{c3} SNR_{k,l,i} & \text{si } SNR_1 \leq SNR_{k,l,i} \leq SNR_2 \\ \alpha_{c4} & \text{si } SNR_{k,l,i} > SNR_2 \end{cases}$$

où α_{c1} , α_{c2} , α_{c3} et α_{c4} sont des constantes prédéterminées, et SNR_1 et SNR_2 sont des seuils prédéterminés.

5. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel l'étape e) consiste à déterminer le signal débruité $\hat{X}_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$ des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$ et éventuellement des phases $\theta_{k,l}$ extraites lors de l'étape a), de sorte que le signal débruité de sortie $SD_{k,l}$ correspond au signal débruité $\hat{X}_{k,l}$, soit $SD_{k,l} = X_{k,l}$.

6. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel l'étape f.2) consiste à reconstruire le signal vocal débruité de sortie $sd(t)$ uniquement à partir des signaux de sortie Sd_l issus de l'étape f.1), lesdits signaux de sortie Sd_l correspondant aux transformées de Fourier inverse du signal débruité de sortie $SD_{k,l}$ propres à chaque trame temporelle l .

7. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel l'étape f.2) consiste à, pour chaque trame temporelle l :

- g) calculer un rapport moyen signal à bruit r_l propre à la trame temporelle l à partir du module carré $|Y_{k,l}|^2$ et de la composante de bruit $\hat{D}_{k,l}$;
- h) comparer le rapport moyen signal à bruit r_l avec un seuil ψ_{TH} prédéterminé ;
- i) reconstruire le signal vocal débruité de sortie $sd(t)$ à partir en considérant que :

- si le rapport moyen signal à bruit r_l est inférieur audit seuil ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction temporelle pour cette trame temporelle l correspond au signal de sortie Sd_l issu de l'étape f.1) ;
- si le rapport moyen signal à bruit r_l est supérieur audit seuil ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction temporelle pour cette trame temporelle l correspond au signal échantillonné y_l issu de l'étape de découpage de l'étape a).

8. Procédé selon la revendication 7, dans lequel l'étape g) met en oeuvre l'algorithme de calcul suivant, pour chaque trame temporelle l :

- g.1) calcul d'une composante de bruit moyenne $\bar{D}_l$ à partir de la composante de bruit $\hat{D}_{k,l}$ estimée lors de l'étape b) et répondant à l'équation :

$$\bar{D}_l = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,l}$$

où M est une constante prédéterminée, de préférence égale à N ou à $N/2$, N étant le nombre de points d'échantillonnage de la transformée de Fourier ;

- g.2) calcul d'un module carré moyen $|Y_{k,l}|^2$ du signal bruité $Y_{k,l}$ répondant à l'équation :

$$\overline{|Y_{k,l}|^2} = \frac{1}{M} \sum_{k=0}^{M-1} |Y_{k,l}|^2$$

g.3) calcul d'une composante filtrée P_l du module carré moyen $|Y_{k,l}|^2$ répondant à l'équation :

$$P_l = \lambda P_{l-1} + (1 - \lambda) \overline{|Y_{k,l}|^2}$$

où - λ est une constante prédéterminée caractéristique d'un filtre passe-bas, de préférence compris entre 0,80 et 0,99 ;

$$- P_0 = \bar{D}_0 = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,0} \text{ pour initialiser l'algorithme.}$$

g.4) calcul du rapport moyen signal à bruit r_l répondant à l'équation :

$$- \text{si } \bar{D}_l > 0 \text{ alors } r_l = \frac{P_l}{\bar{D}_l},$$

$$- \text{si } \bar{D}_l \leq 0 \text{ alors } r_l = 0.$$

9. Procédé selon l'une quelconque des revendications précédentes, dans lequel les étapes a) et f) de conversion mettent en oeuvre une méthode de recouvrement et addition dite « OLA », avec :

- pour l'étape a), un découpage du signal acoustique bruité $y(t)$ en trames temporelles avec un recouvrement entre les trames temporelles successives ;
- pour l'étape f.2), la reconstruction du signal vocal débruité de sortie $sd(t)$ est réalisée par les additions successives des parties en recouvrement des signaux de deux trames temporelles successives.

10. Système (1) de réduction du bruit sur un signal acoustique bruité $y(t)$ issu d'un microphone opérant dans un milieu bruité, comportant :

- une unité de conversion (2) du signal acoustique bruité $y(t)$ dans le domaine temporel en un signal bruité $Y_{k,l}$ dans le domaine fréquentiel, comportant :

- un module de découpage (21) du signal acoustique bruité $y(t)$ en signaux échantillonnés y_l dans des trames temporelles l successives ;
- en sortie du module de découpage (21), un module de fenêtrage (22) des signaux échantillonnés y_l par application d'une fenêtre de pondération ;
- en sortie du module de fenêtrage (22), un module de calcul (23) d'une transformée de Fourier discrète qui délivre en sortie le signal bruité $Y_{k,l}$;

- une unité de traitement numérique (3) dans le domaine fréquentiel comportant, en sortie de l'unité de conversion :

- un premier module d'extraction (31) du module carré $|Y_{k,l}|^2$ du signal bruité $Y_{k,l}$; et éventuellement un deuxième module d'extraction (32) de la phase $\theta_{k,l}$ du signal bruité $Y_{k,l}$;
- en sortie du premier module d'extraction (31), un module d'estimation (33), dit « MCRA », d'une composante de bruit $\hat{D}_{k,l}$ contenue dans le signal bruité $Y_{k,l}$ à partir du module carré $|Y_{k,l}|^2$ issu du premier module d'extraction (31), par un algorithme d'estimation de la densité spectrale de puissance de la composante de bruit selon une méthode de moyennage récursif des minima contrôlés dite « MCRA » ;
- en sortie du premier module d'extraction (31), un module de découpage (34) de la bande fréquentielle en plusieurs sous-bandes fréquentielles $SB_i = [e_i, b_i]$, notamment du type banc de filtres ;
- en sortie du module d'estimation « MCRA » (33) et du module de découpage (34) de la bande fréquentielle, un module d'estimation (35), dit « SSMB », du module carré $|\hat{X}_{k,l,i}|^2$ d'une composante débruitée de sous-bande $\hat{X}_{k,l,i}$ propre à chaque sous-bande SB_i d'un signal débruité $\hat{X}_{k,l}$, par un algorithme de soustraction spectrale multi-bandes à partir de modules carrés de sous-bande $|Y_{k,l,i}|^2$ et de composantes de bruit de

sous-bande $\hat{D}_{k,l,i}$;

- en sortie du module d'estimation « SSMB » (35), et éventuellement du deuxième module d'extraction (32), un module de détermination (36) d'un signal débruité de sortie $SD_{k,l}$ à partir des modules carré $|\hat{X}_{k,l,i}|^2$, et éventuellement des phases $\theta_{k,l}$;

- une unité de conversion (4) dans le domaine temporel comportant, en sortie de l'unité de traitement numérique (3) :

- un module de calcul (41) d'un signal de sortie Sd_l propre à chaque trame temporelle l par application d'une transformée de Fourier inverse du signal débruité de sortie $SD_{k,l}$; et

- un module de reconstruction (42) d'un signal vocal débruité de sortie $sd(t)$ dans le domaine temporel à partir desdits signaux de sortie Sd_l ;

dans lequel le module de détermination (36) d'un signal débruité de sortie $SD_{k,l}$ comprend :

- en sortie du premier module d'extraction (31), un sous-module d'amplification (363) selon un coefficient d'amplification γ , préférentiellement compris entre 0,01 et 0,1, afin de délivrer un signal amplifié $\gamma |Y_{k,l}|^2$;

- en sortie du module d'estimation « SSMB » (35), un sous-module additionneur (364) propre à additionner

le signal amplifié $\gamma |Y_{k,l}|^2$ et les modules carré $|\hat{X}_{k,l,i}|^2$, afin de délivrer en sortie les module carré $|\tilde{X}_{k,l,i}|^2$

de composantes débruitées combinées de sous-bande $\tilde{X}_{k,l,i}$ propres à chaque sous-bande SBi d'un

signal débruité combiné $\tilde{X}_{k,l}$, répondant à l'équation correspondante :

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2 ;$$

- un sous-module racine-carré (361) pour calculer le module $|\tilde{X}_{k,l,i}|$ des composantes débruitées combinées de sous-bande $\tilde{X}_{k,l,i}$; et

- un sous-module de recombinaison (362) des composantes combinées de sous-bande $\tilde{X}_{k,l,i}$ pour obtenir le signal débruité combiné $\tilde{X}_{k,l}$, à partir des modules $|\tilde{X}_{k,l,i}|$, et éventuellement des phases $\theta_{k,l}$, de

sorte que le signal débruité de sortie $SD_{k,l}$ correspond au signal débruité combiné $\tilde{X}_{k,l}$, soit $SD_{k,l} = \tilde{X}_{k,l}$.

11. Système (1) selon la revendication 10, dans lequel le module de détermination (36) du signal débruité de sortie $SD_{k,l}$ comprend :

- un sous-module racine-carré (361) pour calculer le module $|\hat{X}_{k,l,i}|$ des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$, et

- un sous-module de recombinaison (362) des composantes débruitées de sous-bande $\hat{X}_{k,l,i}$ pour obtenir le signal débruité $\hat{X}_{k,l}$ à partir des modules $|\hat{X}_{k,l,i}|$, et éventuellement des phases $\theta_{k,l}$, de sorte que le signal débruité de sortie $SD_{k,l}$ corresponde au signal débruité $\hat{X}_{k,l}$, soit $SD_{k,l} = \hat{X}_{k,l}$.

12. Système (1) selon l'une quelconque des revendications 10 et 11, dans lequel l'unité de traitement numérique (3) comprend en outre, en sortie du module d'estimation « MCRA » (33), un module de détection du bruit (37) comprenant

- un module de calcul d'un rapport moyen signal à bruit r_l propre à chaque trame temporelle l à partir du module carré $|Y_{k,l}|^2$ et de la composante de bruit $\hat{D}_{k,l}$;

- un module de comparaison du rapport moyen signal à bruit r_l propre à chaque trame temporelle l avec un seuil ψ_{TH} prédéterminé ;

- un module de contrôle du module de reconstruction (42) du signal vocal débruité de sortie $sd(t)$ qui est conçu pour que :

- si le rapport moyen signal à bruit r_l est inférieur audit seuil ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction pour cette trame temporelle l correspond au signal de sortie Sd_l issu du module de calcul dudit signal de sortie Sd_l ;
- si le rapport moyen signal à bruit r_l est supérieur audit seuil ψ_{TH} pour la trame temporelle l , alors le signal considéré avant reconstruction pour cette trame temporelle l correspond au signal échantillonné y_l issu du module de découpage du signal acoustique bruité $y(t)$.

Patentansprüche

1. Verfahren zur Reduzierung des Rauschens auf einem verrauschten akustischen System $y(t)$, das von einem Mikrofon stammt, das in einem verrauschten Umfeld funktioniert, umfassend die folgenden aufeinander folgenden Schritte:

a) Umwandeln des verrauschten akustischen Signals $y(t)$ im Zeitbereich in ein verrauschtes Signal $Y_{k,l}$ im Frequenzbereich durch zeitliches Schneiden des verrauschten akustischen Signals $y(t)$ in abgetastete Signale y_l in aufeinander folgenden Zeitrastern l , Fensterung der abgetasteten Signale y_l durch Anwenden eines Gewichtungsfensters und Anwendung einer diskreten Fourier-Transformation mit Extraktion des viereckigen Moduls $|Y_{k,l}|^2$, und eventuell der Phase $\theta_{k,l}$, des verrauschten Signals $Y_{k,l}$;

b) Schätzen einer Rauschkomponente $\hat{D}_{k,l}$, enthalten die im verrauschten Signal $Y_{k,l}$, ausgehend vom viereckigen Modul $|Y_{k,l}|^2$, durch einen Algorithmus zur Schätzung der Leistungsspektraldichte der Rauschkomponente gemäß einem rekursiven Mittelungsverfahren der gesteuerten Minima, genannt "MCRA".

dadurch gekennzeichnet, dass es aussenden, nach dem Schritt b), die folgenden aufeinander folgenden Schritte umfasst:

c) Schneiden des Frequenzbandes in mehrere Unter-Frequenzbänder $SB_i = [e_i, b_i]$, gefolgt von einer Mehrband-Zersetzung des viereckigen Moduls $|Y_{k,l}|^2$ und des Rauschbestandteils $\hat{D}_{k,l}$, bestehend aus der Zersetzung des viereckigen Moduls $|Y_{k,l}|^2$ und die Rauschkomponente $\hat{D}_{k,l}$ in jeweils mehrere viereckige Unterband-Module $|Y_{k,l,i}|^2$ und mehrere Unterband-Geräuschbestandteile $\hat{D}_{k,l,i}$, die jedem der Unterbänder SB_i eigen sind;

d) Schätzen, für jedes der Unterbänder SB_i , des viereckigen Moduls $|\hat{X}_{k,l,i}|^2$ einer entrauschten Komponente des Unterbandes $\hat{X}_{k,l,i}$, die jedem Unterband SB_i eines entrauschten Signals $\hat{X}_{k,l}$, eigen ist, durch einen Algorithmus zur spektralen Mehrband-Subtraktion, genannt "SSMB", ausgehend von den viereckigen Unterband-Modulen $|Y_{k,l,i}|^2$ und den Unterband-Geräuschbestandteilen $\hat{D}_{k,l,i}$;

e) Bestimmen eines entrauschten Ausgangssignals $SD_{k,l}$ ausgehend von den viereckigen Modulen $|\hat{X}_{k,l,i}|^2$, die aus dem Schritt d) stammen, und eventuell den Phasen $\theta_{k,l}$, extrahiert im Schritt a);

f) Umwandeln des entrauschten Ausgangssignals $SD_{k,l}$ in ein entraushtes Ausgangs-Sprachsignal $sd(t)$ im Zeitbereich durch einen Schritt f.1) zur Berechnung eines abgetasteten Ausgangssignal Sd_l , das jedem Zeitraster l eigen ist, durch die Anwendung einer inversen Fourier-Transformation des entrauschten Ausgangssignals $SD_{k,l}$, gefolgt von einem Schritt f.2) zur zeitlichen Rekonstruktion des entrauschten Ausgangs-Sprachsignals $sd(t)$ ausgehend von den abgetasteten Ausgangssignalen Sd_l , und wobei der Schritt e) aus Folgendem besteht:

- Bestimmen, für jedes der Unterbänder SB_i , des viereckigen Moduls $|\hat{X}_{k,l,i}|^2$ einer kombinierten entrauschten Komponente des Unterbandes $\tilde{X}_{k,l,i}$, die jedem Unterband SB_i eines kombinierten entrauschten Signals $\tilde{X}_{k,l}$, eigen ist, das der entsprechenden Gleichung entspricht:

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2$$

wobei γ ein vorbestimmter Erweiterungskoeffizient ist, der vorzugsweise zwischen 0,01 und 0,1 liegt;

- Bestimmen eines kombinierten entrauschten Signals $\tilde{X}_{k,l}$ ausgehend von den viereckigen Modulen

$|\tilde{X}_{k,l,i}|^2$ der kombinierten entrauschten Komponenten des Unterbandes $\tilde{X}_{k,l,i}$, und eventuell den Phasen $\theta_{k,l}$, extrahiert im Schritt a), so dass das entrauschte Ausgangssignal $SD_{k,l}$ dem kombinierten entrauschten

Signal $\tilde{X}_{k,l}$ entspricht,

5

$$\text{d.h. } SD_{k,l} = \tilde{X}_{k,l}.$$

10

2. Verfahren nach Anspruch 1, in dem der Algorithmus zur Schätzung der Leistungsspektraldichte der Rauschkomponente gemäß dem rekursiven Mittelungsverfahren der gesteuerten Minima, genannt "MCRA" beim Schritt b) die folgenden Berechnungsphasen durchführt:

b.1) Berechnung einer gefilterten verrauschten Komponente $S_{k,l}$ die der folgenden Gleichung entspricht:

15

$$S_{k,l} = \alpha_s S_{k,l-1} + (1 - \alpha_s) (1 - a_s) |Y_{k,l}|^2$$

wobei α_s eine charakteristische vorbestimmte Konstante eines Tiefpassfilters ist;

20

b.2) Berechnung einer Wahrscheinlichkeitsdichte der Anwesenheit von Sprache $\tilde{p}_{k,l}$ durch die Durchführung der folgenden progressiven Berechnung:

(i) Berechnung einer minimalen spektralen Komponente $S_{\min,k,l}$ mit

25

- wenn $\text{rem}(k,l) = 0$, dann $S_{\min,k,l} = \min(S_{\min,k,l-1}; S_{k,l})$ und $\text{Stmp}_{k,l} = S_{k,l}$
 - wenn $\text{rem}(k,l) \neq 0$, dann $S_{\min,k,l} = \min(\text{Stmp}_{k,l-1}; S_{k,l})$ und $\text{Stmp}_{k,l} = \min(\text{Stmp}_{k,l}; S_{k,l})$
- wobei $\text{rem}(k,l)$ der Rest der ganzzahligen Division von k durch l ist, dann

(ii) Berechnung einer spektralen Beziehung $S_{r,k,l}$, die der folgenden Gleichung entspricht:

30

$$S_{r,k,l} = \frac{S_{k,l}}{S_{\min,k,l}}$$

(iii) Berechnung einer Anzeigevariablen $I_{k,l}$ mit:

35

- wenn $S_{r,k,l} > \delta_{TH}$, dann $I_{k,l} = 1$
 - wenn $S_{r,k,l} \leq \delta_{TH}$, dann $I_{k,l} = 0$
- wobei δ_{TH} ein vorbestimmter Parameter mit fester Schwelle zum Nachweis von Sprache ist;

40

(iv) Berechnung einer Wahrscheinlichkeitsdichte der Anwesenheit von Sprache $\tilde{p}_{k,l}$ mit:

45

$$\tilde{p}_{k,l} = \alpha_p \tilde{p}_{k,l-1} + (1 - \alpha_p) I_{k,l}$$

wobei α_p eine vorbestimmte Konstante ist;

b.3) Berechnung eines Koeffizienten $\tilde{\alpha}_{k,l}$, der der folgenden Gleichung entspricht:

50

$$\tilde{\alpha}_{k,l} = \alpha + (1 - \alpha) \tilde{p}_{k,l},$$

wobei α eine vorbestimmte Konstante ist;

55

b.4) Berechnung der Rauschkomponente $\hat{D}_{k,l}$, die der folgenden Gleichung entspricht:

$$\hat{D}_{k,l} = \tilde{\alpha}_{k,l} \hat{D}_{k,l-1} + (1 - \tilde{\alpha}_{k,l}) |Y_{k,l}|^2.$$

3. Verfahren nach einem der Ansprüche 1 und 2, wobei der Algorithmus zur spektralen Mehrband-Subtraktion, genannt "SSMB" von Schritt d) die folgenden Berechnungsphasen für jedes der Unterbänder SB_i durchführt:

d.1) Berechnung einer Beziehung Signal - Rauschen $SNR_{k,l,i}$, die jedem Unterband SB_i eigen ist, die der folgenden Formel entspricht:

$$SNR_{k,l,i} = 10 \cdot \log_{10} \left(\frac{\sum_{k=ei}^{bi} |Y_{k,l,i}|^2}{\sum_{k=ei}^{bi} |\hat{D}_{k,l,i}|^2} \right)$$

d.2) Berechnung des viereckigen Moduls $|\hat{X}_{k,l,i}|^2$ der entrauschten Komponente des Unterbands $\hat{X}_{k,l,i}$, die jedem Unterband SB_i eigen ist, gemäß der folgenden Gleichung:

$$|\hat{X}_{k,l,i}|^2 = \begin{cases} |Y_{k,l,i}|^2 - \alpha_i \delta_i |\hat{D}_{k,l,i}| & \text{si } |Y_{k,l,i}|^2 > \alpha_i \delta_i |\hat{D}_{k,l,i}| \\ \beta |Y_{k,l,i}|^2 & \text{si } |Y_{k,l,i}|^2 \leq \alpha_i \delta_i |\hat{D}_{k,l,i}| \end{cases}$$

wobei δ_i ein variabler Parameter in Funktion des entsprechenden Unterbandes SB_i ist, der verschiedene Werte von einem zum anderen Unterband annimmt;

- α_i ein variabler Parameter ist, der vom Wert der Beziehung zwischen Signal - Rauschen $SNR_{k,l,i}$ abhängt, berechnet im entsprechenden Unterband SB_i ; und
- β eine Konstante ist.

4. Verfahren nach Anspruch 3, in dem die Parameter α_i den folgenden Gleichungen entsprechen:

$$\alpha_i = \begin{cases} \alpha_{c1} & \text{si } SNR_{k,l,i} < SNR_1 \\ \alpha_{c2} + \alpha_{c3} SNR_{k,l,i} & \text{si } SNR_1 \leq SNR_{k,l,i} \leq SNR_2 \\ \alpha_{c4} & \text{si } SNR_{k,l,i} > SNR_2 \end{cases}$$

wobei α_{c1} , α_{c2} , α_{c3} und α_{c4} vorbestimmte Konstanten sind und SNR_1 und SNR_2 vorbestimmte Schwellen sind.

5. Verfahren nach einem der Ansprüche 1 und 4, wobei der Schritt e) aus der Bestimmung des entrauschten Signals $\hat{X}_{k,l}$ ausgehend von den viereckigen Modulen $|\hat{X}_{k,l,i}|^2$ der entrauschten Komponenten des Unterbands $\hat{X}_{k,l,i}$, und eventuell den Phasen $\theta_{k,l,i}$, extrahiert in Schritt a), besteht, so dass das entrauschte Ausgangssignal $SD_{k,l}$ dem entrauschten Signal $\hat{X}_{k,l}$ entspricht, d.h. $SD_{k,l} = \hat{X}_{k,l}$.

6. Verfahren nach einem der Ansprüche 1 bis 5, wobei der Schritt f.2) aus der Rekonstruktion des entrauschten Ausgangs-Sprachsignals $sd(t)$ lediglich ausgehend von den Ausgangssignalen SD_1 , die aus Schritt f.1) stammen, besteht, wobei die Ausgangssignale SD_1 den inversen Fourier-Transformationen des entrauschten Ausgangssignals $SD_{k,l}$ entsprechen, die jedem Zeitraster l eigen sind.

7. Verfahren nach einem der Ansprüche 1 bis 5, wobei der Schritt f.2) für jedes Zeitraster l aus Folgendem besteht:

- g) Berechnen einer Beziehung Signal - Rauschen r_i , die dem Zeitraster l eigen ist, ausgehend von dem viereckigen Modul $|Y_{k,l}|^2$ und der Rauschkomponente $\hat{D}_{k,l}$;
- h) Vergleichen der Beziehung Signal - Rauschen r_i mit einer vorbestimmten Schwelle ψ_{TH} ;
- i) Rekonstruieren des entrauschten Ausgangs-Sprachsignals $sd(t)$ ausgehend unter Berücksichtigung des Fol-

genden:

- wenn die mittlere Beziehung Signal - Rauschen r_i geringer als die Schwelle ψ_{TH} für das Zeitraster I ist, dann entspricht das Signal, berücksichtigt von der zeitlichen Rekonstruktion für dieses Zeitraster I, dem Ausgangssignal Sd_i , das aus Schritt f.1) stammt;
- wenn die mittlere Beziehung Signal Rauschen r_i größer als die Schwelle ψ_{TH} für das Zeitraster I ist, dann entspricht das Signal, berücksichtigt von der Rekonstruktion für dieses Zeitraster I, dem abgetasteten Signal y_i , das aus dem Schritt des Schneidens von Schritt a) stammt.

8. Verfahren nach Anspruch 7, in dem der Schritt g) den folgenden Berechnungsalgorithmus für jedes Zeitraster I durchführt:

g.1) Berechnung einer mittleren Rauschkomponente $\bar{D}_I$ ausgehend von der Rauschkomponente $\hat{D}_{k,I}$, die im Schritt b) geschätzt wird und der folgenden Gleichung entspricht:

$$\bar{D}_I = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,I}$$

wobei M eine vorbestimmte Konstante ist, vorzugsweise gleich N oder N/2, wobei N die Anzahl der Abtastpunkte der Fourier-Transformation ist;

g.2) Berechnung eines mittleren viereckigen Moduls $\overline{|Y_{k,I}|^2}$ des verrauschten Signals $Y_{k,I}$, das der folgenden Gleichung entspricht:

$$\overline{|Y_{k,I}|^2} = \frac{1}{M} \sum_{k=0}^{M-1} |Y_{k,I}|^2$$

g3) Berechnung einer gefilterten Komponente P_I des mittleren viereckigen Moduls $\overline{|Y_{k,I}|^2}$ das der folgenden Formel entspricht:

$$P_I = \lambda P_{I-1} + (1 - \lambda) \overline{|Y_{k,I}|^2}$$

wobei λ eine charakteristische vorbestimmte Konstante eines Tiefpassfilters ist, die vorzugsweise zwischen 0,80 und 0,99 liegt;

$P_0 = \bar{D}_0 = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,0}$ um den Algorithmus zu initiieren.

g.4) Berechnen der mittleren Beziehung Signal - Rauschen r_i , die der folgenden Formel entspricht:

$$\text{- si } \bar{D}_I > 0 \text{ alors } r_I = \frac{P_I}{\bar{D}_I},$$

$$\text{- si } \bar{D}_I \leq 0 \text{ alors } r_I = 0.$$

9. Verfahren nach einem der vorhergehenden Ansprüche, in dem die Schritte a) und f) zur Umwandlung ein Verfahren zur Beschichtung und Zugabe, genannt "OLA" durchführen, umfassend:

- für den Schritt a) ein Schneiden des verrauschten akustischen Signals $y(t)$ in zeitliche Raster mit einer Beschichtung zwischen den aufeinander folgenden zeitlichen Rastern;
- für den Schritt f.2) wird die Rekonstruktion des entrauschten Ausgangs-Sprachsignals $sd(t)$ durch die aufeinander folgenden Zugaben der Teile mit Bedeckung der Signale mit zwei zeitlichen Rastern durchgeführt.

10. System (1) zur Reduzierung des Rauschens auf einem verrauschten akustischen Signal $y(t)$, das von einem Mikrofon stammt, das in einem verrauschten Umfeld funktioniert, umfassend:

- eine Umwandlungseinheit (2) des verrauschten akustischen Signals $y(t)$ im Zeitbereich in ein verrauschtes

Signal $Y_{k,l}$ im Frequenzbereich, umfassend:

- ein Modul zum Schneiden (21) des verrauschten akustischen Signals $y(t)$ in abgetastete Signale y_i in aufeinander folgenden Zeitrastern l ;

- am Ausgang des Schneidomoduls (21) ein Fensterungsmodul (22) der abgetasteten Signale y_i durch Anwendung eines Gewichtungsfensters;

- am Ausgang des Fensterungsmoduls (22) ein Berechnungsmodul (23) einer diskreten Fourier-Transformation, die am Ausgang das verrauschte Modul $Y_{k,l}$ liefert,

- eine numerische Behandlungseinheit (3) im Frequenzbereich, umfassend, am Ausgang der Umwandlungseinheit:

- ein erstes Extraktionsmodul (31) des viereckigen Moduls $|Y_{k,l}|^2$ des verrauschten Signals $Y_{k,l}$; und eventuell ein zweites Extraktionsmodul (32) der Phase $\theta_{k,l}$ des verrauschten Signals $Y_{k,l}$,

- am Ausgang des ersten Extraktionsmoduls (31), ein Schätzungsmodul (33), genannt "MCRA", einer Rauschkomponente $\hat{D}_{k,l}$, enthalten im verrauschten Signal $Y_{k,l}$, ausgehend vom viereckigen Modul $|Y_{k,l}|^2$, das aus dem ersten Extraktionsmodul (31) stammt, durch einen Algorithmus zur Schätzung der Leistungsspektraldichte der Rauschkomponente gemäß einem rekursiven Mittelungsverfahren der gesteuerten Minima, genannt "MCRA",

- am Ausgang des ersten Extraktionsmoduls (31), ein Modul zum Schneiden (34) des Frequenzbandes in mehrere Frequenzunterbänder $SB_i = [e_i, b_i]$, insbesondere vom Typ Filterbank;

- am Ausgang des Schätzungsmoduls "MCRA" (33) und des Schneidomoduls (34) des Frequenzbandes ein Schätzungsmodul (35), genannt "SSMB", des viereckigen Moduls $|\hat{X}_{k,l,i}|^2$ einer entrauschten Komponente des Unterbands $\hat{X}_{k,l,i}$, die jedem Unterband SB_i eines entrauschten Signals $\hat{X}_{k,l}$ eigen ist, durch einen Algorithmus zur spektralen Mehrband-Subtraktion ausgehend von viereckigen Unterbandmodulen $|Y_{k,l,i}|^2$ und Unterband-Geräuschbestandteilen $D_{k,l,i}$;

- am Ausgang des Schätzungsmoduls "SSMB" (35), und eventuell des zweiten Extraktionsmoduls (32) ein Modul zum Bestimmen (36) eines entrauschten Ausgangssignals $SD_{k,l}$ ausgehend von den viereckigen Modulen $|\hat{X}_{k,l,i}|^2$ und eventuell den Phasen $\theta_{k,l}$;

- eine Umwandlungseinheit (4) im Zeitbereich, umfassend am Ausgang der numerischen Behandlungseinheit (3):

- ein Berechnungsmodul (41) eines Ausgangssignals Sd_i , das jedem Zeitraster l eigen ist, durch das Anwenden einer inversen Fourier-Transformation des entrauschten Ausgangssignals $SD_{k,l}$; und

- ein Rekonstruktionsmodul (42) eines entrauschten Ausgangs-Sprachsignals $sd(t)$ im Zeitbereich, ausgehend von den Ausgangssignalen Sd_i , in dem das Modul zum Bestimmen (36) eines entrauschten Ausgangssignals $SD_{k,l}$ Folgendes umfasst:

- am Ausgang des ersten Extraktionsmoduls (31), ein Erweiterungsuntermodul (363), gemäß einem Erweiternskoeffizienten γ , der vorzugsweise zwischen 0,01 und 0,1 liegt; um ein erweitertes Signal $\gamma |Y_{k,l}|^2$ zu liefern;

- am Ausgang des Schätzungsmoduls "SSMB" (35) ein Addier-Untermodul (364), das geeignet ist, um das erweiterte Modul $\gamma |Y_{k,l}|^2$ und die viereckigen Module $|\hat{X}_{k,l,i}|^2$ zu addieren, um am Ausgang das viereckige Modul

$|\tilde{X}_{k,l,i}|^2$ von entrauschten Komponenten des Unterbands $\tilde{X}_{k,l,i}$ zu liefern, die jedem Unterband SB_i eines

entrauschten Signals $\tilde{X}_{k,l}$ eigen sind,

das der entsprechenden Gleichung entspricht:

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2;$$

- ein wurzelquadratisches Untermodul (361), um das Modul $|\tilde{X}_{k,l,i}|$ der entrauschten Komponenten des Unterbands $\tilde{X}_{k,l,i}$ zu berechnen; und

- ein Neukombinations-Untermodul (362) der kombinierten Komponenten des Unterbands $X_{k,l,i}$, um das kombinierte entrauschte Signal $\tilde{X}_{k,l}$ ausgehend von den Modulen $|\tilde{X}_{k,l,i}|$, und eventuell den Phasen $\theta_{k,l}$ zu erhalten, so dass das entrauschte Ausgangssignal $SD_{k,l}$ dem kombinierten entrauschten Signal $\tilde{X}_{k,l}$ entspricht,

$$\text{d.h. } SD_{k,l} = \hat{X}_{k,l}.$$

11. System (1) nach Anspruch 10, in dem das Modul zum Bestimmen (36) des entrauschten Ausgangssignals $SD_{k,l}$ Folgendes umfasst:

- ein wurzelquadratisches Untermodul (361), um das viereckige Modul $|\hat{X}_{k,l,i}|$ der entrauschten Komponenten des Unterbands $\hat{X}_{k,l,i}$ zu berechnen und
- ein Untermodul zur Neukombination (362) der entrauschten Komponenten des Unterbands $\hat{X}_{k,l,i}$ um das entrauschte Signal $\hat{X}_{k,l}$ ausgehend von den Modulen $|\hat{X}_{k,l,i}|$ und eventuell den Phasen $\theta_{k,l}$ zu erhalten, so dass das entrauschte Ausgangssignal $SD_{k,l}$ dem kombinierten entrauschten Signal $\hat{X}_{k,l}$ entspricht, d.h. $SD_{k,l} = \hat{X}_{k,l}$.

12. System (1) nach einem der Ansprüche 10 bis 11, wobei die numerische Behandlungseinheit (3) außerdem am Ausgang des Schätzmoduls "MCRA" (33) ein Modul zum Nachweis des Rauschens (37) umfasst, umfassend:

- ein Modul zur Berechnung einer Beziehung Signal - Rauschen r_i , das jedem Zeitraster l eigen ist, ausgehend vom viereckigen Modul $|Y_{k,l}|^2$ und der Rauschkomponente $\hat{D}_{k,l}$.
- ein Modul zur Berechnung der Beziehung Signal - Rauschen r_i , die jedem Zeitraster l eigen ist, mit einer vorbestimmten Schwelle ψ_{TH} ;
- ein Steuermodul des Rekonstruktionsmoduls (42) des entrauschten Ausgangs-Sprachsignals $sd(t)$, das entworfen ist, dass
- wenn die mittlere Beziehung Signal - Rauschen r_i geringer als die Schwelle ψ_{PTH} für das Zeitraster l ist, dann entspricht das Signal, berücksichtigt vor der Rekonstruktion für dieses Zeitraster l , dem Ausgangssignal Sd_l , das aus dem Berechnungsmodul des Ausgangssignals Sd_i stammt,
- wenn die mittlere Beziehung Signal - Rauschen r_i höher als die Schwelle ψ_{PTH} für das Zeitraster l ist, dann entspricht das Signal, berücksichtigt vor der Rekonstruktion für dieses Zeitraster l , dem abgetasteten Signal y_l , das aus dem Schneidemodul des verrauschten akustischen Signals $y(t)$ stammt.

Claims

1. A method for reducing noise on a noisy acoustic signal $y(t)$ from a microphone operating in a noisy environment, including the following successive steps:

- a) converting the noisy acoustic signal $y(t)$ in the time domain into a noisy signal $Y_{k,l}$ in the frequency domain, by time slicing the noisy acoustic signal $y(t)$ into sampled signals y_l into successive time frames l , windowing the sampled signals y_l by applying a weighting window, and applying a discrete Fourier transform, with extraction of the square module $|Y_{k,l}|^2$, and possibly the phase $\theta_{k,l}$, from the noisy signal $Y_{k,l}$;
- b) estimating a noise component $\hat{D}_{k,l}$ contained in the noisy signal $Y_{k,l}$ from the square module $|Y_{k,l}|^2$, by an algorithm of power spectral density estimation of the noise component according to a method of minima controlled recursive averaging called "MCRA";

characterized in that it further comprises, after the step b), the following successive steps:

- c) slicing the frequency band into several frequency sub-bands $SB_i = [e_i, b_i]$, followed by multi-band decomposing the square module $|Y_{k,l}|^2$ and the noise component $\hat{D}_{k,l}$, consisting in decomposing the square module $|Y_{k,l}|^2$ and the noise component $\hat{D}_{k,l}$ into respectively several square modules of sub-band $|Y_{k,l,i}|^2$ and several noise components of sub-band $\hat{D}_{k,l,i}$ specific to each sub-band SB_i ;

- d) estimating, for each sub-band SB_i , the square module $|\hat{X}_{k,l,i}|^2$ of a denoised component of sub-band $\hat{X}_{k,l,i}$ specific to each sub-band SB_i of a denoised signal $\hat{X}_{k,l}$, by an algorithm of multi-band spectral subtraction called "SSMB" from the square modules of sub-band $|Y_{k,l,i}|^2$ and noise components of sub-band $\hat{D}_{k,l,i}$;

- e) determining an output denoised signal $SD_{k,l}$ from the square modules $|\hat{X}_{k,l,i}|^2$ from step d), and possibly the phases $\theta_{k,l}$ extracted during step a);

- f) converting the output denoised signal $SD_{k,l}$ into an output denoised vocal signal $sd(t)$ in the time domain, by a step f.1) of calculating an output sampled signal sd_l specific to each time frame l by applying an inverse Fourier transform of the output denoised signal $SD_{k,l}$, followed by a step f.2) of time reconstructing the output denoised vocal signal $sd(t)$ from the output sampled signals sd_l ;
- and wherein the step e) consists of:

- determining, for each sub-band SB_i , the square module $|\tilde{X}_{k,l,i}|^2$ of a combined denoised component of sub-band $\tilde{X}_{k,l,i}$ specific to each sub-band SB_i of a combined denoised signal $\tilde{X}_{k,l}$, satisfying the corresponding equation:

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2$$

where γ is a predetermined amplification coefficient, preferably ranging between 0.01 and 0.1;

- determining a combined denoised signal $\tilde{X}_{k,l}$ from the square modules

$|\tilde{X}_{k,l,i}|^2$ of the combined denoised components of sub-band $\tilde{X}_{k,l,i}$, and possibly the phases $\theta_{k,l}$ extracted during the step a), such that the output denoised signal $SD_{k,l}$ corresponds to the combined denoised signal $\tilde{X}_{k,l}$, namely $SD_{k,l} = \tilde{X}_{k,l}$.

2. The method according to claim 1, wherein the algorithm of power spectral density estimation of the noise component according to the method of minima controlled recursive averaging called "MCRA" during step b) implements the following calculation phases:

b.1) calculating a filtered noisy component $S_{k,l}$ satisfying the equation:

$$S_{k,l} = \alpha_s S_{k,l-1} + (1 - \alpha_s) |Y_{k,l}|^2$$

where α_s is a predetermined constant characteristic of a low-pass filter;

b.2) calculating a speech presence probability density $\tilde{p}_{k,l}$ by implementing the following progressive calculation:

(i) calculating a spectral minimum component $S_{min,k,l}$ with:

- if $\text{rem}(k,l) = 0$
then $S_{min,k,l} = \min(S_{min,k,l-1}; S_{k,l})$ and $Stmp_{k,l} = S_{k,l}$
- if $\text{rem}(k,l) \neq 0$
then $S_{min,k,l} = \min(Stmp_{k,l-1}; S_{k,l})$ and
 $Stmp_{k,l} = \min(Stmp_{k,l-1}; S_{k,l})$
where $\text{rem}(k,l)$ is the rest of the integer division of k by l , then

(ii) calculating a spectral ratio $Sr_{k,l}$ satisfying the equation:

$$Sr_{k,l} = \frac{S_{k,l}}{S_{min,k,l}}$$

(iii) calculating an indicator variable $I_{k,l}$ with:

- if $Sr_{k,l} > \delta_{TH}$ then $I_{k,l} = 1$
- if $Sr_{k,l} \leq \delta_{TH}$ then $I_{k,l} = 0$
where δ_{TH} is a fixed threshold predetermined parameter of speech detection;

(iv) calculating the speech presence probability density $\tilde{p}_{k,l}$ with:

$\tilde{p}_{k,l} = \alpha_p \tilde{p}_{k,l-1} + (1 - \alpha_p) I_{k,l}$
where α_p is a predetermined constant;

b.3) calculating a coefficient $\tilde{\alpha}_{k,l}$ satisfying the following equation:

$$\tilde{\alpha}_{k,l} = \alpha + (1 - \alpha) \tilde{p}_{k,l},$$

where α is a predetermined constant;

b.4) calculating the noise component $\hat{D}_{k,l}$ satisfying the following equation:

$$\hat{D}_{k,l} = \tilde{\alpha}_{k,l} \hat{D}_{k,l-1} + (1 - \tilde{\alpha}_{k,l}) |Y_{k,l}|^2.$$

3. The method according to any one of claims 1 and 2, wherein the algorithm of multi-band spectral subtraction called "SSMB" of step d) implements the following calculation phases, for each sub-band SB_i :

d.1) calculating a signal-to-noise ratio $SNR_{k,l,i}$ specific to each sub-band SB_i satisfying the following equation:

$$SNR_{k,l,i} = 10 \cdot \log_{10} \left(\frac{\sum_{k=ei}^{bi} |Y_{k,l,i}|^2}{\sum_{k=ei}^{bi} |\hat{D}_{k,l,i}|} \right)$$

d.2) calculating the square module $|\hat{X}_{k,l,i}|^2$ of the denoised component of sub-band $\hat{X}_{k,l,i}$ specific to each sub-band SB_i , according to the equation:

$$|\hat{X}_{k,l,i}|^2 = \begin{cases} |Y_{k,l,i}|^2 - \alpha_i \delta_i |\hat{D}_{k,l,i}| & \text{if } |Y_{k,l,i}|^2 > \alpha_i \delta_i |\hat{D}_{k,l,i}| \\ \beta |Y_{k,l,i}|^2 & \text{if } |Y_{k,l,i}|^2 \leq \alpha_i \delta_i |\hat{D}_{k,l,i}| \end{cases}$$

where δ_i is a variable parameter according to the corresponding sub-band SB_i , taking distinct values from one sub-band to another;

- α_i is a variable parameter which depends on the value of the signal-to-noise ratio $SNR_{k,l,i}$ calculated in the corresponding sub-band SB_i ; and
- β is a constant.

4. The method according to claim 3, wherein the parameters α_i satisfy the following equations:

$$\alpha_i = \begin{cases} \alpha_{c1} & \text{if } SNR_{k,l,i} < SNR_1 \\ \alpha_{c2} + \alpha_{c3} SNR_{k,l,i} & \text{if } SNR_1 \leq SNR_{k,l,i} \leq SNR_2 \\ \alpha_{c4} & \text{if } SNR_{k,l,i} > SNR_2 \end{cases}$$

where α_{c1} , α_{c2} , α_{c3} and α_{c4} are predetermined constants, and SNR_1 and SNR_2 are predetermined thresholds.

5. The method according to any one of claims 1 to 4, wherein the step e) consists of determining the denoised signal $X_{k,l}$ from the square modules $|\hat{X}_{k,l,i}|^2$ of the denoised components of sub-band $X_{k,l,i}$, and possibly of phases $\theta_{k,l}$ extracted during step a), such that the denoised output signal $SD_{k,l}$ corresponds to the denoised signal $\hat{X}_{k,l}$, namely $SD_{k,l} = \hat{X}_{k,l}$.

6. The method according to any one of claims 1 to 5, wherein the step f.2) consists of reconstructing the output denoised vocal signal $sd(t)$ only from the output signals Sd_i from the step f.1), said output signals Sd_i corresponding to the inverse Fourier transforms of the output denoised signal $SD_{k,l}$ specific to each time frame l .

7. The method according to any one of claims 1 to 5, wherein the step f.2) consists, for each time frame l of:

- g) calculating a mean signal-to-noise ratio r_l specific to the time frame l from the square module $|Y_{k,l}|^2$ and the noise component $\hat{D}_{k,l}$;
h) comparing the mean signal-to-noise ratio r_l with a predetermined threshold ψ_{TH} ;
i) reconstructing the output denoised vocal signal $sd(t)$ by taking into consideration that:

- if the mean signal-to-noise ratio r_l is lower than said threshold ψ_{TH} for the time frame l , then the considered signal before time reconstruction for this time frame l corresponds to the output signal Sd_l from step f.1);
- if the mean signal-to-noise ratio r_l is higher than said threshold ψ_{TH} for the time frame l , then the considered signal before time reconstruction for this time frame l corresponds to the sampled signal y_l from the slicing step of step a).

8. The method according to claim 7, wherein the step g) implements the following calculation algorithm, for each time frame l :

- g.1) calculating a mean noise component $\bar{D}_l$ from the noise component $\hat{D}_{k,l}$ estimated during the step b) and satisfying the equation:

$$\bar{D}_l = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,l}$$

where M is a predetermined constant, preferably equal to N or to $N/2$, N being the number of sampling points of the Fourier transform;

- g.2) calculating a mean square module $\overline{|Y_{k,l}|^2}$ of the noisy signal $Y_{k,l}$ satisfying the equation:

$$\overline{|Y_{k,l}|^2} = \frac{1}{M} \sum_{k=0}^{M-1} |Y_{k,l}|^2$$

- g.3) calculating a filtered component P_l of the mean square module $\overline{|Y_{k,l}|^2}$ satisfying the equation:

$$P_l = \lambda P_{l-1} + (1 - \lambda) \overline{|Y_{k,l}|^2}$$

where λ is a predetermined constant characteristic of a low-pass filter, preferably ranging between 0.80 and 0.99;

$$- P_0 = \bar{D}_0 = \frac{1}{M} \sum_{k=0}^{M-1} \hat{D}_{k,0}$$

for initializing the algorithm.

- g.4) calculating the mean signal-to-noise ratio r_l satisfying the equation:

$$\begin{aligned} & \text{- if } \bar{D}_l > 0 \text{ then } r_l = \frac{P_l}{\bar{D}_l} \\ & \text{- if } \bar{D}_l \leq 0 \text{ then } r_l = 0. \end{aligned}$$

9. The method according to any one of the preceding claims, wherein the steps a) and f) of conversion implement an overlap-add method called "OLA", with:

- for step a), a slicing of the noisy acoustic signal $y(t)$ into time frames with an overlapping between the successive time frames;
- for step f.2), the reconstructing of the output denoised vocal signal $sd(t)$ is achieved by the successive additions of the overlapping portions of the signals of two successive time frames.

10. A system (1) for reducing noise on a noisy acoustic signal $y(t)$ from a microphone operating in a noisy environment, comprising:

- a unit (2) for converting the noisy acoustic signal $y(t)$ in the time domain into a noisy signal $Y_{k,l}$ in the frequency domain, including:
 - a module (21) for slicing the noisy acoustic signal $y(t)$ into sampled signals y_l in successive time frames l ;
 - at the output of the slicing module (21), a module (22) for windowing the sampled signals y_l by applying a weighting window;
 - at the output of the windowing module (22), a module (23) for calculating a discrete Fourier transform which outputs the noisy signal $Y_{k,l}$;
 - a digital processing unit (3) in the frequency domain including, at the output of the conversion unit:
 - a first module (31) for extracting the square module $|Y_{k,l}|^2$ from the noisy signal $Y_{k,l}$; and possibly a second module (32) for extracting the phase $\theta_{k,l}$ from the noisy signal $Y_{k,l}$;
 - at the output of the first extraction module (31), an estimation module (33) called "MCRA", for estimating a noise component $\hat{D}_{k,l}$ contained in the noisy signal $Y_{k,l}$ from the square module $|Y_{k,l}|^2$ from the first extraction module (31), by an algorithm of power spectral density estimation of the noise component according to a minima controlled recursive averaging method called "MCRA";
 - at the output of the first extraction module (31), a module (34) for slicing the frequency band into several frequency sub-bands $SB_i = [e_i, b_i]$, in particular of filter bank type;
 - at the output of the estimation module "MCRA" (33) and the module (34) for slicing the frequency band, an estimation module (35), called "SSMB", for estimating the square module $|\hat{X}_{k,l,i}|^2$ of a denoised component of sub-band $\hat{X}_{k,l,i}$ specific to each sub-band SB_i of a denoised signal $\hat{X}_{k,l}$, by a multi-band spectral subtraction algorithm from the square modules of sub-bands $|Y_{k,l,i}|^2$ and noise components of sub-band $\hat{D}_{k,l,i}$;
 - at the output of the estimation module "SSMB" (35), and possibly the second extraction module (32), a module (36) for determining an output denoised signal $SD_{k,l}$ from the square module $|\hat{X}_{k,l,i}|^2$ and possibly the phases $\theta_{k,l}$;
 - a conversion unit (4) in the time domain including, at the output of the digital processing unit (3):
 - a module (41) for calculating an output signal sd_l specific to each time frame l by applying an inverse Fourier transform of the output denoised signal $SD_{k,l}$; and
 - a module (42) for reconstructing an output denoised vocal signal $sd(t)$ in the time domain from said output signals sd_l ;

wherein the module (36) for determining an output denoised signal $SD_{k,l}$ comprises:

- at the output of the first extraction module (31), an amplification sub-module (363) according to an amplification coefficient γ , preferably ranging between 0.01 and 0.1 in order to output an amplified signal $\gamma |Y_{k,l}|^2$;
- at the output of the estimation module "SSMB" (35), an adder sub-module (364) suitable for adding the amplified signal $\gamma |Y_{k,l}|^2$ and the square modules $|X_{k,l,i}|^2$, in order to output the square module $|X_{k,l,i}|^2$ of combined denoised components of sub-band $\tilde{X}_{k,l,i}$ specific to each sub-band SB_i of a combined denoised signal $\tilde{X}_{k,l,i}$, satisfying the corresponding equation:

$$|\tilde{X}_{k,l,i}|^2 = |\hat{X}_{k,l,i}|^2 + \gamma |Y_{k,l,i}|^2 ;$$

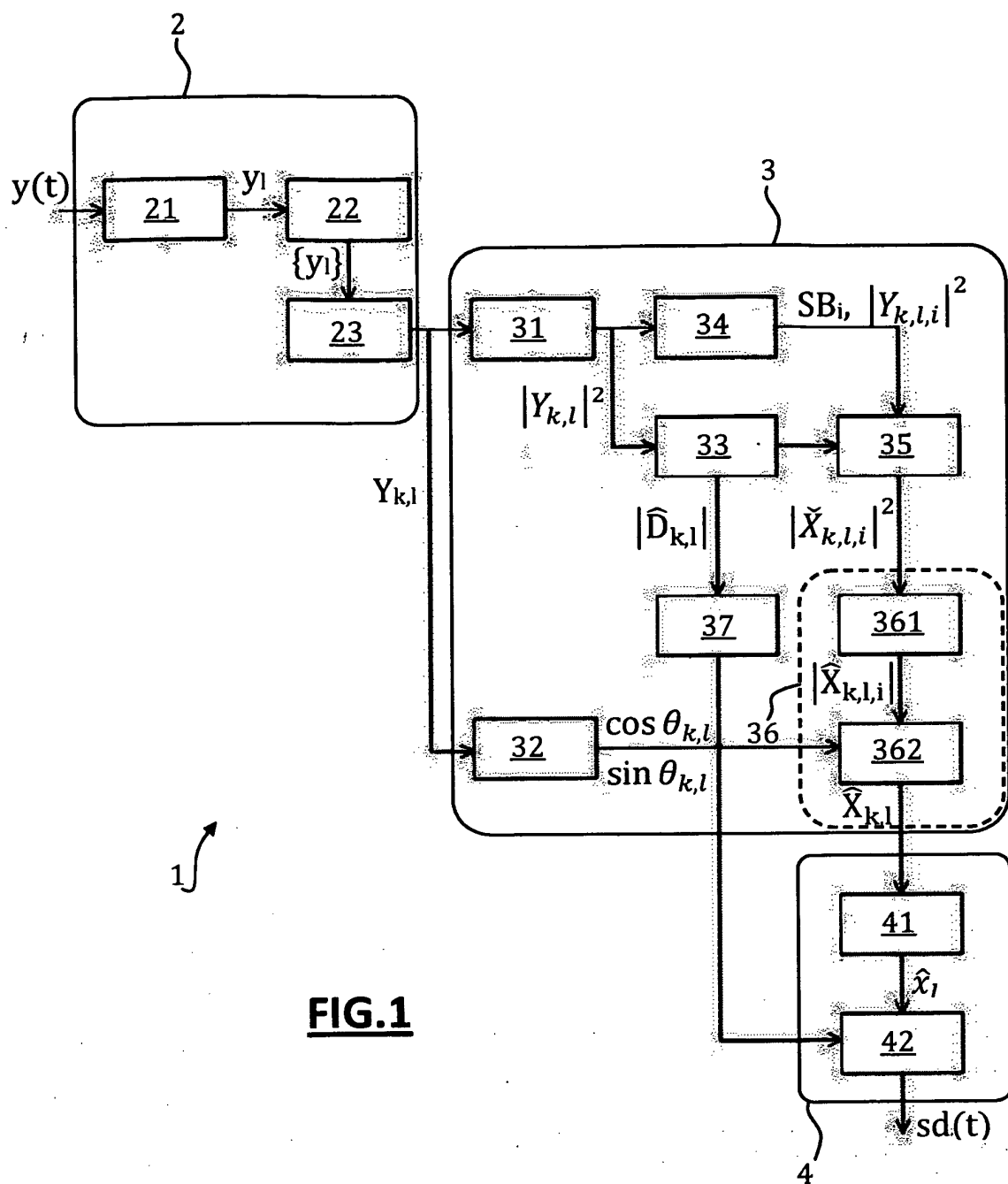
- a square root sub-module (361) for calculating the module $|\tilde{X}_{k,l,i}|$ of the combined denoised components of sub-band $\tilde{X}_{k,l,i}$; and
- a sub-module (362) for recombining combined components of sub-band $\tilde{X}_{k,l,i}$ in order to obtain the combined denoised signal $\tilde{X}_{k,l}$ from the modules $|\tilde{X}_{k,l,i}|$, and possibly the phases $\theta_{k,l}$, such that the output denoised signal $SD_{k,l}$ corresponds to the combined denoised signal $\tilde{X}_{k,l}$, namely $SD_{k,l} = \tilde{X}_{k,l}$.

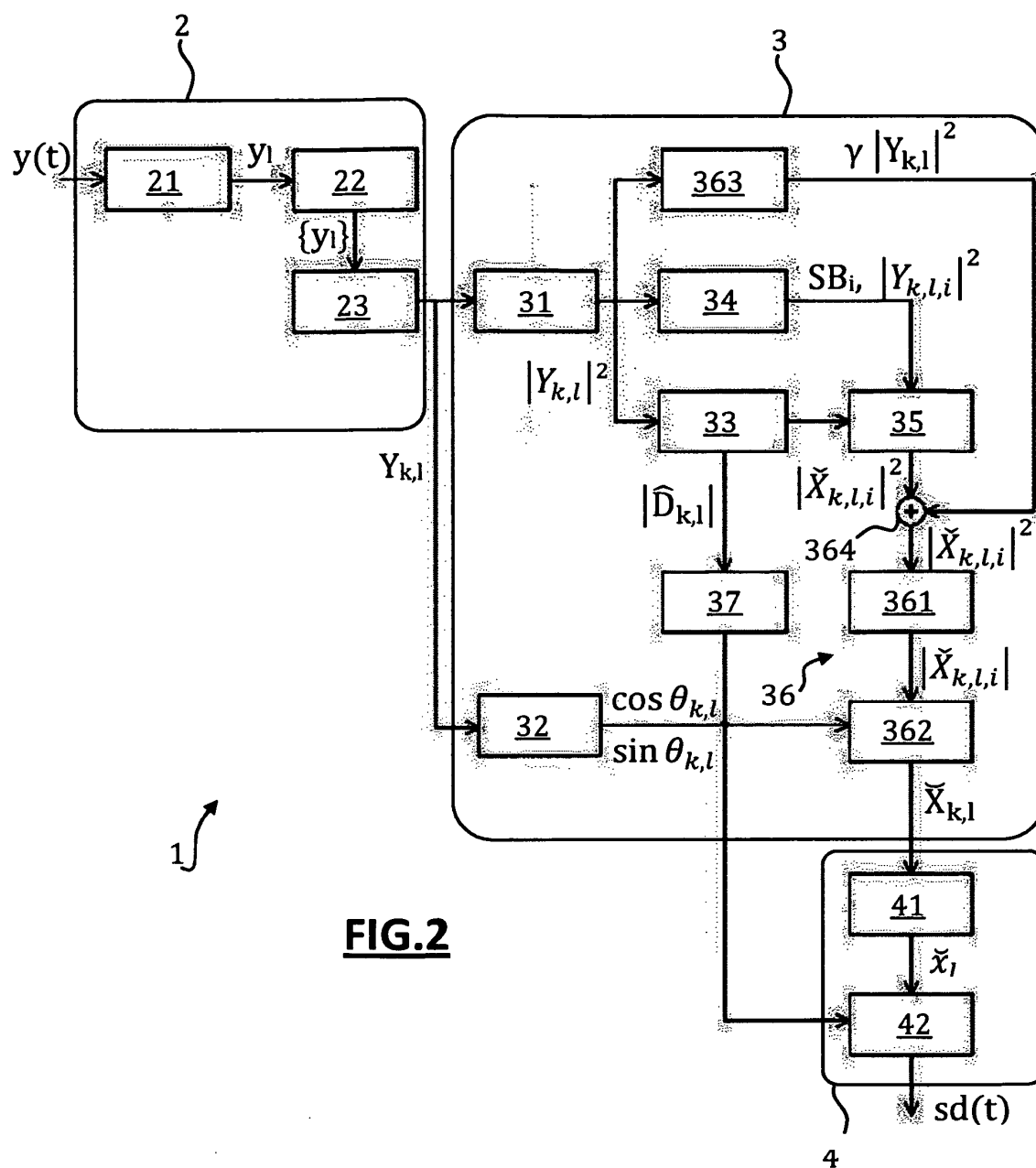
11. The system (1) according to claim 10, wherein the module (36) for determining the output denoised signal $SD_{k,l}$ comprises:

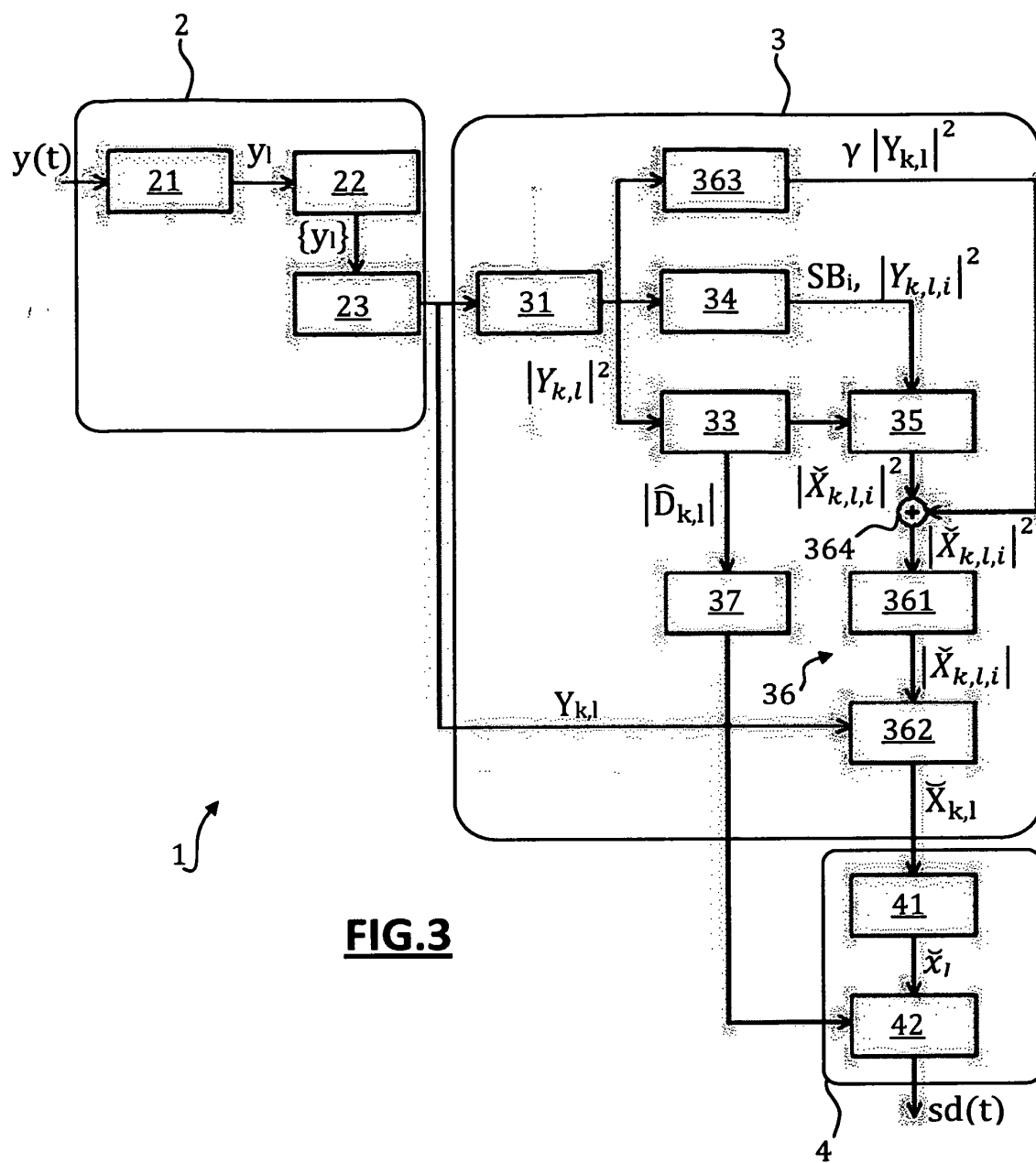
- a square root sub-module (361) for calculating the module $|\hat{X}_{k,l,i}|$ of the denoised components of sub-band $\hat{X}_{k,l,i}$; and
- a sub-module (362) for recombining the denoised components of sub-band $\hat{X}_{k,l,i}$ in order to obtain the denoised signal $R_{k,l}$ based on the modules $|\hat{X}_{k,l,i}|$, and possibly the phases $\theta_{k,l}$, such that the output denoised signal $SD_{k,l}$ corresponds to the denoised signal $R_{k,l}$, namely $SD_{k,l} = \hat{X}_{k,l}$.

12. The system (1) according to any one of claims 10 and 11, wherein the digital processing unit (3) further comprises, at the output of the estimation module "MCRA" (33), a noise detection module (37) comprising

- a module for calculating a mean signal-to-noise ratio r_l specific to each time frame l from the square module $|Y_{k,l}|^2$ and the noise component $\hat{D}_{k,l}$;
- a module for comparing the mean signal-to-noise ratio r_l specific to each time frame l with a predetermined threshold ψ_{TH} ;
- a module for controlling the module (42) for reconstructing the output denoised vocal signal $sd(t)$ which is designed in such a manner that:
 - if the mean signal-to-noise ratio r_l is lower than said threshold ψ_{TH} for the time frame l , then the signal considered before reconstruction for this time frame l corresponds to the output signal Sd_l from the calculation module of said output signal Sd_l ;
 - if the mean signal-to-noise ratio r_l is higher than said threshold ψ_{TH} for the time frame l , then the signal considered before reconstruction for this time frame l corresponds to the sampled signal y_l from the module for slicing the noisy acoustic signal $y(t)$.

**FIG.1**





RÉFÉRENCES CITÉES DANS LA DESCRIPTION

Cette liste de références citées par le demandeur vise uniquement à aider le lecteur et ne fait pas partie du document de brevet européen. Même si le plus grand soin a été accordé à sa conception, des erreurs ou des omissions ne peuvent être exclues et l'OEB décline toute responsabilité à cet égard.

Littérature non-brevet citée dans la description

- **I. COHEN ; B. BERDUGO.** Speech enhancement for non-stationary noise environments. *Signal Processing*, 2001, vol. 81, 2403-2418 [0007]
- **I. COHEN ; B. BERDUGO.** Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement. *IEEE Signal Processing Letters*, Janvier 2002, vol. 9 (1), 12-15 [0007]
- **X. BING-YIN et al.** A Modified Spectral Subtraction Method For Speech Enhancement Based on Masking Property of Human Auditory System. *Wireless Communications & Signals Processing, International Conference on IEEE*, 13 Novembre 2009, 1-5 [0008]
- Multi-band Spectral Subtraction for Enhancing Speech Corrupted by Colored Noise. *IEEE International Conference on Acoustics, Speech And Signal Processing. Proceedings*, 2002, vol. 4 [0010]
- **S.D. KAMATH ; P.C. LOIZOU.** A Multi-Band Spectral Subtraction Method for Enhancing Speech Corrupted by Colored Noise. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2002, vol. 4 [0011]