

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
8 août 2002 (08.08.2002)

PCT

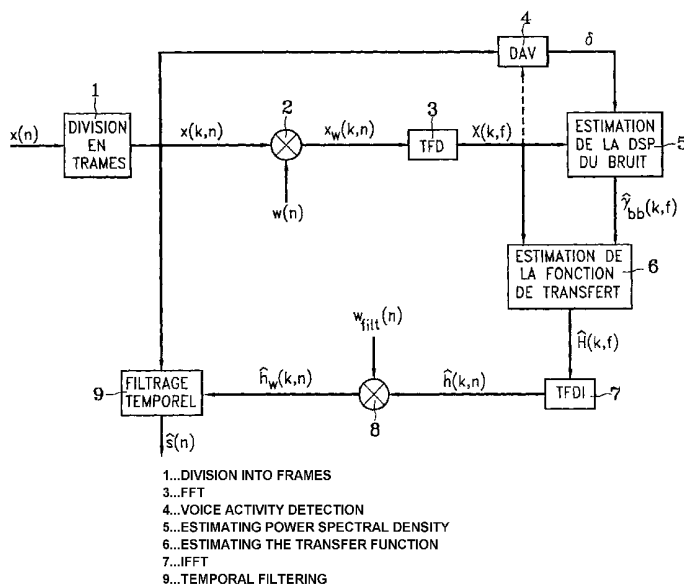
(10) Numéro de publication internationale  
WO 02/061731 A1

- (51) Classification internationale des brevets<sup>7</sup> : G10L 21/02 (71) Déposant (pour tous les États désignés sauf US) : FRANCE TELECOM [FR/FR]; 6, place d'Alleray, F-75015 Paris (FR).
- (21) Numéro de la demande internationale : PCT/FR01/03624
- (22) Date de dépôt international : 19 novembre 2001 (19.11.2001)
- (25) Langue de dépôt : français
- (26) Langue de publication : français
- (30) Données relatives à la priorité : 01/01220 30 janvier 2001 (30.01.2001) FR
- (72) Inventeurs; et (75) Inventeurs/Déposants (pour US seulement) : SCALART, Pascal [FR/FR]; Impasse Poull Ar C'Hog, F-22560 Trebeurden (FR). MARRO, Claude [FR/FR]; 5, rue des Troènes, F-22220 Plouguiel (FR). MAUARY, Laurent [FR/FR]; 16, rue Yves Marie-Herviou, Serval, F-22300 Lannion (FR).
- (74) Mandataires : LOISEL, Bertrand etc.; Cabinet Plasseraud, 84, rue d'Amsterdam, F-75440 Paris Cedex 09 (FR).

[Suite sur la page suivante]

(54) Title: NOISE REDUCTION METHOD AND DEVICE

(54) Titre : PROCÉDE ET DISPOSITIF DE REDUCTION DE BRUIT



(57) Abstract: The invention concerns a method which consists, when analysing an input signal in the frequency domain, in determining a noise level estimator and a useful signal level estimator in an input signal frame, thereby enabling to calculate the transfer function of a first noise-reducing filter; carrying out a second pass to fine-tune the useful signal level estimator, by combining the signal spectrum and the first filter transfer function, then to calculate the transfer function of a second noise-reducing filter on the basis of the fine-tuned useful signal level estimator and the noise level estimator. Said second noise-reducing filter is then used to reduce the noise level in the frame.

(57) Abrégé : Dans une analyse d'un signal d'entrée dans le domaine fréquentiel, on détermine un estimateur de niveau de bruit et un estimateur de niveau de signal utile dans une trame du signal d'entrée, ce qui permet de calculer la fonction de transfert d'un premier filtre réducteur de bruit. Une seconde passe est effectuée pour affiner

[Suite sur la page suivante]

WO 02/061731 A1



(81) **États désignés (national)** : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) **États désignés (régional)** : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ,

CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Déclaration en vertu de la règle 4.17 :**

— *relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement*

**Publiée :**

— *avec rapport de recherche internationale*

*En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.*

---

l'estimateur de niveau de signal utile, en combinant le spectre du signal et la fonction de transfert du premier filtre réducteur de bruit, puis pour calculer la fonction de transfert d'un second filtre réducteur de bruit sur la base de l'estimateur affiné du niveau de signal utile et de l'estimateur de niveau de bruit. Ce second filtre réducteur de bruit est alors utilisé pour réduire le niveau de bruit dans la trame.

## PROCEDE ET DISPOSITIF DE REDUCTION DE BRUIT

La présente invention concerne les techniques de traitement de signal servant à diminuer le niveau de bruit présent dans un signal d'entrée.

Un domaine d'application important est celui du traitement du signal audio (parole ou musique), comprenant de façon non-limitative :

- téléconférence et visioconférence en milieu bruité (en salle dédiée ou bien à partir d'ordinateurs multimédia, etc.) ;
- téléphonie : traitement au niveau des terminaux, fixes ou portables et/ou dans les réseaux de transport ;
- 10 • terminaux mains-libres, notamment de bureau, de véhicule ou portables ;
- prise de son dans les lieux publics (gare, aéroport, etc.) ;
- prise de son mains-libres dans les véhicules ;
- reconnaissance de parole robuste à l'environnement acoustique ;
- prise de son pour le cinéma et les médias (radio, télévision, par exemple
- 15 pour le journalisme sportif ou les concerts, ...).

L'invention est également applicable à tout domaine où l'on cherche à extraire une information utile à partir d'une observation bruitée. On peut notamment citer les domaines suivants : imagerie sous-marine, télédétection sous-marine, traitement des signaux biomédicaux (EEG, ECG, imagerie

20 biomédicale, etc.).

Un problème caractéristique de la prise de son concerne l'environnement acoustique dans lequel est placé le microphone de prise de son et plus particulièrement le fait que, en raison de l'impossibilité de maîtriser complètement cet environnement, un signal perturbateur (qualifié de bruit) est également présent au sein du signal d'observation.

25

Afin d'améliorer la qualité du signal, des systèmes de réduction de bruit sont développés dans le but d'extraire l'information utile en effectuant un traitement sur le signal d'observation bruité. Lorsque le signal audio est un signal de parole transmis à distance, ces systèmes permettent d'accroître son

30 intelligibilité et de réduire la fatigue du correspondant. En plus de ces applications de communication parlée, l'amélioration de la qualité du signal de parole s'avère également utile pour la reconnaissance vocale, dont les performances sont fortement altérées lorsque l'utilisateur est dans un

environnement bruyant.

Le choix d'une technique de traitement du signal pour réaliser l'opération de réduction de bruit dépend en premier lieu du nombre d'observations disponibles à l'entrée du traitement. Dans la présente description, nous considérerons le cas où un seul signal d'observation est disponible. Les méthodes de réduction de bruit adaptées à cette problématique mono-capteur reposent principalement sur des techniques de traitement du signal telles que le filtrage adaptatif par avance/retard, le filtrage paramétrique de Kalman, ou bien encore le filtrage par modification spectrale à court-terme.

Cette dernière famille (filtrage par modification spectrale à court-terme) regroupe pratiquement l'ensemble des solutions utilisées dans les équipements industriels en raison de la simplicité des concepts mis en jeu et de la grande disponibilité d'outils de base (par exemple la transformée de Fourier discrète) nécessaires à leur programmation. Cependant, l'essor de ces techniques de réduction de bruit repose beaucoup sur la possibilité d'effectuer facilement ces traitements en temps réel sur un processeur de traitement du signal, sans introduire de distorsions majeures sur le signal disponible en sortie du traitement. Dans les méthodes de cette famille, le traitement consiste le plus souvent à estimer une fonction de transfert d'un filtre réducteur de bruit, puis à réaliser le filtrage à partir d'une multiplication dans le domaine spectral, ce qui permet d'opérer la réduction de bruit par atténuation spectrale à court-terme, avec un traitement par blocs.

On note  $x(n)$ , le signal d'observation bruité, issu du mélange du signal désiré  $s(n)$  et du bruit de perturbation  $b(n)$ ,  $n$  désignant l'indice temporel en temps discret. Le choix d'une représentation en temps discret est lié à une mise en œuvre orientée vers le traitement numérique du signal, mais on notera que les méthodes décrites ci-après s'appliquent aussi aux signaux en temps continu. Le signal est analysé par segments successifs ou trames d'indice  $k$  de longueur constante. Des notations couramment utilisées pour des représentations dans les domaines discrets temps et fréquence sont :

- $X(k,f)$  : transformée de Fourier ( $f$  est l'indice de fréquence) de la  $k$ -ième trame ( $k$  est l'indice de trame) du signal analysé  $x(n)$  ;
- $S(k,f)$  : transformée de Fourier de la  $k$ -ième trame du signal désiré  $s(n)$  ;

- $\hat{v}$  : estimation d'une grandeur (temporelle ou fréquentielle)  $v$  ; par exemple,  $\hat{S}(k,f)$  est l'estimation de la transformée de Fourier du signal désiré ;
- $\gamma_{uu}(f)$ : densité spectrale de puissance (DSP) d'un signal  $u(n)$ .

5 Dans la plupart des techniques de réduction de bruit le signal bruité  $x(n)$  subit un filtrage dans le domaine fréquentiel pour produire un signal utile estimé  $\hat{s}(n)$  le plus proche possible du signal original  $s(n)$  dépourvu de toute perturbation. Comme indiqué précédemment, cette opération de filtrage consiste à réduire chaque composante fréquentielle  $f$  du signal bruité compte  
 10 tenu du rapport signal à bruit (RSB) estimé dans cette composante. Ce RSB fonction de la fréquence  $f$  est noté ici  $\eta(k,f)$  pour la trame  $k$ .

Pour chacune des trames, le signal est d'abord multiplié par une fenêtre de pondération permettant d'améliorer l'estimation ultérieure des grandeurs spectrales nécessaires au calcul du filtre réducteur de bruit. Chaque  
 15 trame ainsi fenêtrée est ensuite analysée dans le domaine spectral (généralement à l'aide de la transformée de Fourier discrète dans sa version rapide). Cette opération est appelée transformation de Fourier à court terme (TFCT). Cette représentation fréquentielle  $X(k,f)$  du signal observé permet à la fois d'estimer la fonction de transfert  $H(k,f)$  du filtre réducteur de bruit, et  
 20 d'appliquer ce filtre dans le domaine spectral par simple multiplication entre cette fonction de transfert et le spectre à court terme du signal bruité, soit :

$$\hat{S}(k,f) = H(k,f).X(k,f) \quad (1)$$

Le retour dans le domaine temporel du signal ainsi obtenu est ensuite effectué par simple transformée spectrale inverse. La synthèse du signal  
 25 débruité est généralement effectuée par une technique de recouvrement et addition de blocs (OLA, « overlap add ») ou une technique de sauvegarde de blocs (OLS, « overlap save »). Cette opération de reconstruction du signal dans le domaine temporel est appelée transformation de Fourier à court terme inverse (TFCTI).

30 On trouvera une description détaillée des méthodes d'atténuation spectrale à court terme dans les références : J.S. Lim, A.V. Oppenheim, "Enhancement and bandwidth compression of noisy speech", Proceedings of

the IEEE, vol. 67, pp. 1586-1604, 1979 ; et R.E. Crochiere, L.R. Rabiner, "Multirate digital signal processing", Prentice Hall, 1983.

Les principales tâches réalisées par un tel système de réduction de bruit sont :

- 5       • la détection d'activité vocale (DAV) ;
- l'estimation de la densité spectrale de puissance (DSP) de bruit pendant les instants d'inactivité vocale ;
- l'application d'une atténuation spectrale à court terme évaluée à partir d'une règle de suppression des composantes spectrales de bruit ;
- 10       • la synthèse du signal traité à partir d'une technique de type OLS ou OLA.

Le choix de la règle de suppression des composantes de bruit est important puisqu'il détermine la qualité du signal transmis. Ces règles de suppression ne modifient généralement que l'amplitude  $|X(k, f)|$  des composantes spectrales du signal bruité, et non leur phase. On fait en général

15 les hypothèses suivantes :

- le bruit et le signal utile sont statistiquement décorrélés ;
- le signal utile est intermittent (présence de périodes de silence dans lesquelles le bruit peut être estimé) ;
- l'oreille humaine n'est pas sensible à la phase du signal (voir D.L. Wang, J.S. Lim, "The unimportance of phase in speech enhancement ", IEEE
- 20       Trans. on ASSP, vol. 30, N° 4, pp. 679-681, 1982).

L'atténuation spectrale à court-terme  $H(k, f)$ , appliquée au signal d'observation  $X(k, f)$  sur la trame d'indice  $k$  à la composante fréquentielle  $f$ , est généralement déterminée à partir de l'estimation du rapport signal à bruit local

25  $\eta(k, f)$ . Une caractéristique commune à l'ensemble des règles de suppression est leur comportement asymptotique, donné par :

$$\begin{aligned} H(k, f) &\approx 1 \quad \text{pour } \eta(k, f) \gg 1 \\ H(k, f) &\approx 0 \quad \text{pour } \eta(k, f) \ll 1 \end{aligned} \quad (2)$$

Les règles de suppression couramment employées sont :

- 30       • la soustraction spectrale en puissance (voir l'article précité de J.S. Lim et A.V. Oppenheim), pour laquelle la fonction de transfert  $H(k, f)$  du filtre réducteur de bruit s'écrit :

- 5 -

$$H(k,f) = \sqrt{\frac{\gamma_{ss}(k,f)}{\gamma_{bb}(k,f) + \gamma_{ss}(k,f)}} \quad (3)$$

- la soustraction spectrale en amplitude (voir S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction", IEEE Trans. on Audio, Speech and Signal Processing, vol. 27, N°. 2, pp. 113-120, Avril 1979), pour laquelle la fonction de transfert  $H(k,f)$  s'écrit :

$$H(k,f) = 1 - \sqrt{\frac{\gamma_{bb}(k,f)}{\gamma_{bb}(k,f) + \gamma_{ss}(k,f)}} \quad (4)$$

- la mise en œuvre directe du filtre de Wiener (cf. article précité de J.S. Lim et A.V. Oppenheim), pour laquelle la fonction de transfert  $H(k,f)$  s'écrit :

$$H(k,f) = \frac{\gamma_{ss}(k,f)}{\gamma_{bb}(k,f) + \gamma_{ss}(k,f)} \quad (5)$$

10 Dans ces expressions,  $\gamma_{ss}(k,f)$  et  $\gamma_{bb}(k,f)$  représentent respectivement les densités spectrales de puissance du signal utile et du bruit présent au sein de la composante fréquentielle  $f$  du signal d'observation  $X(k,f)$  sur la trame d'indice  $k$ .

A partir des expressions (3)-(5), il est possible d'étudier, en fonction du rapport signal à bruit local mesuré sur une composante fréquentielle donnée  $f$ , le comportement de l'atténuation spectrale appliquée au signal bruité. On constate que toutes les règles donnent lieu à une atténuation identique lorsque le rapport signal à bruit local est important. La règle de soustraction en puissance est optimale au sens du maximum de vraisemblance pour des modèles gaussiens (voir O. Cappé, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor", IEEE Trans. on Speech and Audio Processing, vol. 2, N°. 2, pp. 345-349, Avril 1994). Mais c'est celle pour laquelle la puissance du bruit reste la plus importante en sortie du traitement. Pour l'ensemble des règles de suppression, on constate qu'une faible variation du rapport signal à bruit local autour de la valeur de coupure suffit à faire passer du cas de l'atténuation totale ( $H(k,f) \approx 0$ ) au cas d'une modification spectrale négligeable ( $H(k,f) \approx 1$ ).

Cette dernière propriété constitue l'une des causes du phénomène qualifié de « bruit musical ». En effet, le bruit ambiant, caractérisé à la fois par des composantes déterministes et aléatoires, ne peut être caractérisé que

- 6 -

durant les périodes de non-activité vocale. Du fait de la présence de ces composantes aléatoires, il existe de très fortes variations entre la contribution réelle d'une composante fréquentielle  $f$  du bruit pendant les périodes d'activité vocale et son estimation en moyenne réalisée sur plusieurs trames pendant les  
5 instants de non-activité vocale. Du fait de cette différence, l'estimation du rapport signal à bruit local peut fluctuer autour du niveau de coupure, soit donc engendrer en sortie de traitement des composantes spectrales qui apparaissent, puis disparaissent et dont la durée de vie moyenne ne dépasse pas statistiquement l'ordre de grandeur de la fenêtre d'analyse considérée. La  
10 généralisation de ce comportement sur l'ensemble de la bande passante introduit un bruit résiduel audible et gênant, qualifié de « bruit musical ».

Plusieurs études se sont attachées à réduire l'influence de ce bruit. Les solutions préconisées se déclinent suivant plusieurs axes :

- moyennage des estimations à court-terme (cf. article précité S.F. Boll) ;
- 15 • surestimation du spectre de puissance du bruit (voir M. Berouti et al, "Enhancement of speech corrupted by acoustic noise", Int. Conf. on Speech, Signal Processing, pp. 208-211, 1979; et P. Lockwood, J. Boudy, "Experiments with a non-linear spectral subtractor, hidden Markov models and the projection for robust speech recognition in cars", Proc. of  
20 EUSIPCO'91, pp. 79-82, 1991);
- poursuite des minima de la densité spectrale de bruit (voir R. Martin, "Spectral subtraction based on minimum statistics", in Signal Processing VII: Theories and Applications, EUSIPCO'94, pp. 1182-1185, Sept. 1994).

25 Plusieurs études ont également porté sur l'établissement de nouvelles règles de suppression basées sur des modèles statistiques des signaux de parole et de bruit additif. Ces études ont permis d'introduire de nouveaux algorithmes à « décisions souples » car ils possèdent un degré de liberté supplémentaire par rapport aux méthodes classiques (voir R.J. Mac Aulay,  
30 M.L. Malpass, "Speech enhancement using a soft-decision noise suppression filter", IEEE trans. on Audio, Speech and Signal Processing, vol. 28, N°. 2, pp. 137-145, Avril 1980, Y. Ephraim, D. Malah, "Speech enhancement using optimal non-linear spectral amplitude estimation", Int. Conf. on Speech, Signal



Processing, pp. 1118-1121, 1983, Y. Ephraim, D. Malha, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator", IEEE Trans. on ASSP, vol. 32, N° 6, pp. 1109-1121, 1984).

5 Les règles de modification spectrale à court terme précitées présentent les caractéristiques suivantes :

- Le calcul de l'atténuation spectrale à court-terme repose sur l'estimation du rapport signal à bruit sur chacune des composantes spectrales, les équations (3)-(5) faisant chacune intervenir la quantité :

$$10 \quad \eta(k,f) = \frac{\gamma_{ss}(k,f)}{\gamma_{bb}(k,f)} \quad (6)$$

Ainsi, les performances de la technique de réduction de bruit (distorsions, réduction effective du niveau de bruit) sont gouvernées par la pertinence de cet estimateur du rapport signal à bruit.

- Ces techniques sont basées sur un traitement par blocs (avec possibilité de recouvrement entre les blocs successifs) qui consiste à filtrer  
15 l'ensemble des échantillons d'une trame donnée, présente à l'entrée du dispositif de réduction de bruit, par une seule atténuation spectrale. Cette propriété réside dans le fait que l'application du filtre est effectuée par une multiplication dans le domaine spectral. Ceci est particulièrement  
20 contraignant lorsque le signal présent sur la trame courante ne vérifie pas les hypothèses de stationnarité au second ordre, par exemple dans le cas d'un début ou fin de mot, ou bien dans le cas d'une trame mixte voisé/non-voisé.
- La multiplication effectuée dans le domaine spectral correspond en  
25 réalité à une opération de convolution cyclique. En pratique, pour éviter les distorsions, l'opération qu'on cherche à effectuer est une convolution linéaire, ce qui nécessite à la fois d'ajouter un certain nombre d'échantillons nuls à chaque trame d'entrée (technique dite du « zero padding ») et d'effectuer un traitement supplémentaire visant à limiter le  
30 support temporel de la réponse impulsionnelle du filtre réducteur de bruit. Satisfaire à la contrainte de convolution temporelle augmente ainsi nécessairement l'ordre de la transformée spectrale et, par suite, la

complexité arithmétique du traitement de réduction de bruit. La technique la plus utilisée pour limiter le support temporel de la réponse impulsionnelle du filtre réducteur de bruit consiste à introduire une contrainte dans le domaine temporel, ce qui nécessite (i) une première transformation spectrale « inverse » permettant d'obtenir la réponse impulsionnelle  $h(k,n)$  à partir de la connaissance de la fonction de transfert du filtre  $H(k,f)$ , (ii) une limitation du nombre de points de cette réponse impulsionnelle, conduisant à un filtre temporel tronqué  $h'(k,n)$ , puis (iii) une seconde transformation spectrale « directe » permettant d'obtenir la fonction de transfert modifiée  $H'(k,f)$  à partir de la réponse impulsionnelle tronquée  $h'(k,n)$ .

• En pratique, chaque trame d'analyse est multipliée par une fenêtre d'analyse  $w(n)$  avant d'effectuer l'opération de transformée spectrale. Lorsque le filtre réducteur de bruit est de type passe-tout (c'est-à-dire  $H(k,f) \approx 1, \forall f$ ), la fenêtre d'analyse doit satisfaire à la condition suivante

$$\sum_k w(n-k.D) = 1 \quad (7)$$

si l'on souhaite que la condition de reconstitution parfaite soit vérifiée. Dans cette équation, le paramètre  $D$  représente le décalage (en nombre d'échantillons) entre deux trames d'analyse successives. D'un autre côté, le choix de la fenêtre de pondération  $w(n)$  (typiquement de type Hanning, Hamming, Blackman, etc.) conditionne la largeur du lobe principal de  $W(f)$  ainsi que l'amplitude des lobes secondaires (relativement à celle du lobe principal). Si le lobe principal est large, les transitions rapides de la transformée du signal original sont très mal approchées. Si l'amplitude relative des lobes secondaires est grande, l'approximation obtenue possède des oscillations gênantes, surtout autour des discontinuités. Il est donc difficile de satisfaire à la fois à la contrainte d'analyse spectrale pertinente (choix de la largeur du lobe principal, et de l'amplitude des lobes latéraux) et à la contrainte de faible retard introduit par le processus de filtrage de réduction de bruit (décalage temporel entre le signal en entrée et en sortie du traitement). Respecter la seconde contrainte amène à utiliser des trames successives sans aucun

recouvrement et donc une fenêtre d'analyse de type rectangulaire, ce qui ne conduit pas à effectuer une analyse spectrale pertinente. Le seul moyen de respecter ces deux contraintes en même temps est d'effectuer une analyse spectrale à partir d'une première transformation spectrale réalisée sur une trame pondérée par une fenêtre d'analyse adéquate (pour effectuer une bonne estimation spectrale), et de réaliser en parallèle une seconde transformée spectrale sur les données non-fenêtrées (pour effectuer l'opération de convolution par multiplication spectrale). Une telle technique se révèle en pratique beaucoup trop coûteuse en termes de complexité arithmétique.

EP-A-0 710 947 décrit un dispositif de réduction de bruit couplé à un annuleur d'écho. La réduction de bruit est opérée par un filtrage par blocs dans le domaine temporel, au moyen d'une réponse impulsionnelle obtenue par transformation de Fourier inverse de la fonction de transfert  $H(k,f)$  estimée en fonction du rapport signal à bruit au cours de l'analyse spectrale.

Un but principal de la présente invention est d'améliorer les performances des méthodes de réduction de bruit.

L'invention propose ainsi un procédé de réduction de bruit dans des trames successives d'un signal d'entrée, comprenant les étapes suivantes pour certaines au moins des trames :

- calculer un spectre du signal d'entrée par transformation dans le domaine fréquentiel ;
- obtenir un estimateur de niveau de bruit en fonction de la fréquence ;
- calculer un premier estimateur de niveau de signal utile dans la trame en fonction de la fréquence ;
- calculer la fonction de transfert d'un premier filtre réducteur de bruit sur la base du premier estimateur de niveau de signal utile et de l'estimateur de niveau de bruit ;
- calculer un second estimateur de niveau de signal utile dans la trame en fonction de la fréquence, en combinant le spectre du signal d'entrée et la fonction de transfert du premier filtre réducteur de bruit ;

- calculer la fonction de transfert d'un second filtre réducteur de bruit sur la base du second estimateur de niveau de signal utile et de l'estimateur de niveau de bruit ; et
- utiliser la fonction de transfert du second filtre réducteur de bruit dans une opération de filtrage de la trame pour produire un signal à bruit réduit.

Les niveaux de bruit et de signal utile qu'on estime sont typiquement des DSP, ou plus généralement des grandeurs corrélées à ces DSP.

Le calcul en deux passes, dont la particularité réside dans une mise à jour plus rapide de la DSP du signal utile  $\gamma_{ss}(k,f)$ , confère au second filtre réducteur de bruit deux avantages appréciables par rapport aux méthodes antérieures. D'une part, on obtient un suivi plus rapide des non-stationnarités du signal utile, en particulier lors des variations rapides de son enveloppe temporelle (par exemple des attaques ou des extinctions pour du signal de parole lors d'une transition silence/parole). D'autre part, le filtre réducteur de bruit est mieux estimé, ce qui se traduit par un accroissement des performances de la méthode (réduction de bruit plus importante et dégradation du signal utile réduite).

Le procédé est généralisable aux cas où on effectue plus de deux passes. A partir de la p-ième fonction de transfert obtenue ( $p \geq 2$ ), on recalcule alors l'estimateur de niveau de signal utile, et on réévalue une (p+1)-ième fonction de transfert pour la réduction de bruit. La définition ci-dessus du procédé vaut aussi pour de tels cas où on fait  $P > 2$  passes : il suffit de considérer que le « premier estimateur de niveau de signal utile » selon cette définition est celui obtenu au cours de la (P-1)-ième passe. En pratique, on observe de bonnes performances du procédé avec  $P = 2$ .

Dans une réalisation avantageuse du procédé, le calcul du spectre comporte une pondération de la trame de signal d'entrée par une fonction de fenêtrage et une transformation de la trame pondérée vers le domaine fréquentiel, la fonction de fenêtrage étant dissymétrique de manière à appliquer une pondération plus forte à la moitié la plus récente de la trame qu'à la moitié la moins récente de la trame.

Le choix d'une telle fonction de fenêtrage permet de concentrer le

5 poids de l'estimation spectrale vers les échantillons les plus récents tout en permettant une fenêtre ayant de bonnes propriétés spectrales (remontée contrôlée des lobes secondaires). Ceci permet de suivre rapidement les variations du signal. Il est à noter que ce mode de calcul du spectre pour l'analyse en fréquence est également applicable lorsque l'estimation de la fonction de transfert du filtre réducteur de bruit est réalisée en une seule passe.

Le procédé est utilisable lorsque le filtrage du signal d'entrée est effectué par blocs dans le domaine fréquentiel, par les méthodes précitées d'atténuation spectrale à court terme. Le signal débruité est alors produit sous la forme de ses composantes spectrales  $\hat{S}(k,f)$ , qui peuvent être exploitées directement (par exemple dans une application de codage ou de reconnaissance de la parole) ou transformées vers le domaine temporel pour obtenir explicitement le signal  $\hat{s}(n)$ .

Toutefois, dans une réalisation préférée du procédé, on détermine une réponse impulsionnelle de filtre réducteur de bruit pour la trame courante à partir d'une transformation vers le domaine temporel de la fonction de transfert du second filtre réducteur de bruit, et on réalise l'opération de filtrage de la trame dans le domaine temporel au moyen de la réponse impulsionnelle déterminée pour ladite trame.

Avantageusement, la détermination de la réponse impulsionnelle de filtre réducteur de bruit pour la trame courante comporte alors les étapes suivantes :

- transformer vers le domaine temporel la fonction de transfert du second filtre réducteur de bruit pour obtenir une première réponse impulsionnelle ; et
- tronquer la première réponse impulsionnelle à une longueur de troncature correspondant à un nombre d'échantillons sensiblement plus petit (typiquement au moins cinq fois plus petit) que le nombre de points de la transformation vers le domaine temporel.

Cette limitation du support temporel du filtre réducteur de bruit présente un double avantage. D'une part elle permet d'éviter les problèmes de repliement temporel (respect de la convolution linéaire). D'autre part, elle assure un lissage permettant d'éviter les effets d'un filtre trop agressif, qui

pourraient dégrader le signal utile. Elle peut être accompagnée d'une pondération de la réponse impulsionnelle tronquée par une fonction de fenêtrage sur un nombre d'échantillons correspondant à la longueur de troncature. Il est à noter que cette limitation du support temporel du filtre est également applicable lorsque l'estimation de la fonction de transfert est réalisée en une seule passe.

Lorsque le filtrage est réalisé dans le domaine temporel, il est avantageux de subdiviser la trame courante en plusieurs sous-trames et de calculer pour chaque sous-trame une réponse impulsionnelle interpolée à partir de la réponse impulsionnelle de filtre réducteur de bruit déterminée pour la trame courante et de la réponse impulsionnelle de filtre réducteur de bruit déterminée pour au moins une trame précédente. L'opération de filtrage de la trame comporte alors un filtrage du signal de chaque sous-trame dans le domaine temporel conformément à la réponse impulsionnelle interpolée calculée pour ladite sous-trame.

Ce traitement en sous-trames permet d'appliquer un filtre réducteur de bruit variant au sein d'une même trame, et donc bien adapté aux non-stationnarités du signal traité. Dans le cas du traitement d'un signal vocal, ceci se rencontre en particulier sur les trames mixtes, c'est-à-dire présentant des sons voisés et non-voisés. Il est à noter que ce traitement en sous-trames est également applicable lorsque l'estimation de la fonction de transfert du filtre est réalisée en une seule passe.

Un autre aspect de la présente invention se rapporte à un dispositif de réduction de bruit adapté à la mise en œuvre du procédé ci-dessus.

D'autres particularités et avantages de la présente invention apparaîtront dans la description ci-après d'exemples de réalisation non limitatifs, en référence aux dessins annexés, dans lesquels :

- la figure 1 est un schéma synoptique d'un dispositif de réduction de bruit apte à la mise en œuvre du procédé selon l'invention ;
- la figure 2 est un schéma synoptique d'une unité d'estimation de la fonction de transfert d'un filtre réducteur de bruit utilisable dans un dispositif selon la figure 1 ;

- la figure 3 est un schéma synoptique d'une unité de filtrage temporel utilisable dans un dispositif selon la figure 1 ; et
- la figure 4 est un graphique d'une fonction de fenêtrage utilisable dans une réalisation particulière du procédé.

5 Les figures 1 à 3 donnent une représentation d'un dispositif selon l'invention sous la forme d'unités distinctes. Dans une implémentation typique du procédé, les opérations de traitement de signal sont effectuées, comme il est usuel, par un processeur de signal numérique exécutant des programmes dont les différents modules fonctionnels correspondent aux unités précitées.

10 En référence à la figure 1, un dispositif de réduction de bruit selon l'invention comporte une unité 1 qui distribue le signal entrée  $x(n)$ , tel qu'un signal audio numérique, en trames successives de longueur  $L$  échantillons (indexées par un entier  $k$ ). Chaque trame d'indice  $k$  est pondérée (multiplieur 2) en la multipliant par une fonction de fenêtrage  $w(n)$ , ce qui produit le signal  
15  $x_w(k,n) = w(n).x(k,n)$  pour  $0 \leq n < L$ .

Le passage dans le domaine fréquentiel est réalisé en appliquant la transformée de Fourier discrète (TFD) aux trames pondérées  $x_w(k,n)$  au moyen d'une unité 3 qui délivre la transformée de Fourier  $X(k,f)$  de la trame courante.

Pour les passages dans les domaines temps-fréquence et vice-versa  
20 mis en jeu dans l'invention, la TFD et la transformée inverse vers le domaine temporel (TFDI) utilisée en aval le cas échéant (unité 7) sont avantageusement une transformée de Fourier rapide (ou FFT, « Fast Fourier Transform ») et transformée de Fourier inverse rapide (ou IFFT, « Inverse Fast Fourier Transform »), respectivement. D'autres transformations temps-fréquence, telles  
25 que la transformée en ondelettes, sont également utilisables.

Une unité 4 de détection d'activité vocale (DAV) permet de discriminer les trames de bruit seul et les trames de parole, et délivre une indication binaire d'activité vocale  $\delta$  pour la trame courante. Toute méthode connue de DAV est utilisable, qu'elle opère dans le domaine temporel sur la base du signal  $x(k,n)$   
30 ou, comme indiqué en trait interrompu, dans le domaine fréquentiel sur la base du signal  $X(k,f)$ .

La DAV contrôle l'estimation de la DSP du bruit par l'unité 5. Ainsi, pour chaque trame  $k_b$  de "bruit seul" détectée par l'unité 4 ( $\delta = 0$ ), la densité

- 14 -

spectrale de puissance de bruit  $\hat{\gamma}_{bb}(k_b, f)$  est estimée par l'expression récursive suivante :

$$\begin{cases} \hat{\gamma}_{bb}(k_b, f) = \alpha(k_b) \cdot \hat{\gamma}_{bb}(k_b - 1, f) + (1 - \alpha(k_b)) |X(k_b, f)|^2 \\ \hat{\gamma}_{bb}(k, f) = \hat{\gamma}_{bb}(k_b, f) \end{cases} \quad (10)$$

où  $k_b$  est soit la trame de bruit courante si  $\delta = 0$ , soit la dernière trame de bruit si  $\delta = 1$  ( $k$  est détectée comme trame de signal utile), et  $\alpha(k_b)$  est un paramètre de lissage pouvant varier au cours du temps.

On notera que le mode de calcul de  $\hat{\gamma}_{bb}(k_b, f)$  n'est pas limité à cet estimateur à lissage exponentiel, tout autre estimateur de DSP pouvant être utilisé par l'unité 5.

10 A partir du spectre  $X(k, f)$  de la trame courante et de l'estimation de niveau de bruit  $\hat{\gamma}_{bb}(k_b, f)$ , une autre unité 6 estime la fonction de transfert (FT) du filtre réducteur de bruit  $\hat{H}(k, f)$ . L'unité 7 applique la TFDI à cette FT pour obtenir la réponse impulsionnelle correspondante  $\hat{h}(k, n)$ .

Une fonction de fenêtrage  $w_{\text{filt}}(n)$  est appliquée à cette réponse impulsionnelle  $\hat{h}(k, n)$  par un multiplieur 8 afin d'obtenir la réponse impulsionnelle  $\hat{h}_w(k, n)$  du filtre temporel du dispositif de réduction de bruit. L'opération effectuée par l'unité de filtrage 9 pour produire le signal temporel débruité  $\hat{s}(n)$  est, dans son principe, une convolution du signal d'entrée par la réponse impulsionnelle  $\hat{h}_w(k, n)$  déterminée pour la trame courante.

20 La fonction de fenêtrage  $w_{\text{filt}}(n)$  a un support nettement plus court que la longueur d'une trame. En d'autres termes, la réponse impulsionnelle  $\hat{h}(k, n)$  résultant de la TFDI est tronquée avant de lui appliquer la pondération par la fonction  $w_{\text{filt}}(n)$ . De préférence, la longueur de troncature  $L_{\text{filt}}$ , exprimée en nombre d'échantillons, est au moins cinq fois plus petite que la longueur de la trame. Elle est typiquement de l'ordre de grandeur du dixième de cette longueur de trame.

Les  $L_{\text{filt}}$  coefficients les plus significatifs de la réponse impulsionnelle font l'objet de la pondération par la fenêtre  $w_{\text{filt}}(n)$ , qui est par exemple une



fenêtre de Hamming ou de Hanning de longueur  $L_{\text{filt}}$  :

$$\hat{h}_w(k, n) = w_{\text{filt}}(n) \cdot \hat{h}(k, n) \quad \text{pour } 0 \leq n < L_{\text{filt}} \quad (11)$$

La limitation du support temporel du filtre réducteur de bruit permet d'éviter les problèmes de repliement temporel, afin de respecter la convolution  
5 linéaire. Elle assure en outre un lissage permettant d'éviter les effets d'un filtre trop agressif qui pourraient dégrader le signal utile.

La figure 2 illustre une organisation préférée de l'unité 6 d'estimation de la fonction de transfert  $\hat{H}(k, f)$  du filtre réducteur de bruit, qui dépend de la DSP du bruit  $b(n)$  et de celle du signal utile  $s(n)$ .

10 On a décrit comment l'unité 5 peut estimer la DSP du bruit  $\hat{\gamma}_{bb}(k, f)$ . Mais la DSP  $\gamma_{ss}(k, f)$  du signal utile ne peut être obtenue directement du fait du mélange du signal et du bruit pendant les périodes d'activité vocale. Pour la pré-estimer, le module 11 de l'unité 6 de la figure 2 utilise par exemple un estimateur à décision dirigée (voir Y. Ephraim, D. Malha, "Speech  
15 enhancement using a minimum mean square error short-time spectral amplitude estimator", IEEE Trans. on ASSP, vol. 32, N° 6, pp. 1109-1121, 1984), conformément à l'expression suivante :

$$\hat{\gamma}_{ss1}(k, f) = \beta(k) \cdot \left| \hat{S}(k-1, f) \right|^2 + (1 - \beta(k)) P \left[ \left| X(k, f) \right|^2 - \hat{\gamma}_{bb}(k, f) \right] \quad (12)$$

où  $\beta(k)$  est un paramètre barycentrique pouvant varier au cours du temps et  
20  $\hat{S}(k-1, f)$  est le spectre du signal utile estimé relativement à la trame précédente d'indice  $k-1$  (par exemple  $\hat{S}(k-1, f) = \hat{H}(k-1, f) \cdot X(k-1, f)$ , obtenu par le multiplieur 12 de la figure 2). La fonction  $P$  assure le seuillage de la grandeur  $\left| X(k, f) \right|^2 - \hat{\gamma}_{bb}(k, f)$  qui risque d'être négative en cas d'erreur d'estimation. Elle est donnée par :

$$25 \quad P[z(k, f)] = \begin{cases} z(k, f) & \text{si } z(k, f) > 0 \\ \hat{\gamma}_{bb}(k, f) & \text{sinon} \end{cases} \quad (13)$$

Il est à noter que le calcul de  $\hat{\gamma}_{ss1}(k, f)$  n'est pas limité à cet estimateur à décision dirigée. En effet un estimateur à lissage exponentiel ou tout autre estimateur de densité spectrale de puissance peut-être utilisé.

Une pré-estimation de la FT du filtre réducteur de bruit pour la trame courante est calculée par le module 13, comme une fonction des DSP estimées  $\hat{\gamma}_{ss1}(k,f)$  et  $\hat{\gamma}_{bb}(k,f)$  :

$$\hat{H}_1(k,f) = F(\hat{\gamma}_{ss1}(k,f), \hat{\gamma}_{bb}(k,f)) \quad (14)$$

5 Ce module 13 peut notamment mettre en œuvre la règle de soustraction spectrale en puissance ( $F(y, z) = \sqrt{\frac{y}{y+z}}$ , selon (3)), en amplitude ( $F(y, z) = 1 - \sqrt{\frac{z}{y+z}}$ , selon (4)), ou encore celle du filtre de Wiener en boucle ouverte ( $F(y, z) = \frac{y}{y+z}$ , selon (5)).

Habituellement, la fonction de transfert finale du filtre réducteur de bruit  
10 est obtenue par l'équation (14). Pour améliorer les performances du filtre, on propose de l'estimer par une procédure itérative en deux passes. La première passe consiste en les opérations effectuées par les modules 11 à 13.

La fonction de transfert  $\hat{H}_1(k,f)$  ainsi obtenue est réutilisée pour affiner l'estimation de la DSP du signal utile. L'unité 6 (multiplieur 14 et module 15)  
15 calcule pour cela la grandeur  $\hat{\gamma}_{ss}(k,f)$  donnée par :

$$\hat{\gamma}_{ss}(k,f) = |\hat{H}_1(k,f) \cdot X(k,f)|^2 \quad (15)$$

La deuxième passe consiste alors pour le module 16 à calculer l'estimateur final  $\hat{H}(k,f)$  de la fonction de transfert du filtre réducteur de bruit à partir de l'estimation affinée de la DSP du signal utile :

$$20 \quad \hat{H}(k,f) = F(\hat{\gamma}_{ss}(k,f), \hat{\gamma}_{bb}(k,f)) \quad (16)$$

la fonction  $F$  pouvant être la même que celle employée par le module 13.

Ce calcul en deux passes permet une mise à jour plus rapide de la DSP du signal utile  $\hat{\gamma}_{ss}(k,f)$  et une meilleure estimation du filtre.

La figure 3 illustre une organisation préférée de l'unité de filtrage  
25 temporel 9, fondée sur une subdivision de la trame courante en  $N$  sous-frames et permettant ainsi d'appliquer une fonction de réduction de bruit capable d'évoluer au sein d'une même trame de signal.

Un module 21 effectue une interpolation de la réponse impulsionnelle

tronquée et pondérée  $\hat{h}_w(k, n)$  afin d'obtenir un jeu de  $N \geq 2$  réponses impulsionnelles de filtres de sous-trames  $\hat{h}_w^{(i)}(k, n)$  pour  $i$  allant de 1 à  $N$ .

Le filtrage en sous-trame peut être implémenté à l'aide d'un filtre transverse 23 de longueur  $L_{\text{filt}}$  dont les coefficients  $\hat{h}_w^{(i)}(k, n)$  ( $0 \leq n < L_{\text{filt}}$ ,  $1 \leq i \leq N$ ) sont présentés en cascade par le sélecteur 22 sur la base de l'index  $i$  de la sous-trame courante. Les sous-trames du signal à filtrer sont obtenues par une subdivision de la trame d'entrée  $x(k, n)$ . Le filtre transverse 23 calcule ainsi le signal à bruit réduit  $\hat{s}(n)$  par convolution du signal d'entrée  $x(n)$  avec les coefficients  $\hat{h}_w^{(i)}(k, n)$  relatifs à la sous-trame courante.

Les réponses  $\hat{h}_w^{(i)}(k, n)$  des filtres de sous-trames peuvent être calculées par le module 21 comme des sommes pondérées de la réponse impulsionnelle  $\hat{h}_w(k, n)$  déterminée pour la trame courante et de celle  $\hat{h}_w(k-1, n)$  déterminée pour la trame précédente. Lorsque les sous-trames correspondent à un découpage régulier de la trame, la fonction de mélange pondéré peut notamment être :

$$\hat{h}_w^{(i)}(k, n) = \left(\frac{N-i}{N}\right)\hat{h}_w(k-1, n) + \left(\frac{i}{N}\right)\hat{h}_w(k, n) \quad (17)$$

On remarquera que le cas où l'on applique directement le filtre  $\hat{h}_w(k, n)$  correspond à  $N = 1$  (pas de sous-trames).

### Exemple 1

Cet exemple de dispositif est adapté à une application à la communication parlée, notamment en pré-traitement d'un codeur de parole à bas débit.

On utilise des fenêtres non-recouvrantes pour réduire au maximum théorique le retard introduit par le traitement tout en offrant la possibilité à l'utilisateur de choisir un fenêtrage adapté à l'application. Ceci est possible puisque le fenêtrage du signal d'entrée du dispositif n'est pas soumis à une contrainte de reconstruction parfaite.

Dans une telle application, la fonction de fenêtrage  $w(n)$  appliquée par le multiplieur 2 est avantageusement dissymétrique afin de réaliser une

pondération plus forte sur la moitié la plus récente de la trame que sur la moitié la moins récente.

Comme illustré par la figure 4, la fenêtre d'analyse dissymétrique  $w(n)$  peut être construite à l'aide de deux demi-fenêtres de Hanning de tailles différentes  $L_1$  et  $L_2$  :

$$w(n) = \begin{cases} 0,5 - 0,5 \cdot \cos\left(\frac{\pi n}{L_1}\right) & \text{pour } 0 \leq n < L_1 \\ 0,5 + 0,5 \cdot \cos\left(\frac{\pi(n-L_1+1)}{L_2}\right) & \text{pour } L_1 \leq n < L_1 + L_2 = L \end{cases} \quad (18)$$

Beaucoup de codeurs de parole pour les mobiles utilisent des trames de longueur 20 ms et fonctionnent à la fréquence d'échantillonnage  $F_e = 8$  kHz (soit 160 échantillons par trame). Dans l'exemple représenté sur la figure 4, on a choisi  $L = 160$ ,  $L_1 = 120$  et  $L_2 = 40$ .

Le choix d'une telle fenêtre permet de concentrer le poids de l'estimation spectrale vers les échantillons les plus récents, tout en garantissant une bonne fenêtre spectrale. La méthode proposée permet un tel choix du fait que l'on n'a aucune contrainte de reconstruction parfaite du signal à la synthèse (signal reconstruit en sortie par filtrage temporel).

Pour avoir une meilleure résolution fréquentielle, les unités 3 et 7 utilisent une FFT de longueur  $L_{\text{FFT}} = 256$ . Ce choix est également motivé car la FFT est numériquement optimale lorsqu'elle s'applique sur des trames de longueur en puissance de 2. Il est donc nécessaire au préalable de prolonger le bloc fenêtré  $x_w(k,n)$  par  $L_{\text{FFT}} - L = 96$  échantillons nuls (« zero-padding ») :

$$x_w(k,n) = 0 \quad \text{pour } L \leq n < L_{\text{FFT}} \quad (19)$$

La détection d'activité vocale utilisée dans cet exemple est une méthode classique basée sur des comparaisons d'énergies à court terme et à long terme du signal. La mise à jour de l'estimation de la densité spectrale de puissance de bruit  $\hat{\gamma}_{bb}(k,f)$  est réalisée par estimation à lissage exponentiel, conformément à l'expression (10) avec  $\alpha(k_b) = 0,8553$ , correspondant à une constante de temps de 128 ms, jugée suffisante pour assurer un compromis entre une estimation fiable et un suivi des variations temporelles de la statistique du bruit.

La FT du filtre de réduction de bruit  $\hat{H}_f(k,f)$  est pré-estimée conformément à la formule (5) (filtre de Wiener en boucle ouverte), après avoir pré-estimé la DSP du signal utile selon l'estimateur à décision-dirigée défini en (12) avec  $\beta(k) = 0,98$ . La même fonction  $F$  est réutilisée par le module 16 pour  
5 produire l'estimation finale  $\hat{H}(k,f)$  de la FT.

Du fait que la FT  $\hat{H}(k,f)$  est à valeurs réelles, le filtre temporel est rendu causal par :

$$\begin{cases} \hat{h}_{\text{caus}}(k, n) = \hat{h}(k, n+L/2) & \text{pour } 0 \leq n < L/2 \\ \hat{h}_{\text{caus}}(k, n) = \hat{h}(k, n-L/2) & \text{pour } L/2 \leq n < L \end{cases} \quad (20)$$

On sélectionne ensuite les  $L_{\text{filt}} = 21$  coefficients de ce filtre, qu'on  
10 pondère par une fenêtre de Hanning  $w_{\text{filt}}(n)$  de longueur  $L_{\text{filt}}$ , valeur correspondant aux échantillons significatifs pour cette application :

$$\hat{h}_w(k, n) = w_{\text{filt}}(n) \cdot \hat{h}_{\text{caus}}(k, n + \frac{L}{2} - \frac{L_{\text{filt}} - 1}{2}) \quad \text{pour } 0 \leq n < L_{\text{filt}} \quad (21)$$

avec : 
$$w_{\text{filt}}(n) = 0,5 - 0,5 \cdot \cos\left(\frac{2\pi n}{L_{\text{filt}} - 1}\right) \quad \text{pour } 0 \leq n < L_{\text{filt}} \quad (22)$$

Le filtrage temporel est réalisé par  $N = 4$  filtres de sous-trames  $\hat{h}_w^{(i)}(k, n)$   
15 obtenus par les fonctions de mélange pondéré données par (17). Ces quatre filtres sont ensuite appliqués à l'aide d'un filtrage transverse de longueur  $L_{\text{filt}} = 21$  aux quatre sous-trames du signal d'entrée  $x^{(i)}(k, n)$ , celles-ci étant obtenues par une extraction contiguë de quatre sous-trames de taille  $L/4 = 40$  échantillons du signal d'observation  $x(k, n)$  :

20 
$$x^{(i)}(k, n) = x(k, n) \quad \text{pour } (i-1) \cdot L/N \leq n < i \cdot L/N \quad (22)$$

### Exemple 2

Cet exemple de dispositif est adapté à une application à la reconnaissance de parole robuste (dans un environnement bruité).

Dans cet exemple, on utilise des trames d'analyse de longueur  $L$  qui  
25 présentent des recouvrements mutuels de  $L/2$  échantillons entre deux trames successives, et la fenêtre utilisée est de type Hanning :

- 20 -

$$w(n) = 0,5 - 0,5 \cdot \cos\left(\frac{2\pi n}{L-1}\right) \quad \text{pour } 0 \leq n < L \quad (23)$$

La longueur des trames est fixée à 20 ms, soit  $L = 160$  à la fréquence d'échantillonnage  $F_e = 8$  kHz, et les trames sont complétées par 96 échantillons nuls (« zero-padding ») pour la FFT.

5 Dans cet exemple, le calcul de la FT du filtre réducteur de bruit est basé sur un rapport de racines carrées des densités spectrales de puissance du bruit  $\hat{\gamma}_{bb}(k,f)$  et du signal utile  $\hat{\gamma}_{ss}(k,f)$ , et par conséquent sur les modules de l'estimée du bruit  $|\hat{B}(k,f)| = \sqrt{\hat{\gamma}_{bb}(k,f)}$  et du signal utile  $|\hat{S}(k,f)| = \sqrt{\hat{\gamma}_{ss}(k,f)}$ .

10 La détection d'activité vocale utilisée dans cet exemple est une méthode classique existante basée sur des comparaisons d'énergies à court terme et à long terme du signal. La mise à jour de l'estimation du module du signal de bruit  $|\hat{B}(k,f)| = \sqrt{\hat{\gamma}_{bb}(k,f)}$  est réalisée par estimation à lissage exponentiel :

$$\begin{cases} |\hat{B}(k_b, f)| = \alpha |\hat{B}(k_b - 1, f)| + (1 - \alpha) |X(k_b, f)| \\ |\hat{B}(k, f)| = |\hat{B}(k_b, f)| \end{cases} \quad (24)$$

15 où  $k_b$  est la trame de bruit courante ou la dernière trame de bruit (si  $k$  est détectée comme trame de signal utile). La grandeur de lissage  $\alpha$  est choisie constante et égale à 0,99, soit une constante de temps de 1,6 s.

La FT du filtre de réduction de bruit  $\hat{H}_1(k,f)$  est pré-estimée par le module 13 selon :

$$20 \quad \hat{H}_1(k,f) = F\left(|\hat{S}(k,f)|, |\hat{B}(k,f)|\right) \quad (25)$$

$$\text{avec :} \quad F(y, z) = \frac{y}{y + z} \quad (26)$$

Le fait de calculer une racine carrée permet d'effectuer des estimations sur les modules, qui sont liés au RSB  $\eta(k,f)$  par :

$$\eta(k,f) = \frac{|\hat{S}(k,f)|^2}{|\hat{B}(k,f)|^2} \quad (27)$$

25 L'estimateur du signal utile en module  $|\hat{S}(k,f)|$  est obtenu par :

- 21 -

$$|\hat{S}(k,f)| = \beta |\hat{S}(k-1,f)|^2 + (1 - \beta) P[|X(k,f)| - |\hat{B}(k,f)|] \quad (28)$$

avec  $\beta(k) = 0,98$ .

Le multiplicateur 14 fait le produit de la FT pré-estimée  $\hat{H}_1(k,f)$  par le spectre  $X(k,f)$ , et le module du résultat (et non son carré) est obtenu en 15 pour  
 5 fournir l'estimation affinée de  $|\hat{S}(k,f)|$ , à partir de laquelle le module 16 produit l'estimation finale  $\hat{H}(k,f)$  de la FT en utilisant la même fonction  $F$  que dans (25).

La réponse temporelle  $\hat{h}_w(k,n)$  est ensuite obtenue exactement de la même manière que dans l'exemple 1 (passage en domaine temporel, restitution de la causalité, sélection des échantillons significatifs et fenêtrage).  
 10 La seule différence réside dans le choix du nombre de coefficients sélectionné  $L_{\text{filt}}$ , qui est fixé à  $L_{\text{filt}} = 17$  dans cet exemple.

Le filtrage de la trame d'entrée  $x(k,n)$  est réalisé en lui appliquant directement la réponse temporelle du filtre de réduction de bruit obtenue  $\hat{h}_w(k,n)$ . Ne pas faire de filtrage en sous-trames revient à prendre  $N = 1$  dans  
 15 l'expression (17).

## REVENDICATIONS

1. Procédé de réduction de bruit dans des trames successives d'un signal d'entrée ( $x(n)$ ), comprenant les étapes suivantes pour certaines au moins des trames :
  - 5 - calculer un spectre ( $X(k,f)$ ) du signal d'entrée par transformation dans le domaine fréquentiel ;
  - obtenir un estimateur de niveau de bruit en fonction de la fréquence ;
  - calculer un premier estimateur de niveau de signal utile dans la trame en fonction de la fréquence ;
  - 10 - calculer la fonction de transfert ( $\hat{H}_1(k,f)$ ) d'un premier filtre réducteur de bruit sur la base du premier estimateur de niveau de signal utile et de l'estimateur de niveau de bruit ;
  - calculer un second estimateur de niveau de signal utile dans la trame en fonction de la fréquence, en combinant le spectre du signal d'entrée et la
  - 15 fonction de transfert du premier filtre réducteur de bruit ;
  - calculer la fonction de transfert ( $\hat{H}(k,f)$ ) d'un second filtre réducteur de bruit sur la base du second estimateur de niveau de signal utile et de l'estimateur de niveau de bruit ; et
  - utiliser la fonction de transfert du second filtre réducteur de bruit dans
  - 20 une opération de filtrage de la trame pour produire un signal à bruit réduit.
  
2. Procédé selon la revendication 1, dans lequel le calcul du spectre ( $X(k,f)$ ) comporte une pondération de la trame de signal d'entrée par une fonction de fenêtrage ( $w(n)$ ) et une transformation de la trame pondérée vers le
- 25 domaine fréquentiel, la fonction de fenêtrage étant dissymétrique de manière à appliquer une pondération plus forte à la moitié la plus récente de la trame qu'à la moitié la moins récente de la trame.
  
3. Procédé selon la revendication 1 ou 2, dans lequel on détermine une réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_w(k,n)$ ) pour la trame
- 30 courante à partir d'une transformation vers le domaine temporel de la fonction



- 23 -

de transfert ( $\hat{H}(k,f)$ ) du second filtre réducteur de bruit, et on réalise l'opération de filtrage de la trame dans le domaine temporel au moyen de la réponse impulsionnelle déterminée pour ladite trame.

4. Procédé selon la revendication 3, dans lequel la détermination de la  
5 réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_w(k,n)$ ) pour la trame courante comporte les étapes suivantes :

- transformer vers le domaine temporel la fonction de transfert ( $\hat{H}(k,f)$ ) du second filtre réducteur de bruit pour obtenir une première réponse impulsionnelle ; et
- 10 - tronquer la première réponse impulsionnelle à une longueur de troncature correspondant à un nombre d'échantillons sensiblement plus petit que le nombre de points de la transformation vers le domaine temporel.

5. Procédé selon la revendication 4, dans lequel la détermination de la  
15 réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_w(k,n)$ ) pour la trame courante comporte en outre l'étape suivante :

- pondérer la réponse impulsionnelle tronquée par une fonction de fenêtrage ( $w_{\text{filt}}(n)$ ) sur un nombre d'échantillons correspondant à ladite longueur de troncature.

20 6. Procédé selon l'une quelconque des revendications 3 à 5, dans lequel on subdivise la trame courante en plusieurs sous-trames et on calcule pour chaque sous-trame une réponse impulsionnelle interpolée ( $\hat{h}_w^{(i)}(k,n)$ ) à partir de la réponse impulsionnelle de filtre réducteur de bruit déterminée pour la trame courante et de la réponse impulsionnelle de filtre réducteur de bruit  
25 déterminée pour au moins une trame précédente, et dans lequel l'opération de filtrage de la trame comporte un filtrage du signal de chaque sous-trame dans le domaine temporel conformément à la réponse impulsionnelle interpolée calculée pour ladite sous-trame.

- 24 -

7. Procédé selon la revendication 6, dans lequel les réponses impulsionnelles interpolées ( $\hat{h}_W^{(i)}(k, n)$ ) sont calculées pour les différentes sous-frames de la trame courante comme des sommes pondérées de la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k, n)$ ) déterminée pour la trame courante et de la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k-1, n)$ ) déterminée pour la trame précédente.
8. Procédé selon la revendication 7, dans lequel la réponse impulsionnelle interpolée ( $\hat{h}_W^{(i)}(k, n)$ ) calculée pour la i-ième sous-trame de la trame courante ( $1 \leq i \leq N$ ) est égale à  $(N-i)/N$  fois la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k-1, n)$ ) déterminée pour la trame précédente plus  $i/N$  fois la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k, n)$ ) déterminée pour la trame courante,  $N$  étant le nombre de sous-frames de la trame courante.
9. Procédé selon l'une quelconque des revendications précédentes, dans lequel le signal d'entrée ( $x(n)$ ) est un signal audio.
10. Dispositif de réduction de bruit dans un signal d'entrée ( $x(n)$ ), comprenant :
- des moyens (1-3) de calcul d'un spectre ( $X(k, f)$ ) d'une trame du signal d'entrée par transformation dans le domaine fréquentiel ;
  - des moyens (5) d'obtention d'un estimateur de niveau de bruit en fonction de la fréquence ;
  - des moyens (11) de calcul d'un premier estimateur de niveau de signal utile dans la trame en fonction de la fréquence ;
  - des moyens (13) de calcul de la fonction de transfert ( $\hat{H}_f(k, f)$ ) d'un premier filtre réducteur de bruit sur la base du premier estimateur de niveau de signal utile et de l'estimateur de niveau de bruit ;
  - des moyens (14-15) de calcul d'un second estimateur de niveau de signal utile dans la trame en fonction de la fréquence, en combinant le

- 25 -

spectre du signal d'entrée et la fonction de transfert du premier filtre réducteur de bruit ;

- des moyens (16) de calcul de la fonction de transfert ( $\hat{H}(k,f)$ ) d'un second filtre réducteur de bruit sur la base du second estimateur de niveau de signal utile et de l'estimateur de niveau de bruit ; et
- des moyens (7-9) de filtrage de la trame au moyen de la fonction de transfert du second filtre réducteur de bruit pour produire un signal à bruit réduit.

11. Dispositif selon la revendication 10, dans lequel les moyens de calcul de spectre comprennent des moyens (2) de pondération de la trame de signal d'entrée ( $x(n)$ ) par une fonction de fenêtrage ( $w(n)$ ) et des moyens (3) de transformation de la trame pondérée vers le domaine fréquentiel, la fonction de fenêtrage étant dissymétrique de manière à appliquer une pondération plus forte à la moitié la plus récente de la trame qu'à la moitié la moins récente de la trame.

12. Dispositif selon la revendication 10 ou 11, comprenant des moyens (7-8) de détermination d'une réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_w(k,n)$ ) pour la trame courante à partir d'une transformation vers le domaine temporel de la fonction de transfert ( $\hat{H}(k,f)$ ) du second filtre réducteur de bruit, dans lequel les moyens de filtrage (9) opèrent dans le domaine temporel au moyen de la réponse impulsionnelle déterminée pour la trame courante.

13. Dispositif selon la revendication 12, dans lequel les moyens de détermination de la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_w(k,n)$ ) comprennent des moyens (7) de transformation vers le domaine temporel de la fonction de transfert ( $\hat{H}(k,f)$ ) du second filtre réducteur de bruit, pour obtenir une première réponse impulsionnelle, et des moyens (8) de troncature de la première réponse impulsionnelle à une longueur de troncature correspondant à un nombre d'échantillons sensiblement plus petit que le nombre de points de la transformation vers le domaine temporel.

- 26 -

14. Dispositif selon la revendication 13, dans lequel les moyens de détermination de la réponse impulsionnelle de filtre réducteur de bruit comprennent des moyens (8) de pondération de la réponse impulsionnelle tronquée par une fonction de fenêtrage ( $w_{\text{filt}}(n)$ ) sur un nombre d'échantillons  
5 correspondant à ladite longueur de troncature.

15. Dispositif selon l'une quelconque des revendications 12 à 14, comprenant des moyens pour subdiviser la trame courante en plusieurs sous-trames et des moyens (21) de calcul d'une réponse impulsionnelle interpolée ( $\hat{h}_W^{(i)}(k, n)$ ) pour chaque sous-trame à partir de la réponse impulsionnelle de  
10 filtre réducteur de bruit ( $\hat{h}_W(k, n)$ ) déterminée pour la trame courante et de la réponse impulsionnelle de filtre réducteur de bruit déterminée pour au moins une trame précédente, dans lequel les moyens de filtrage (9) comprennent un filtre (23) pour filtrer le signal de chaque sous-trame dans le domaine temporel conformément à la réponse impulsionnelle interpolée calculée pour ladite sous-  
15 trame.

16. Dispositif selon la revendication 15, dans lequel les moyens de calcul de réponse impulsionnelle interpolée sont agencés pour calculer les réponses impulsionnelles interpolées ( $\hat{h}_W^{(i)}(k, n)$ ) pour les différentes sous-trames de la trame courante comme des sommes pondérées de la réponse  
20 impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k, n)$ ) déterminée pour la trame courante et de la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k-1, n)$ ) déterminée pour la trame précédente.

17. Dispositif selon la revendication 16, dans lequel la réponse impulsionnelle interpolée ( $\hat{h}_W^{(i)}(k, n)$ ) calculée pour la  $i$ -ième sous-trame de la  
25 trame courante ( $1 \leq i \leq N$ ) est égale à  $(N-i)/N$  fois la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k-1, n)$ ) déterminée pour la trame précédente plus  $i/N$  fois la réponse impulsionnelle de filtre réducteur de bruit ( $\hat{h}_W(k, n)$ ) déterminée pour la trame courante,  $N$  étant le nombre de sous-trames de la trame courante.

- 27 -

18. Dispositif selon l'une quelconque des revendications 10 à 17, dans lequel le signal d'entrée  $x(n)$  est un signal audio.

FIG. 1

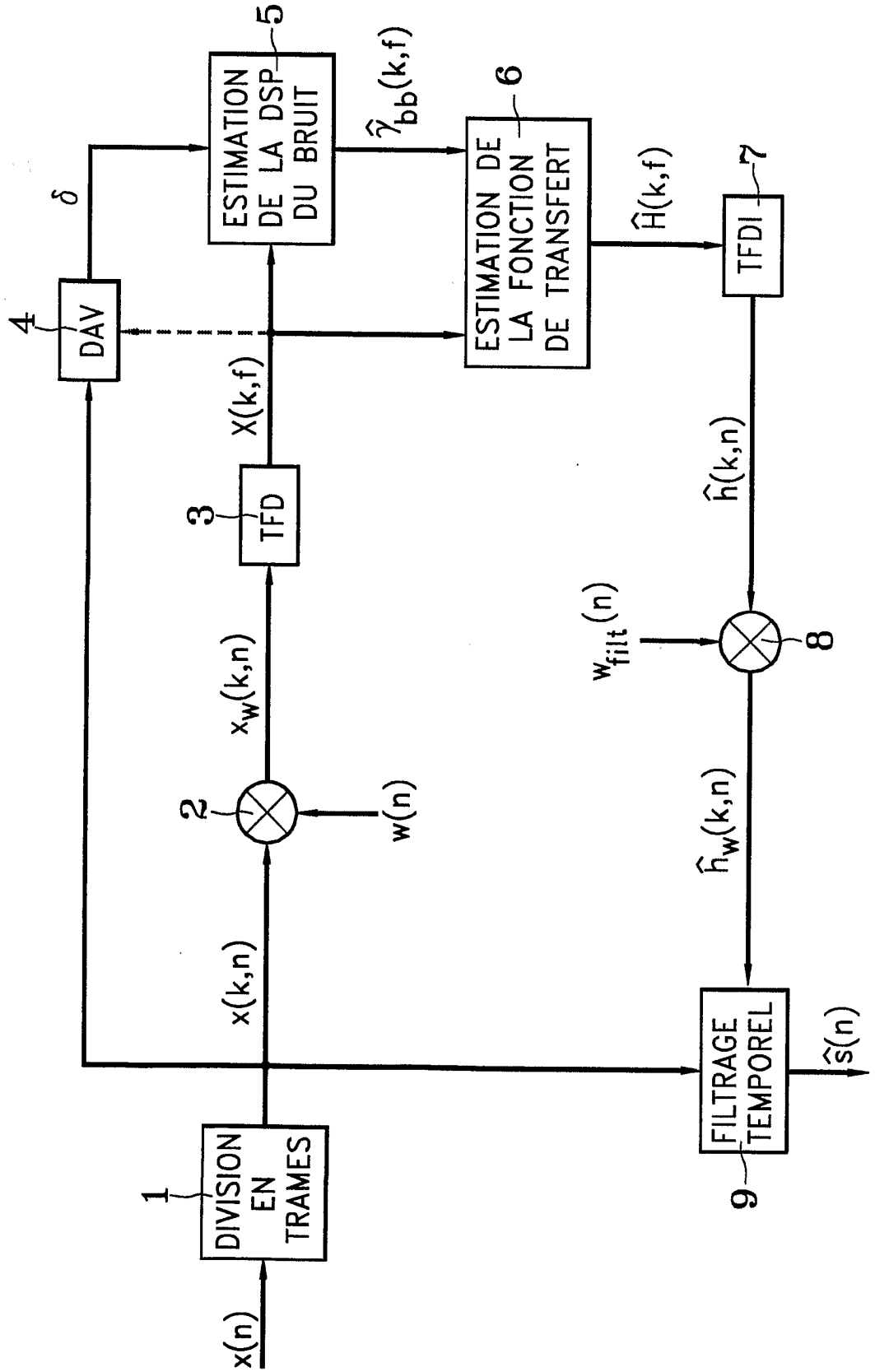


FIG. 2

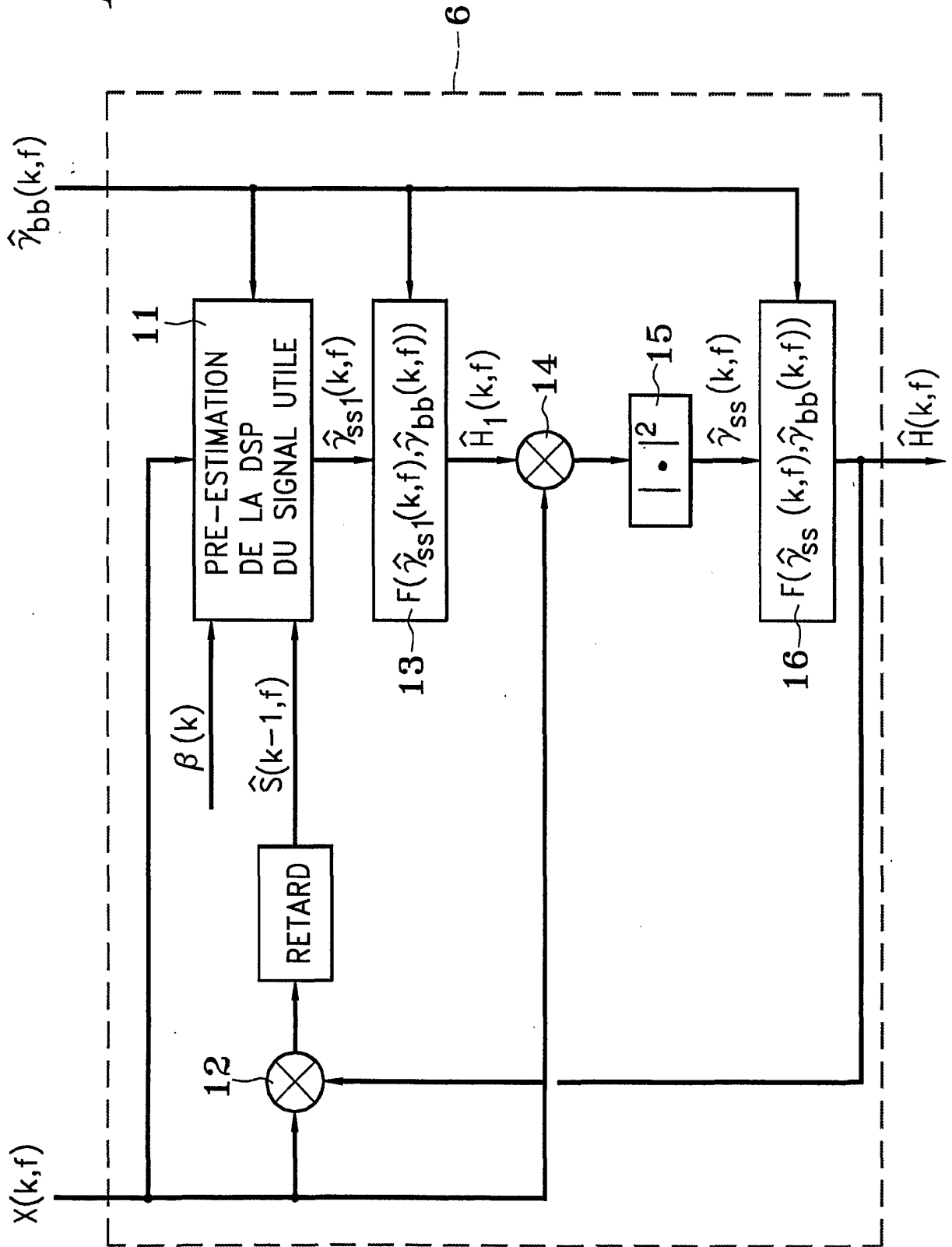


FIG. 3

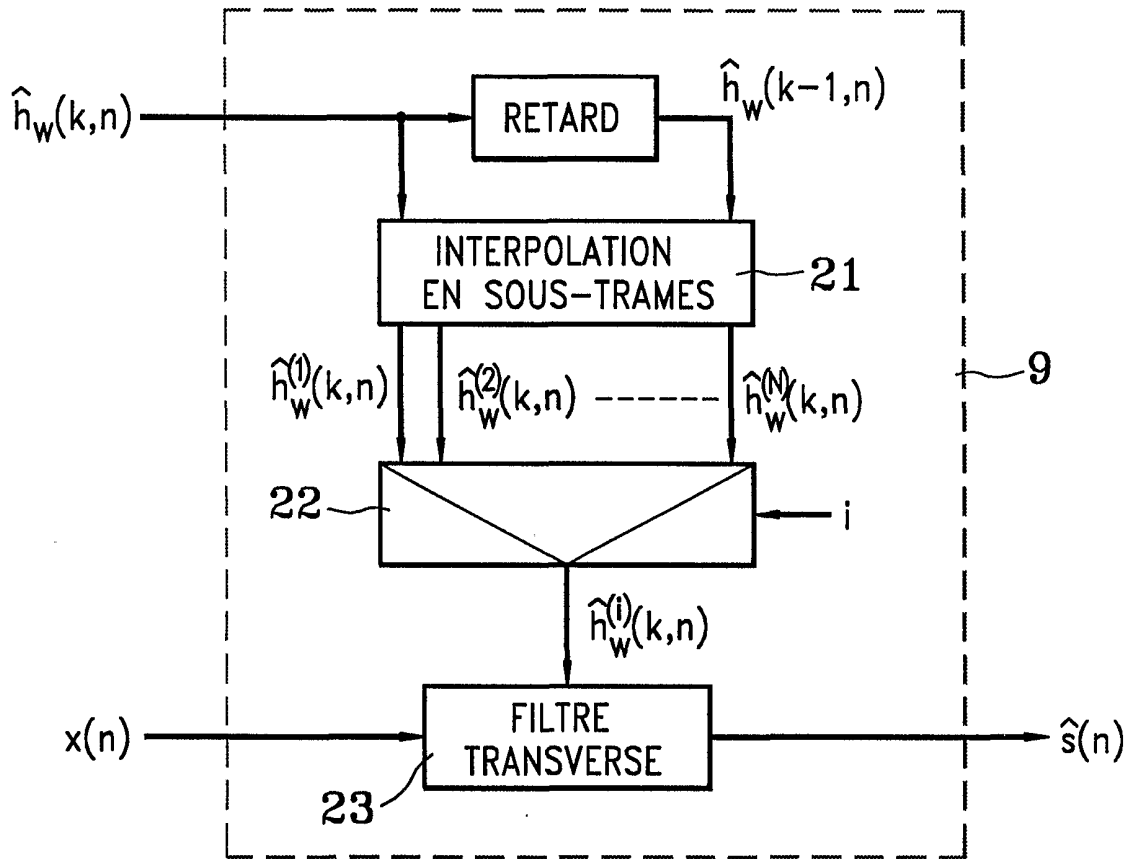
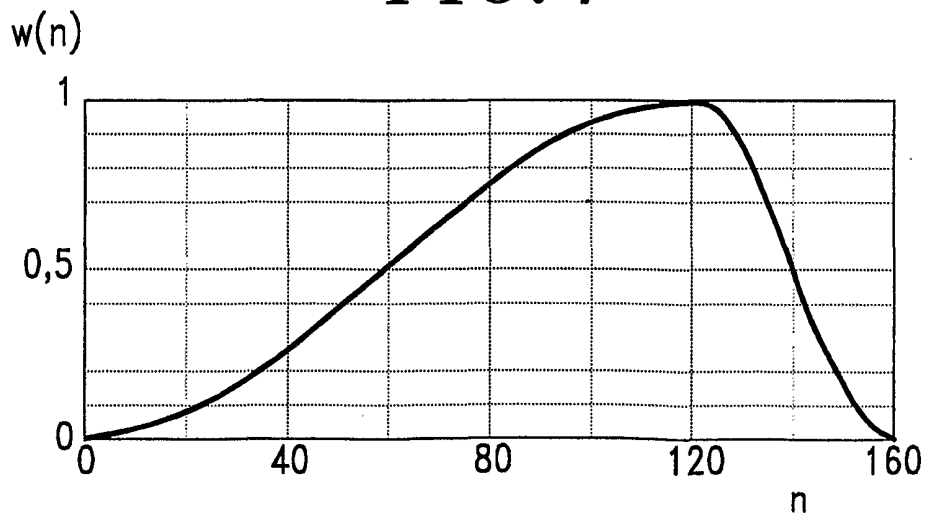


FIG. 4





# INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 01/03624

**A. CLASSIFICATION OF SUBJECT MATTER**  
 IPC 7 G10L21/02

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
 IPC 7 G10L

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

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

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 730 262 A (NIPPON ELECTRIC CO) 4 September 1996 (1996-09-04) abstract ---	1,10
A	EP 0 918 317 A (SEXTANT AVIONIQUE) 26 May 1999 (1999-05-26) abstract; figures 1,8 ---	1,10
A	SIM B L ET AL: "A PARAMETRIC FORMULATION OF THE GENERALIZED SPECTRAL SUBTRACTION METHOD" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE INC. NEW YORK, US, vol. 6, no. 4, 1 July 1998 (1998-07-01), pages 328-336, XP000785363 ISSN: 1063-6676 abstract paragraphs '002A!', '!', '002D! ---	1,10
	-/--	

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

25 February 2002

Date of mailing of the international search report

04/03/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Van Doremalen, J

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 01/03624

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 856 833 A (NIPPON ELECTRIC CO) 5 August 1998 (1998-08-05) abstract page 4, line 4 - line 5 -----	1,10

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 01/03624

Information on patent family members

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0730262	A	04-09-1996	JP 2760373 B2	28-05-1998
			JP 8241086 A	17-09-1996
			AU 693648 B2	02-07-1998
			AU 4588696 A	12-09-1996
			CA 2170811 A1	04-09-1996
			EP 0730262 A2	04-09-1996
			US 5644641 A	01-07-1997
-----				
EP 0918317	A	26-05-1999	FR 2771542 A1	28-05-1999
			EP 0918317 A1	26-05-1999
			JP 11265198 A	28-09-1999
-----				
EP 0856833	A	05-08-1998	JP 2874679 B2	24-03-1999
			JP 10215193 A	11-08-1998
			AU 736904 B2	02-08-2001
			AU 5278698 A	06-08-1998
			CA 2228097 A1	29-07-1998
			EP 0856833 A2	05-08-1998
			US 6266422 B1	24-07-2001
-----				

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No  
PCT/FR 01/03624

<b>A. CLASSEMENT DE L'OBJET DE LA DEMANDE</b> CIB 7 G10L21/02		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
<b>B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE</b>		
Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 7 G10L		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés) EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB		
<b>C. DOCUMENTS CONSIDERES COMME PERTINENTS</b>		
Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	EP 0 730 262 A (NIPPON ELECTRIC CO) 4 septembre 1996 (1996-09-04) abrégé	1,10
A	EP 0 918 317 A (SEXTANT AVIONIQUE) 26 mai 1999 (1999-05-26) abrégé; figures 1,8	1,10
-/--		
<input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents		
<input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe		
° Catégories spéciales de documents cités:		
*A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent *E* document antérieur, mais publié à la date de dépôt international ou après cette date *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée	*T* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention *X* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément *Y* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier *&* document qui fait partie de la même famille de brevets	
Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale	
25 février 2002	04/03/2002	
Nom et adresse postale de l'administration chargée de la recherche internationale	Fonctionnaire autorisé	
Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Van Doremalen, J	

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No  
PCT/FR 01/03624

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>SIM B L ET AL: "A PARAMETRIC FORMULATION OF THE GENERALIZED SPECTRAL SUBTRACTION METHOD"                      IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE INC. NEW YORK, US, vol. 6, no. 4,                      1 juillet 1998 (1998-07-01), pages 328-336, XP000785363                      ISSN: 1063-6676                      abrégé                      alinéas '002A!, '!, '002D!                      -----</p>	1,10
A	<p>EP 0 856 833 A (NIPPON ELECTRIC CO)                      5 août 1998 (1998-08-05)                      abrégé                      page 4, ligne 4 - ligne 5                      -----</p>	1,10

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande Internationale No

PCT/FR 01/03624

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)		Date de publication
EP 0730262	A	04-09-1996	JP	2760373 B2	28-05-1998
			JP	8241086 A	17-09-1996
			AU	693648 B2	02-07-1998
			AU	4588696 A	12-09-1996
			CA	2170811 A1	04-09-1996
			EP	0730262 A2	04-09-1996
			US	5644641 A	01-07-1997
EP 0918317	A	26-05-1999	FR	2771542 A1	28-05-1999
			EP	0918317 A1	26-05-1999
			JP	11265198 A	28-09-1999
EP 0856833	A	05-08-1998	JP	2874679 B2	24-03-1999
			JP	10215193 A	11-08-1998
			AU	736904 B2	02-08-2001
			AU	5278698 A	06-08-1998
			CA	2228097 A1	29-07-1998
			EP	0856833 A2	05-08-1998
			US	6266422 B1	24-07-2001