

19 RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

⑪ **N° de publication :**
(à n'utiliser que pour les
commandes de reproduction)

2 980 620

②¹ N° d'enregistrement national : **11 58521**

⑤ Int Cl⁸: **G 10 L 19/028** (2013.01)

(12) **DEMANDE DE BREVET D'INVENTION**

A1

②② **Date de dépôt : 23.09.11.**

③④ **Priorité :**

71 Demandeur(s) : FRANCE TELECOM Société anonyme — FR.

43 **Date de mise à la disposition du public de la demande : 29.03.13 Bulletin 13/13.**

(56) Liste des documents cités dans le rapport de recherche préliminaire : Se reporter à la fin du présent fascicule

72 Inventeur(s) : O'LEARY SEAN, RAGOT STEPHANE
et DANIEL JEROME.

⑥ Références à d'autres documents nationaux apparentés :

⁽⁷³⁾ **Titulaire(s) :** FRANCE TELECOM Société anonyme.

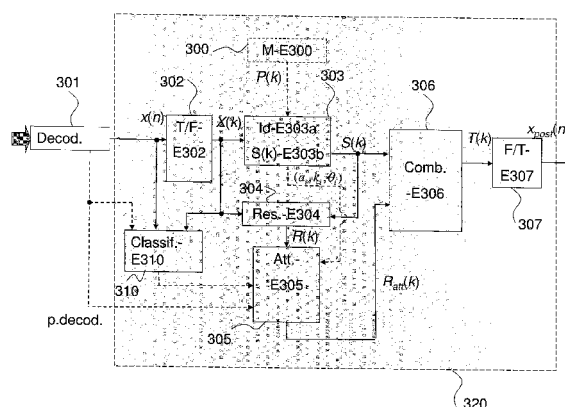
74 Mandataire(s) : FRANCE TELECOM Société anonyme.

54 TRAITEMENT D'AMELIORATION DE LA QUALITE DES SIGNAUX AUDIOFREQUENCES DECODEES.

(57) L'invention se rapporte à un procédé de traitement d'un signal audiofréquence décodé. Ce procédé est tel qu'il comporte les étapes suivantes:

- identification (E303a) de composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel;
- calcul (E303b) d'un spectre des composantes tonales identifiées à partir d'un modèle sinusoïdal;
- obtention (E304,) d'un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales; atténuation spectrale (E305,) du résidu; et
- combinaison (E306,) du résidu atténué et du signal issu du spectre des composantes tonales.

Le procédé de traitement améliore la qualité des signaux audiofréquences traités. L'invention vise également un dispositif de traitement mettant en oeuvre le procédé ainsi décrit et un décodeur comportant ce dispositif.



FR 2 980 620 - A1



Traitement d'amélioration de la qualité des signaux audiofréquences décodés

La présente invention se rapporte au domaine du traitement des signaux audiofréquences et notamment des signaux de parole ou de musique qui ont été codés et
5 décodés par codeurs et décodeurs de parole.

Le dispositif de traitement audio selon l'invention est adapté notamment pour la transmission et/ou le stockage des signaux audiofréquences. L'invention concerne plus particulièrement le post-traitement des signaux décodés pour améliorer la qualité des signaux musicaux décodés.

10 Différentes techniques existent pour convertir sous forme numérique et compresser un signal audiofréquence (parole, musique, etc.). Les techniques les plus courantes dans les services de télécommunications sont les méthodes de codage de forme d'onde, telles que le codage MIC (pour "Modulation par Impulsions Codées") et MICDA (pour "Modulation par Impulsion et Codage Différentiel adaptatif") dits aussi "PCM" ou "ADPCM" en anglais, les
15 méthodes de codage paramétrique par analyse par synthèse comme le codage CELP (pour "Code Excited Linear Prediction" en anglais), et les méthodes de codage perceptuel en sous-bandes ou par transformée.

Ces techniques traitent le signal d'entrée de façon séquentielle échantillon par échantillon (MIC ou MICDA) ou par blocs d'échantillons dits trames (CELP, codage par
20 transformée).

On s'intéresse ici plus particulièrement au cas du codage CELP comme exemple de codeur-décodeur de parole, cependant l'invention s'applique au cas général des codeurs de parole (ex: MIC, MICDA, CELP, etc...) .

25 Le codage CELP - dont sa variante appelée ACELP (pour "Algebraic CELP") utilisée par exemple dans les normes 3GPP AMR et AMR-WB - est un codage prédictif fondé sur le modèle source-filtre de production de la parole. Le filtre correspond en général à un filtre tout-pôle de fonction de transfert $1/A(z)$ obtenu par prédiction linéaire LPC (pour "Linear Predictive Coding" en anglais). Le signal est synthétisé à l'aide de la version quantifiée,
30 $1/\hat{A}(z)$, du filtre $1/A(z)$. La source - c'est-à-dire l'excitation du filtre linéaire prédictif $1/\hat{A}(z)$ - est en général représentée comme la combinaison d'une excitation adaptative (obtenue par prédiction long-terme modélisant la vibration des cordes vocales), et d'une excitation fixe (ou innovation) codée efficacement sous la forme de dictionnaires d'impulsions (ACELP), de dictionnaires de bruit, etc. La recherche de l'excitation "optimale" est réalisée par
35 minimisation d'un critère d'erreur quadratique dans le domaine du signal pondéré par un filtre de fonction de transfert $W(z)$ dérivée du filtre de prédiction linéaire $A(z)$, de la forme $W(z) = A(z/\gamma_1)/A(z/\gamma_2)$.

On peut noter que dans le codec 3GPP AMR-WB, qui est décrit dans l'article de B. Bessette et *al.*, intitulé "The Adaptive Multirate Wideband Speech Coder (AMR-WB)", IEEE Transactions on Speech, Audio and Language Processing, Nov. 2002, le codage ACELP est en fait appliqué par trames de 20 ms non pas au signal direct échantillonné à 16 kHz mais à un
 5 signal pré-accentué échantillonné à 12,8 kHz dans une bande audio réduite (0-6400 Hz); la pré-accentuation (ou pré-emphase) est réalisée par le filtre de fonction de transfert $1 - \alpha z^{-1}$ avec $\alpha = 0,68$. Le filtre de pondération perceptuelle utilisé pour la mise en forme de bruit de codage est de la forme $W(z) = A(z/\gamma)/(1 - \alpha z^{-1})$ avec $\gamma = 0,92$.

Le codage CELP repose sur une approche de codage temporelle et prédictive à
 10 partir d'un modèle de signal (LPC); ce type de codage est très efficace sur la parole (propre), mais il donne à bas débit une qualité souvent médiocre pour les cas des signaux s'éloignant des hypothèses du modèle de production de la parole. Ainsi, pour les signaux musicaux ayant une structure composée de composantes tonales qui sont mal codées par un modèle CELP (ex: signaux harmoniques, multipitch, ou inharmoniques, tels que les sons d'orgue, de piano,
 15 etc.), il est bien connu qu'un codage par transformée est beaucoup plus adapté.

Néanmoins, les codeurs de parole de type CELP sont historiquement déployés dans les applications de téléphonie fixe et mobile car ils fournissent une meilleure qualité à bas débit pour les signaux de parole qui sont les signaux les plus importants pour les applications de téléphonie. Cependant, les signaux musicaux ou les contenus mixtes (mélange de parole et musique) représentent malgré tout une classe de signaux importante dans certains cas
 20 d'usage tels que la musique d'attente, la tonalité de retour, etc. Il est donc pertinent et important de chercher à améliorer pour la musique et les contenus mixtes la qualité des codeurs de parole qui sont déjà déployés dans les services.

La **figure 1** montre l'exemple concret d'un signal musical (un son d'orgue)
 25 échantillonné à 16 kHz, préfiltré par un masque P.341 (50-7000 Hz) et analysé par FFT à court-terme sur un support de 512 échantillons (32 ms). Le spectre du signal ('sig') sur la trame considérée fait apparaître un certain nombre de composantes tonales (ct). Après codage du signal par le codeur AMR-WB à 12,65 kbit/s, le spectre de bruit ('err') correspondant apparaît comme relativement fort entre les composantes tonales. Cette
 30 caractéristique du bruit de codage s'explique par le fait que le codage CELP est temporel et met en forme le bruit en principe selon la réponse en fréquence du filtre $W(z)^{-1}$, contrairement à un codage par transformée opérant dans un domaine fréquentiel, le codage CELP ne peut pas "creuser" entre les harmoniques (ou composantes tonales).

35 Un exemple de technique améliorant le décodage CELP pour les signaux musicaux est présenté dans l'article de T. Vaillancourt et *al.* intitulé "Inter-tone noise reduction in a low bit rate CELP decoder", Proc. ICASSP 2009. Cette technique est utilisée dans le mode "bande

étroite" de la norme UIT-T G.718. Le principe de cette technique consiste à réaliser un post-traitement dans le domaine fréquentiel (par FFT court-terme) du signal CELP décodé et de réduire le bruit de codage CELP entre les partiels (composantes tonales) par une fonction d'atténuation spectrale et un ajustement de gain (ou d'énergie). Ce post-traitement s'apparente à une réduction de bruit classique par atténuation spectrale à court-terme. Il comporte les étapes suivantes illustrées à la **figure 2**:

- Une pré-emphase (bloc 201) et une transformation de Fourier discrète (FFT) à court-terme (bloc 202) sont appliquées par trames de 20 ms sur le signal décodé échantillonné à 16 kHz – la FFT a un support temporel de 30 ms
- Une classification du signal CELP décodé (bloc 203) afin d'évaluer le niveau de stationnarité dans la trame courante et ajuster la zone de fréquence du post-traitement et le niveau de réduction de bruit maximal.
- Une atténuation du bruit de codage (bloc 204): à chaque trame d'indice temporel t , le niveau de bruit de codage $N_{CB}^{(t)}(i)$ est estimé par sous-bandes critiques d'indice i (CB pour "Critical Bands" en anglais). Un rapport signal à bruit $SNR(f)$ a posteriori est ensuite estimé par raie fréquentielle, celui-ci étant défini comme le rapport entre l'énergie du signal décodé à la raie de fréquence f et l'énergie du bruit $N_{CB}^{(t)}(i)$ dans la bande critique incluant la raie de fréquence f . Une fonction d'atténuation spectrale $g_s(f)$ est finalement calculée raie par raie, en fonction du signal à bruit, et le gain $g_s(f)$ est ensuite lissé.
- La correction de l'atténuation (bloc 205) par un gain de correction $g_{corr}(f)$ pour compenser la perte d'énergie due à l'atténuation spectrale du bloc 204 ainsi que l'atténuation des hautes fréquences par le modèle CELP.
- La synthèse temporelle (bloc 206) par addition-recouvrement (bloc 207) et la dé-emphase (bloc 208).

La technique de Vaillancourt et *al.* est de complexité raisonnable et elle occasionne un retard additionnel (10 ms) compatible avec les applications conversationnelles. Cependant, elle présente plusieurs inconvénients:

- La découpe en "bandes critiques" divise le spectre en régions adjacentes arbitraires; l'estimation du bruit de codage par sous-bandes est simpliste et imprécise. Or, la qualité du traitement de réduction de bruit dépend fortement de la fiabilité de l'estimation du niveau de bruit et donc du rapport signal à bruit. Cette technique peut produire des artefacts tels que du bruit musical.
- La réduction de bruit entre les composantes tonales (bloc 204) est réalisée en appliquant un facteur d'atténuation (gain) à chaque raie du spectre d'amplitude

mélangeant signal et bruit de codage. L'atténuation spectrale peut ainsi affecter à la fois le signal, et plus particulièrement sa structure harmonique (dont les lobes associés à chaque composante tonale), et le bruit de codage, alors qu'il serait préférable de ne pas perturber le signal et d'atténuer uniquement le bruit de codage.

5

Il existe donc un besoin, pour le post-traitement des signaux musicaux, de préserver le signal et plus particulièrement sa structure harmonique court-terme tout en réduisant efficacement le bruit de codage. Cette technique doit s'appliquer dans le cas où il n'est pas nécessaire de transmettre (du codeur au décodeur) d'information supplémentaire pour le traitement – on considère en effet ici le cas d'un post-traitement en aveugle au décodeur.

La présente invention vient améliorer la situation.

Elle propose à cet effet, un procédé de traitement d'un signal audiofréquence décodé tel qu'il comporte les étapes suivantes:

- identification de composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel;
- calcul d'un spectre des composantes tonales identifiées à partir d'un modèle sinusoïdal;
- obtention d'un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales;
- atténuation spectrale du résidu; et
- combinaison du résidu atténué et du signal issu du spectre des composantes tonales.

Ainsi, le modèle sinusoïdal utilisé pour obtenir un spectre de composantes tonales est adapté à représenter la structure harmonique (ou structure tonale) court-terme des signaux musicaux. Le fait d'effectuer l'atténuation spectrale sur le résidu obtenu par la différence entre le signal décodé et celui issu du spectre des composantes tonales, permet de ne pas affecter la structure harmonique (ou structure tonale) du signal lors du traitement.

Les différents modes particuliers de réalisation mentionnés ci-après peuvent être ajoutés indépendamment ou en combinaison les uns avec les autres, aux étapes du procédé d'allocation défini ci-dessus.

Dans un premier mode de réalisation, les étapes d'obtention du résidu, d'atténuation et de combinaison sont effectués dans le domaine fréquentiel.

Ainsi, le résidu est obtenu par la différence entre le spectre du signal décodé et le spectre des composantes tonales.

Dans un deuxième mode de réalisation les étapes d'obtention du résidu, d'atténuation et de combinaison sont effectuées dans le domaine temporel.

Le résidu est ici obtenu par la différence entre le signal décodé et le signal temporel issu du spectre des composantes tonales. La transformée inverse temps-fréquence est alors ici effectuée après l'obtention du spectre des composantes tonales.

Selon un mode particulier de réalisation, l'identification des composantes tonales s'effectue par une méthode d'interpolation sur un nombre prédéterminé de raies fréquentielles.

Ainsi, l'estimation de certains paramètres (fréquence, amplitude) des composantes tonales s'effectue par une méthode d'interpolation à partir d'un nombre prédéterminé de raies fréquentielles.

L'interpolation permet d'identifier les composantes tonales avec une précision dépassant la résolution inter-raie de la transformée temps-fréquence.

Dans un mode de réalisation particulier, l'identification des composantes tonales, plus particulièrement l'estimation des paramètres (amplitude, phase), utilise un motif spectral pré-calculé et stocké en mémoire.

Ce motif spectral permet d'effectuer une analyse sinusoïdale en boucle fermée, c'est-à-dire de tenir compte de la synthèse par motif à l'analyse. Cette méthode est avantageuse pour assurer une correspondance totale entre analyse et synthèse dans le modèle sinusoïdal.

De façon avantageuse, le calcul du spectre des composantes tonales utilise un motif spectral pré-calculé et stocké en mémoire.

L'utilisation d'un motif spectral permet ici d'obtenir le spectre sinusoïdal associé à un pic fréquentiel avec une précision dépassant la résolution inter-raie de la transformée temps-fréquence.

Dans un mode possible de réalisation, l'atténuation spectrale s'effectue en fonction d'une donnée de classification du signal audiofréquence.

Le post-traitement étant destiné à améliorer la qualité sur les signaux musicaux, il est essentiel de l'appliquer sur des signaux classifiés comme étant de la musique.

Ainsi, dans un mode de réalisation possible, lorsque le signal n'est pas détecté comme un signal musical, le gain d'atténuation est différent. Il peut par exemple être égal à 1 pour un signal qui n'est pas détecté comme musical, annulant ainsi l'atténuation, afin de préserver les signaux tels que la parole ou les transitoires.

Dans le cas d'une atténuation dans le domaine temporel, l'atténuation spectrale s'effectue par un filtrage du résidu dans le domaine temporel.

La présente invention vise également un dispositif de traitement d'un signal audiofréquence décodé, tel qu'il comporte:

- un module d'identification de composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel;
- un module de calcul d'un spectre des composantes tonales identifiées à partir d'un modèle sinusoïdal;

- un module d'obtention d'un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales;
- un module d'atténuation spectrale du résidu; et
- un module de combinaison du résidu atténué et du signal issu du spectre des composantes tonales.

Ce dispositif présente les mêmes avantages que le procédé décrit précédemment, qu'il met en œuvre.

L'invention vise aussi un décodeur de signal audiofréquence comportant un dispositif de traitement tel que décrit ci-dessus.

Elle vise un programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de traitement tel que décrit, lorsque ces instructions sont exécutées par un processeur.

Enfin l'invention se rapporte à un support de stockage, lisible par un processeur, intégré ou non au dispositif de traitement, éventuellement amovible, mémorisant un programme informatique mettant en œuvre un procédé de traitement tel que décrit précédemment.

D'autres caractéristiques et avantages de l'invention apparaîtront plus clairement à la lecture de la description suivante, donnée uniquement à titre d'exemple non limitatif, et faite en référence aux dessins annexés, sur lesquels :

- la figure 1 décrite précédemment, illustre un exemple de spectre court-terme d'un signal musical et du bruit de codage CELP associé;
- la figure 2 représente un organigramme de la méthode de réduction de bruit de l'état de l'art, telle que décrite précédemment;
- la figure 3 illustre sous forme de bloc diagramme, à la fois le dispositif de traitement et le procédé de traitement qu'il met en œuvre selon un premier mode de réalisation de l'invention;
- les figures 4a et 4b montrent des exemples d'analyse fréquentielle court-terme d'un signal sinusoïdal utilisés dans un mode de réalisation de l'invention;
- la figure 5 illustre le principe d'interpolation parabolique utilisé dans un mode de réalisation particulier de l'invention;
- les figures 6a, 6b et 6c illustrent des exemples de fenêtres temporelles utilisées pour la mise en œuvre des étapes de transformée temps-fréquence ou de transformée inverse, dans un mode de réalisation de l'invention;
- les figures 7a et 7b illustrent un exemple de spectre de composantes tonales dans un mode de réalisation de l'invention et d'un motif fréquentiel calculé et mémorisé dans un mode de réalisation de l'invention;

- la figure 8 illustre sous forme d'organigramme les étapes mises en œuvre pour le calcul d'un spectre de composantes tonales selon un mode de réalisation de l'invention;
- la figure 9 illustre sous forme d'organigramme le détail d'une étape de calcul du spectre de composantes tonales selon un mode de réalisation de l'invention;
- les figures 10a et 10b illustrent des exemples de facteurs d'atténuation utilisés dans le procédé de traitement selon l'invention;
- la figure 11 illustre les étapes de fenêtrages et de synchronisation du signal d'entrée et du signal de sortie du dispositif de traitement, dans un mode de réalisation de l'invention;
- la figure 12 illustre sous forme de bloc diagramme, à la fois le dispositif de traitement et le procédé de traitement qu'il met en œuvre selon un deuxième mode de réalisation de l'invention; et
- la figure 13 illustre un exemple de représentation matérielle d'un dispositif de traitement selon l'invention.

La **figure 3** selon un premier mode de réalisation de l'invention, représente un dispositif de traitement 320 d'amélioration de la qualité d'un signal audiofréquence décodé. Ce dispositif de traitement peut être intégré à un décodeur, comprenant alors le module 301 de décodage, ou peut être indépendant du décodeur et recevoir ainsi en entrée un signal audiofréquence décodé ($x(n)$) – l'intérêt d'intégrer le module de traitement 320 au décodeur est qu'il peut bénéficier de paramètres internes au décodeur, comme par exemple les paramètres décodés.

La figure 3 sert également à illustrer les principales étapes mises en œuvre par le procédé de traitement selon ce premier mode de réalisation de l'invention.

Un signal audiofréquence, par exemple, en provenance d'un codeur de parole de type normalisé 3GPP AMR-WB est donc d'abord décodé par un module de décodage 301. Dans ce mode de réalisation particulier, sans perte de généralité, on considère ici que les signaux audiofréquences sont échantillonnés à $F=16$ kHz et divisés en trames de $L=320$ échantillons (20 ms).

Dans une variante de réalisation, le traitement peut s'appliquer également au signal décodé filtré par un filtre perceptuel $W(z)$ ou par un filtre de prédiction linéaire. Dans ce cas, l'homme de l'art comprendra que le filtre inverse devra être utilisé après le post-traitement 320.

Dans l'exemple de réalisation illustré, ce bloc de décodage 301 correspond au décodeur AMR-WB tel que décrit dans l'article de B. Bessette et *al.*, intitulé "The Adaptive Multirate Wideband Speech Coder (AMR-WB)", IEEE Transactions on Speech, Audio and Language Processing, Nov. 2002.

Il réalise un décodage ACELP à la fréquence d'échantillonnage de 12,8 kHz, une extension de bande pour générer les fréquences de 6,4-7 kHz et une combinaison de la bande basse (0-6400 Hz) et de la bande haute (6400-7000 Hz).

Le bloc 301 fonctionne par trame de 20 ms ($L=320$ échantillons à 16 kHz) et synthétise un signal noté $x(n)$, dont les échantillons sont par convention numérotés par les indices $n=0, \dots, L-1$, avec $L=320$, dans la trame courante – suivant cette convention la trame décodée précédente correspond donc aux indices $n=-N, \dots, -1$.

On rappelle ici que le décodeur AMR-WB traite un train binaire toutes les 20 ms, et que ce train binaire contient la représentation binaire des paramètres décodés suivants:

- 10 ○ Coefficients LPC codés sous forme d'ISF (pour "Immittance Spectral Frequencies" en anglais) pour un ordre LPC de 16 - pour chaque trame de 20 ms
- Retard de pitch - pour chaque sous-trame de 5 ms
- Indicateur binaire (dit "pitch flag" en anglais) indiquant si la contribution de pitch est filtrée passe-bas ou non - pour chaque sous-trame de 5 ms
- 15 ○ Indices de dictionnaire fixe pour chaque sous-trame - pour chaque sous-trame de 5 ms
- Gains de dictionnaires adaptatif et fixe - pour chaque sous-trame de 5 ms

On ne liste pas ici les paramètres décodés uniquement au débit de 23,85 kbit/s concernant la bande haute (6400-7000 Hz) car le dispositif de post-traitement est plutôt dédié à l'amélioration du signal décodé par le modèle CELP dans la bande basse (0-6400 Hz).

Dans une variante du mode de réalisation, le post-traitement 320 est appliqué au signal décodé en bande basse par le décodeur ACELP d'AMR-WB. Dans ce cas, le post-traitement 320 fonctionne avec des trames de 20 ms (256 échantillons) à la fréquence d'échantillonnage de 12,8 kHz. La bande haute (6,4-7 kHz) du décodeur AMR-WB est alors décalée temporellement d'un nombre prédéterminé d'échantillons, afin d'aligner la bande haute synthétisée (6400-7000 Hz) avec la bande basse (0-6400 Hz) décodée et post-traitée.

Cette variante du mode de réalisation, où le signal décodé est échantillonné à 12,8 kHz d'échantillonnage, pourrait également comprendre le cas où le signal décodé est pré-filtré avant d'appliquer le post-traitement 320, le pré-filtrage étant réalisé par exemple par un filtre de pondération perceptuelle de la forme $F(z) = \hat{A}(z/\gamma)/(1-\alpha z^{-1})$, où $\hat{A}(z)$ est le filtre LPC décodé dans chaque sous-trame CELP. Ce pré-filtrage permettrait d'appliquer le post-traitement 320 dans le domaine du "signal pondéré" correspondant au domaine de l'analyse par synthèse du codage CELP. Cette variante implique d'appliquer un post-filtrage par $1/F(z)$ sur le résultat $x_{post}(n)$ du post-traitement.

Le bloc 302 réalise une analyse fréquentielle court-terme (E302) du signal $x(n)$, $n=-(N-L), \dots, L-1$, incluant les échantillons de la trame décodée courante ($n=0, \dots, L-1$) et une partie de la trame décodée précédente ($n=-(N-L), \dots, -1$).

L'analyse fréquentielle du module 302 est mise en œuvre, dans le mode de réalisation décrit, par fenêtrage et transformation de Fourier rapide (FFT):

- Une fenêtre d'analyse est appliquée au signal pour obtenir le signal fenêtré

$$x_w(m) = w_a(m).x(m - (N - L)), \quad m = 0, \dots, N - 1$$

5 où la fenêtre d'analyse $w_a(m)$ est définie comme une fenêtre de Hamming de longueur $N = 512$:

$$w_a(m) = 0.54 - 0.46 \cos\left(2\pi \frac{m}{N-1}\right), \quad m = 0, \dots, N - 1$$

Cette fenêtre est illustrée à la **figure 6a**.

10 Bien entendu, d'autres fenêtres d'analyse peuvent être prévues, par exemple des fenêtres asymétriques qui permettraient de mieux conditionner la fenêtre de synthèse utilisée par la suite en 307, mais qui possèdent des propriétés spectrales moins intéressantes à l'analyse.

- Pour des questions d'efficacité de calcul justifiées plus loin, on préconise en pratique d'intégrer à la transformée temps/fréquence (puis son inverse, pour la synthèse) un « recentrage de phase ». Celui-ci consiste à opérer une permutation circulaire du signal fenêtré de $N/2$ échantillons vers la gauche (pour l'analyse), soit la demi-longueur de la fenêtre, avant d'effectuer la Transformée de Fourier. Alternativement, cela revient à appliquer une correction du spectre de phase résultant de la Transformée de Fourier par ajout d'une phase linéaire $\pi k N / K$. Par l'un ou l'autre moyen, cela revient à changer la définition de la transformée en :

$$X(k) = \sum_{m=0}^{K-1} x_w(m).e^{-j\frac{2\pi k}{K}(m-\frac{N}{2})}, \quad k = 0, \dots, K - 1$$

Dans une variante du mode de réalisation, le signal fenêtré $x_w(m)$, $m = 0, \dots, N - 1$, est transformé par FFT de longueur $K = N = 512$ pour obtenir le spectre discret $X(k)$:

$$25 \quad X(k) = \sum_{m=0}^{K-1} x_w(m).e^{-j\frac{2\pi k}{K}m}, \quad k = 0, \dots, K - 1$$

Dans ce cas, l'homme de l'art comprendra facilement comment adapter le mode de réalisation principal pour tenir compte du déphasage impliqué par la permutation circulaire de $N/2$ échantillons.

30 Dans les deux modes de réalisation décrits pour la transformée temps/fréquence (avec ou sans recentrage de phase), la FFT est calculée suivant la méthode décrite dans l'article de H.V. Sorensen et al., intitulé "Real-valued fast Fourier transform algorithm," IEEE Trans. on Signal Processing, Vol.35, No.6, pp 849-863, 1987.

Dans une variante supplémentaire du mode de réalisation, d'autres transformée temps-fréquence pourront être utilisées, par exemple la "Modulated Complex Lapped Transform" (MCLT) combinant des transformées en cosinus et sinus discrètes modifiées mises en quadrature. Pour que le modèle sinusoïdal s'applique il est important d'utiliser une transformée complexe afin d'estimer l'amplitude et la phase des composantes tonales. Des techniques d'analyse temps-fréquence telles que les représentations temps-fréquences parcimonieuses (*sparse* en anglais) s'appliquent également pour mettre en œuvre le modèle sinusoïdal.

Dans le mode de réalisation privilégié, la longueur de FFT est choisie comme étant identique à celle de la fenêtre d'analyse, ce qui se justifie parfaitement dans le cas où $K=N=512$ est une puissance de 2. Cependant, le choix de $K=N$ n'est pas restrictif au sens de l'invention, et l'invention s'applique également au cas où la longueur K de FFT est rendue supérieure à celle N du signal fenêtré au moyen d'un bourrage de zéros et éventuellement décalage temporel du signal – cette alternative est en général intéressante dans le cas où la longueur N du signal fenêtré n'est pas une puissance de 2.

Le signal $x_w(m)$ étant réel, le spectre $X(k)$ est à symétrie hermitienne et seules les fréquences positives $k = 0, \dots, K/2$ sont nécessaires. Le spectre obtenu $X(k)$, $k = 0, \dots, K/2$ est mis en coordonnées polaires en séparant amplitude $|X(k)|$ et phase $\angle X(k)$. Comme on le verra plus loin, dans un certain mode de réalisation, il n'est pas nécessaire de calculer ce terme de phase, économisant une certaine quantité d'opération.

On décrit à présent une méthode standard d'analyse fréquentielle de signaux composés de sinusoïdes pures.

La sinusoïde discrète d'amplitude A_0 , de fréquence normalisée ω_0 et de phase θ_0 , $v(n) = A_0 \cos(\omega_0 n + \theta_0)$, a pour transformée de Fourier (à temps discret):

$$V(\omega) = \frac{A_0}{2} \left(\delta(\omega + \omega_0) e^{j\theta_0} + \delta(\omega - \omega_0) e^{-j\theta_0} \right), \omega \in [-\pi, \pi]$$

Pour réaliser une analyse fréquentielle court-terme, les signaux sont en général pondérés par une fenêtre $w(n)$ de longueur entière $N > 0$ telle que $w(n) \neq 0$ pour $n = 0, \dots, N-1$ et $w(n) = 0$ sinon. Dans ce cas la sinusoïde fenêtrée $v_w(n) = w(n)v(n)$ a pour transformée de Fourier:

$$V_w(\omega) = \frac{A_0}{2} \left(W(\omega + \omega_0) e^{j\theta_0} + W(\omega - \omega_0) e^{-j\theta_0} \right)$$

où il est important de noter que le spectre $W(\omega)$ est périodique avec une période de 2π .

Ainsi la sinusoïde fenêtrée a pour spectre le résultat de la convolution de la transformée périodisée de la fenêtre $W(\omega)$ par deux Diracs centrés sur $\pm\omega_0$. Ce résultat s'étend facilement au cas où le signal est une combinaison linéaire de I sinusoïdes. Par linéarité, si

$$5 \quad v(n) = \sum_{i=0}^{I-1} A_i \cos(\omega_i n + \theta_i),$$

alors

$$V_w(\omega) = \sum_{i=0}^{I-1} \frac{A_i}{2} (W(\omega + \omega_i) e^{j\theta_i} + W(\omega - \omega_i) e^{-j\theta_i}).$$

Lorsque le signal est analysé par transformée de Fourier discrète de longueur $K \geq N$, la transformée devient $V_w\left(\omega = \frac{2\pi k}{K}\right)$, $k = 0, \dots, K-1$; pour alléger les notations, la

10 transformée de Fourier discrète de $v_w(n)$ est notée $V_w(k)$ au lieu de $V_w\left(\frac{2\pi k}{K}\right)$. On obtient alors:

$$V_w(k) = \sum_{i=0}^{I-1} \frac{A_i}{2} (W(k + k_i) e^{j\theta_i} + W(k - k_i) e^{-j\theta_i}), \quad k = 0, \dots, K-1$$

avec $k_i = \frac{K}{2\pi} \omega_i$.

Le spectre $W_K(k)$ (également nommé $W(k)$, ici) de longueur $K \geq N$ de la fenêtre $w(m)$ est
15 défini par l'équation :

$$W_K(k) = \sum_{m=0}^{K-1} w_{zp}(m) e^{-j\frac{2\pi k}{K}m}, \quad k = 0, \dots, K-1;$$

la fenêtre $w_{zp}(m)$ étant obtenue par un bourrage de zéros ("zero padding" en anglais) autour de la fenêtre $w(m)$

$$w_{zp}(m) = \begin{cases} w(m) & m = 0, \dots, N-1 \\ 0 & m \notin [0, N[\end{cases}$$

20 Avantagemement et de façon cohérente avec la transformée du signal fenêtré $x_w(m)$, on préconise d'utiliser la version modifiée de la transformée temps-fréquence intégrant le « recentrage de phase » décrit plus haut, et qui s'écrit mathématiquement :

$$W_K(k) = \sum_{m=0}^{K-1} w_{zp}(m) e^{-j\frac{2\pi k}{K}(m - \frac{N}{2})}, \quad k = 0, \dots, K-1$$

L'intérêt est qu'ayant à faire à une fenêtre symétrique, le « recentrage de phase » lui
25 assure un spectre réel (c'est-à-dire que la partie imaginaire du spectre est nulle, ou encore que sa phase est nulle à π près), et cela quel que soit le sur-échantillonnage fréquentiel choisi (lequel dépend du bourrage de zéros).

Un exemple d'analyse fréquentielle à court-terme est montré aux **figures 4a et 4b** pour les conditions suivantes :

- $I=2$ sinusoïdes avec

$$\begin{cases} A_0 = 1 \\ \omega_0 = 2\pi \frac{4}{32} \\ \theta_0 = 0 \end{cases} \quad \text{et} \quad \begin{cases} A_1 = 0,5 \\ \omega_1 = 2\pi \frac{12,5}{32} \\ \theta_1 = 0 \end{cases}$$

- 5 ○ Fenêtrage de Hamming (normalisé) de longueur $N=32$:

$$w(n) = \frac{w_{ham}(n)}{\sum_{n=0}^{N-1} w_{ham}(n)}, \quad n = 0, \dots, N-1$$

avec

$$w_{ham}(n) = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N-1}\right), \quad n = 0, \dots, N-1$$

- 10 On peut noter que la normalisation par la moyenne de la fenêtre n'est ici qu'un exemple, la fenêtre pourrait ne pas être normalisée ou un autre facteur de normalisation pourrait être utilisé, ce qui ne fait que translater en dB les valeurs spectrales discutées ici.

- FFT de longueur $K=N=32$ (figure 4a) ou $K=8N=256$ (figure 4b)

15

Les spectres d'amplitude (en dB) obtenus par transformées de Fourier discrètes de longueurs respectives $K=32$ et $K=256$ sont donc représentés aux figures 4a et 4b; seule la partie correspondant aux fréquences positives est montrée.

On distingue bien 2 pics (de forme $\approx \left| \frac{A_i}{2} W_K(k+k_i) \right|$) autour de k_0 et k_1 . A la figure

- 20 4a la seconde sinusoïde a une position k_1 fractionnaire, alors que sur la figure 4b les deux sinusoïdes sont localisées sur des raies entières ("bins" en anglais); on vérifie par la même occasion que la transformée de longueur $K=256$ est une version interpolée de celle de longueur $K=32$ (signalée par les points 'o').

- 25 Le procédé de traitement selon l'invention permet de déterminer, à partir d'un spectre d'amplitude de résolution limitée (comme à la figure 4a), le nombre et les paramètres de sinusoïdes présentes à court-terme dans un signal audio et d'extraire spectralement ces composantes sinusoïdales par soustraction complexe du spectre discret de la forme

$$\frac{A_i}{2} \left(W_K(k+k_i) e^{j\theta_i} + W_K(k-k_i) e^{-j\theta_i} \right).$$

L'exemple des figures 4a et 4b montre bien que la précision d'une transformation de Fourier discrète de longueur K est en général insuffisante pour obtenir directement une estimation de la fréquence des composantes tonales.

Il est important de noter que dans le domaine de la transformée de Fourier à temps discret l'amplitude du "pic" associé à la i -ème sinusoïde est $\frac{A_i}{2} W(0)$ à la position $\pm k_i$. Dans le domaine de la transformée de Fourier rapide de longueur K , le pic à la position $\pm k_i$ n'est pas forcément directement observable à cause de la précision inter-raie; en supposant que la position correspond à une raie de Fourier, l'amplitude du "pic" associé à la sinusoïde devient:

$$a_i = \frac{A_i}{2} W_K(0)$$

Par la suite, la terminologie "amplitude" d'une sinusoïde fera essentiellement référence à l'amplitude dans le domaine de Fourier, a_i , et non l'amplitude de la sinusoïde dans le domaine temporel, A_i . On peut remarquer que la valeur de a_i intègre naturellement la longueur de la fenêtre K ; si la fenêtre d'analyse est normalisée de façon adéquate, il est possible de s'assurer que $W_K(0) = 1$.

15

Dans un mode particulier de réalisation, une méthode d'interpolation est utilisée pour améliorer la précision de cette estimation. Celle-ci est réalisée dans le module 303 de la figure 3, à une étape E603a d'identification des composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel ($X(k)$).

L'interpolation est utile pour estimer précisément les paramètres des sinusoïdes (fréquence, amplitude et phase) à partir d'un spectre d'amplitude issu d'une transformée de Fourier discrète de longueur N relativement limitée.

Le bourrage de zéros ("zero padding") n'améliore pas la résolution de la transformée de Fourier, mais il permet d'échantillonner la transformée de Fourier à temps discret sur plus de raies et par suite d'obtenir une meilleure précision. Dans un mode de réalisation privilégié, le bourrage de zéros n'est pas utilisé pour l'analyse du signal, mais seulement pour le pré-stockage de la transformée de la fenêtre d'analyse afin de faciliter la synthèse de sinusoïdes dont la fréquence est estimée avec une précision supérieure à la résolution inter-raie de la transformée de Fourier discrète d'analyse.

L'estimation des paramètres (a_i, k_i , qui sont l'amplitude maximale et la fréquence fractionnaire de la sinusoïde) est réalisée selon l'invention par interpolation parabolique du spectre d'amplitude, telle que décrite dans l'article de J.O. Smith et X. Serra, intitulé "PARSHL: A Program for the Analysis/Synthesis of Sounds Based on a Sinusoidal Representation", Proceedings of the International Computer Music Conference, Urbana-Champaign, Illinois, 1987.

Le principe de l'interpolation parabolique pour l'analyse spectrale est rappelé à la **figure 5**, où l'exemple décrit aux figures 4 est repris. L'interpolation parabolique repose sur le fait qu'une parabole est complètement déterminée par les valeurs en 3 points distincts; ainsi, étant données les amplitudes (en dB) $Y(l_i - 1)$, $Y(l_i)$, $Y(l_i + 1)$ de 3 raies de Fourier adjacentes d'indices $l_i - 1$, l_i , $l_i + 1$, il est facile de déterminer les paramètres α_i , d_i et β_i de la parabole d'équation générale :

$$Y(k) = \alpha_i (k - (l_i + d_i))^2 + \beta_i$$

où α_i est la courbure de la parabole, $l_i + d_i$ donne la position du sommet de la parabole, et β_i est la valeur de la parabole à son sommet (car $\beta_i = Y(l_i + d_i)$).

A partir de l'équation générale de la parabole et étant données les valeurs $Y(k - 1)$, $Y(k)$, $Y(k + 1)$, on obtient :

$$\begin{cases} Y(l_i - 1) = \alpha_i (1 + d_i)^2 + \beta_i \\ Y(l_i) = \alpha_i d_i^2 + \beta_i \\ Y(l_i + 1) = \alpha_i (1 - d_i)^2 + \beta_i \end{cases}$$

En calculant $Y(l_i + 1) + Y(l_i - 1) - 2Y(l_i)$, à partir des 3 équations précédentes on trouve :

$$\alpha_i = \frac{1}{2} (Y(l_i + 1) + Y(l_i - 1) - 2Y(l_i))$$

Ensuite, en calculant $Y(l_i - 1) - Y(l_i + 1)$, on trouve:

$$d_i = \frac{Y(l_i - 1) - Y(l_i + 1)}{4\alpha_i}$$

Enfin, le dernier paramètre se calcule comme étant :

$$Y(l_i) = \alpha_i d_i^2 + \beta_i \Rightarrow \beta_i = Y(l_i) - \alpha_i d_i^2$$

En combinant les expressions de d_i et β_i , on peut également écrire :

$$\beta_i = Y(l_i) - (\alpha_i d_i) d_i$$

ce qui donne :

$$\beta_i = Y(l_i) - \frac{Y(l_i - 1) - Y(l_i + 1)}{4} d_i$$

Pour l'exemple de la figure 5, l'interpolation parabolique sur les 2 "pics" centrés en $l_0 = 4$ et $l_1 = 12$ donne les paramètres estimés suivants:

$$\begin{cases} \alpha_0 = -7,1964 \\ d_0 = -0,0001 \\ \beta_0 = -5,9779 \end{cases} \quad \text{et} \quad \begin{cases} \alpha_1 = -8,2003 \\ d_1 = -0,4926 \\ \beta_1 = -11,6735 \end{cases}$$

Sur cet exemple, l'interpolation parabolique donne une estimation de la fréquence de la i -ème sinusoïde sous la forme $l_i + d_i$; cette estimation est précise dans cet exemple. De plus l'amplitude estimée en dB, $\beta_i = 10 \log_{10}(a_i)$ où $a_i = \frac{A_i}{2}$ (la fenêtre d'analyse étant normalisée telle que $W_K(0) = 1$, pour cet exemple), est proche des valeurs exactes de -6 dB et -12 dB.

Ainsi dans le mode de réalisation illustré à la figure 3, le spectre du signal $x(n)$ est analysé dans le bloc 303 afin d'extraire les composantes tonales qui sont à préserver par le post-traitement.

L'analyse comporte les étapes E603a d'identification des composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel et E603b de calcul d'un spectre des composantes tonales identifiées, à partir d'un modèle sinusoïdal tel que décrit précédemment.

A noter que l'identification des composantes sinusoïdales est réalisée en combinant la détection de sinusoïdes et l'estimation de leurs paramètres (fréquence, amplitude, phase); la détection sert à déterminer si un pic correspond à une sinusoïde "valide" (ou "sélectionnée"). L'ensemble des sinusoïdes ainsi détectées et estimées constitue le modèle sinusoïdal du signal dans la trame courante. Ce modèle sinusoïdal est représenté de façon équivalente par le spectre complexe combinant les différentes sinusoïdes identifiées. Ce modèle sert à l'extraction des composantes tonales du signal. L'extraction est ici réalisée en retranchant au spectre du signal courant le spectre issu du modèle sinusoïdal, pour obtenir un résidu qui sera post-traité.

L'étape E603a comporte une détection des pics fréquentiels dans le spectre d'amplitude du signal.

Un pic est détecté à une raie d'indice k dans le spectre d'amplitude $|X(k)|$ si le critère suivant est vérifié:

$$|X(k)| > |X(k-1)| \text{ et } |X(k)| \geq |X(k+1)|, \quad k = 1, \dots, K/2 - 1$$

Les indices k localisant les différents pics ainsi trouvés sont rassemblés dans un vecteur $(l_1, \dots, l_i, \dots, l_I)$, où I est le nombre de pics détectés. Dans une variante du mode de réalisation, on pourra s'assurer que I ne dépasse pas un nombre prédéterminé de pics, par exemple 100, afin de réduire la complexité des opérations qui dépendent du nombre de pics.

Suivant cette notation, l_i correspond au numéro de raie de Fourier localisant le pic d'indice i .

Afin d'extraire les composantes tonales par ordre d'importance, les pics détectés sont triés en amplitude, de sorte qu'ils vérifient:

$$|X(l_1)| \geq |X(l_2)| \geq \dots \geq |X(l_I)|$$

Les pics détectés de position l_i , $i = 1, \dots, I$ servent à l'identification et l'extraction de composantes tonales dans le spectre $X(k)$.

Un pic d'indice l_i est supposé être associé à une sinusoïde dont l'amplitude a_i et la fréquence k_i sont estimés par interpolation parabolique du spectre d'amplitude $|X(k)|$ à partir des 3 raies de Fourier contiguës d'indices $l_i - 1$, l_i et $l_i + 1$.

La courbure de la parabole est calculée comme décrit précédemment suivant l'équation:

$$\alpha_i = \frac{1}{2} (X_{dB}(l_i - 1) - 2X_{dB}(l_i) + X_{dB}(l_i + 1))$$

où $X_{dB}(k) = 20 \log_{10}(|X(k)|)$.

La fréquence de la i -ème sinusoïde correspond à la position du sommet de la parabole :

$$k_i = l_i + \frac{X_{dB}(l_i - 1) - X_{dB}(l_i + 1)}{4\alpha_i}$$

ce qui peut également être écrit:

$$k_i = l_i + \frac{1}{2} \frac{X_{dB}(l_i - 1) - X_{dB}(l_i + 1)}{X_{dB}(l_i - 1) - 2X_{dB}(l_i) + X_{dB}(l_i + 1)}$$

L'amplitude de la i -ème sinusoïde correspond à la valeur au sommet de la parabole :

$$a_i = 10^{\beta_i/20}$$

avec

$$\beta_i = X_{dB}(l_i) - \frac{X_{dB}(l_i - 1) - X_{dB}(l_i + 1)}{4} (k_i - l_i)$$

Une sinusoïde associée au pic d'indice i est sélectionnée, autrement dit considérée comme "valide", si elle vérifie les critères suivants:

$$i \leq 20$$

ou

$$0,75 < \frac{\alpha_{ref}}{\alpha_i} < 1,5 \text{ (avec } \alpha_{ref} = -7.3904)$$

Ainsi, on choisit dans le mode de réalisation décrit ici, de sélectionner systématiquement les sinusoïdes associées aux 20 premiers pics détectés car le codage AMR-WB préserve relativement bien les composantes tonales les plus importantes (en termes d'amplitude). Pour les pics suivants on considère qu'une sinusoïde est valide si la courbure α_i estimée par interpolation ne s'écarte pas trop d'une courbure de référence notée α_{ref} . Le

choix de sélectionner par défaut les 20 premiers pics pourrait naturellement être ajusté pour ne sélectionner par exemple que le pic le plus important, sans altérer l'invention. La valeur $\alpha_{ref} = -7.3904$ est dérivée d'une interpolation parabolique sur le spectre d'amplitude de la fenêtre d'analyse $w_a(m)$ (de longueur M); cette valeur doit être en général ajustée en

5 fonction des paramètres de l'analyse fréquentielle mise en œuvre.

Si la sinusoïde de fréquence estimée k_i et associée au pic localisé sur la raie l_i est "valide" (suivant les critères ci-dessus), la phase de cette sinusoïde est estimée par interpolation. L'estimation de la phase θ_i est fonction des phases des 2 raies contiguës d'indice k_i^- et k_i^+ :

$$10 \quad k_i^- = \lfloor k_i \rfloor \text{ et } k_i^+ = k_i + 1$$

Si $|\angle X(k_i^+) - \angle X(k_i^-)| > \pi$, $\angle X(k_i^+)$ est d'abord modifiée de la façon suivante:

$$\angle X(k_i^+) = \begin{cases} \angle X(k_i^+) + 2\pi & \text{si } \angle X(k_i^-) \geq 0 \\ \angle X(k_i^+) - 2\pi & \text{si } \angle X(k_i^-) < 0 \end{cases}$$

La phase θ_i est estimée par l'interpolation linéaire suivante :

$$15 \quad \theta_i = \frac{\angle X(k_i^+) - \angle X(k_i^-)}{k_i - k_i^-}$$

Dans une variante du mode de réalisation, l'estimation des paramètres pourra être réalisée autrement que par interpolation parabolique, dans le domaine logarithmique ($\log_{10}$) du spectre d'amplitude et interpolation linéaire du spectre de phase. Par exemple, l'estimation de la fréquence pourra employer une des méthodes de détection de maximum décrite dans l'article de R.B. Fischer et D.K Naidu, "A Comparison of Algorithms for Subpixel Peak Detection", qui considère le contexte d'une détection de pics dans une image avec une résolution inférieure au pixel mais s'adapte facilement au problème de l'estimation de fréquence. Plusieurs méthodes alternatives d'interpolation de fréquence peuvent être tirées de l'article de Fischer et Naidu, comme par exemple:

25 - l'approximation gaussienne qui correspond à une interpolation parabolique sur le spectre d'amplitude ramené dans le domaine logarithmique naturel ($\ln$ ou $\log$) – autrement dit, le spectre d'amplitude dans le domaine $\log$ est approximé par une gaussienne, ce qui donne

$$k_i = l_i + \frac{1}{2} \frac{\ln |X(l_i - 1)| - \ln |X(l_i + 1)|}{\ln |X(l_i - 1)| - 2 \ln |X(l_i)| + \ln |X(l_i + 1)|}.$$

- l'interpolation parabolique sur le signal direct (ici le spectre d'amplitude) avec.

$$k_i = l_i + \frac{1}{2} \frac{|X(l_i - 1)| - |X(l_i + 1)|}{|X(l_i - 1)| - 2|X(l_i)| + |X(l_i + 1)|}$$

- l'interpolation linéaire avec

$$k_i = \begin{cases} l_i + \frac{1}{2} \frac{|X(l_i + 1)| - |X(l_i - 1)|}{|X(l_i)| - |X(l_i - 1)|} & \text{si } |X(l_i + 1)| > |X(l_i - 1)| \\ l_i + \frac{1}{2} \frac{|X(l_i + 1)| - |X(l_i - 1)|}{|X(l_i)| - |X(l_i + 1)|} & \text{autrement} \end{cases}$$

- le centre de gravité avec

$$5 \quad k_i = l_i + \frac{|X(l_i + 1)| - |X(l_i - 1)|}{|X(l_i - 1)| + |X(l_i)| + |X(l_i + 1)|}$$

Il va de soi que les variantes d'estimation de fréquence ci-dessus peuvent également être adaptées en remplaçant $|X(l_i)|$ par une valeur dérivée, comme $|X(l_i)|^2$.

De même, étant donnée la fréquence estimée, l'estimation de l'amplitude et la phase sus-décrite peut être avantageusement remplacée par celle d'un facteur multiplicatif complexe (synthétisant donc le rôle de l'amplitude et de la phase) $c_i = a_i \cdot e^{j\theta_i}$ en utilisant la méthode suivante.

La fréquence du sinus détecté étant maintenant connue, on peut donc définir, comme décrit plus loin, la portion Ψ_i de spectre de référence (un "motif") qui sera additionnée au spectre de modélisation sinusoïdale, après application du facteur multiplicatif complexe c_i qu'il s'agit ici d'évaluer. Dans la variante proposée ici, le coefficient c_i est calculé de sorte de minimiser l'erreur de reconstruction de la portion de spectre concernée au sens des moindres carrés. Cette optimisation dite LS (pour "Least Square" en anglais) se résume en la forme bien connue d'un ratio entre un produit scalaire et une norme au carré :

$$c_i = \Psi_i^* \cdot \mathbf{X} / \|\Psi_i\|^2 = \Psi_i^* \cdot \mathbf{X} / (\Psi_i^* \cdot \Psi_i)$$

où la notation $*$, appliquée ici à un vecteur colonne de coefficients complexes, correspond au transposé conjugué du dit vecteur, et la multiplication de vecteurs (vecteur ligne à gauche, vecteur colonne à droite) s'entend comme un produit scalaire.

En pratique, chaque élément du motif – défini précisément plus loin – comprend 2η éléments; chaque produit scalaire implique les produits et somme de η paires de termes complexes, chaque produit complexe impliquant 2 multiplications et une addition.

Il est important de noter que le calcul du coefficient c_i est applicable que lorsque la fréquence de la sinusoïde est au moins à distance $\eta/2$ de la fréquence nulle et de la fréquence de Nyquist, afin d'éviter le repliement spectral. Dans le cas contraire, l'homme de l'art saura adapter la formule en tenant compte du repliement fréquentiel, ou il pourra utiliser l'interpolation linéaire de la phase. Par convention, dans la suite du document ainsi que dans les figures, on pourra désigner par (a_i, k_i, θ_i) les paramètres de la sinusoïde i , de façon équivalente aux paramètres (c_i, k_i) même si c'est le coefficient complexe c_i qui est estimé et utilisé, et que notamment la phase θ_i n'a pas lieu d'être calculée explicitement.

Dans le cas d'une fenêtre d'analyse symétrique et d'une « transformée à phase recentrée » telle que décrite plus haut, le "motif" stocké est un spectre réel. On peut donc réduire de moitié le nombre de multiplications et d'additions. Dans une variante recommandée de l'invention, on aura stocké en mémoire une version du motif normalisée de telle sorte que la norme du motif calculé sur les η sous-échantillons fréquentiels soit toujours égale à 1

Cette normalisation réduit le nombre d'opérations.

L'estimation de l'amplitude et de la phase à l'aide du motif est une alternative avantageuse à l'estimation séparée de l'amplitude et de la phase par interpolation parabolique et linéaire, en ce qu'elle évite un certain nombre d'opérations coûteuses (divisions, calculs de phase par arc-tangente...). Cette estimation par motif est une forme d'analyse par synthèse, car l'estimation des paramètres des sinusoïdes repose sur le motif utilisé à la synthèse.

A noter que la raie d'indice $k=0$ est également comptée comme une sinusoïde valide avec : $A_i = |X(0)|$, $k_i = 0$, $\theta_i = 0$ ou π suivant que la valeur de $X(0)$ est respectivement positive ou négative.

Le bloc 303 effectue ensuite une étape E303b de calcul du spectre $(S(k))$ des composantes tonales ou sinusoïdes "valides" ainsi identifiées. Pour cela, il utilise dans un exemple de réalisation un motif précalculé et mis en mémoire à une étape E300 par un module 300.

Selon le mode de réalisation décrit ici du procédé de traitement de l'invention, le spectre d'une sinusoïde est calculé en tenant compte de la fenêtre d'analyse utilisée pour réaliser l'analyse spectrale de Fourier à court-terme dans le bloc 302.

Comme rappelé précédemment, le spectre d'une sinusoïde fenêtrée s'obtient par convolution du spectre de la fenêtre considérée avec des Diracs ce qui est équivalent à des opérations de translation (modulo 2π) et addition/recouvrement. Etant donné que la fréquence k_i peut prendre n'importe quelle valeur dans $[-\frac{K}{2}, \frac{K}{2}]$, y compris des valeurs

- 5 fractionnaires, la translation du spectre de la fenêtre par un décalage fréquentiel arbitraire est mise en œuvre à l'aide d'une interpolation du spectre de la fenêtre sur une longueur M , où $M > K$.

Avantageusement, le dispositif suivant l'invention utilise une méthode rapide de calcul du spectre court-terme de sinusoïdes à partir d'un "motif" $P(k)$ fréquentiel pré-calculé et stocké en mémoire en 300.

On suppose ici que la fenêtre d'analyse de longueur N appliquée à chaque trame dans le bloc 302 est une fenêtre de Hamming normalisée, telle que décrite plus haut mais en posant $N=512$, et que nous noterons $w_a(m)$.

Cependant, l'invention s'applique à n'importe quelle fenêtre de pondération de longueur N .

- 15 Dans le mode de réalisation privilégié de l'invention, on utilise une transformée de longueur $K=N=512$ et un facteur d'interpolation $M/K = 128$, soit $M=65536$, ce qui nous fournit un spectre $W_M(k)$ qui correspond au spectre discret (normalisé) de la fenêtre $w_a(m)$ avec une interpolation d'un facteur $M/K=128$. L'interpolation garantit une précision suffisante pour représenter des fréquences (fractionnaires) k_i avec une précision inter-raie de $1/128$
- 20 (par rapport à une transformée de Fourier discrète de longueur N). En fait pour $K=512$, la résolution de la FFT (voir la description du bloc 302) est de $16000/512=31,25$ Hz. Le spectre $W_M(k)$ est quant à lui représenté avec une résolution de $31,25/128 \approx 0,24$ Hz.

- Une partie du spectre $W_M(k)$ est illustré à la **figure 7a** où les fréquences négatives sont incluses à des fins d'illustration. Il apparaît que l'atténuation du lobe secondaire est de l'ordre de -43 dB ce qui justifie d'utiliser essentiellement le lobe principal pour représenter le spectre de la fenêtre.
- 25

Pour réduire le nombre de valeurs de $W_M(k)$ à stocker et la complexité de la soustraction complexe (voir la description du bloc 304), on définit le "motif" (normalisé) comme:

30
$$P(k) = \frac{W_M(k)}{|W_M(0)|}, \quad k = 0, \dots, \left\lceil \frac{\eta}{2} \right\rceil \frac{M}{K}$$

en ne retenant que l'intervalle $\left[0, \left\lceil \frac{\eta}{2} \right\rceil\right]$ pour la longueur de transformée N .

Ce motif ne correspond qu'aux fréquences positives car les fréquences négatives s'obtiennent par symétrie hermitienne.

Dans le mode de réalisation décrit ici, on utilise $\eta = 4$ raies (pour la longueur de transformée K).

Le choix de $\eta = 4$, étant donné le facteur d'interpolation $M/K = 128$, limite avantageusement le stockage du "motif" $P(k)$ à $\left\lceil \frac{\eta}{2} \right\rceil \frac{M}{K} + 1 = 257$ valeurs. Le motif $P(k)$

5 est illustré à la **figure 7b** représentant son amplitude, sa phase étant nulle dans le cas privilégié d'une fenêtre symétrique et d'une transformée à « phase recentrée ».

Ainsi, le "motif" $P(k)$, $k = 0, \dots, 256$, est calculé préalablement aux étapes mises en œuvre dans le procédé de traitement et pré-stocké en mémoire ROM dans le bloc 300. Ce motif spectral peut aussi bien être stocké en représentation polaire (amplitude, phase) qu'en
10 représentation cartésienne (partie réelle, partie imaginaire), selon les modalités de calcul choisies.

Il va de soi que l'invention couvre la variante où ce motif $P(k)$ est stocké de façon redondante en incluant également les valeurs correspondant aux fréquences négatives; cette variante n'apporte cependant qu'un avantage négligeable en termes de complexité de calcul.

15

L'étape E303b du bloc 303 effectue alors le calcul du spectre combiné. Le spectre combiné des sinusoides est d'abord initialisé à la valeur nulle :

$$S(k) = 0, \quad k = 0, \dots, K/2$$

On considère par la suite que ces valeurs sont rassemblées dans un vecteur de
20 longueur $K/2+1$:

$$\mathbf{S} = (S(0), S(1), \dots, S(K/2))$$

Le vecteur $\mathbf{S}$ est mis à jour de façon itérative en ajoutant le spectre discret de chaque sinusoïde (ou composante tonale) d'indice i . On approxime le spectre discret d'une sinusoïde de paramètres (a_i, k_i, θ_i) par translation fréquentielle de $\pm k_i$ raies fréquentielles
25 (avec la précision inter-raie de $K/M=1/128$) et addition-recouvrement du "motif" pré-calculé $P(k)$, $k = 0, \dots, 256$, et multiplication par c_i ou $a_i e^{j\theta_i}$.

Ce calcul est illustré en détail à la **figure 8**, pour l'itération i , c'est-à-dire pour le calcul du spectre discret de la sinusoïde d'indice i (>0) dans le cas où cette sinusoïde est valide.

30 Ainsi, le bloc E801 indique que l'organigramme de la figure 8 met en œuvre la construction (const) du spectre à partir des paramètres estimés a_i, k_i, θ_i de la sinusoïde.

Une étape E802 initialise le résultat de la sommation, laquelle est mise en œuvre sous la forme d'une boucle (E803, E804, E807, e808).

Les termes de la somme sont calculés aux étapes E805 et E806, dont les détails sont
35 présentés à la figure 9.

Ainsi, pour la k -ème sinusoïde, les écarts k_i^- et k_i^+ entre la fréquence k_i et les indices entiers des raies proches pour la transformée de longueur K sont convertis par homothétie en indices Δ_i^- et Δ_i^+ relatifs à la transformée en longueur M , de la façon suivante:

$$\begin{aligned} \Delta_i^- &= \frac{M}{K} (k_i - \lfloor k_i \rfloor) = \frac{M}{K} (k_i - k_i^-) \\ \Delta_i^+ &= \frac{M}{K} - \frac{M}{K} (\lfloor k_i \rfloor - k_i) = \frac{M}{K} (k_i^+ - k_i) \end{aligned}$$

où $\lfloor \cdot \rfloor$ représente l'arrondi à l'entier inférieur. Par définition

$$\Delta_i^- + \Delta_i^+ = \frac{M}{K}$$

Le spectre de la k -ème sinusoïde est construit en ne retenant que $\eta = 4$ raies du motif $P(k)$ (pour les fréquences positives). L'organigramme de la figure 8 met en œuvre la sommation selon l'équation suivante :

$$\mathbf{S} = \mathbf{S} + \sum_{p=0}^{\frac{\eta-1}{2}} \text{spec} \left(k_i^+ - p - 1, a_i e^{j\theta_i} P \left(\left\lceil p \cdot \frac{M}{K} + \Delta_i^- \right\rceil \right) \right) + \text{spec} \left(k_i^+ + p, a_i e^{j\theta_i} P \left(\left\lceil p \cdot \frac{M}{K} + \Delta_i^+ \right\rceil \right) \right)$$

où la fonction $\text{spec}(k, V)$ est maintenant définie à travers l'organigramme présenté à la **figure 9**.

Afin de compléter la description faite plus haut de la méthode dite "LS" d'estimation du paramètre d'amplitude complexe c_i équivalent aux paramètres d'amplitude a_i et de phase θ_i , on reformule cette équation ainsi :

$$\mathbf{S} = \mathbf{S} + c_i \Psi_i = \mathbf{S} + a_i e^{j\theta_i} \Psi_i$$

$$\Psi_i = \sum_{p=0}^{\frac{\eta-1}{2}} \text{spec} \left(k_i^+ - p - 1, P \left(\left\lceil p \cdot \frac{M}{K} + \Delta_i^- \right\rceil \right) \right) + \text{spec} \left(k_i^+ + p, P \left(\left\lceil p \cdot \frac{M}{K} + \Delta_i^+ \right\rceil \right) \right)$$

Cette factorisation est exacte sous l'hypothèse (généralement vérifiée, sauf cas particuliers) que la sinusoïde n'est pas trop proche des fréquences nulle et de Nyquist (d'indices 0 et $K/2$), de sorte que le motif échantillonné s'additionne seulement sur le demi-spectre de synthèse, sans repliement spectral.

Le bloc E901 indique que l'organigramme de la figure 9 met en œuvre la mise à jour partielle du spectre (spec) dans l'intervalle $[0, K/2]$ pour la sinusoïde considérée. Cette mise à jour prend pour paramètre la raie de Fourier k et la valeur spectrale V associée à cette raie.

L'étape E902 initialise le résultat du calcul. Ensuite 4 cas sont distingués:

- $k < 0$ (étape E903): la symétrie hermitienne est mise à profit à l'étape E904, ce qui revient à écrire le conjugué complexe de V à la raie de position $-k$;

- $k=0$ ou $K/2$ (étape E905): les raies DC et de Nyquist étant par définition à valeur réelle et sans facteur $1/2$, la valeur considérée est $2.\text{Re}(V)$ où $\text{Re}(\cdot)$ donne la partie réelle (étape E906);
- $k>K/2$ (étape E907): la périodicité du spectre discret et la symétrie hermitienne pour les fréquences négatives est mise à profit à l'étape E908, ce qui revient à écrire le conjugué complexe de V à la raie de position $K-k$;
- Dans les autres cas (étape E909): la valeur V est directement additionnée à la position k .

10 Le calcul de $\mathbf{S}$ présenté précédemment est répété pour chaque sinusoïde valide d'indice i , ce qui revient à calculer:

$$\mathbf{S} = \sum_{i=0}^I \left[\sum_{p=0}^{\frac{\eta-1}{2}} \text{spec} \left(k_i^+ - p - 1, a_i e^{j\theta_i} P \left(\left[p \cdot \frac{M}{K} + \Delta_i^- \right] \right) \right) + \text{spec} \left(k_i^+ + p, a_i e^{j\theta_i} P \left(\left[p \cdot \frac{M}{K} + \Delta_i^+ \right] \right) \right) \right]$$

où I est le nombre de sinusoïdes valides, ou composantes tonales identifiées. On suppose ici, sans perte de généralité (car il suffit d'une indexation appropriée), que l'ensemble des sinusoïdes non valides ne sont pas comptées.

On peut noter que pour la sinusoïde d'indice 0, qui correspond à la raie DC (composante continue) est prise en compte dans la formule ci-dessus. Cette opération permet de ne pas modifier la valeur DC (composante continue) par post-traitement.

20 Dans une variante du mode de réalisation, le processus de détection de pics est répété de façon itérative, pour ne détecter à la i -ème itération que le pic de position l_i correspondant au maximum du spectre $X(k)$ auquel est retranché à chaque itération le spectre des sinusoïdes déjà extraites. Cette variante correspond à une forme d'analyse connue dans l'état de l'art sous le nom de "matching pursuit".

25

Le bloc 304 de la figure 3 calcule à l'étape E304 un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales. Dans le cas présent de la figure 3, l'obtention de ce résidu s'effectue dans le domaine fréquentiel par la différence entre le spectre du signal audiofréquence et le spectre des composantes tonales identifiées. On obtient ainsi le spectre résiduel après extraction (soustraction) des sinusoïdes synthétisées. Le résidu s'obtient dans le domaine complexe comme suit:

$$R(k) = X(k) - S(k), \quad k = 0, \dots, K/2$$

Dans le cas idéal où le signal original est constitué comme une combinaison de sinusoïdes qui sont préservées par codage-décodage, le résidu correspond alors au bruit de codage. Cependant en général, le résidu obtenu comporte :

35

- une partie correspondant au résidu de modélisation sinusoïdale du signal original (incluant les composantes non-tonales du signal ainsi que les erreurs de modélisation)
- une partie correspondant au bruit de codage.

5 Dans le bloc 305 de la figure 3, une étape E305 d'atténuation spectrale est mise en œuvre. Cette opération utilise en général toute l'information disponible. Elle peut par exemple être fonction de la classification des composantes tonales extraites, du signal résiduel et des paramètres décodés.

10 Le bloc 305 réalise une atténuation spectrale du spectre du résidu sous la forme suivante:

$$R_{att}(k) = g(k).R(k), \quad k = 0, \dots, K/2$$

où $g(k)$ est un gain adaptatif vérifiant de façon générale : $0 \leq g(k) \leq 1$.

Dans une variante du mode de réalisation privilégié, le gain appliqué au résidu subira un lissage temporel afin d'atténuer les variations inter-frames et éviter certains artefacts.

15 Dans ce cas le bloc 305 effectuera le calcul

$$R_{att}(k) = \bar{g}(k).R(k), \quad k = 0, \dots, K/2$$

où le gain lissé $\bar{g}(k)$ s'obtient par exemple par lissage exponentiel du $g(k)$ suivant l'équation

$$\bar{g}(k) = \zeta(k)\bar{g}(k) + (1 - \zeta(k))g(k)$$

20 avec un facteur d'oubli $\zeta(k)$ qui peut être soit fixe (par exemple de valeur 0,75) soit adaptatif (par exemple de valeur $\zeta(k) = 1 - g(k)$).

Dans le mode de réalisation particulier, on utilise les informations de classification obtenues par le bloc 310. Ce bloc 310 sera décrit ultérieurement.

Ainsi, le gain d'atténuation est par exemple forcé à la valeur $g(k) = 1, \quad k = 0, \dots, N/2$

25 lorsque la trame courante est classifiée comme "autre".

Lorsque la trame est classifiée comme "musique", dans le mode de réalisation privilégié, le gain $g(k)$ est calculé comme le gain d'un filtre de Wiener fréquentiel (contraint) sous la forme:

$$g(k) = \max\left(\frac{RSB(k)}{1 + RSB(k)}, g_{\min}\right), \quad k = 0, \dots, N/2$$

30 où $g_{\min}$ définit un niveau de réduction maximale (avec par exemple $g_{\min}=0,25$) et le rapport signal sur bruit $RSB(k)$ est estimé par l'équation

$$RSB(k) = \frac{|X(k)|^2}{B(k)}$$

$|X(k)|^2$ correspondant à l'énergie de la raie d'indice k et $B(k)$ étant une estimation du niveau du bruit de codage calculée à partir du signal résidu $R(k)$ suivant l'équation :

$$B^{(t)}(k) = 0,2.B^{(t-1)}(k) + 0,6.E^{(t)}(k) + 0,2.E^{(t-1)}(k)$$

où t et $t-1$ correspondent aux trames courante et précédente, respectivement, et

$$E^{(t)}(k) = \frac{1}{stop(k) - start(k)} \sum_{i=start(k)}^{stop(k)-1} |X(i)|^2$$

est l'énergie moyenne dans la bande de Bark (incomplète proche de $k = 0$ et $k = K/2$) centrée autour de la raie k . La bande de Bark est définie par les raies dans l'intervalle $[start(k), stop(k)-1]$, où les tableaux de début (start) et fin de bande (stop) sont définis ci-dessous pour l'exemple $K=512$:

10

```
start={0, 0, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,
16, 17, 18, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31,
32, 33, 34, 35, 36, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47,
48, 49, 50, 51, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63,
64, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 76, 77, 78,
79, 80, 81, 82, 83, 84, 85, 86, 87, 87, 88, 89, 90, 91, 92, 93, 94,
95, 96, 97, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 106, 107,
108, 109, 110, 111, 112, 113, 114, 115, 115, 116, 117, 118, 119, 120,
121, 122, 123, 124, 124, 125, 126, 127, 128, 129, 130, 131, 132, 132,
133, 134, 135, 136, 137, 138, 139, 139, 140, 141, 142, 143, 144, 145,
146, 147, 147, 148, 149, 150, 151, 152, 153, 154, 154, 155, 156, 157,
158, 159, 160, 161, 161, 162, 163, 164, 165, 166, 167, 168, 168, 169,
170, 171, 172, 173, 174, 174, 175, 176, 177, 178, 179, 180, 180, 181,
182, 183, 184, 185, 186, 186, 187, 188, 189, 190, 191, 192, 193, 194,
195, 196, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209,
210, 211, 212, 213, 213, 214, 215, 216, 217, 218, 219, 220, 220, 221,
222, 223, 224, 225, 226, 226, 227, 228, 229, 230, 231, 232, 233}
```

```
stop={3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,
20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37,
38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55,
56, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 72, 73, 74,
75, 76, 77, 78, 79, 80, 81, 82, 83, 85, 86, 87, 88, 89, 90, 91, 92,
93, 94, 95, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 108, 109,
110, 111, 112, 113, 114, 115, 116, 118, 119, 120, 121, 122, 123, 124,
125, 126, 128, 129, 130, 131, 132, 133, 134, 135, 136, 138, 139, 140,
141, 142, 143, 144, 145, 147, 148, 149, 150, 151, 152, 153, 155, 156,
157, 158, 159, 160, 161, 162, 164, 165, 166, 167, 168, 169, 170, 172,
173, 174, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 186, 188,
189, 190, 191, 192, 193, 195, 196, 197, 198, 199, 200, 202, 203, 204,
205, 206, 207, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219,
219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232,
233, 234, 236, 237, 238, 239, 240, 241, 242, 244, 245, 246, 247, 248,
249, 251, 252, 253, 254, 255, 256, 257, 257, 257, 257, 257, 257,
257, 257, 257, 257, 257, 257, 257, 257, 257, 257, 257, 257,
257, 257, 257}
```

A noter que d'autres coefficients de lissage (de valeur 0,2, 0,6 et 0,2) peuvent être ajustés différemment.

Dans ce mode de réalisation, le niveau de bruit de codage $B(k)$ est donc estimé à l'aide de bandes critiques "flottantes" dans le sens où chaque raie de Fourier d'indice k se voit associée à un intervalle d'intégration $[\text{start}(k), \text{stop}(k)-1]$. Les tableaux donnés ci-dessus correspondent à un intervalle de $\frac{1}{2}$ Bark de part et d'autre de chaque raie d'indice k , autrement dit l'intervalle $[\text{start}(k), \text{stop}(k)-1]$ a une largeur (approximative) de 1 Bark. Dans une variante du mode de réalisation, l'estimation du bruit de codage pourra utiliser des intervalles de taille différente, par exemple 2 Barks ou $\frac{1}{2}$ Bark.

Par ailleurs, le niveau d'atténuation maximale, $g_{\min}$, pourra être rendu adaptatif en fonction de la classification, comme dans la technique de Vaillancourt et al. De plus le gain de Wiener pourra être remplacée par une autre fonction de réduction utilisant le rapport signal à bruit $RSB(k)$.

Le rapport signal à bruit $RSB(k)$ est par essence relativement élevé au voisinage des sinusoïdes extraites – en effet dans ce voisinage le résidu de modélisation sinusoïdal est de niveau inférieur à celui du signal puisqu'une sinusoïde a été retranchée ; par contre, pour les raies d'indice k situées dans les zones fréquentielles inter-tones, le résidu a un niveau correspondant à celui du signal et le rapport $RSB(k)$ aura une valeur proche de 0 dB. Dans une variante du mode de réalisation, on pourra artificiellement baisser la valeur du rapport signal à bruit en lui appliquant un facteur (par exemple 1/16) afin de forcer une atténuation plus forte que le gain de Wiener associé à un RSB de 0 dB.

Dans une variante du mode de réalisation, le gain $g(k)$ est calculé à partir des coefficients du filtre LPC décodé $\hat{A}(z)$ associés à la dernière sous-trame (5 ms) de la trame courante. Ces coefficients sont fournis par le bloc 301 de décodage. Soit donc le filtre de fonction de transfert :

$$H(z) = \frac{\hat{A}(z/\gamma_1)}{\hat{A}(z/\gamma_2)}$$

avec $\gamma_1=0,6$ et $\gamma_2=0,95$. On rappelle que les coefficients du filtre $\hat{A}(z)$ et donc la fonction de transfert $H(z)$, sont définis pour la fréquence d'échantillonnage de 12,8 kHz. Dans la présente description, le signal traité est échantillonné à 16 kHz. On définit alors $K' = \lfloor K \times 12,8/16 \rfloor = 410$, qui est l'indice fréquentiel correspondant approximativement à la fréquence 12,8 kHz. De là, le gain $g(k)$ est calculé comme suit :

$$g(k) = \begin{cases} g_0 \cdot \left| H\left(e^{j2\pi k/K'}\right) \right| & k = 0, \dots, K'/2 \\ g(K'/2) & k = K'/2 + 1, \dots, K/2 \end{cases}$$

où g_0 est un facteur de normalisation tel que $\sum_{k=0}^{K'/2} g(k) = 1$. Il est bien évident qu'une transposition de l'invention à un domaine d'échantillonnage identique à celui où est défini

$\hat{A}(z)$ impliquerait une valeur $K'=K$. Afin d'éviter une atténuation trop forte, la valeur du gain

calculé pourra être saturée de sorte que:

$$g(k) = \begin{cases} \max\left(g_0 \cdot \left|H\left(e^{j2\pi k/K'}\right)\right|, g_{\min}\right) & k = 0, \dots, K'/2 \\ g(K'/2) & k = K'/2 + 1, \dots, K/2 \end{cases}$$

Cette variante qui utilise le filtre de prédiction linéaire décodé a l'avantage de fournir une
 5 fonction de pondération relativement lisse et variant lentement d'une trame à l'autre pour un signal stationnaire; par contre elle a le désavantage de colorer le résidu par la réponse (étendue) du filtre $H(z)$.

D'autres réalisations du calcul du gain sont possibles, sans changer le principe de l'invention.

10 Dans une variante du mode de réalisation ainsi décrit, le gain est calculé en fonction du modèle sinusoïdal. Les composantes tonales qui sont extraites et synthétisées ont pour paramètres (a_i, k_i, θ_i) avec $i \in \{1, \dots, I\}$, où I est le nombre de sinusoïdes valides (après élimination des sinusoïdes non valides qui ne sont pas comptées et après indexation appropriée) et $k_1 < k_2 < \dots < k_I$. Dans ces conditions on introduit un paramètre d'atténuation
 15 en dB $\chi(k)$, $k = 0, \dots, N/2$, tel que:

$$\chi(k) = \sin^2\left(\pi \frac{k - k_i}{k_{i+1} - k_i}\right), \text{ avec } i \text{ tel que } k_i \leq k \leq k_{i+1}$$

avec les conventions suivantes: $k_0 = -k_1$ et $k_{I+1} = K - k_I$.

Un exemple pour ce paramètre d'atténuation est montré à la **figure 10a** pour l'exemple suivant avec $I = 6$: $k_{i=0, \dots, 6} = \{5.25, 15.1, 35.5, 120, 180, 240\}$. On voit que le
 20 paramètre d'atténuation a une valeur qui reste proche de zéro aux alentours des fréquences k_i .

En général le paramètre d'atténuation pourra être n'importe quelle fonction positive et normalisée à une valeur maximale de 1 (sans perte de généralité car il suffit d'une normalisation) avec un minimum (proche de 0) aux extrémités de $[k_i, k_{i+1}]$ et un maximum
 25 unique à l'intérieur de l'intervalle $[k_i, k_{i+1}]$.

Dans une variante supplémentaire, le paramètre d'atténuation en dB est donné par :

$$\chi(k) = \frac{(k_{i+1} - k_i)^2}{4} (k - \lceil k_i \rceil)(\lfloor k_{i+1} \rfloor - k) \text{ si } k \in [k_i, k_{i+1}]$$

avec les conventions suivantes: $k_0 = -k_1$ et $k_{I+1} = K - k_I$. Dans ce cas, le paramètre d'atténuation est une fonction parabolique par morceaux, où les sommets de chaque
 30 parabole ont une amplitude de 1 et les valeurs sont nulles pour les raies k qui correspondent

(exactement) à la fréquence d'une des sinusoïdes. Un exemple pour ce paramètre d'atténuation est montré à la **figure 10b** pour les mêmes fréquences k_i que précédemment.

Dans un mode de réalisation avantageux, le gain $g(k)$ est donné par:

$$g(k) = 10^{-\frac{\lambda \mathcal{X}(k)}{20}}$$

5 où λ est le niveau d'atténuation en dB désirée – on prend par exemple $\lambda = 6$ dB.

Le bloc 310 effectue donc une classification (étape E310) du signal audiofréquence $x(n)$.

La classification sert à détecter les signaux qui sont visés par le post-traitement, essentiellement les signaux musicaux avec une structure harmonique ; ainsi sur les autres
10 signaux, le post-traitement doit rester "transparent", en particulier pour les signaux de parole et les signaux transitoires (attaques d'instruments de musique par exemple).

Les détails de la classification dépassent l'objet de l'invention décrite ici. On présente néanmoins plusieurs exemples de mise en œuvre de classification.

Dans le mode de réalisation principal, la classification comprend deux classes
15 possibles: "musique" et "autre". Cette classification est obtenue à partir de la recommandation UIT-T G.720.1 "Generic sound activity detector (GSAD)" avec les adaptations suivantes:

- La norme G.720.1 est utilisée en bande élargie (à 16 kHz) ; l'algorithme GSAD fonctionne avec des trames de 10 ms et sans pré-exploration ou
20 "lookahead" en anglais. Il est appliqué au signal décodé sur 4 trames (soit 40 ms) afin de couvrir un support temporel incluant le signal analysé.
- Si les 4 décisions obtenues dans la trame courante sont "musique", le bloc 310 classe la trame courante comme "musique", sinon comme "autre".

Dans une variante du mode de réalisation, la classification sera mise en œuvre
25 comme dans l'article cité précédemment de Vaillancourt et al. Dans ce cas, le signal est par défaut supposé être un signal de musique, et le niveau d'atténuation maximale $g_{\min}$ est adapté en fonction d'un paramètre de variation d'énergie σ_E calculé selon le principe de l'article de T. Vaillancourt et al. cité précédemment afin de quantifier le degré de variation de l'énergie spectrale dans le domaine FFT:

$$30 \quad \sigma_E = \sqrt{\sum_{t=-Nb}^0 (\bar{E}_{\text{var}}^{t-1} - \bar{E}_{\text{var}}^t)^2}$$

où $\bar{E}_{\text{var}}^{t-1}$ est une moyenne des écarts d'énergie entre trames adjacentes. Un seuillage (adaptatif) est ensuite appliqué sur le paramètre σ_E pour déterminer le niveau de stabilité du signal. Les détails du calcul de σ_E ne sont pas repris ici, mais ils peuvent être trouvés dans l'article de T. Vaillancourt et al.

On suppose ici que le calcul de σ_E conduit à un niveau d'atténuation maximale compris entre 0 et 12 dB, comme dans l'article de T. Vaillancourt et al.. Dans ce cas la valeur de λ est donnée par le niveau d'atténuation maximale fourni par la classification.

5 Dans une autre variante du mode de réalisation, l'information de classification pourrait être disponible au décodeur en tant que paramètre multiplexé dans le train binaire.

Le bloc 306 de la figure 3 effectue une étape de combinaison du résidu atténué et du signal issu du spectre des composantes tonales.

10 Dans ce mode de réalisation, l'étape E306 recombine le spectre résiduel atténué et le spectre des sinusoides identifiées comme suit:

$$T(k) = S(k) + R_{att}(k), \quad k = 0, \dots, K/2$$

Le spectre du signal post-traité $T(k)$, $k = 0, \dots, K/2$, est étendu en $T(k)$, $k = 0, \dots, K-1$ avec

$$T(k) = T^*(K-k), \quad k = K/2+1, \dots, K-1$$

15 où $*$ est le conjugué complexe. Pour obtenir le signal temporel $t(m)$, le signal $T(k)$ est transformé par le bloc 307 qui comprend une FFT inverse (étape E307) de longueur $K=512$, suivie d'une permutation circulaire de $N/2$ échantillons temporels vers la droite effectuée pour compenser le « recentrage de phase » lorsque celui-ci est appliqué à l'analyse, comme cela décrit plus haut comme mode de réalisation principal. Cela revient à écrire la Transformée de Fourier inverse ainsi ::

$$t(m) = \frac{1}{K} \sum_{k=0}^{K-1} T(k) \cdot e^{j \frac{2\pi k}{K} (m - \frac{N}{2})}, \quad m = 0, \dots, K-1$$

Dans une variante de l'invention où le recentrage de phase n'est pas appliqué à l'analyse, la transformée fréquence-temps (bloc 307) pour la synthèse consiste en une simple FFT inverse :

$$25 \quad t(m) = \frac{1}{K} \sum_{k=0}^{K-1} T(k) \cdot e^{j \frac{2\pi k}{K} m}, \quad m = 0, \dots, K-1$$

Dans les deux modes de réalisation décrits, la FFT inverse est calculée suivant la méthode décrite dans l'article de H.V. Sorensen et al., intitulé "Real-valued fast Fourier transform algorithm," IEEE Trans. on Signal Processing, Vol.35, No.6, pp 849-863, 1987.

La fenêtre de synthèse $w_s(m)$ est définie comme

$$30 \quad w_s(m) = \begin{cases} 0 & m = 0, \dots, D-1 \\ \frac{w_{ola}(m-D)}{w_a(m)} & m = D, \dots, N-1 \end{cases}, \quad m = 0, \dots, N-1$$

où l'on a introduit par commodité la quantité $D = N - L - L_{ola}$ d'échantillons ignorés et rendus nuls par le fenêtrage de synthèse, et où la fenêtrage $w_{ola}(m)$, $m = 0, \dots, L + L_{ola} - 1$ d'addition-recouvrement (OLA pour Overlap-Add en anglais) est définie par:

$$w_{ola}(m) = \begin{cases} \frac{1}{2} \left(1 - \cos \left(\frac{\pi}{L_{ola}} (m + 0.5) \right) \right) & m = 0, \dots, L_{ola} - 1 \\ 1 & m = L_{ola}, \dots, L - 1 \\ \frac{1}{2} \left(1 + \cos \left(\frac{\pi}{L_{ola}} ((m - L) + 0.5) \right) \right) & m = L, \dots, L + L_{ola} - 1 \end{cases}$$

5 avec $L_{ola} = 40$ et $L = 320$.

Les fenêtrages $w_{ola}(m)$ et $w_s(m)$ sont par ailleurs illustrées aux **figures 6b et 6c** respectivement.

Comme illustré par la **figure 11**, la fenêtrage de synthèse a vocation à traduire
10 simultanément le redressement d'enveloppe temporelle et préparer le procédé d'addition-recouvrement. On définit alors le vecteur d'échantillons résultant de ce fenêtrage :

$$t_w(m) = w_s(m)t(m).$$

En pratique, les D premiers échantillons n'ont pas besoin d'être calculés car nuls. Les
15 L_{ola} derniers échantillons, affectés par la pente décroissante de la fenêtrage $w_{ola}(m)$, ne sont exploitables que pour la synthèse de la trame suivante, et sont stockés à cet effet dans un vecteur

$$t_{save}(n) = t_w(n + N - D), \quad n = 0, \dots, L_{ola}$$

Pour la synthèse de la trame courante, on combine les échantillons $t_{save}^{(prev)}(n)$
sauvegardés à la trame précédente (ou bien mis à zéro si le post-traitement n'a pas été
20 appliqué à la trame précédente) et une portion du vecteur $t_w(m)$, de la façon suivante, achevant le procédé d'addition-recouvrement :

$$x_{post}(n) = \begin{cases} t_w(n + D) + t_{save}^{(prev)}(n) & n = 0, \dots, L_{ola} - 1 \\ t_w(n + D) & n = L_{ola}, \dots, L - 1 \end{cases}$$

On vérifie bien que compte-tenu du processus d'addition-recouvrement, les L
échantillons de signal post-traité $x_{post}(n)$ sont obtenus avec un retard de L_{ola} échantillons par
25 rapport à la trame courante de L échantillons considérée en entrée du post-traitement. Ainsi, dans le mode de réalisation décrit, le post-traitement induit un retard algorithmique supplémentaire de $L_{ola} = 40$ échantillons, soit 2.5 ms.

Le mode de réalisation principal repose sur une atténuation spectrale du résidu de modélisation sinusoïdale ; l'atténuation décrite précédemment correspond à une convolution

circulaire (multiplication du résidu par un gain dans le domaine de la transformée Fourier discrète). Or, il est connu de l'homme de l'art qu'une fonction de gain ayant une dynamique trop forte peut entraîner un résultat de convolution circulaire très différent d'une convolution linéaire. Dans une variante du mode de réalisation, un bourrage à zéro et une technique

5 d'addition-recouvrement/mémorisation ("overlapp-add/save" en anglais) pourra être mise en œuvre pour prendre en compte la longueur non nulle de la réponse impulsionnelle associée à la fonction de gain réalisant l'atténuation spectrale. En général cette variante occasionne un retard temporel supplémentaire.

10 Dans un deuxième mode de réalisation, les étapes d'obtention du résidu, d'atténuation et de combinaison du procédé de traitement tel que décrit en référence à la figure 3, peuvent être effectuées dans le domaine temporel.

La **figure 12** illustre ce procédé ainsi que le dispositif de traitement 1200 correspondant. Sur cette figure, les blocs 300, 301, 302, 303 et 310 restent identiques à ceux

15 décrits précédemment en référence à la figure 3.

Le signal modélisé en somme de sinus encore appelé spectre des composantes tonales ($S(k)$) est dans ce mode de réalisation d'abord restitué dans le domaine temporel (bloc 1201) par une étape E1201 de transformée inverse et son enveloppe temporelle redressée par un fenêtrage approprié, en laissant de côté la portion située sur la zone de

20 recouvrement (les L_{ola} derniers échantillons) pour le traitement de la trame suivante, et en combinant de façon appropriée la même portion sauvegardée à la trame précédente pour reconstituer le début de la trame (avec un retard L_{ola}).

Le résidu est alors calculé à l'étape E1202 par le bloc 1202 par simple soustraction entre le signal audiofréquence décodé ($x(n)$) retardé de L_{ola} échantillons et le signal issu du

25 spectre des composantes tonales ($s(n)$) après transformation en 1201.

A ce résidu, est appliquée alors une atténuation spectrale (bloc 1203) par un filtrage temporel (étape E1203) dont le design des paramètres est piloté par le module de modélisation sinusoïdale (bloc 303) et/ou des paramètres décodés (p.decod.) à partir du train binaire et/ou le module de classification (bloc 310).

30 L'étape de combinaison E1204 du résidu ainsi atténué ($r_{att}(n)$) et du signal temporel des composantes tonales ($s(n)$), est effectuée par le bloc 1204 pour obtenir le signal temporel traité ($x_{post}(n)$).

De la même façon que le procédé décrit en référence à la figure 3, le traitement effectué sur le signal permet de réduire le bruit de codage qui pouvait subsister sur les

35 signaux musicaux lors du décodage effectué par des décodeurs plus adaptés à des signaux de parole. Cette méthode permet donc d'obtenir des signaux à la fois de parole et de musique de meilleure qualité, tout en gardant les codeurs et décodeurs de parole existants.

La **figure 13** représente un exemple de réalisation matérielle d'un dispositif de traitement selon l'invention. Celui-ci peut faire partie intégrante d'un décodeur audiofréquence ou d'un équipement recevant des signaux audiofréquences décodés.

5 Ce type de dispositif comporte un processeur PROC coopérant avec un bloc mémoire BM comportant une mémoire de stockage et/ou de travail MEM.

10 Le bloc mémoire peut avantageusement comporter un programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de traitement au sens de l'invention, lorsque ces instructions sont exécutées par le processeur PROC, et notamment les étapes d'identification de composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel, de calcul d'un spectre des composantes tonales identifiées à partir d'un modèle sinusoïdal, d'obtention d'un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales, d'atténuation spectrale du résidu et de combinaison du résidu atténué et du signal issu du spectre des composantes tonales.

15 Typiquement, la description de la figure 3 ou de la figure 12 reprend les étapes d'un algorithme d'un tel programme informatique. Le programme informatique peut également être stocké sur un support mémoire lisible par un lecteur du dispositif ou téléchargeable dans l'espace mémoire de celui-ci.

20 La mémoire MEM enregistre le motif $P(k)$ utilisé dans le procédé selon un mode de réalisation de l'invention et de manière générale, toutes les données nécessaires à la mise en œuvre du procédé.

Un tel dispositif comporte un module d'entrée E apte à recevoir un signal audiofréquence ($x(n)$) et un module de sortie S apte à transmettre le signal traité $x_{\text{post}}(n)$.

25 Dans un mode possible de réalisation, le dispositif ainsi décrit peut également comporter les fonctions de décodage en plus des fonctions de traitement selon l'invention.

REVENDEICATIONS

- 5 1. Procédé de traitement d'un signal audiofréquence décodé, caractérisé en ce qu'il comporte les étapes suivantes:
- identification (E303a) de composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel;
 - calcul (E303b) d'un spectre des composantes tonales identifiées à partir d'un modèle sinusoïdal;
 - 10 - obtention (E304, E1202) d'un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales;
 - atténuation spectrale (E305, E1203) du résidu; et
 - combinaison (E306, E1204) du résidu atténué et du signal issu du spectre des composantes tonales.
- 15 2. Procédé selon la revendication 1, caractérisé en ce que les étapes d'obtention du résidu, d'atténuation et de combinaison sont effectués dans le domaine fréquentiel.
- 20 3. Procédé selon la revendication 1, caractérisé en ce que les étapes d'obtention du résidu, d'atténuation et de combinaison sont effectués dans le domaine temporel.
- 25 4. Procédé selon la revendication 1, caractérisé en ce que l'identification des composantes tonales s'effectue par une méthode d'interpolation sur un nombre prédéterminé de raies fréquentielles.
- 30 5. Procédé selon la revendication 1, caractérisé en ce que l'identification des composantes tonales utilise un motif spectral ($P(k)$) pré-calculé et stocké en mémoire.
- 35 6. Procédé selon la revendication 1, caractérisé en ce que le calcul du spectre des composantes tonales utilise un motif spectral ($P(k)$) pré-calculé et stocké en mémoire.
7. Procédé selon la revendication 1, caractérisé en ce que l'atténuation spectrale s'effectue en fonction d'une donnée de classification du signal audiofréquence.
8. Procédé selon la revendication 3, caractérisé en ce que l'atténuation spectrale s'effectue par un filtrage du résidu dans le domaine temporel.

- 5 9. Dispositif de traitement d'un signal audiofréquence décodé, caractérisé en ce qu'il comporte:
- un module d'identification (303) de composantes tonales dans le signal audiofréquence représenté dans le domaine fréquentiel;
 - un module de calcul (303) d'un spectre des composantes tonales identifiées à partir d'un modèle sinusoïdal;
 - un module d'obtention (304, 1202) d'un résidu par la différence entre le signal audiofréquence décodé et un signal issu du spectre des composantes tonales;
 - 10 - un module d'atténuation spectrale (305, 1203) du résidu; et
 - un module de combinaison (306, 1204) du résidu atténué et du signal issu du spectre des composantes tonales.
- 15 10. Décodeur de signal audiofréquence caractérisé en ce qu'il comporte un dispositif de traitement selon la revendication 9.
- 20 11. Programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de traitement selon l'une des revendications 1 à 8, lorsque ces instructions sont exécutées par un processeur.
- 25

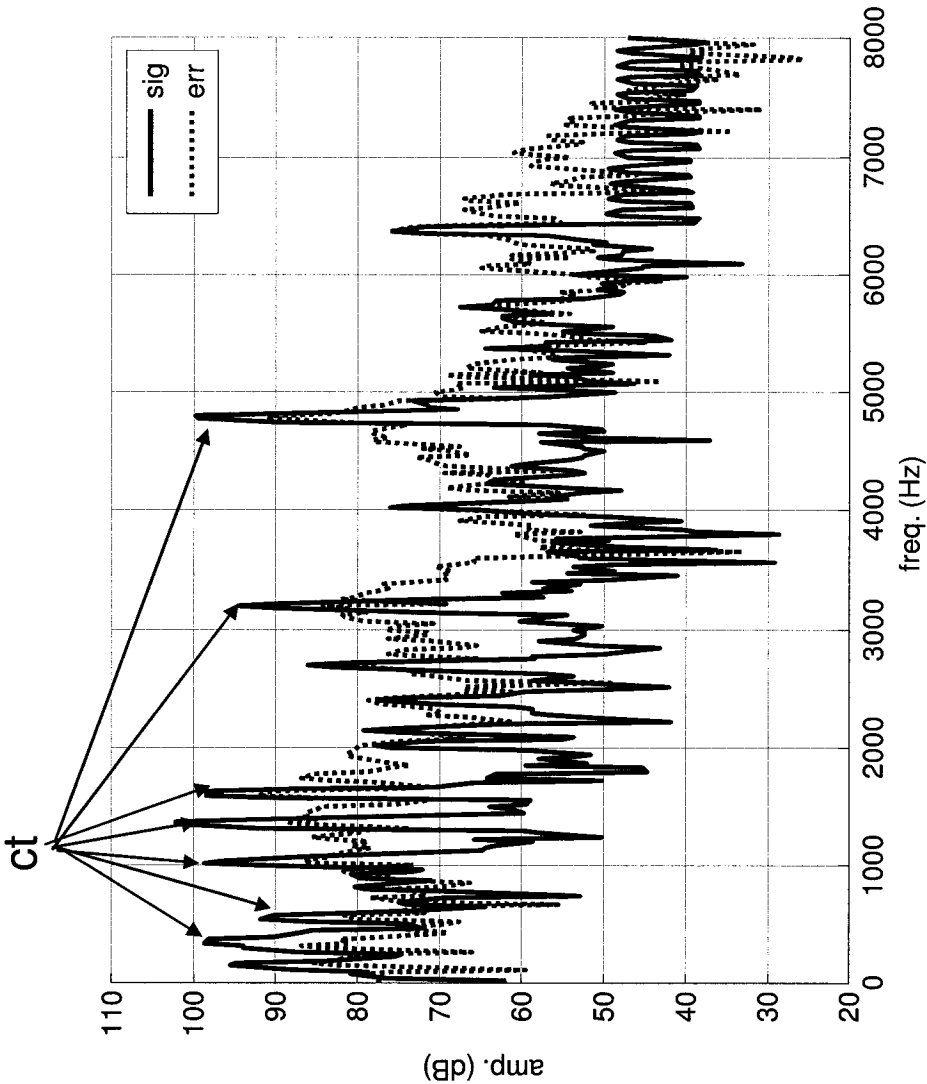


Fig. 1

2/13

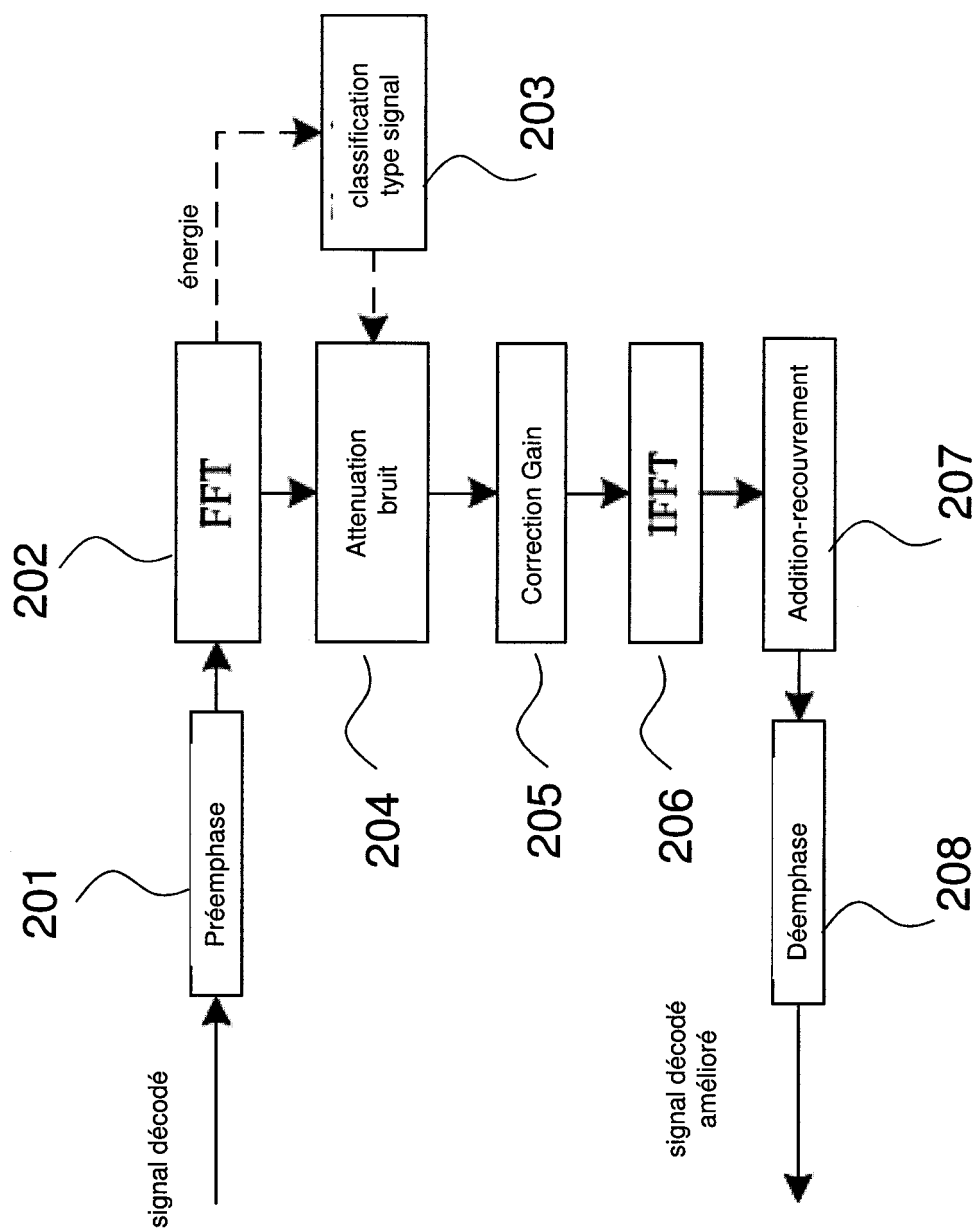
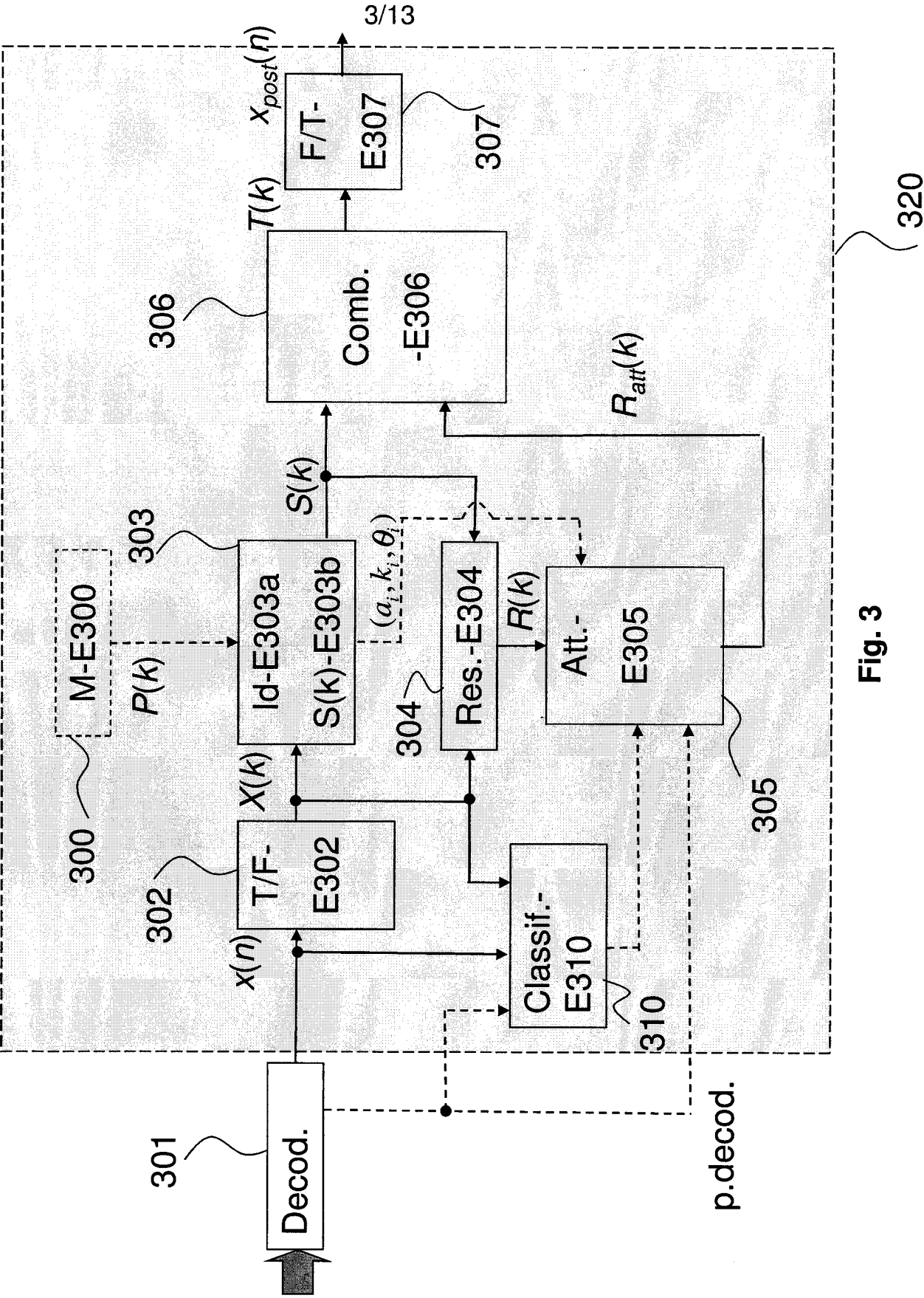
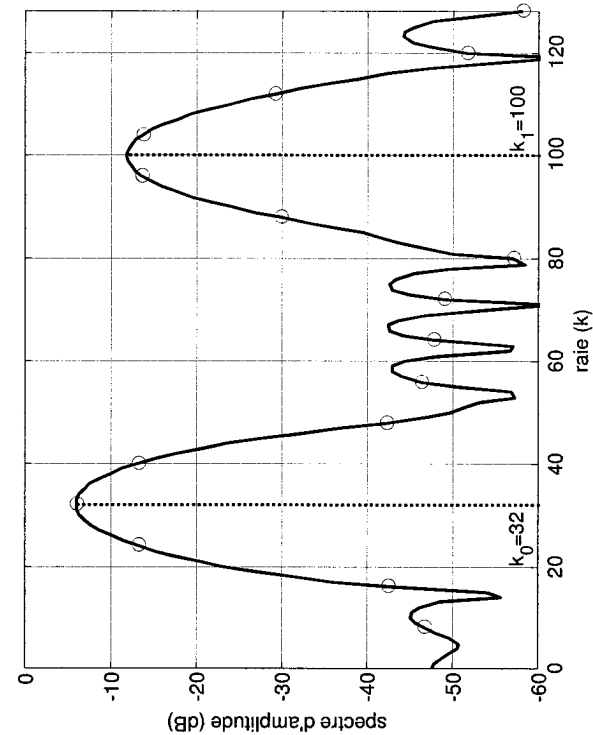


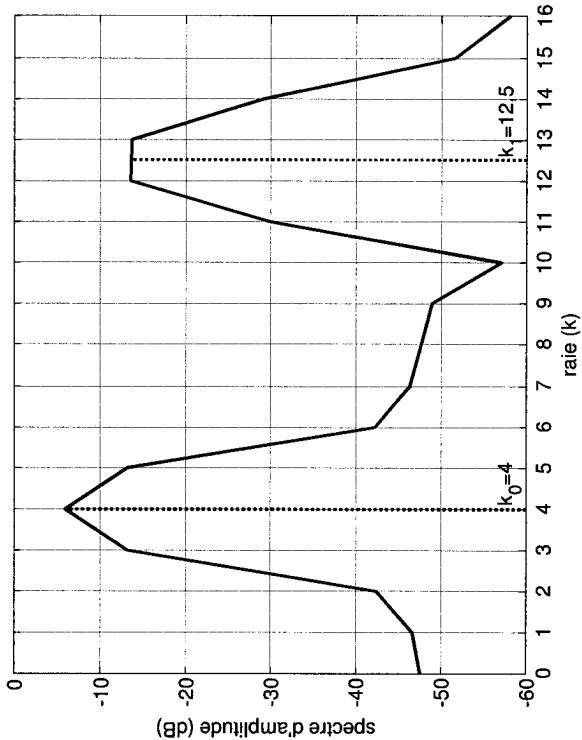
Fig. 2 (état de l'art)





$N=32$
 $K=8N=256$

Fig. 4b



$N=32$
 $K=N=32$

Fig. 4a

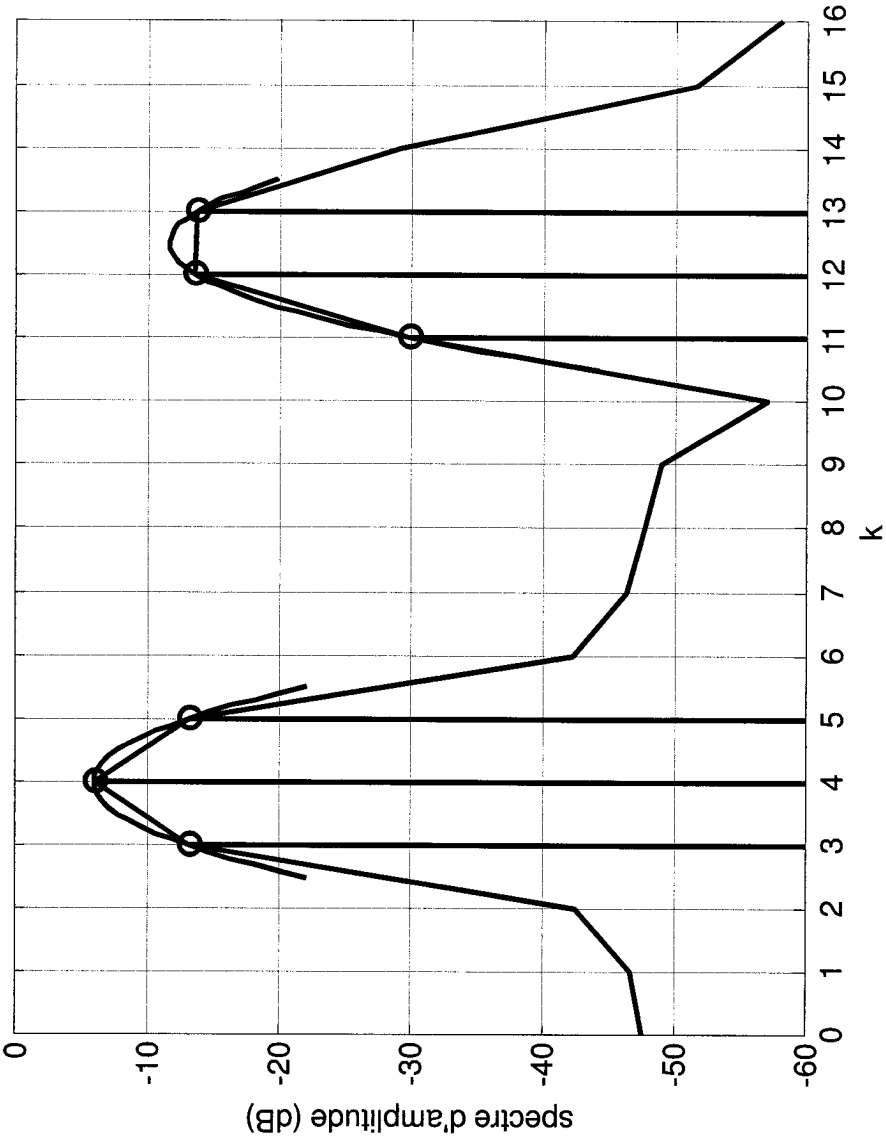


Fig. 5

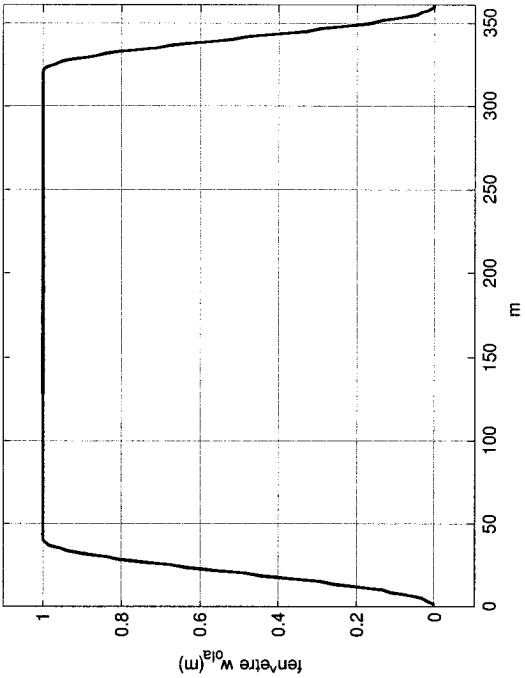


Fig. 6b

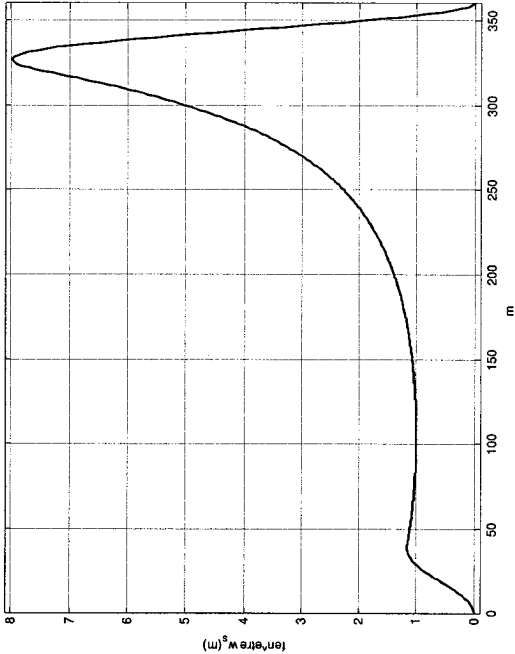


Fig. 6c

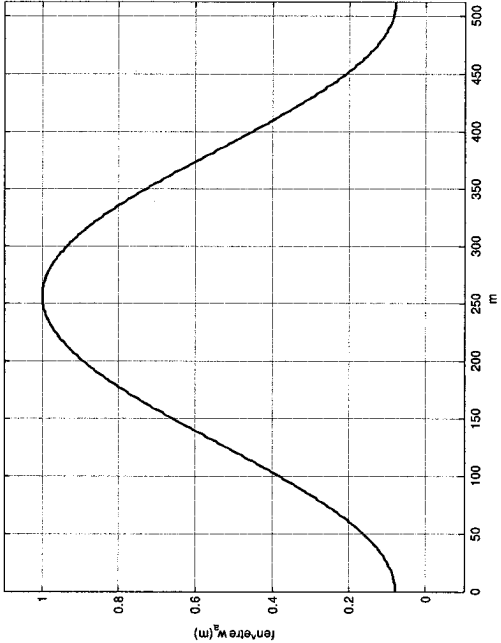
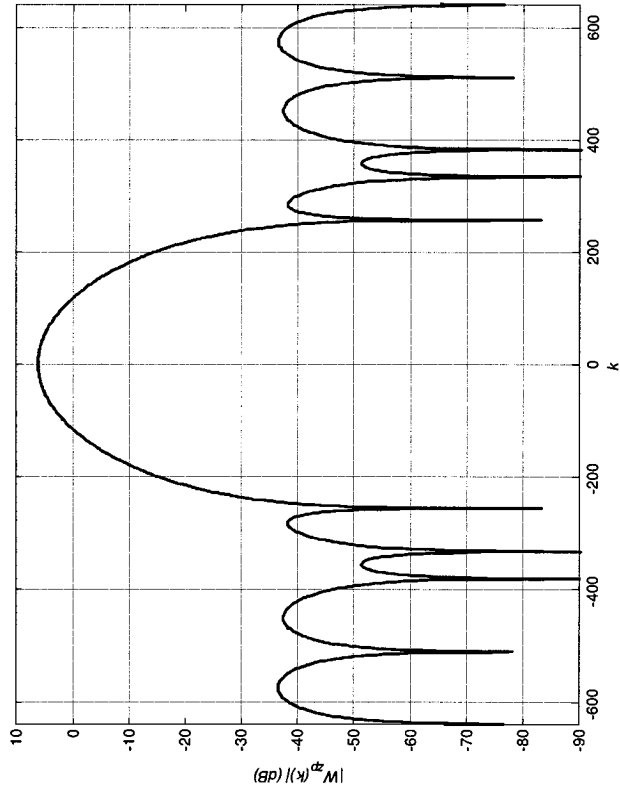
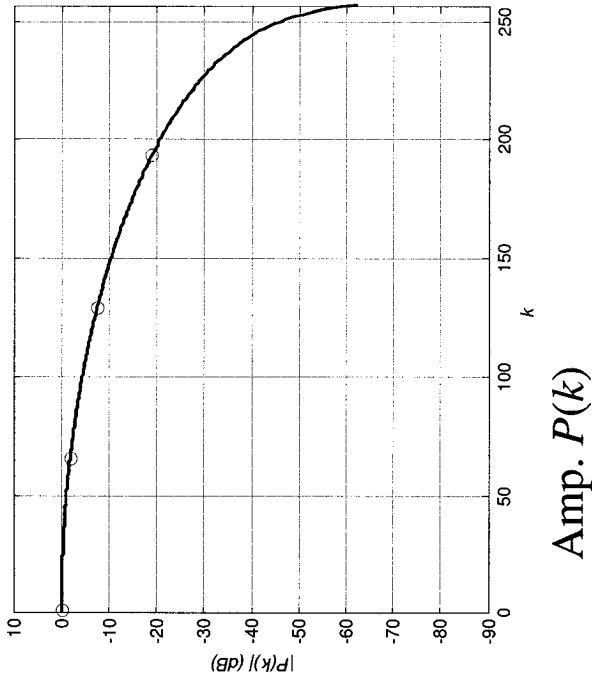


Fig. 6a



$$W_M(k)$$

Fig. 7a



$$\text{Amp. } P(k)$$

Fig. 7b

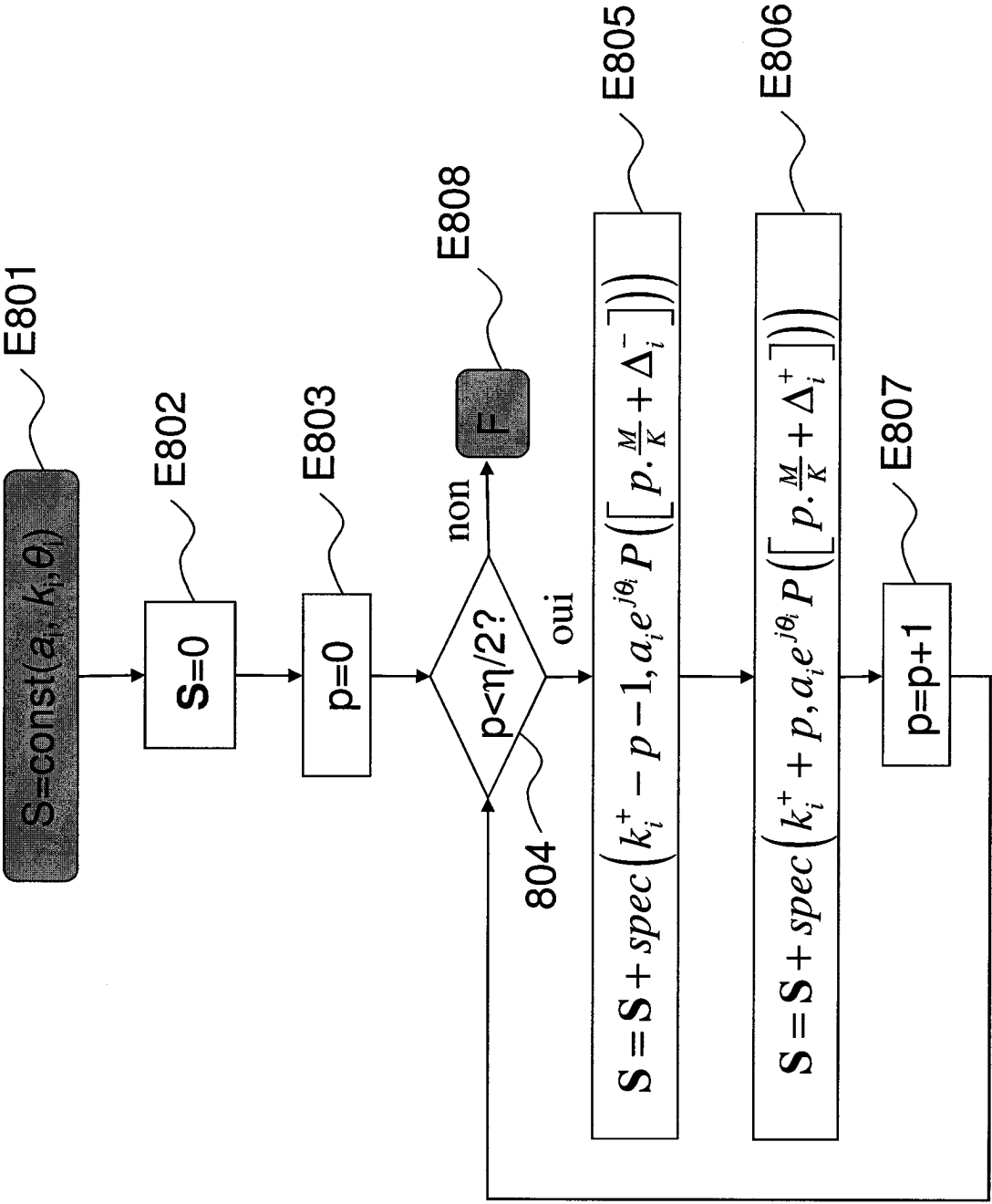


Fig. 8

9/13

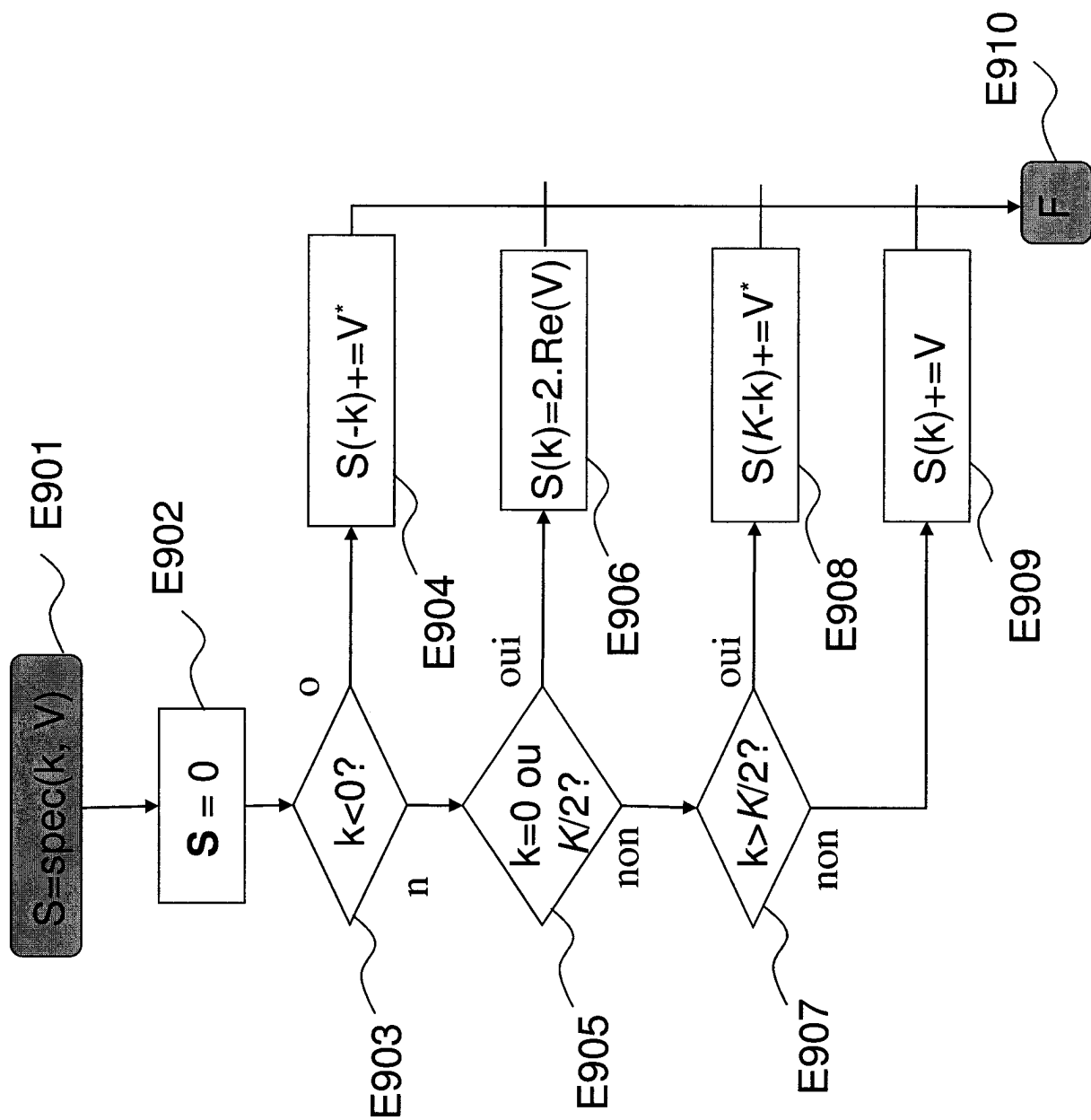


Fig. 9

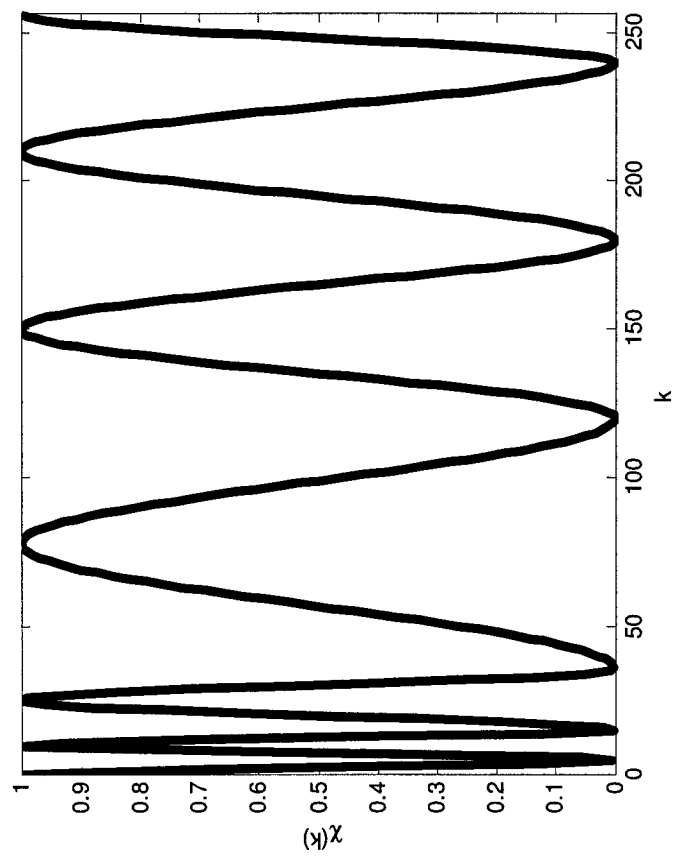


Fig. 10a

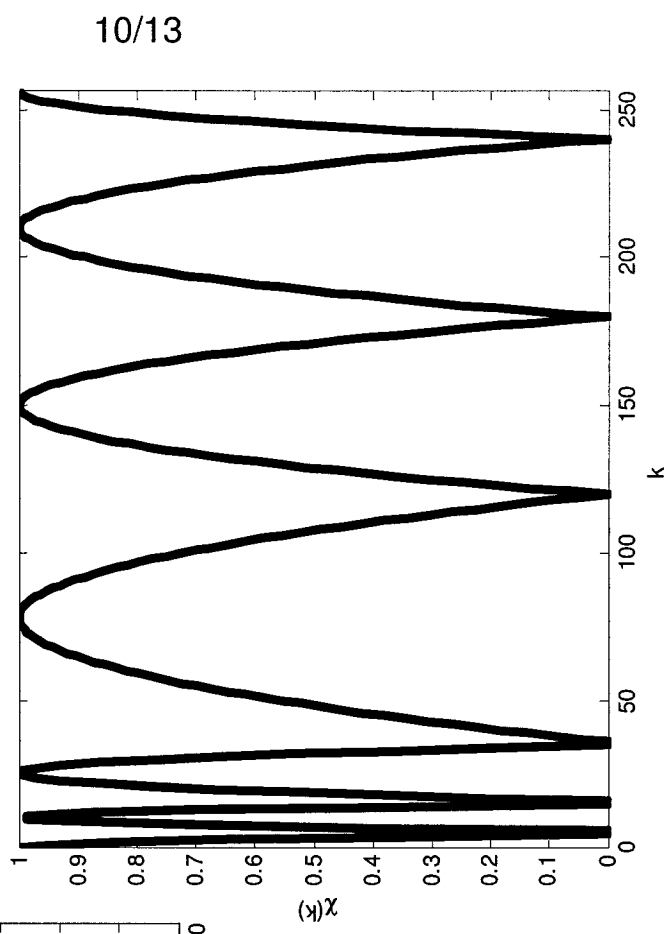


Fig. 10b

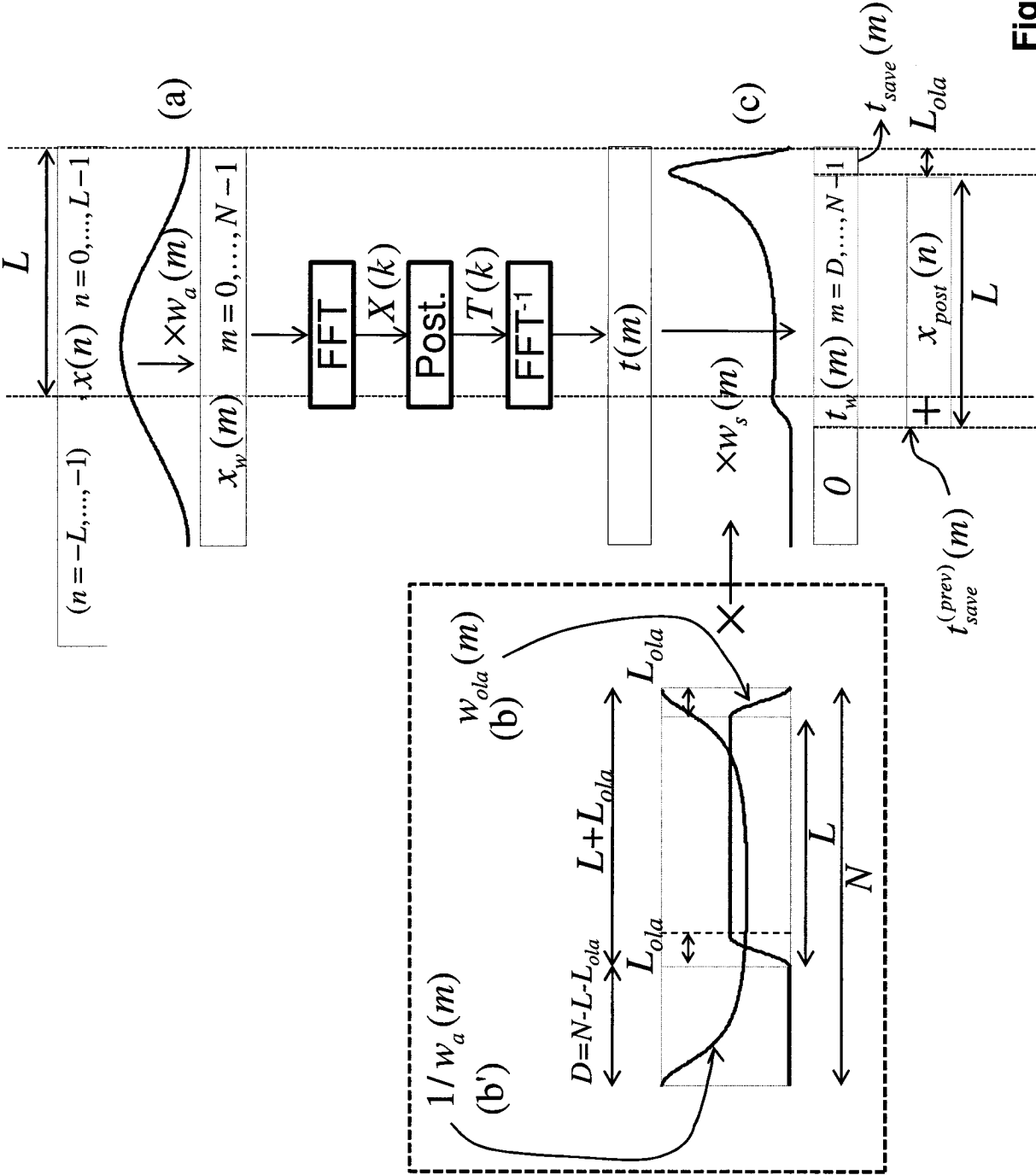


Fig. 11

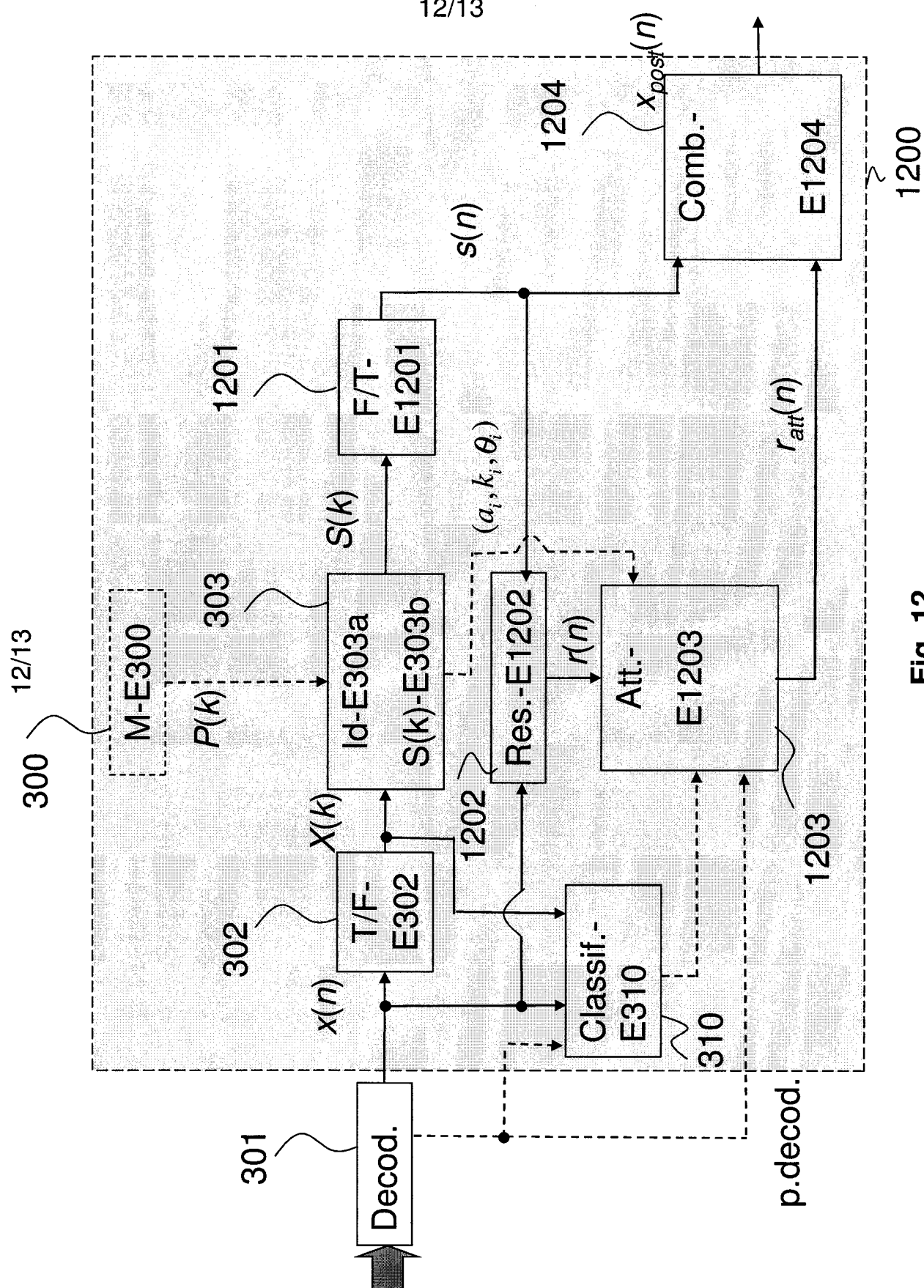


Fig. 12

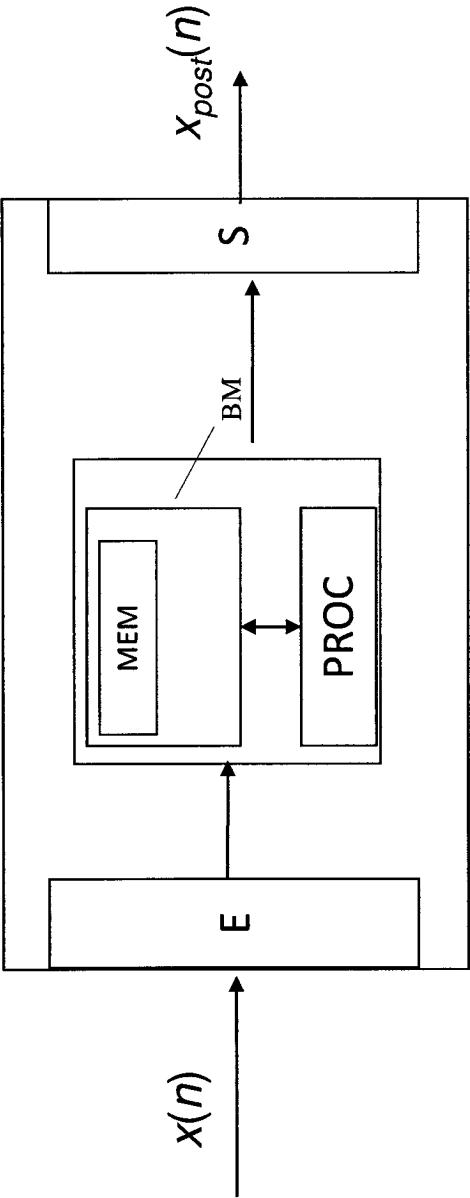


Fig. 13



RAPPORT DE RECHERCHE PRÉLIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 757211
FR 1158521

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	EP 1 744 305 A2 (SAMSUNG ELECTRONICS CO LTD [KR]) 17 janvier 2007 (2007-01-17) * alinéa [0001] * * alinéa [0025] * * page 4, ligne 33-53 * * page 6, ligne 7-34 * -----	1-11	G10L19/02
X	HARDWICK J ET AL: "Speech enhancement using the dual excitation speech model", 1993 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1993. ICASSP-93; [PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP)], PISCATAWAY, NJ, USA, vol. 2, 27 avril 1993 (1993-04-27), pages 367-370, XP010110470, DOI: 10.1109/ICASSP.1993.319314 ISBN: 978-0-7803-0946-3 * pages II-367, colonne de droite, dernier alinéa - pages II-368, colonne de gauche * * pages II-369, colonne de droite * -----	1-11	DOMAINES TECHNIQUES RECHERCHÉS (IPC)
X	EP 1 511 011 A2 (MICROSOFT CORP [US]) 2 mars 2005 (2005-03-02) * alinéas [0030] - [0043] * ----- -/--	1-11	G10L
Date d'achèvement de la recherche		Examineur	
22 février 2012		Ramos Sánchez, U	
CATÉGORIE DES DOCUMENTS CITÉS		T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire			



RAPPORT DE RECHERCHE PRÉLIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 757211
FR 1158521

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
A,D	JULIUS O SMITH III AND XAVIER SERRA: "PARSHL: An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation", PROCEEDINGS OF THE 1987 INTERNATIONAL COMPUTER MUSIC CONFERENCE, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, USA, AUGUST 23-26, 1987,, 1 août 1987 (1987-08-01), pages 290-297, XP009130237, * page 293, colonne de gauche - colonne de droite, alinéa 2 *	4	
A	US 2005/283361 A1 (YOSHII KAZUYOSHI [JP] ET AL) 22 décembre 2005 (2005-12-22) * alinéas [0074], [0075], [0077], [0080] * * alinéas [0100] - [0101] * * alinéa [0107] * * alinéas [0109] - [0112] *	5,6	
			DOMAINES TECHNIQUES RECHERCHÉS (IPC)
Date d'achèvement de la recherche		Examineur	
22 février 2012		Ramos Sánchez, U	
CATÉGORIE DES DOCUMENTS CITÉS X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant			

ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 1158521 FA 757211

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.

Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du **22-02-2012**

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
EP 1744305 A2	17-01-2007	EP 1744305 A2	17-01-2007
		US 2007010997 A1	11-01-2007

EP 1511011 A2	02-03-2005	AT 347162 T	15-12-2006
		CN 1591574 A	09-03-2005
		DE 602004003439 T2	29-03-2007
		EP 1511011 A2	02-03-2005
		JP 4731855 B2	27-07-2011
		JP 2005070779 A	17-03-2005
		KR 20050022371 A	07-03-2005
		US 2005049857 A1	03-03-2005

US 2005283361 A1	22-12-2005	JP 4318119 B2	19-08-2009
		JP 2006005807 A	05-01-2006
		US 2005283361 A1	22-12-2005
