



(11) **EP 3 014 611 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention
de la délivrance du brevet:
13.03.2019 Bulletin 2019/11

(51) Int Cl.:
G10L 21/038 ^(2013.01) **G10L 19/06** ^(2013.01)
G10L 19/08 ^(2013.01)

(21) Numéro de dépôt: **14742262.0**

(86) Numéro de dépôt international:
PCT/FR2014/051563

(22) Date de dépôt: **24.06.2014**

(87) Numéro de publication internationale:
WO 2014/207362 (31.12.2014 Gazette 2014/53)

(54) **EXTENSION AMÉLIORÉE DE BANDE DE FRÉQUENCE DANS UN DÉCODEUR DE SIGNAUX AUDIOFRÉQUENCES**

VERBESSERTE FREQUENZBANDERWEITERUNG IN EINEM AUDIOSIGNALDECODIERER
IMPROVED FREQUENCY BAND EXTENSION IN AN AUDIO SIGNAL DECODER

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

(30) Priorité: **25.06.2013 FR 1356100**

(43) Date de publication de la demande:
04.05.2016 Bulletin 2016/18

(73) Titulaire: **Orange**
75015 Paris (FR)

(72) Inventeurs:
• **KANIEWSKA, Magdalena**
3001 Leuven (BE)

• **RAGOT, Stéphane**
Lannion 22300 (FR)

(56) Documents cités:
WO-A2-2013/066238 US-A1- 2003 050 786

- **WOLTERS M ET AL: "A CLOSER LOOK INTO MPEG-4 HIGH EFFICIENCY AAC", PREPRINTS OF PAPERS PRESENTED AT THE AES CONVENTION, XX, XX, vol. 115, 10 octobre 2003 (2003-10-10), XP008063876,**
- **NEUENDORF MAX ET AL: "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types", AES CONVENTION 132; APRIL 2012, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 26 avril 2012 (2012-04-26), XP040574618,**

EP 3 014 611 B1

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

Description

[0001] La présente invention se rapporte au domaine du codage/décodage et du traitement de signaux audiofréquences (comme des signaux de parole, de musique ou autres) pour leur transmission ou leur stockage.

[0002] Plus particulièrement, l'invention concerne un procédé et un dispositif d'extension de bande de fréquence dans un décodeur ou un processeur réalisant une amélioration de signal audiofréquence.

[0003] De nombreuses techniques existent pour compresser (avec perte) un signal audiofréquence comme la parole ou la musique.

[0004] Les méthodes classiques de codage pour les applications conversationnelles sont en général classifiées en *codage de forme d'onde* (MIC pour "Modulation par Impulsion et codage", MICDA pour "Modulation par Impulsion et Codage Différentiel Adaptatif", codage par transformée...), *codage paramétrique* (LPC pour "Linear Predictive Coding" en anglais, codage sinusoïdal...) et codage hybride paramétrique avec une quantification des paramètres par "analyse par synthèse" dont le codage CELP (pour "Code Excited Linear Prediction" en anglais) est l'exemple le plus connu.

[0005] Pour les applications non conversationnelles, l'état de l'art en codage de signal audio (mono) est constitué par le codage perceptuel par transformée ou en sous-bandes, avec un codage paramétrique des hautes fréquences par réplication de bande.

Une revue des méthodes classiques de codage de parole et audio se trouve dans les ouvrages W.B. Kleijn and K.K. Paliwal (Eds.), *Speech Coding and Synthesis*, Elsevier, 1995 ; M. Bosi, R.E. Goldberg, *Introduction to Digital Audio Coding and Standards*, Springer 2002 ; J. Benesty, M.M. Sondhi, Y. Huang (Eds.), *Handbook of Speech Processing*, Springer 2008.

[0006] On s'intéresse ici plus particulièrement au codec (codeur et décodeur) normalisé 3GPP AMR-WB (pour "Adaptive Multi-Rate Wideband" en anglais) qui fonctionne à une fréquence d'entrée/sortie de 16 kHz et dans lequel le signal est divisé en deux sous-bandes, la bande basse (0-6.4 kHz) qui est échantillonnée à 12.8 kHz et codée par modèle CELP et la bande haute (6.4-7 kHz) qui est reconstruite de façon paramétrique par « *extension de bande* » (ou BWE pour "Bandwidth Extension" en anglais) avec ou sans information supplémentaire selon le mode de la trame courante. On peut noter ici que la limitation de la bande codée du codec AMR-WB à 7kHz est essentiellement liée au fait que la réponse en fréquence en émission des terminaux en bande élargie a été approximée au moment de la normalisation (ETSI/3GPP puis UIT-T) selon le masque fréquentiel défini dans la norme UIT-T P.341 et plus précisément en utilisant un filtre dit « P341 » défini dans la norme UIT-T G.191 qui coupe les fréquences au-dessus de 7 kHz (ce filtre respecte le masque défini dans P.341). Cependant, en théorie, il est bien connu qu'un signal échantillonné à 16 kHz peut avoir une bande audio définie de 0 à 8000 Hz ; le codec AMR-WB introduit donc une limitation de la bande haute en comparaison à la largeur de bande théorique de 8 kHz.

[0007] Le codec de parole 3GPP AMR-WB a été normalisé en 2001 principalement pour les applications de téléphonie en mode circuit (CS) sur GSM (2G) et UMTS (3G). Ce même codec a été aussi normalisé en 2003 à l'UIT-T en tant que recommandation G.722.2 "Wideband coding speech at around 16kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)".

[0008] Il comprend neuf débits, appelés modes, de 6.6 à 23.85 kbit/s, et comprend des mécanismes de transmission continue (DTX pour "Discontinuous Transmission") avec détection d'activité vocale (VAD pour "Voice Activity Detection") et génération de bruit de confort (CNG pour "Confort Noise Generation") à partir de trames de description de silence (SID pour "Silence Insertion Descriptor"), ainsi que des mécanismes de correction de trames perdues (FEC pour "Frame Erasure Concealment", parfois appelé PLC pour "Packet Loss Concealment").

[0009] On ne reprend pas ici les détails de l'algorithme de codage et de décodage AMR-WB, une description détaillée de ce codec se trouve dans les spécifications 3GPP (TS 26.190, 26.191, 26.192, 26.193, 26.194, 26.204) et UIT-T-G.722.2 (et les Annexes et Appendice correspondantes) ainsi que dans l'article de B. Bessette et al. intitulé « The adaptive multirate wideband speech codec (AMR-WB) », IEEE Transactions on Speech and Audio Processing, vol. 10, no. 8, 2002, pp. 620-636 et les codes sources des standards 3GPP et UIT-T associés.

[0010] Le principe de l'extension de bande dans le codec AMR-WB est assez rudimentaire. En effet, la bande haute (6.4-7 kHz) est générée en mettant en forme un bruit blanc par le biais d'une enveloppe temporelle (appliquée sous la forme de gains par sous-trame) et fréquentielle (par l'application d'un filtre de synthèse de prédiction linéaire ou LPC pour "Linear Predictive Coding"). Cette technique d'extension de bande est illustrée à la **figure 1**.

[0011] Un bruit blanc, $u_{HB1}(n)$, $n = 0, \dots, 79$, est généré à 16 kHz par sous-trame de 5 ms par générateur congruentiel linéaire (bloc 100). Ce bruit $u_{HB1}(n)$ est mis en forme dans le temps par application de gains par sous-trame ; cette opération est décomposée en deux étapes de traitement (blocs 102, 106 ou 109) :

- Un premier facteur est calculé (bloc 101) pour mettre le bruit blanc $u_{HB1}(n)$ (bloc 102) à un niveau semblable à celui de l'excitation, $u(n)$, $n = 0, \dots, 63$, décodée à 12.8 kHz dans la bande basse :

$$u_{HB2}(n) = u_{HB1}(n) \sqrt{\frac{\sum_{l=0}^{63} u(l)^2}{\sum_{l=0}^{79} u_{HB1}(l)^2}}$$

[0012] On peut noter ici que la normalisation des énergies se fait en comparant des blocs de taille différente (64 pour $u(n)$ et 80 pour $u_{HB1}(n)$), sans compensation des différences de fréquences d'échantillonnage (12.8 ou 16 kHz).

- L'excitation dans la bande haute est ensuite obtenue (bloc 106 ou 109) sous la forme :

$$u_{HB}(n) = \hat{g}_{HB} u_{HB2}(n)$$

où le gain \hat{g}_{HB} est obtenu différemment selon le débit. Si le débit de la trame actuelle est <23.85 kbit/s, le gain \hat{g}_{HB} est estimé « en aveugle » (c'est-à-dire sans information supplémentaire); dans ce cas, le bloc 103 filtre le signal décodé en bande basse par un filtre passe-haut ayant une fréquence de coupure à 400 Hz pour obtenir un signal $\hat{s}_{hp}(n)$, $n = 0, \dots, 63$ - ce filtre passe-haut élimine l'influence des très basses fréquences qui peuvent biaiser l'estimation faite dans le bloc 104 - puis on calcule le «tilt» (indicateur de pente spectrale) noté e_{tilt} du signal $\hat{s}_{hp}(n)$ par autocorrélation normalisée (bloc 104):

$$e_{tilt} = \frac{\sum_{n=1}^{63} \hat{s}_{hp}(n) \hat{s}_{hp}(n-1)}{\sum_{n=0}^{63} \hat{s}_{hp}(n)^2}$$

et enfin on calcule \hat{g}_{HB} sous la forme :

$$\hat{g}_{HB} = w_{SP} g_{SP} + (1 - w_{SP}) g_{BG}$$

où $g_{SP} = 1 - e_{tilt}$ est le gain appliqué dans les trames actives de parole (SP pour speech), $g_{BG} = 1.25 g_{SP}$ est le gain appliqué dans les trames inactives de parole associées à un bruit de fond (BG pour Background) et w_{SP} est une fonction de pondération qui dépend de la détection d'activité vocale (VAD). On comprend que l'estimation du tilt (e_{tilt}) permet d'adapter le niveau de la bande haute en fonction de la nature spectrale du signal ; cette estimation est particulièrement importante quand la pente spectrale du signal décodé CELP est telle que l'énergie moyenne décroît quand la fréquence augmente (cas d'un signal voisé où e_{tilt} est proche de 1, donc $g_{SP} = 1 - e_{tilt}$ est ainsi réduit). A noter aussi que le facteur \hat{g}_{HB} dans le décodage AMR-WB est borné pour prendre des valeurs dans l'intervalle [0.1, 1.0].

[0013] A 23.85 kbit/s, une information de correction est transmise par le codeur AMR-WB et décodée (blocs 107, 108) afin d'affiner le gain estimé par sous-trame (4 bits toutes les 5ms, soit 0.8 kbit/s).

[0014] L'excitation artificielle $u_{HB}(n)$ est ensuite filtrée (bloc 111) par un filtre de synthèse LPC (bloc 111) de fonction de transfert $1/A_{HB}(z)$ et fonctionnant à la fréquence d'échantillonnage de 16 kHz. La réalisation de ce filtre dépend du débit de la trame courante:

- A 6.6 kbit/s, le filtre $1/A_{HB}(z)$ est obtenu en pondérant par un facteur $\gamma=0.9$ un filtre LPC d'ordre 20, $1/\hat{A}^{ext}(z)$ qui « extrapole » le filtre LPC d'ordre 16, $1/\hat{A}(z)$, décodé dans la bande basse (à 12.8 kHz) - les détails de l'extrapolation dans le domaine des paramètres ISF (pour "Imittance Spectral Frequency" en anglais) sont décrits dans la norme G.722.2 à la section 6.3.2.1; dans ce cas,

$$1/A_{HB}(z) = 1/\hat{A}^{ext}(z/\gamma)$$

- Aux débits > 6.6 kbit/s, le filtre $1/A_{HB}(z)$ est d'ordre 16 et correspond simplement à :

$$1/A_{HB}(z) = 1/\hat{A}(z/\gamma)$$

où $\gamma=0.6$. A noter que dans ce cas le filtre $1/\hat{A}(z/\gamma)$ est utilisé à 16 kHz, ce qui résulte en un étalement (par homothétie) de la réponse en fréquence de ce filtre de [0, 6.4 kHz] à [0, 8 kHz].

[0015] Le résultat, $s_{HB}(n)$, est enfin traité par un filtre passe-bande (bloc 112) de type FIR ("Finite Impulse Response"), pour ne garder que la bande 6 - 7 kHz ; à 23.85 kbit/s, un filtre passe-bas également de type FIR (bloc 113) se rajoute au traitement pour atténuer encore plus les fréquences supérieures à 7 kHz. La synthèse en hautes fréquences (HF) est finalement additionnée (bloc 130) à la synthèse en basses fréquences (BF) obtenue avec les blocs 120 à 123 et ré-échantillonnée à 16 kHz (bloc 123). Ainsi même si la bande haute s'étend en théorie de 6.4 à 7 kHz dans le codec AMR-WB, la synthèse HF est plutôt comprise dans la bande 6-7 kHz avant addition avec la synthèse BF.

[0016] On peut identifier plusieurs inconvénients à la technique d'extension de bande du codec AMR-WB :

- Le signal dans la bande haute est un bruit blanc mis en forme (par gains temporels par sous-trame, par filtrage par $1/A_{HB}(z)$ et filtrage passe-bande), ce qui n'est pas un bon modèle général du signal dans la bande 6.4-7 kHz. Il existe par exemple des signaux de musique très harmoniques pour lesquels la bande 6.4-7 kHz contient des composantes sinusoïdales (ou tones) et aucun bruit (ou peu de bruit), pour ces signaux l'extension de bande du codec AMR-WB dégrade fortement la qualité.
- Le filtre passe-bas à 7 kHz (bloc 113) introduit un décalage de près de 1 ms entre les bandes basses et hautes, ce qui peut potentiellement dégrader la qualité de certains signaux en désynchronisant légèrement les deux bandes à 23.85 kbit/s - cette désynchronisation peut également poser problème lors d'une commutation de débit de 23.85 kbit/s à d'autres modes.
- L'estimation de gains par sous-trame (bloc 101, 103 à 105) n'est pas optimale. Pour partie, elle se base sur une égalisation de l'énergie « absolue » par sous-trame (bloc 101) entre des signaux à des fréquences différentes : l'excitation artificielle à 16 kHz (bruit blanc) et un signal à 12.8 kHz (excitation ACELP décodée). On peut noter en particulier que cette approche induit implicitement une atténuation de l'excitation bande haute (par un ratio $12.8/16=0.8$) ; en fait, on notera également qu'aucune désaccentuation (ou déemphasis) n'est effectuée sur la bande haute dans le codec AMR-WB, ce qui induit implicitement une amplification relative proche de 0.6 (qui correspond à la valeur de la réponse en fréquence de $1/(1-0.68z^{-1})$ à 6400 Hz).

[0017] En fait, les facteurs de 1/0.8 et de 0.6 se compensent approximativement.

- Sur la parole, les tests de caractérisation du codec 3GPP AMR-WB documentés dans le rapport 3GPP TR 26.976 ont montré que le mode à 23.85 kbit/s a une qualité moins bonne qu'à 23.05 kbit/s, sa qualité est en fait similaire à celle du mode à 15.85 kbit/s. Ceci montre en particulier que le niveau du signal HF artificiel doit être contrôlé de façon très prudente, car la qualité est dégradée à 23.85 kbit/s alors que les 4 bits par trame sont sensés permettre de mieux approcher l'énergie des hautes fréquences originales.
- La limitation de la bande codée à 7 kHz résulte de l'application d'un modèle strict de la réponse en émission des terminaux acoustiques (filtre P.341 dans la norme UIT-T G.191). Or, pour une fréquence d'échantillonnage de 16 kHz, les fréquences dans la bande 7-8 kHz restent importantes, en particulier pour les signaux de musique, pour assurer un bon niveau de qualité.

[0018] L'algorithme de décodage AMR-WB a été amélioré en partie avec le développement du codec scalable UIT-T G.718 qui a été normalisé en 2008.

[0019] La norme UIT-T G.718 comprend un mode dit interopérable, pour lequel le codage coeur est compatible avec le codage G.722.2 (AMR-WB) à 12.65 kbit/s ; de plus, le décodeur G.718 a la particularité de pouvoir décoder un train binaire AMR-WB/G.722.2 à tous les débits possibles du codec AMR-WB (de 6.6 à 23.85 kbit/s).

[0020] Le décodeur interopérable G.718 en mode bas délai (« low delay » en anglais) (G.718-LD) est illustré à la figure 2. On liste ci-dessous les améliorations apportées à la fonctionnalité de décodage de train binaire AMR-WB dans le décodeur G.718, avec des références à la figure 1 lorsque nécessaire :

- L'extension de bande (décrite par exemple dans la clause 7.13.1 de la recommandation G.718, bloc 206) est identique à celle du décodeur AMR-WB, sauf que le filtre passe-bande 6-7 kHz et le filtre de synthèse $1/A_{HB}(z)$ (blocs 111 et 112) sont en ordre inversé. De plus, à 23.85 kbit/s les 4 bits transmis par sous-frames par le codeur AMR-WB ne sont pas utilisés dans le décodeur G.718 interopérable ; la synthèse des hautes fréquences (HF) à 23.85 kbit/s est donc identique à 23.05 kbit/s ce qui évite le problème connu de qualité du décodage AMR-WB à

23.85 kbit/s. A fortiori, le filtre passe-bas à 7 kHz (bloc 113) n'est pas utilisé, et le décodage spécifique du mode à 23.85 kbit/s est omis (blocs 107 à 109).

- Un post-traitement de la synthèse à 16 kHz (voir clause 7.14 de G.718) est mis en oeuvre dans G.718 par "noise gate" dans le bloc 208 (pour « améliorer » la qualité des silences par réduction du niveau), filtrage passe-haut (bloc 209), post-filtre de basses fréquences (dit « *bass postfilter* ») dans le bloc 210 atténuant le bruit inter-harmonique en basses fréquences et une conversion en entiers 16 bits avec contrôle de saturation (avec contrôle de gain ou AGC) dans le bloc 211.

[0021] Cependant l'extension de bande dans les codecs AMR-WB et/ou G.718 (mode interopérable) reste encore limitée sur plusieurs aspects.

- En particulier, la synthèse de hautes fréquences par bruit blanc mis en forme (par une approche temporelle de type source-filtre LPC) est un modèle très limité du signal dans la bande des fréquences supérieures à 6.4 kHz.
- Seule la bande 6.4-7 kHz est re-synthétisée de façon artificielle, alors qu'en pratique une bande plus large (jusqu'à 8 kHz) est en théorie possible à la fréquence d'échantillonnage de 16 kHz, ce qui peut potentiellement améliorer la qualité des signaux, s'ils ne sont pas pré-traités par un filtre de type P.341 (50-7000 Hz) tel que définie dans la *Software Tool Library* (norme G.191) de l'UIT-T.

[0022] Un autre exemple d'extension de bande en aveugle est divulgué par l'art antérieur US 2003/050786 A1 qui comprend une analyse LPC du signal à bande étroite, l'utilisation de ces coefficients afin d'obtenir une enveloppe du signal en bande élargie, l'extension du signal d'excitation en bande étroite filtré, un filtrage permettant d'obtenir un signal en bande haute qui est finalement combiné avec le signal en bande étroite pour fournir un signal de bande élargie.

[0023] Il existe donc un besoin pour améliorer l'extension de bande dans un codec de type AMR-WB ou une version interopérable de ce codec ou plus généralement pour améliorer l'extension de bande d'un signal audio.

[0024] La présente invention vient améliorer la situation.

[0025] L'invention propose à cet effet, un procédé d'extension de bande de fréquence d'un signal audiofréquence lors d'un processus de décodage ou d'amélioration comportant une étape de décodage ou d'extraction, dans une première bande de fréquence dite bande basse, d'un signal d'excitation et des coefficients d'un filtre de prédiction linéaire. Le procédé est tel qu'il comporte les étapes suivantes:

- obtention d'un signal étendu dans au moins une deuxième bande de fréquence supérieure à la première bande de fréquence à partir du signal d'excitation sur-échantillonné et étendu dans la au moins une deuxième bande de fréquence;
- mise à l'échelle du signal étendu par un gain défini par sous-trame en fonction d'un ratio entre le ratio entre l'énergie par sous-trame et l'énergie par trame du signal d'excitation de la bande basse et le ratio entre l'énergie par sous-trame et l'énergie par trame du signal étendu;
- filtrage dudit signal étendu mis à l'échelle par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse.

[0026] Ainsi, la prise en compte du signal d'excitation (issu du décodage de la bande basse ou d'une extraction du signal en bande basse) permet d'effectuer l'extension de bande avec un modèle de signal plus adapté pour certains types de signaux comme les signaux de musique.

[0027] En effet, le signal d'excitation décodé ou estimé dans la bande basse comporte dans certains cas des harmoniques, qui quand elles existent, peuvent être transposées en haute fréquence de telle sorte que cela permet d'assurer un certain niveau d'harmonicité dans la bande haute reconstruite.

[0028] L'extension de bande selon le procédé permet donc d'améliorer la qualité pour ce type de signaux.

[0029] De plus, l'extension de bande selon le procédé est réalisée en étendant d'abord un signal d'excitation et en appliquant ensuite une étape de filtrage de synthèse ; cette approche exploite le fait que l'excitation décodée dans la bande basse est un signal dont le spectre est relativement plat, ce qui évite les traitements de blanchiment du signal décodé qui peuvent exister dans les méthodes connues d'extension de bande dans le domaine fréquentiel dans l'état de l'art.

[0030] On remarquera que même si l'invention est motivée par l'amélioration de la qualité de l'extension de bande dans le contexte du codage AMR-WB interopérable, les différents modes de réalisation s'appliquent au cas plus général de l'extension de bande d'un signal audio, en particulier dans un dispositif d'amélioration effectuant une analyse du signal audio pour extraire les paramètres nécessaires à l'extension de bande.

[0031] Le fait de prendre en compte l'énergie au niveau de la trame courante et celle de la sous-trame dans le signal en bande basse (première bande de fréquence) permet d'ajuster le ratio entre l'énergie par sous-trame et l'énergie par trame dans la bande haute (deuxième bande de fréquence) et d'ajuster ainsi des ratios d'énergie plutôt que des énergies

absolues. Cela permet de garder dans la bande haute le même ratio d'énergie entre sous-trame et trame que dans la bande basse, ce qui est particulièrement bénéfique lorsque l'énergie des sous-frames varie beaucoup, par exemple dans le cas de sons transitoires, d'attaques.

[0032] 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'extension défini ci-dessus.

[0033] Dans un mode de réalisation, le procédé comporte en outre une étape de filtrage passe-bande adaptatif en fonction du débit de décodage de la trame courante.

[0034] Ce filtrage adaptatif permet d'optimiser la largeur de bande étendue en fonction du débit, et donc la qualité du signal reconstruit après extension de bande. En effet, pour les bas débits (typiquement à 6.6 et 8.85 kbit/s pour AMR-WB), la qualité générale du signal décodé en bande basse (par le codec AMR-WB ou une version interopérable) n'étant pas très bonne, il est préférable de ne pas trop étendre la bande décodée et donc de limiter l'extension de bande en adaptant la réponse en fréquence du filtre passe-bande associé pour couvrir par exemple une bande approximative de 6 à 7 kHz ; cette limitation est d'autant plus avantageuse que le signal d'excitation lui-même est relativement mal codé et il est préférable de ne pas en utiliser une sous-bande trop large pour l'extension des hautes fréquences. A l'opposé, pour les débits supérieurs (12.65 kbit/s et supérieurs pour AMR-WB), la qualité peut être améliorée avec une synthèse HF couvrant une bande plus large, par exemple approximativement de 6 à 7.7 kHz. La limite haute de 7.7 kHz (au lieu de 8 kHz) est un exemple de réalisation, qui pourra être ajustée à des valeurs proches de 7.7 kHz. Cette limite est ici justifiée par le fait que l'extension est faite dans l'invention sans information auxiliaire et qu'une extension jusqu'à 8 kHz (même si elle est en théorie possible) pourrait résulter en des artefacts pour des signaux particuliers. De plus, cette limitation à 7.7 kHz tient compte du fait que typiquement les filtres anti-repliement en conversion analogique/numérique et les filtres de ré-échantillonnage entre 16 kHz et d'autres fréquences ne sont pas parfaits et qu'ils introduisent typiquement une réjection aux fréquences inférieures à 8 kHz.

[0035] Dans un mode possible de réalisation le procédé comporte une étape de transformée temps-fréquence du signal d'excitation, l'étape d'obtention d'un signal étendu s'effectuant alors dans le domaine fréquentiel et une étape de transformée temps-fréquence inverse du signal étendu avant les étapes de mise à l'échelle et de filtrage.

[0036] La mise en oeuvre de l'extension de bande (du signal d'excitation) dans le domaine fréquentiel permet d'obtenir une finesse d'analyse en fréquence dont on ne dispose pas avec une approche temporelle, et permet aussi d'avoir une résolution fréquentielle suffisante pour détecter des harmoniques et transposer en hautes fréquences des harmoniques du signal (dans la bande basse) pour améliorer la qualité tout en respectant la structure du signal.

[0037] Dans un mode de réalisation détaillé, l'étape de génération d'un signal d'excitation sur-échantillonné et étendu est effectuée selon l'équation suivante:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

avec k l'indice de l'échantillon, $U_{HB1}(k)$ le spectre du signal d'excitation étendu, $U(k)$ le spectre du signal d'excitation obtenu après l'étape de transformée et *start_band* une variable prédéfinie.

[0038] Ainsi, cette fonction comprend bien un ré-échantillonnage du signal d'excitation en ajoutant des échantillons au spectre de ce signal.

[0039] Dans la bande de fréquence correspondant aux échantillons allant de 200 à 239, le spectre original est conservé, pour pouvoir y appliquer une réponse d'atténuation progressive du filtre passe-haut dans cette bande de fréquence et aussi pour ne pas introduire de défauts 9 audibles lors de l'étape d'addition de la synthèse basse fréquence à la synthèse haute fréquence.

[0040] Dans un mode de réalisation particulier, le procédé comporte une étape de filtrage de désaccentuation du signal étendu au moins dans la deuxième bande de fréquence.

[0041] Ainsi, le signal dans la deuxième bande de fréquence est ramené dans un domaine cohérent avec le signal dans la première bande de fréquence.

[0042] Dans un mode particulier de réalisation, le procédé comporte en outre une étape de génération d'un signal de bruit au moins dans la deuxième bande de fréquence le signal étendu étant obtenu par combinaison du signal d'excitation étendu et du signal de bruit.

[0043] En effet, il suffit d'avoir des caractéristiques issues du signal d'excitation sur-échantillonné et étendu dans au moins une deuxième bande de fréquence pour avoir un modèle de signal adapté à certains types de signaux. Celui-ci peut être combiné à un autre signal, par exemple un bruit généré pour obtenir le signal étendu ayant un modèle de signal adapté.

[0044] Dans un mode de réalisation l'étape de combinaison s'effectue par mixage additif adaptatif avec un gain

d'égalisation de niveau entre le signal d'excitation étendu et le signal de bruit.

[0045] L'application de ce gain d'égalisation permet à l'étape de combinaison de s'adapter aux caractéristiques du signal pour optimiser la proportion relative de bruit dans le mélange.

[0046] La présente invention vise également un dispositif d'extension de bande de fréquence d'un signal audiofréquence comportant un étage de décodage ou d'extraction, dans une première bande de fréquence dite bande basse, d'un signal d'excitation et des coefficients d'un filtre de prédiction linéaire. Le dispositif est tel qu'il comporte:

- un module d'obtention d'un signal étendu ($U_{HB2}(k)$, 503) dans au moins une deuxième bande de fréquence supérieure à la première bande de fréquence à partir du signal d'excitation sur-échantillonné et étendu dans la au moins une deuxième bande de fréquence ($U_{HB1}(k)$);
- mise à l'échelle (507) du signal étendu par un gain défini par sous-trame en fonction d'un ratio entre le ratio entre l'énergie par sous-trame et l'énergie par trame du signal d'excitation de la bande basse et le ratio entre l'énergie par sous-trame et l'énergie par trame du signal étendu;
- un module de filtrage (510) dudit signal étendu mis à l'échelle par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse.

[0047] Ce dispositif présente les mêmes avantages que le procédé décrit précédemment, qu'il met en oeuvre.

[0048] L'invention vise un décodeur comportant un dispositif tel que décrit.

[0049] Elle vise un programme informatique comportant des instructions de code pour la mise en oeuvre des étapes du procédé d'extension de bande tel que décrit, lorsque ces instructions sont exécutées par un processeur.

[0050] Enfin l'invention se rapporte à un support de stockage, lisible par un processeur, intégré ou non au dispositif d'extension de bande, éventuellement amovible, mémorisant un programme informatique mettant en oeuvre un procédé d'extension de bande tel que décrit précédemment.

[0051] 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 illustre une partie d'un décodeur de type AMR-WB mettant en oeuvre des étapes d'extension de bande de fréquence de l'état de l'art et tel que décrit précédemment;
- la figure 2 illustre un décodeur de type interopérable G.718-LD à 16kHz selon l'état de l'art et tel que décrit précédemment;
- la figure 3 illustre un décodeur interopérable avec le codage AMR-WB et intégrant un dispositif d'extension de bande selon un mode de réalisation de l'invention;
- la figure 4 illustre sous forme d'organigramme, les étapes principales d'un procédé d'extension de bande selon un mode de réalisation de l'invention;
- la figure 5 illustre un premier mode de réalisation dans le domaine fréquentiel d'un dispositif d'extension de bande selon l'invention;
- la figure 6 illustre un exemple de réponse en fréquence d'un filtre passe-bande utilisée dans un mode particulier de l'invention;
- la figure 7 illustre un deuxième mode de réalisation dans le domaine temporel d'un dispositif d'extension de bande selon l'invention; et
- la figure 8 illustre une réalisation matérielle d'un dispositif d'extension de bande selon l'invention.

[0052] La **figure 3** illustre un exemple de décodeur, compatible avec la norme AMR-WB/G.722.2 dans lequel on retrouve un post-traitement similaire à celui introduit dans G.718 et décrit en référence à la figure 2 et une extension de bande améliorée selon le procédé d'extension de l'invention, mis en oeuvre par le dispositif d'extension de bande illustré par le bloc 309.

[0053] Contrairement au décodage AMR-WB qui fonctionne avec une fréquence d'échantillonnage de sortie de 16 kHz et au décodage G.718 qui fonctionne à 8 ou 16 kHz, on considère ici un décodeur qui peut fonctionner avec un signal de sortie (synthèse) à la fréquence $f_s = 8, 16, 32$ ou 48 kHz. A noter qu'on suppose ici que le codage a été effectué selon l'algorithme AMR-WB avec une fréquence interne de 12.8 kHz pour le codage CELP en bande basse et à 23.85 kbit/s un codage de gain par sous-trame à la fréquence de 16 kHz ; même si l'invention est décrite ici au niveau du décodage, on suppose ici que le codage peut aussi fonctionner avec un signal d'entrée à la fréquence $f_s = 8, 16, 32$ ou 48 kHz et des opérations de ré-échantillonnage adéquate, dépassant le cadre de l'invention, sont mises en oeuvre au codage en fonction de la valeur de f_s . On peut noter que quand $f_s = 8$ kHz, dans le cas d'un décodage compatible avec AMR-WB, il n'est pas nécessaire d'étendre la bande basse 0-6.4 kHz, car la bande audio reconstruite à la fréquence f_s est limitée à 0-4000 Hz.

[0054] A la figure 3, le décodage CELP (BF pour basses fréquences) fonctionne toujours à la fréquence interne de 12.8 kHz, comme dans AMR-WB et G.718, et l'extension de bande (HF pour hautes fréquences) faisant l'objet de

l'invention fonctionne à la fréquence de 16 kHz, les synthèses BF et HF sont combinées (bloc 312) à la fréquence f_s après ré-échantillonnage adéquat (bloc 306 et traitement interne au bloc 311). Dans des variantes de l'invention, la combinaison des bandes basse et haute pourra se faire à 16 kHz, après avoir ré-échantillonnée la bande basse de 12.8 à 16 kHz, avant de ré-échantillonner le signal étendu à la fréquence f_s .

[0055] Le décodage selon la figure 3 dépend du mode (ou débit) AMR-WB associé à la trame courante reçue. A titre indicatif et sans que cela impacte le bloc 309, le décodage de la partie CELP en bande basse comporte les étapes suivantes:

- Démultiplexage des paramètres codés (bloc 300) en cas de trame correctement reçue ($bfi=0$ où bfi est le « *bad frame indicator* » valant 0 pour une trame reçue et 1 pour une trame perdue)
- Décodage des paramètres ISF avec interpolation et conversion en coefficients LPC (bloc 301) comme décrit dans la clause 6.1 de la norme G.722.2.
- Décodage de l'excitation CELP (bloc 302), avec une partie adaptative et fixe pour reconstruire l'excitation (exc ou $u'(n)$) dans chaque sous-trame de longueur 64 à 12.8 kHz:

$$u'(n) = \hat{g}_p v(n) + \hat{g}_c c(n), \quad n = 0, \dots, 63$$

en suivant les notations de la clause 7.1.2.1 de G.718 concernant le décodage CELP, où $v(n)$ et $c(n)$ sont respectivement les mots de code des dictionnaires adaptatif et fixe, et \hat{g}_p et \hat{g}_c sont les gains décodés associés. Cette excitation $u'(n)$ est utilisée dans le dictionnaire adaptatif de la sous-trame suivante ; elle est ensuite post-traitée et on distingue comme dans G.718 l'excitation $u'(n)$ (aussi notée exc) de sa version post-traitée modifiée $u(n)$ (aussi notée $exc2$) qui sert d'entrée au filtre de synthèse, $1/\hat{A}(z)$, dans le bloc 303. Dans des variantes qui peuvent être mises en oeuvre pour l'invention, les post-traitements appliqués à l'excitation peuvent être modifiés (par exemple, la dispersion de phase peut être améliorée) ou ces post-traitements peuvent être étendus (par exemple, une réduction du bruit inter-harmonique peut être mise en oeuvre), sans affecter la nature du procédé d'extension de bande selon l'invention.

- Filtrage de synthèse par $1/\hat{A}(z)$ (bloc 303) où le filtre LPC décodé $\hat{A}(z)$ est d'ordre 16
- Post-traitement bande étroite (bloc 304) selon la clause 7.3 de G.718 si $f_s=8$ kHz.
- Désaccentuation (bloc 305) par le filtre $1 / (1-0.68z^{-1})$
- Post-traitement des basses fréquences (bloc 306) tel que décrit à la clause 7.14.1.1 de G.718. Ce traitement introduit un retard qui est pris en compte dans le décodage de la bande haute (>6.4 kHz).
- Ré-échantillonnage de la fréquence interne de 12.8 kHz à la fréquence de sortie f_s (bloc 307). Plusieurs réalisations sont possibles. Sans perte de généralité, on considère ici à titre d'exemple que si $f_s=8$ ou 16 kHz, le ré-échantillonnage décrit dans la clause 7.6 de G.718 est repris ici, et si $f_s=32$ ou 48 kHz, des filtres à réponse impulsionnelle finie (FIR) supplémentaires sont utilisés.
- Calcul des paramètres du « *noise gate* » (bloc 308) qui est réalisé de façon préférentielle comme décrit dans la clause 7.14.3 de G.718.

[0056] On peut noter que l'utilisation des blocs 306, 308, 314 est optionnelle.

[0057] On notera également que le décodage de la bande basse décrit ci-dessus suppose une trame courante dite « active » avec un débit entre 6.6 et 23.85 kbit/s. En fait, quand le mode DTX (transmission continue en français) est activé, certaines trames peuvent être codées comme « inactives » et dans ce cas on peut soit transmettre un descripteur de silence (sur 35 bits) soit ne rien transmettre. En particulier, on rappelle que la trame SID décrit plusieurs paramètres : paramètres ISF moyennés sur 8 trames, énergie moyenne sur 8 trames, flag de dithering pour la reconstruction de bruit non stationnaire. Dans tous les cas, au décodeur, on retrouve le même modèle de décodage que pour une trame active, avec une reconstruction de l'excitation et d'un filtre LPC pour la trame courante, ce qui permet d'appliquer l'invention même sur des trames inactives. Le même constat s'applique pour le décodage de « trames perdues » (ou FEC, PLC) dans lequel le modèle LPC est appliqué.

[0058] Contrairement au décodage AMR-WB ou G.718, le décodeur selon l'invention permet d'étendre la bande basse décodée (50-6400 Hz en tenant en compte du filtrage passe-haut à 50 Hz au décodeur, 0-6400 Hz dans le cas général) à une bande étendue dont la largeur varie, allant approximativement de 50-6900 Hz à 50-7700 Hz en fonction du mode mis en oeuvre dans la trame courante. On peut ainsi parler d'une première bande de fréquence de 0 à 6400 Hz et d'une deuxième bande de fréquence de 6400 à 8000 Hz. En réalité, dans le mode de réalisation privilégié, l'extension de l'excitation est réalisée dans le domaine fréquentiel dans une bande de 5000 à 8000 Hz, pour permettre un filtrage passe-bande de largeur 6000 à 6900 ou 7700 Hz.

[0059] Dans un mode de réalisation privilégié, à 23.85 kbit/s, comme dans le décodeur G.718 décrit en référence à la figure 2, l'information de correction de gain HF (0.8 kbit/s) transmise à 23.85 kbit/s est ici ignorée. Ainsi à la figure 3,

aucun bloc spécifique à 23.85 kbit/s n'est utilisé.

[0060] La partie décodage bande haute est réalisée dans le bloc 309 représentant le dispositif d'extension de bande selon l'invention et qui est détaillé à la figure 5 dans un premier mode de réalisation et à la figure 7 dans un deuxième mode de réalisation.

[0061] Ce dispositif comprend au moins un module d'obtention d'un signal étendu dans au moins une deuxième bande de fréquence supérieure à la première bande de fréquence à partir d'un signal d'excitation sur-échantillonné et étendu dans au moins une deuxième bande de fréquence ($U_{HB1}(k)$), un module de mise à l'échelle du signal étendu par un gain défini par sous-trame en fonction d'un ratio d'énergie par trame et sous-trame du signal audiofréquence dans la première bande de fréquence et un module de filtrage dudit signal étendu mis à l'échelle par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse.

[0062] Afin d'aligner les bandes basses et hautes décodées, un retard (bloc 310) est introduit dans le premier mode de réalisation pour synchroniser les sorties des blocs 306 et 307 et la bande haute synthétisée à 16 kHz est ré-échantillonnée de 16 kHz à la fréquence f_s (sortie de bloc 311). Par exemple, quand $f_s=16$ kHz le retard $T=30$ échantillons, qui correspond au retard de ré-échantillonnage de 12.8 à 16 kHz de 15 échantillons + retard du post-traitement des basses fréquences de 15 échantillons. La valeur du retard T devra être adaptée pour les autres cas ($f_s=32, 48$ kHz) en fonction des traitements mis en oeuvre. On rappelle que quand $f_s=8$ kHz, il n'est pas nécessaire d'appliquer les blocs 309 à 311 car la bande du signal en sortie du décodeur est limitée à 0-4000 Hz.

[0063] A noter que le procédé d'extension de l'invention mis en oeuvre dans le bloc 309 selon le premier mode de réalisation n'introduit de façon préférentielle aucun retard supplémentaire par rapport à la bande basse reconstruite à 12.8 kHz ; cependant, dans des variantes de l'invention (par exemple en utilisant une transformation temps/fréquence avec recouvrement), un retard pourra être introduit. Ainsi, de façon générale la valeur de T dans le bloc 310 devra être ajustée en fonction de la mise en oeuvre spécifique. Par exemple dans le cas où le post-traitement des basses fréquences (bloc 306) n'est pas utilisé, le retard à introduire pour $f_s=16$ kHz pourra être fixé à $T=15$ échantillons ; de même si l'invention est réalisée suivant la variante du mode de réalisation décrite à la figure 7, la valeur de T se réduit pour compenser le retard introduit par le post-traitement des basses fréquences (bloc 306) s'il est utilisé.

[0064] Les bandes basse et haute sont ensuite combinées (ajoutées) dans le bloc 312 et la synthèse obtenue est post-traitée par filtrage passe-haut à 50 Hz (de type IIR) d'ordre 2 dont les coefficients dépendent de la fréquence f_s (bloc 313) et post-traitement de sortie avec application optionnelle du "noise gate" de façon similaire à G.718 (bloc 314).

[0065] Le dispositif d'extension de bande selon l'invention, illustré par le bloc 309 selon le mode de réalisation du décodeur de la figure 3, met en oeuvre un procédé d'extension de bande décrit maintenant en référence à la **figure 4**.

[0066] Ce dispositif d'extension peut également être indépendant du décodeur et peut mettre en oeuvre le procédé décrit à la figure 4 pour effectuer une extension de bande d'un signal audio existant stocké ou transmis au dispositif, avec une analyse du signal audio pour en extraire une excitation et un filtre LPC.

[0067] Ce dispositif reçoit en entrée un signal d'excitation dans une première bande de fréquence dite bande basse $u(n)$ dans le cas d'une mise en oeuvre dans le domaine temporel ou $U(k)$ dans le cas d'une mise en oeuvre dans le domaine fréquentiel pour laquelle une étape de transformée temps fréquence est alors appliquée.

[0068] Dans le cas d'une application dans un décodeur, ce signal d'excitation reçu est un signal décodé.

[0069] Dans le cas d'un dispositif d'amélioration indépendant du décodeur, le signal d'excitation bande basse est extrait par analyse du signal audio.

[0070] Dans un mode possible de réalisation, le signal audio bande basse est ré-échantillonné avant l'étape d'extraction de l'excitation, si bien que l'excitation extraite du signal audio par prédiction linéaire estimée à partir du signal bande basse (ou de paramètres LPC associés à la bande basse) est déjà ré-échantillonnée. Un exemple de réalisation dans ce cas consiste à prendre un signal bande basse échantillonné à 12.8 kHz dont on dispose d'un filtre LPC bande basse décrivant l'enveloppe spectrale court-terme pour la trame courante, le sur-échantillonner à 16 kHz, et le filtrer par un filtre de prédiction LPC obtenu en extrapolant le filtre LPC. Un autre exemple de réalisation consiste à prendre un signal bande basse échantillonné à 12.8 kHz dont on ne dispose pas de modèle LPC, le sur-échantillonner à 16 kHz, effectuer une analyse LPC sur ce signal à 16 kHz, et filtrer ce signal par un filtre de prédiction LPC obtenu par cette analyse.

[0071] Une étape E401 de génération d'un signal d'excitation sur-échantillonné étendu ($u_{ex}(n)$ ou $U_{HB1}(k)$) dans une deuxième bande de fréquence supérieure à la première bande de fréquence est effectuée. Cette étape de génération peut comporter à la fois une étape de ré-échantillonnage et une étape d'extension ou simplement une étape d'extension en fonction du signal d'excitation obtenu en entrée.

[0072] Cette étape est détaillée ultérieurement dans les modes de réalisation décrits en référence aux figures 5 et 7.

[0073] Ce signal d'excitation sur-échantillonné étendu est utilisé pour obtenir un signal étendu ($U_{HB2}(k)$) dans une deuxième bande de fréquence. Ce signal étendu possède alors un modèle de signal adapté à certains types de signaux grâce aux caractéristiques du signal d'excitation étendu.

[0074] Ce signal étendu peut être obtenu après combinaison du signal d'excitation sur-échantillonné et étendu avec un autre signal, par exemple un signal de bruit.

[0075] Ainsi, dans un mode de réalisation, une étape E402 de génération d'un signal de bruit ($u_{HB}(n)$ ou $U_{HB}(k)$) au

moins dans la deuxième bande de fréquence est effectuée. La deuxième bande de fréquence est par exemple une bande haute fréquence allant de 6000 à 8000 Hz. Par exemple, ce bruit peut être généré de façon pseudo-aléatoire par un générateur congruentiel linéaire. Dans des variantes de l'invention, on pourra remplacer cette génération de bruit par d'autres méthodes, par exemple on pourra définir un signal d'amplitude constante (de valeur arbitraire, telle que 1) et appliquer des signes aléatoires à chaque raie fréquentielle générée.

[0076] Le signal d'excitation étendu est ensuite combiné au signal de bruit à l'étape E403 pour obtenir le signal étendu que l'on pourra également dénommer signal combiné ($u_{HB1}(n)$ ou $U_{HB2}(k)$) dans la bande de fréquence étendue correspondant à toute la bande de fréquence incluant la première et la deuxième bande de fréquence. Ainsi, la combinaison de ces deux types de signaux permet d'obtenir un signal combiné avec des caractéristiques plus adaptées à certains types de signaux comme des signaux musicaux.

[0077] En effet, le signal d'excitation décodé ou estimé dans la bande basse comporte dans certains cas, des harmoniques plus proches des signaux musicaux que le signal de bruit seul. Les harmoniques basses fréquences, si elles existent, peuvent ainsi être transposées en haute fréquence de telle sorte que leur mélange avec du bruit permet d'assurer un certain niveau d'harmonicité ou de niveau relatif de bruit ou de planéité spectrale (« spectral flatness » en anglais) dans la bande haute reconstruite.

[0078] L'extension de bande selon le procédé améliore la qualité pour ce type de signaux par rapport à AMR-WB.

[0079] Le signal combiné (ou étendu) est ensuite filtré en E404 par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse ($\hat{A}(z)$) décodés ou obtenus par analyse et extraction à partir du signal bande basse ou une version sur-échantillonnée de celui-ci. L'extension de bande selon le procédé est donc réalisée en étendant d'abord un signal d'excitation et en appliquant ensuite une étape de filtrage de synthèse par prédiction linéaire (LPC); cette approche exploite le fait que l'excitation LPC décodée dans la bande basse est un signal dont le spectre est relativement plat, ce qui évite des traitements supplémentaires de blanchiment du signal décodé dans l'extension de bande.

[0080] De façon avantageuse, les coefficients de ce filtre peuvent par exemple être obtenus à partir des paramètres décodés du filtre de prédiction linéaire (LPC) en bande basse. Si le filtre LPC utilisé en bande haute échantillonnée 16 kHz est de la forme $1/\hat{A}(z/\gamma)$, où $1/\hat{A}(z)$ est le filtre décodé en bande basse, et γ un facteur de pondération, la réponse en fréquence du filtre $1/\hat{A}(z/\gamma)$ correspond à un étalement de la réponse en fréquence du filtre décodé en bande basse. Dans une variante on pourra étendre le filtre $1/\hat{A}(z)$ à un ordre supérieur (comme à 6.6 kbit/s dans le bloc 111) pour éviter un tel étalement.

[0081] De façon préférentielle mais optionnelle, des étapes supplémentaires de filtrage passe-bande adaptatif en E405 et/ou de mise à l'échelle en E407 peuvent être effectuées pour d'une part améliorer la qualité du signal d'extension selon le débit de décodage et d'autre part pour s'assurer de garder le même ratio d'énergie entre une sous-trame et une trame de signal combiné que dans la bande basse de fréquence.

[0082] Ces étapes seront expliquées plus en détails dans les modes de réalisation des figures 5 et 7.

[0083] Dans un premier mode de réalisation, le dispositif d'extension de bande est décrit maintenant en référence à la figure 5. Ce dispositif met en oeuvre le procédé d'extension de bande décrit précédemment en référence à la figure 4.

[0084] Ainsi, à l'entrée de ce dispositif, un signal d'excitation bande basse décodé ou estimé par analyse est reçu ($u(n)$). L'extension de bande utilise ici l'excitation décodée à 12.8 kHz (exc2 ou $u(n)$) en sortie du bloc 302.

[0085] On notera que dans ce mode de réalisation, la génération de l'excitation sur-échantillonné et étendu s'effectue dans une bande de fréquence allant de 5 à 8 kHz incluant donc une deuxième bande de fréquence (6.4-8kHz) supérieure à la première bande de fréquence (0-6.4 kHz).

[0086] Ainsi, la génération du signal d'excitation étendu s'effectue au moins sur la deuxième bande de fréquence mais aussi sur une partie de la première bande de fréquence.

[0087] Bien évidemment, les valeurs définissant ces bandes de fréquences peuvent être différentes selon le décodeur ou le dispositif de traitement dans lequel l'invention s'applique.

[0088] Pour cet exemple de réalisation, ce signal est transformé pour obtenir un spectre de signal d'excitation $U(k)$ par le module de transformation temps-fréquence 500.

Dans un mode de réalisation particulier, la transformée utilise une DCT-IV (pour "Discrete Cosine Transform" - Type IV en anglais) (bloc 500) sur la trame courante de 20 ms (256 échantillons), sans fenêtrage, ce qui revient à transformer directement $u(n)$ avec $n = 0, \dots, 255$ selon la formule suivante :

$$U(k) = \sum_{n=0}^{N-1} u(n) \cos \left(\frac{\pi}{N} \left(n + \frac{1}{2} \right) \left(k + \frac{1}{2} \right) \right)$$

où $N = 256$ et $k = 0, \dots, 255$.

[0089] On note ici que la transformation sans fenêtrage (ou de façon équivalente avec une fenêtre rectangulaire

implicite de la longueur de la trame) est possible car le traitement est effectué dans le domaine de l'excitation, et non le domaine du signal, si bien qu'aucun artefact (effets de bloc) n'est audible, ce qui constitue un avantage important de ce mode de réalisation de l'invention.

[0090] Dans ce mode de réalisation, la transformation DCT-IV est mise en oeuvre par FFT suivant l'algorithme dit « *Evolved DCT (EDCT)* » décrit dans l'article de D.M. Zhang, H.T. Li, A Low Complexity Transform - Evolved DCT, IEEE 14th International Conference on Computational Science and Engineering (CSE), Aug. 2011, pp. 144-149, et mis en oeuvre dans les normes UIT-T G.718 Annex B et G.729.1 Annex E.

[0091] Dans des variantes de l'invention et sans perte de généralité, la transformation DCT-IV pourra être remplacée par d'autres transformations temps-fréquences court-terme de même longueur et dans le domaine de l'excitation, comme une FFT (pour "*Fast Fourier Transform*" en anglais) ou une DCT-II (*Discrete Cosine Transform* - Type II). De façon alternative, on pourra remplacer la DCT-IV sur la trame par une transformation avec recouvrement-addition et fenêtrage de longueur supérieure à la longueur de la trame courante, par exemple en utilisant une MDCT (pour "*Modified Discrete Cosine Transform*" en anglais). Dans ce cas le retard T dans le bloc 310 de la figure 3, devra être ajusté (réduit) de façon adéquate en fonction du retard additionnel dû à l'analyse/synthèse par cette transformée.

[0092] Le spectre DCT, $U(k)$, de 256 échantillons couvrant la bande 0-6400 Hz (à 12.8 kHz), est ensuite étendu (bloc 501) en un spectre de 320 échantillons couvrant la bande 0-8000 Hz (à 16 kHz) sous la forme suivante :

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

où on prend de façon préférentielle $start_band = 160$.

[0093] Le bloc 501 fonctionne comme module de génération d'un signal d'excitation sur-échantillonné et étendu et réalise l'étape E401 comportant un ré-échantillonnage de 12.8 à 16 kHz dans le domaine fréquentiel, en rajoutant 1/4 d'échantillons ($k = 240, \dots, 319$) au spectre, le ratio entre 16 et 12.8 étant de 5/4.

[0094] De plus, le bloc 501 réalise un filtrage passe-haut implicite dans la bande 0-5000 Hz puisque les 200 premiers échantillons de $U_{HB1}(k)$ sont mis à zéro ; comme expliqué plus tard, ce filtrage passe-haut est également complété par une partie d'atténuation progressive des valeurs spectrales d'indices $k = 200, \dots, 255$ dans la bande 5000-6400 Hz, cette atténuation progressive est mise en oeuvre dans le bloc 504 mais pourrait être réalisée séparément en dehors du bloc 504. De façon équivalente et dans des variantes de l'invention, la mise en oeuvre du filtrage passe-haut séparée en blocs de coefficients d'indice $k = 0, \dots, 199$ mis à zéro, de coefficients $k = 200, \dots, 255$ atténués, dans le domaine transformé, pourra donc être effectué en une seule étape.

[0095] Dans cet exemple de réalisation et selon la définition de $U_{HB1}(k)$, on remarque que la bande 5000-6000 Hz de $U_{HB1}(k)$ (qui correspond aux indices $k = 200, \dots, 239$) est copiée à partir de la bande 5000-6000 Hz de $U(k)$. Cette approche permet de conserver le spectre original dans cette bande et elle évite d'introduire des distorsions dans la bande 5000-6000 Hz lors de l'addition de la synthèse HF avec la synthèse BF - en particulier la phase du signal (implicitement représentée dans le domaine DCT-IV) dans cette bande est préservée.

[0096] La bande 6000-8000 Hz de $U_{HB1}(k)$ est ici définie en copiant la bande 4000-6000 Hz de $U(k)$ puisque la valeur de $start_band$ est fixée préférentiellement à 160.

[0097] Dans une variante du mode de réalisation, la valeur de $start_band$ pourra être rendue adaptative autour de la valeur de 160, sans modifier la nature de l'invention. Les détails de l'adaptation de la valeur $start_band$ ne sont pas décrits ici car ils dépassent le cadre de l'invention sans en changer la portée.

[0098] Pour certains signaux en bande élargie (échantillonnés à 16 kHz), la bande haute (>6 kHz) peut être bruitée, harmonique ou comporter un mélange de bruit et d'harmoniques. De plus, le niveau d'harmonicité dans la bande 6000-8000 Hz est généralement corrélé à celui des bandes de fréquences inférieures. Ainsi, dans un mode particulier de réalisation, le bloc 502 de génération de bruit, met en oeuvre l'étape E402 de la figure 4 et réalise une génération de bruit dans le domaine fréquentiel, $U_{HBN}(k)$ pour $k = 240, \dots, 319$ (80 échantillons) correspondant à une deuxième bande de fréquence dite haute fréquence afin de combiner ensuite ce bruit avec le spectre $U_{HB1}(k)$ dans le bloc 503.

[0099] Dans un mode de réalisation particulier, le bruit (dans la bande 6000-8000 Hz) est généré de façon pseudo-aléatoire avec un générateur congruentiel linéaire sur 16 bits :

$$U_{HBN}(k) = \begin{cases} 0 & k = 0, \dots, 239 \\ 31821 U_{HBN}(k-1) + 13849 & k = 240, \dots, 319 \end{cases}$$

avec la convention que $U_{HBN}(239)$ dans la trame courante correspond à la valeur $U_{HBN}(319)$ de la trame précédente. Dans des variantes de l'invention, on pourra remplacer cette génération de bruit par d'autres méthodes.

[0100] Le bloc 503 de combinaison peut être réalisé de différentes façons. De façon privilégiée, on considère un mixage additif adaptatif de la forme :

$$U_{HB2}(k) = \beta U_{HB1}(k) + \alpha G_{HBN} U_{HBN}(k), \quad k = 240, \dots, 319$$

où G_{HBN} est un facteur de normalisation servant à égaliser le niveau d'énergie entre les deux signaux,

$$G_{HBN} = \sqrt{\frac{\sum_{k=240}^{319} U_{HB1}(k)^2 + \varepsilon}{\sum_{k=240}^{319} U_{HBN}(k)^2 + \varepsilon}}$$

avec $\varepsilon=0.01$, et le coefficient α (compris entre 0 et 1) est ajusté en fonction de paramètres estimés à partir de la bande basse décodée et le coefficient β (compris entre 0 et 1) dépend de α .

[0101] Dans un mode de réalisation privilégié on calcule l'énergie du bruit dans trois bandes : 2000-4000 Hz, 4000-6000 Hz et 6000-8000 Hz, avec

$$E_{N2-4} = \sum_{k \in N(80,159)} U^2(k)$$

$$E_{N4-6} = \sum_{k \in N(160,239)} U^2(k)$$

$$E_{N4-6} = \sum_{k \in N(240,319)} U^2(k)$$

où

$$U'(k) = \begin{cases} \sqrt{\frac{\sum_{k=160}^{239} U^2(k)}{\sum_{k=80}^{159} U^2(k)}} U(k) & k = 80, \dots, 159 \\ U(k) & k = 160, \dots, 239 \\ \sqrt{\frac{\sum_{k=160}^{239} U^2(k)}{\sum_{k=240}^{319} U_{HB1}^2(k)}} U_{HB1}(k) & k = 240, \dots, 319 \end{cases}$$

et $N(k_1, k_2)$ est l'ensemble des indices k pour lesquels le coefficient d'indice k est classifié comme étant associé à du bruit. Cet ensemble peut être par exemple obtenu en détectant les pics locaux dans $U'(k)$ vérifiant $|U'(k)| \geq |U'(k-1)|$ et $|U'(k)| \geq |U'(k+1)|$ et en considérant que ces raies ne sont pas associés à du bruit, soit (en appliquant la négation de la condition précédente):

$$N(a, b) = \{a \leq k \leq b \mid |U'(k)| < |U'(k-1)| \text{ ou } |U'(k)| < |U'(k+1)|\}$$

[0102] On peut noter que d'autres méthodes de calcul de l'énergie du bruit sont possibles, par exemple en prenant la valeur médiane du spectre sur la bande considérée ou en appliquant un lissage à chaque raie fréquentielle avant de calculer l'énergie par bande.

[0103] On fixe α de telle sorte que le ratio entre l'énergie du bruit dans les bandes 4-6 kHz et 6-8 kHz soit le même qu'entre les bandes 2-4 kHz et 4-6 kHz :

$$\alpha = \sqrt{\frac{\rho - E_{N6-8}}{\sum_{k=160}^{239} U^2(k) - E_{N6-8}}}$$

où

$$E_{N4-6} = \max(E_{N4-6}, E_{N2-4}), \quad \rho = \frac{E_{N4-6}^2}{E_{N2-4}}, \quad \rho = \max(\rho, E_{N6-8})$$

où $\max(.,.)$ est la fonction qui donne le maximum des deux arguments.

[0104] Dans des variantes de l'invention, le calcul de α pourra être remplacé par d'autres méthodes. Par exemple, dans une variante, on pourra extraire (calculer) différents paramètres (ou « features » en anglais) caractérisant le signal en bande basse, dont un paramètre « tilt » similaire à celui calculé dans le codec AMR-WB, et on estimera le facteur α en fonction d'une régression linéaire à partir de ces différents paramètres en limitant sa valeur entre 0 et 1. La régression linéaire pourra par exemple être estimée de façon supervisée en estimant le facteur α en se donnant la bande haute originale dans une base d'apprentissage. On notera que le mode de calcul de α ne limite pas la nature de l'invention.

[0105] Dans un mode de réalisation privilégié, on prend

$$\beta = \sqrt{1 - \alpha^2}$$

afin de préserver l'énergie du signal étendu après mixage.

[0106] Dans une variante les facteurs β et α pourront être adaptés pour tenir compte du fait qu'un bruit injecté dans une bande donnée du signal est perçu en général comme plus fort qu'un signal harmonique à la même énergie dans la même bande. Ainsi on pourra modifier les facteurs β et α comme suit:

$$\begin{aligned} \beta &\leftarrow \beta \cdot f(\alpha) \\ \alpha &\leftarrow \alpha \cdot f(\alpha) \end{aligned}$$

où $f(\alpha)$ est une fonction décroissante de α , par exemple $f(\alpha) = b - a\sqrt{\alpha}$, $f(\alpha)$ $b=1.1$, $a = 1.2$, $f(\alpha)$ limité de 0.3 à 1. Il faut remarquer qu'après multiplication par $f(\alpha)$, $\alpha^2 + \beta^2 < 1$ si bien que l'énergie du signal $U_{HB2}(k) = \beta U_{HB1}(k) + \alpha G_{HBN} U_{HBN}(k)$ est plus basse que l'énergie de $U_{HB1}(k)$ (la différence d'énergie dépend de α , plus on rajoute de bruit, plus l'énergie est atténuée). Dans d'autres variantes de l'invention on pourra prendre :

$$\beta = 1 - \alpha$$

ce qui permet de préserver le niveau d'amplitude (quand les signaux combinés sont de même signe) ; cependant cette variante a le désavantage de résulter en une énergie globale (au niveau de $U_{HB2}(k)$) qui n'est pas monotone en fonction de α .

[0107] On remarque donc ici que le bloc 503 réalise l'équivalent du bloc 101 de la figure 1 pour normaliser le bruit blanc en fonction d'une excitation qui est par contre ici dans le domaine fréquentiel, déjà étendue à la cadence de 16 kHz ; de plus, le mixage est limité à la bande 6000-8000 Hz.

[0108] Dans une variante simple, on peut considérer une réalisation du bloc 503, où les spectres, $U_{HB1}(k)$ ou $G_{HBN} U_{HBN}(k)$, sont sélectionnés (commutés) de façon adaptative, ce qui revient à n'autoriser que les valeurs 0 ou 1 pour α ; cette approche revient à classifier le type d'excitation à générer dans la bande 6000-8000 Hz

[0109] Le bloc 504 réalise de façon optionnelle, une double opération d'application de réponse en fréquence de filtre

passer-bande et de filtrage de désaccentuation (ou déemphasis) dans le domaine fréquentiel.

[0110] Dans une variante de l'invention, le filtrage de désaccentuation pourra être réalisé dans le domaine temporel, après le bloc 505 voire avant le bloc 500 ; cependant, dans ce cas, le filtrage passe-bande réalisé dans le bloc 504 peut laisser certaines composantes basses fréquences de très faibles niveaux qui se voient amplifier par désaccentuation, ce qui peut modifier de façon légèrement perceptible la bande basse décodée. Pour cette raison, on préfère ici réaliser la désaccentuation dans le domaine fréquentiel. Dans le mode de réalisation privilégié, les coefficients d'indice $k=0, \dots, 199$ sont mis à zéro, ainsi la désaccentuation est limitée aux coefficients supérieurs.

L'excitation est d'abord désaccentuée selon l'équation suivante :

$$U_{HB2}'(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ G_{deemph}(k-200)U_{HB2}(k) & k = 200, \dots, 255 \\ G_{deemph}(55)U_{HB2}(k) & k = 256, \dots, 319 \end{cases}$$

où $G_{deemph}(k)$ est la réponse en fréquence du filtre $1/(1-0.68z^{-1})$ sur une bande de fréquence discrète restreinte. En prenant en compte les fréquences discrètes (impaires) de la DCT-IV, on définit ici $G_{deemph}(k)$ comme :

$$G_{deemph}(k) = \frac{1}{|e^{j\theta_k} - 0.68|}, \quad k = 0, \dots, 255$$

où

$$\theta_k = \frac{256 - 80 + k + \frac{1}{2}}{256}.$$

Dans le cas où une autre transformation que la DCT-IV est utilisée, la définition de θ_k pourra être ajustée (par exemple pour des fréquences paires).

On note que la désaccentuation est appliquée en deux phases pour $k = 200, \dots, 255$ correspondant à la bande de fréquence 5000-6400 Hz, où la réponse $1/(1-0.68z^{-1})$ est appliquée comme à 12.8 kHz, et pour $k = 256, \dots, 319$ correspondant à la bande de fréquence 6400-8000 Hz, où la réponse est étendue de 16 kHz ici à une valeur constante dans la bande 6.4-8 kHz.

[0111] On peut noter que dans le codec AMR-WB la synthèse HF n'est pas désaccentuée. Dans le mode de réalisation présenté ici, le signal hautes fréquences est au contraire désaccentué de manière à le ramener dans un domaine cohérent avec le signal basses fréquences (0-6.4 kHz) qui sort du bloc 305. Ceci est important pour l'estimation et l'ajustement ultérieur de l'énergie de la synthèse HF.

[0112] Dans une variante du mode de réalisation, afin de réduire la complexité, on pourra fixer $G_{deemph}(k)$ à une valeur constante indépendante de k , en prenant par exemple $G_{deemph}(k) = 0.6$ ce qui correspond approximativement à la valeur moyenne de $G_{deemph}(k)$ pour $k = 200, \dots, 319$ dans les conditions du mode de réalisation décrit ci-dessus.

[0113] Dans une autre variante du mode de réalisation du dispositif d'extension, la désaccentuation pourra être réalisée de façon équivalente dans le domaine temporel après DCT inverse. Une telle réalisation est mise en oeuvre à la figure 7 décrite plus loin.

[0114] En plus de la désaccentuation, un filtrage passe-bande est appliqué avec deux parties séparées : l'une passe-haut fixe, l'autre passe-bas adaptative (fonction du débit).

[0115] Ce filtrage est effectué dans le domaine fréquentiel, et sa réponse en fréquence est illustrée à la **figure 6**. Les fréquences de coupures à 3 dB sont 6000 Hz pour la partie basse et pour la partie haute approximativement 6900, 7300, 7600 Hz à 6.6, 8.86 et aux débits supérieurs à 8.85 kbit/s (respectivement).

[0116] Dans le mode de réalisation privilégiée, on calcule la réponse partielle de filtre passe-bas dans le domaine fréquentiel comme suit :

$$G_{lp}(k) = 1 - 0.999 \frac{k}{N_{lp} - 1}$$

où $N_{lp}=60$ à 6.6 kbit/s, 40 à 8.85 kbit/s, 20 aux débits >8.85 bit/s. Ensuite on applique un filtre passe-bande sous la forme :

$$U_{HB3}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ G_{hp}(k-200)U_{HB2}'(k) & k = 200, \dots, 255 \\ U_{HB2}'(k) & k = 256, \dots, 319 - N_{lp} \\ G_{lp}(k-320 - N_{lp})U_{HB2}'(k) & k = 320 - N_{lp}, \dots, 319 \end{cases}$$

[0117] La définition de $G_{hp}(k)$, $k = 0, \dots, 55$, est donnée par exemple au tableau 1 ci-dessous.

Tableau 1

K	$g_{hp}(k)$	K	$g_{hp}(k)$	K	$g_{hp}(k)$	k	$g_{hp}(k)$
0	0.001622428	14	0.114057967	28	0.403990611	42	0.776551214
1	0.004717458	15	0.128865425	29	0.430149896	43	0.800503267
2	0.008410494	16	0.144662643	30	0.456722014	44	0.823611104
3	0.012747280	17	0.161445005	31	0.483628433	45	0.845788355
4	0.017772424	18	0.179202219	32	0.510787115	46	0.866951597
5	0.023528982	19	0.197918220	33	0.538112915	47	0.887020781
6	0.030058032	20	0.217571104	34	0.565518011	48	0.905919644
7	0.037398264	21	0.238133114	35	0.592912340	49	0.923576092
8	0.045585564	22	0.259570657	36	0.620204057	50	0.939922577
9	0.054652620	23	0.281844373	37	0.647300005	51	0.954896429
10	0.064628539	24	0.304909235	38	0.674106188	52	0.968440179
11	0.075538482	25	0.328714699	39	0.700528260	53	0.980501849
12	0.087403328	26	0.353204886	40	0.726472003	54	0.991035206
13	0.100239356	27	0.378318805	41	0.751843820	55	1.000000000

[0118] On notera que dans des variantes de l'invention les valeurs de $G_{hp}(k)$ pourront être modifiées tout en gardant une atténuation progressive. De même le filtrage passe-bas à largeur de bande variable, $G_{lp}(k)$, pourra être ajusté avec des valeurs ou un support fréquentiel différents, sans changer le principe de cette étape de filtrage.

[0119] On notera aussi que l'exemple de filtrage passe-bande illustré à la figure 6 pourra être adapté en définissant une seule étape de filtrage combinant les filtrages passe-haut et passe-bas.

[0120] Dans un autre mode de réalisation, le filtrage passe-bande pourra être réalisé de façon équivalente dans le domaine temporel (comme dans le bloc 112 de la figure 1) avec différents coefficients de filtre selon le débit, après une étape de DCT inverse. Une telle réalisation est mise en oeuvre à la figure 7 décrite plus loin. Cependant, on notera qu'il est avantageux de réaliser cette étape directement dans le domaine fréquentiel car le filtrage est effectué dans le domaine de l'excitation LPC et donc les problèmes de convolution circulaire et d'effets de bord sont très limités dans ce domaine.

[0121] Le bloc 505 de transformée inverse réalise une DCT inverse sur 320 échantillons pour trouver l'excitation haute-fréquence échantillonnée à 16 kHz. Sa mise en oeuvre est identique au bloc 500, car la DCT-IV est orthonormée, sauf que la longueur de la transformée est de 320 au lieu de 256, et on obtient:

$$u_{HB}(n) = \sum_{k=0}^{N_{16k}-1} U_{HB3}(k) \cos\left(\frac{\pi}{N_{16k}}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right)$$

où $N_{16k} = 320$ et $k=0, \dots, 319$.

[0122] Cette excitation échantillonnée à 16 kHz est ensuite mise à l'échelle par des gains définis par sous-trame de

80 échantillons (bloc 507).

Dans un mode de réalisation privilégié, on calcule d'abord (bloc 506) un gain $g_{HB1}(m)$ par sous-trame par des ratios d'énergie des sous-frames tel que dans chaque sous-trame d'indice $m=0, 1, 2$ ou 3 de la trame courante:

$$g_{HB1}(m) = \sqrt{\frac{e_3(m)}{e_2(m)}}$$

où

$$e_1(m) = \sum_{n=0}^{63} u(n+64m)^2 + \varepsilon$$

$$e_2(m) = \sum_{n=0}^{79} u_{HB}(n+80m)^2 + \varepsilon$$

$$e_3(m) = e_1(m) \frac{\sum_{n=0}^{319} u_{HB}(n)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon}$$

avec $\varepsilon = 0.01$. On peut écrire le gain par sous-trame $g_{HB1}(m)$ sous la forme :

$$g_{HB1}(m) = \sqrt{\frac{\sum_{n=0}^{63} u(n+64m)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon} \frac{\sum_{n=0}^{79} u_{HB}(n+80m)^2 + \varepsilon}{\sum_{n=0}^{319} u_{HB}(n)^2 + \varepsilon}}$$

ce qui montre qu'on assure dans le signal u_{HB} le même ratio entre énergie par sous-trame et énergie par trame que dans le signal $u(n)$.

[0123] Le bloc 507 effectue la mise à l'échelle du signal combiné (ou étendu) (étape E406 de la figure 4) selon l'équation suivante:

$$u'_{HB}(n) = g_{HB1}(m) u_{HB}(n), \quad n = 80m, \dots, 80(m+1) - 1$$

[0124] On notera que la réalisation du bloc 506 diffère de celle du bloc 101 de la figure 1, car l'énergie au niveau de la trame courante est prise en compte en plus de celle de la sous-trame. Cela permet d'avoir le ratio de l'énergie de chaque sous-trame par rapport à l'énergie de la trame. On compare donc des ratios d'énergie (ou énergies relatives) plutôt que les énergies absolues entre bande basse et bande haute.

[0125] Ainsi, cette étape de mise à l'échelle permet de conserver dans la bande haute le ratio d'énergie entre la sous-trame et la trame de la même façon que dans la bande basse.

[0126] De façon optionnelle, le bloc 509 effectue ensuite la mise à l'échelle du signal (étape E407 de la figure 4) selon l'équation suivante:

$$u_{HB}''(n) = g_{HB2}(m)u_{HB}'(n), n = 80m, \dots, 80(m+1) - 1$$

où le gain $g_{HB2}(m)$ est obtenu à partir du bloc 508 en exécutant les blocs 103, 104 et 105 du codec AMR-WB (l'entrée du bloc 103 étant l'excitation décodée en bande basse, $u(n)$). Les blocs 508 et 509 sont utiles pour ajuster le niveau du filtre de synthèse LPC (bloc 510), ici en fonction du tilt du signal. D'autres méthodes de calcul du gain $g_{HB2}(m)$ sont possibles sans changer la nature de l'invention.

[0127] Enfin, l'excitation, $u_{HB}'(n)$ ou $u_{HB}''(n)$, est filtrée (étape E404 de la figure 4) par le module de filtrage 510 qui peut être réalisé ici en prenant comme fonction de transfert $1/\hat{A}(z/\gamma)$, où $\gamma = 0.9$ à 6.6 kbit/s et $\gamma = 0.6$ aux autres débits, ce qui limite l'ordre du filtre à l'ordre 16.

[0128] Dans une variante, ce filtrage pourra être réalisé de la même façon que ce qui est décrit pour le bloc 111 de la figure 1 du décodeur AMR-WB, cependant l'ordre du filtre passe à 20 au débit de 6.6, ce qui ne change pas de façon significative la qualité du signal synthétisé. Dans une autre variante, on pourra effectuer le filtrage de synthèse LPC dans le domaine fréquentiel, après avoir calculé la réponse en fréquence du filtre mis en oeuvre dans le bloc 510.

[0129] Dans des variantes de réalisation de l'invention, le codage de la bande basse (0-6.4 kHz) pourra être remplacé par un codeur CELP autre que celui utilisé dans AMR-WB, comme par exemple le codeur CELP dans G.718 à 8 kbit/s. Sans perte de généralité d'autres codeurs en bande élargie ou fonctionnant à des fréquences supérieures à 16 kHz, dans lesquels le codage de la bande basse fonctionne à une fréquence interne à 12.8 kHz pourraient être utilisés. Par ailleurs, l'invention peut être adaptée de façon évidente à d'autres fréquences d'échantillonnage que 12.8 kHz, lorsqu'un codeur de basses fréquences fonctionne à une fréquence d'échantillonnage inférieure à celle du signal original ou reconstruit. Lorsque le décodage en bande basse n'utilise pas de prédiction linéaire, on ne dispose pas d'un signal d'excitation à étendre, dans ce cas on pourra réaliser une analyse LPC du signal reconstruit dans la trame courante et on calculera une excitation LPC de manière à pouvoir appliquer l'invention.

[0130] Enfin, dans une autre variante de l'invention, l'excitation ($u(n)$) est ré-échantillonnée, par exemple par interpolation linéaire ou "spline" cubique, de 12.8 à 16 kHz avant transformation (par exemple DCT-IV) de longueur 320. Cette variante a le défaut d'être plus complexe, car la transformée (DCT-IV) de l'excitation est alors calculée sur une plus grande longueur et le ré-échantillonnage n'est pas effectué dans le domaine de la transformée.

[0131] De plus, dans des variantes de l'invention, tous les calculs nécessaires à l'estimation des gains (G_{HBN} , $g_{HB1}(m)$, $g_{HB2}(m)$, g_{HBN} , ...) pourront être effectués dans un domaine logarithmique.

[0132] En référence à la **figure 7**, un deuxième mode de réalisation du dispositif d'extension de bande est maintenant décrit. Ce mode de réalisation fonctionne dans le domaine temporel. Ce deuxième mode de réalisation ne fait pas partie de l'invention.

[0133] Comme dans le mode de réalisation de la figure 5, le principe du mode de réalisation avec mixage d'un signal étendu à 16 kHz et d'un signal de bruit est conservé, cependant ce mélange est cette fois réalisé dans le domaine temporel et cette fois-ci la génération principale de l'excitation se fait par sous-trame et non par trame.

[0134] Le signal d'excitation $u(n)$, $n = 0, \dots, 255$, issu du décodage basse fréquence dans la trame courante est d'abord ré-échantillonné sans retard (étape E401 de la figure 4) à 16 kHz (bloc 700) et dans un mode de réalisation particulier, une interpolation linéaire est utilisée pour obtenir le signal d'excitation étendu dans une deuxième bande de fréquence, $u_{ext}(n)$, $n = 0, \dots, 319$. Dans une variante de réalisation, on pourra utiliser d'autres méthodes de ré-échantillonnage, par exemple par "splines" ou par filtrage multi-cadences.

[0135] On s'assure que l'énergie du signal $u_{ext}(n)$ a un niveau similaire à l'excitation $u(n)$ avec les blocs 701 et 702 de la façon suivante:

$$u_{ext}'(n) = u_{ext}(n) \sqrt{\frac{\sum_{l=0}^{63} u(l)^2}{\sum_{l=0}^{79} u_{ext}(l)^2}}$$

[0136] Dans une variante de réalisation on pourra multiplier $u_{ext}'(n)$ par 5/4 pour compenser l'atténuation par le ratio 12.8/16, causée par différentes fréquences d'échantillonnage de signaux $u_{ext}(n)$ et $u(n)$.

Le générateur de bruit dans le bloc 703 met en oeuvre l'étape E402 de la figure 4 et peut être réalisé comme dans le bloc 502 décrit en figure 5, sauf que le signal en sortie correspond à une sous-trame temporelle, $u_{HBN}(n)$, $n=0, \dots, 319$

Le bloc 704 de combinaison peut être réalisé de différentes façons. De façon privilégiée, on considère un mixage additif adaptatif *par sous-trame* de la forme :

$$u_{HB1}(n + 80m) = \beta u_{ext}(n + 80m) + \alpha g_{HBN} u_{HBN}(n + 80m), n = 0, \dots, 79 \text{ où } g_{HBN} \text{ est un facteur de normalisation servant}$$

à égaliser le niveau d'harmonicité des deux signaux combinés

$$g_{HBN} = \sqrt{\frac{\sum_{k=0}^{79} u_{ext}(n)^2 + \varepsilon}{\sum_{k=0}^{79} u_{HBN}(n)^2 + \varepsilon}}$$

et m est l'indice de la sous-trame et les facteurs α et β sont calculés comme dans le premier mode de réalisation. On remarque donc ici que le bloc 704 réalise l'équivalent du bloc 101 de la figure 1. De plus, le calcul du facteur α oblige à calculer la transformée du signal d'excitation décodé (ou le signal décodé lui-même selon le domaine de calcul du niveau relatif de bruit ou de planéité spectrale dite "spectral flatness" en anglais) en bande basse si ce calcul repose sur la planéité spectrale ; dans des variantes, dont l'utilisation d'une régression linéaire décrite précédemment, une telle transformée n'est pas nécessaire.

[0137] Ensuite le signal temporel est désaccentué (bloc 705) par un filtre de la forme $g_{deemph}/(1-0.68z^{-1})$, où g_{deemph} est calculé de sorte à prolonger le filtre $1/(1-0.68z^{-1})$ (défini à 12.8 kHz) à la fréquence d'échantillonnage de 16 kHz $g_{deemph} = [(1-0.68e^{2\pi 6000/16000}) / (1-0.68e^{2\pi 6000/12800})]$, puis traité par un filtrage passe-bande de largeur de bande variable (bloc 706) dont l'ordre est fixe (de valeur 30) mais les coefficients changent en fonction du débit de décodage de la trame courante.

[0138] Un exemple de réalisation d'un tel filtrage passe-bande adaptatif de type FIR est donné aux tableaux ci-dessous définissant la réponse impulsionnelle du filtre FIR selon le débit.

Tableau 2a (6.6 kbit/s)

n	$h(n)$	n	$h(n)$	n	$h(n)$	n	$h(n)$
0	-0.0002581	8	0.0306285	16	-0.1451668	24	-0.0114595
1	0.0003791	9	-0.0716116	17	0.0626279	25	0.0090482
2	0.0002581	10	0.0995869	18	0.0286124	26	-0.0029758
3	-0.0002177	11	-0.0885791	19	-0.0885791	27	-0.0002177
4	-0.0029758	12	0.0286124	20	0.0995869	28	0.0002581
5	0.0090482	13	0.0626279	21	-0.0716116	29	0.0003791
6	-0.0114595	14	-0.1451668	22	0.0306285	30	-0.0002581
7	0	15	0.1783678	23	0	-	-

Tableau 2b (8.85 kbit/s)

n	$h(n)$	n	$h(n)$	n	$h(n)$	n	$h(n)$
0	0.0019706	8	0.0312161	16	-0.1720177	24	-0.0030672
1	-0.0064291	9	-0.0709664	17	0.0817478	25	-0.0041966
2	0.0124179	10	0.0980678	18	0.0181018	26	0.0132058
3	-0.0160589	11	-0.0842625	19	-0.0842625	27	-0.0160589
4	0.0132058	12	0.0181018	20	0.0980678	28	0.0124179
5	-0.0041966	13	0.0817478	21	-0.0709664	29	-0.0064291
6	-0.0030672	14	-0.1720177	22	0.0312161	30	0.0019706
7	-0.0036671	15	0.2083360	23	-0.0036671	-	-

Tableau 2c (débits >8.85 kbit/s)

n	$h(n)$	n	$h(n)$	n	$h(n)$	n	$h(n)$
0	0.0013312	8	0.0606146	16	-0.1916778	24	0.0221682
1	-0.0047346	9	-0.0860005	17	0.1093354	25	-0.0180046
2	0.0098657	10	0.0924138	18	-0.0129187	26	0.0171709
3	-0.0147045	11	-0.0607694	19	-0.0607694	27	-0.0147045
4	0.0171709	12	-0.0129187	20	0.0924138	28	0.0098657
5	-0.0180046	13	0.1093354	21	-0.0860005	29	-0.0047346
6	0.0221682	14	-0.1916778	22	0.0606146	30	0.0013312
7	-0.0360130	15	0.2240719	23	-0.0360130	-	-

[0139] L'étape de mise à l'échelle (E407 à la figure 4) est effectuée par les blocs 508 et 509 identiques à la figure 5.

[0140] L'étape de filtrage (E404 de la figure 4) est effectuée par le module de filtrage (bloc 510) identique à celui décrit en référence à la figure 5.

[0141] Il n'est pas utile ici de mettre en oeuvre une étape de mise à l'échelle comme effectué dans le mode de réalisation de la figure 5 par les blocs 506 et 507 puisque l'excitation est générée par sous-trames. La cohérence du ratio d'énergie au niveau de la trame est déjà assurée.

[0142] Dans des variantes de l'invention, l'excitation en bande basse $u(n)$ et le filtre LPC $1/\hat{A}(z)$ seront estimés par trame, par analyse LPC d'un signal en bande basse dont la bande doit être étendue. Le signal d'excitation bande basse est alors extrait par analyse du signal audio.

[0143] Dans un mode possible de réalisation de cette variante, le signal audio bande basse est ré-échantillonné avant l'étape d'extraction de l'excitation, si bien que l'excitation extraite du signal audio (par prédiction linéaire) est déjà ré-échantillonnée.

L'invention illustrée à la figure 5, ou le deuxième mode de réalisation ne faisant pas partie de l'invention et décrit à la figure 7, s'applique dans ce cas à une bande basse qui n'est pas décodée mais analysée.

[0144] La figure 8 représente un exemple de réalisation matérielle d'un dispositif de d'extension de bande 800 selon l'invention. Celui-ci peut faire partie intégrante d'un décodeur de signal audiofréquence ou d'un équipement recevant des signaux audiofréquences décodés ou non.

[0145] Ce type de dispositif comporte un processeur PROC coopérant avec un bloc mémoire BM comportant une mémoire de stockage et/ou de travail MEM.

Un tel dispositif comporte un module d'entrée E apte à recevoir un signal audio d'excitation décodé ou extrait dans une première bande de fréquence dite bande basse ($u(n)$ ou $U(k)$) et les paramètres d'un filtre de synthèse de prédiction linéaire ($\hat{A}(z)$). Il comporte un module de sortie S apte à transmettre le signal haute fréquence synthétisé (HF_syn) par exemple à un module d'application d'un retard comme le bloc 310 de la figure 3 ou à un module de ré-échantillonnage comme le module 311.

[0146] Le bloc mémoire peut avantageusement comporter un programme informatique comportant des instructions de code pour la mise en oeuvre des étapes du procédé de d'extension de bande au sens de l'invention, lorsque ces instructions sont exécutées par le processeur PROC, et notamment les étapes d'obtention d'un signal étendu dans au moins une deuxième bande de fréquence supérieure à la première bande de fréquence à partir d'un signal d'excitation sur-échantillonné et étendu dans au moins une deuxième bande de fréquence, de mise à l'échelle du signal étendu par un gain défini par sous-trame en fonction d'un ratio d'énergie d'une trame et d'une sous-trame et de filtrage dudit signal étendu mis à l'échelle par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse.

[0147] Typiquement, la description de la figure 4 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.

[0148] La mémoire MEM enregistre de manière générale, toutes les données nécessaires à la mise en oeuvre du procédé.

[0149] Dans un mode possible de réalisation, le dispositif ainsi décrit peut également comporter les fonctions de décodage bande basse et autre fonctions de traitement décrites par exemple en figure 3 en plus des fonctions d'extension de bande selon l'invention.

Revendications

1. Procédé d'extension de bande de fréquence d'un signal audiofréquence lors d'un processus de décodage ou d'amélioration comportant une étape de décodage ou d'extraction, dans une première bande de fréquence dite bande basse, d'un signal d'excitation et des coefficients d'un filtre de prédiction linéaire, le procédé étant **caractérisé en ce qu'il** comporte les étapes suivantes:

- obtention d'un signal étendu ($U_{HB2}(k)$, E403)) dans au moins une deuxième bande de fréquence supérieure à la première bande de fréquence à partir du signal d'excitation sur-échantillonné et étendu dans la au moins une deuxième bande de fréquence ($U_{HB1}(k)$, E401);
- mise à l'échelle (E406) du signal étendu par un gain défini par sous-trame en fonction d'un ratio entre le ratio entre l'énergie par sous-trame et l'énergie par trame du signal d'excitation de la bande basse et le ratio entre l'énergie par sous-trame et l'énergie par trame du signal étendu ;
- filtrage (E404) dudit signal étendu mis à l'échelle par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse.

2. Procédé selon la revendication 1, **caractérisé en ce qu'il** comporte en outre une étape de filtrage (E405) passe-bande adaptatif en fonction du débit de décodage de la trame courante, du signal étendu avant l'étape de mise à l'échelle.

3. Procédé selon la revendication 1, **caractérisé en ce qu'il** comporte une étape de transformée temps-fréquence du signal d'excitation, l'étape d'obtention d'un signal étendu s'effectuant alors dans le domaine fréquentiel et une étape de transformée temps-fréquence inverse du signal étendu avant les étapes de mise à l'échelle et de filtrage.

4. Procédé selon la revendication 3, **caractérisé en ce que** l'étape de génération d'un signal d'excitation sur-échantillonné et étendu est effectuée selon l'équation suivante:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

avec k l'indice de l'échantillon, $U_{HB1}(k)$ le spectre du signal d'excitation étendu, $U(k)$ le spectre du signal d'excitation obtenu après l'étape de transformée et *start_band* une variable prédéfinie.

5. Procédé selon l'une des revendications 1 à 4, **caractérisé en ce qu'il** comporte une étape de filtrage de désaccentuation du signal étendu au moins dans la deuxième bande de fréquence.

6. Procédé selon la revendication 1, **caractérisé en ce qu'il** comporte en outre une étape de génération (E402) d'un signal de bruit au moins dans la deuxième bande de fréquence le signal étendu ($U_{HB2}(k)$) étant obtenu par combinaison (E403) du signal d'excitation étendu et du signal de bruit.

7. Procédé selon la revendication 6, **caractérisé en ce que** l'étape de combinaison s'effectue par mixage additif adaptatif avec un gain d'égaleisation de niveau entre le signal d'excitation étendu et le signal de bruit.

8. Dispositif d'extension de bande de fréquence d'un signal audiofréquence comportant une étape de décodage ou d'extraction, dans une première bande de fréquence dite bande basse, d'un signal d'excitation et des coefficients d'un filtre de prédiction linéaire, le dispositif étant **caractérisé en ce qu'il** comporte:

- un module d'obtention d'un signal étendu ($U_{HB2}(k)$, 503)) dans au moins une deuxième bande de fréquence supérieure à la première bande de fréquence à partir du signal d'excitation sur-échantillonné et étendu dans la au moins une deuxième bande de fréquence ($U_{HB1}(k)$);
- mise à l'échelle (507) du signal étendu par un gain défini par sous-trame en fonction d'un ratio entre le ratio entre l'énergie par sous-trame et l'énergie par trame du signal d'excitation de la bande basse et le ratio entre l'énergie par sous-trame et l'énergie par trame du signal étendu;
- un module de filtrage (510) dudit signal étendu mis à l'échelle par un filtre de prédiction linéaire dont les coefficients sont dérivés des coefficients du filtre de la bande basse.

9. Décodeur de signal audiofréquence **caractérisé en ce qu'il** comporte un dispositif d'extension de bande de fréquence conforme à la revendication 8.
10. Programme informatique comportant des instructions de code pour la mise en oeuvre des étapes du procédé d'extension de bande de fréquence selon l'une des revendications 1 à 7, lorsque ces instructions sont exécutées par un processeur.
11. Support de stockage lisible par un dispositif d'extension de bande de fréquence sur lequel est enregistré un programme informatique comprenant des instructions de code pour l'exécution des étapes du procédé d'extension de bande de fréquence selon l'une des revendications 1 à 7.

Patentansprüche

1. Verfahren zur Frequenzbänderweiterung eines Audiosignals bei einem Decodierungs- oder Verbesserungsprozess, umfassend einen Schritt der Decodierung oder Extraktion, in einem ersten Frequenzband, Tiefband genannt, eines Anregungssignals und der Koeffizienten eines linearen Vorhersagefilters, wobei das Verfahren **dadurch gekennzeichnet ist, dass** es die folgenden Schritte umfasst:
- Erhalten eines erweiterten Signals ($U_{HB2}(k)$, E403) in mindestens einem zweiten Frequenzband, das über dem ersten Frequenzband liegt, ausgehend von dem in dem mindestens einen zweiten Frequenzband überabgetasteten und erweiterten Anregungssignal ($U_{HB1}(k)$, E401);
 - Skalieren (E406) des erweiterten Signals durch einen Verstärkungsfaktor, der je Teilrahmen in Abhängigkeit von einem Verhältnis zwischen dem Verhältnis zwischen der Energie je Teilrahmen und der Energie je Rahmen des Anregungssignals des Tiefbandes und dem Verhältnis zwischen der Energie je Teilrahmen und der Energie je Rahmen des erweiterten Signals definiert wird;
 - Filterung (E404) des skalierten erweiterten Signals durch ein lineares Prädiktionsfilter, dessen Koeffizienten von den Koeffizienten des Filters des Tiefbandes abgeleitet sind.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es ferner vor dem Schritt des Skalierens einen Schritt der adaptiven Bandpassfilterung (E405) des erweiterten Signals in Abhängigkeit von der Decodierungsleistung des aktuellen Rahmens umfasst.
3. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es einen Schritt der Zeit-Frequenz-Transformation des Anregungssignals, wobei der Schritt des Erhaltens eines erweiterten Signals in dem Frequenzbereich stattfindet, und einen Schritt der inversen Zeit-Frequenz-Transformation des erweiterten Signals vor den Schritten des Skalierens und der Filterung umfasst.
4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der Schritt des Generierens eines überabgetasteten und erweiterten Anregungssignals nach der folgenden Gleichung durchgeführt wird:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

wobei k der Index des Samples, $U_{HB1}(k)$ das Spektrum des erweiterten Anregungssignals, $U(k)$ das Spektrum des nach dem Transformationsschritt erhaltenen Anregungssignals und $start_band$ eine vorgegebene Variable ist.

5. Verfahren nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet, dass** es einen Schritt der Deakzentuierungsfaltung des erweiterten Signals mindestens in dem zweiten Frequenzband umfasst.
6. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es ferner einen Schritt des Erzeugens (E402) eines Rauschsignals mindestens in dem zweiten Frequenzband umfasst, wobei das erweiterte Signal ($U_{HB2}(k)$) durch Kombinieren (E403) des erweiterten Anregungssignals und des Rauschsignals erhalten wird.
7. Verfahren nach Anspruch 6, **dadurch gekennzeichnet, dass** der Schritt des Kombinierens durch adaptives addi-

tives Mischen mit einem Entzerrungsverstärkungsfaktor zwischen dem erweiterten Anregungssignal und dem Rauschsignal durchgeführt wird.

8. Vorrichtung zur Frequenzbänderweiterung eines Audiosignals, umfassend eine Stufe zur Decodierung oder Extraktion, in einem ersten Frequenzband, Tiefband genannt, eines Anregungssignals und der Koeffizienten eines linearen Prädiktionsfilters, wobei die Vorrichtung **dadurch gekennzeichnet ist, dass** sie umfasst:

- ein Modul zum Erhalten eines erweiterten Signals ($U_{HB2}(k)$, 503) in mindestens einem zweiten Frequenzband, das über dem ersten Frequenzband liegt, ausgehend von dem in dem mindestens einen zweiten Frequenzband überabgetasteten und erweiterten Anregungssignal ($U_{HEB1}(k)$) ;
- Skalieren (507) des erweiterten Signals durch einen Verstärkungsfaktor, der je Teilrahmen in Abhängigkeit von einem Verhältnis zwischen dem Verhältnis zwischen der Energie je Teilrahmen und der Energie je Rahmen des Anregungssignals des Tiefbandes und dem Verhältnis zwischen der Energie je Teilrahmen und der Energie je Rahmen des erweiterten Signals definiert wird;
- ein Modul zur Filterung (510) des skalierten erweiterten Signals durch ein lineares Prädiktionsfilter, dessen Koeffizienten von den Koeffizienten des Filters des Tiefbandes abgeleitet sind.

9. Audiosignaldecoder, **dadurch gekennzeichnet, dass** er eine Vorrichtung zur Frequenzbänderweiterung nach Anspruch 8 umfasst.

10. Computerprogramm mit Codeanweisungen zum Umsetzen der Schritte des Frequenzbänderweiterungsverfahrens nach einem der Ansprüche 1 bis 7, wenn diese Anweisungen von einem Prozessor ausgeführt werden.

11. Speichermedium, das von einer Vorrichtung zur Frequenzbänderweiterung gelesen werden kann, auf dem ein Computerprogramm gespeichert ist, das Codeanweisungen zum Umsetzen der Schritte des Verfahrens zur Frequenzbänderweiterung nach einem der Ansprüche 1 bis 7 umfasst.

Claims

1. Method for extending a frequency band of an audio frequency signal in a decoding or improvement process, including a step of decoding or extracting, in a first frequency band, called low band, an excitation signal and coefficients of a linear prediction filter, the method being **characterized in that** it includes the following steps:

- obtaining an extended signal ($U_{HB2}(k)$, E403) in at least one second frequency band higher than the first frequency band from the excitation signal upsampled and extended into the at least one second frequency band ($U_{HB1}(k)$, E401);
- scaling (E406) the extended signal by a defined gain per subframe as a function of a ratio between the ratio between the energy per subframe and the energy per frame of the excitation signal of the low band and the ratio between the energy per subframe and the energy per frame of the extended signal;
- filtering (E404) said scaled extended signal using a linear prediction filter whose coefficients are derived from the coefficients of the low band filter.

2. Method according to Claim 1, **characterized in that** it furthermore includes a step (E405) of adaptive bandpass filtering, as a function of the decoding bit rate of the current frame, on the extended signal before the scaling step.

3. Method according to Claim 1, **characterized in that** it includes a time-frequency transform step on the excitation signal, the step of obtaining an extended signal then being performed in the frequency domain, and an inverse time-frequency transform step on the extended signal before the scaling and filtering steps.

4. Method according to Claim 3, **characterized in that** the step of generating an upsampled and extended excitation signal is performed using the following equation:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

where k is the index of the sample, $U_{HB1}(k)$ is the spectrum of the extended excitation signal, $U(k)$ is the spectrum of the excitation signal obtained after the transform step and $start_band$ is a predefined variable.

5. Method according to one of Claims 1 to 4, **characterized in that** it includes a de-emphasis filtering step on the extended signal at least in the second frequency band.
6. Method according to Claim 1, **characterized in that** it furthermore includes a step (E402) of generating a noise signal at least in the second frequency band, the extended signal ($U_{HB2}(k)$) being obtained by combining (E403) the extended excitation signal and the noise signal.
7. Method according to Claim 6, **characterized in that** the combination step is performed by adaptive additive mixing with an equalization gain of a level between the extended excitation signal and the noise signal.
8. Device for extending a frequency band of an audio frequency signal including a stage for decoding or extracting, in a first frequency band, called low band, an excitation signal and coefficients of a linear prediction filter, the device being **characterized in that** it includes:
 - a module for obtaining an extended signal ($U_{HB2}(k)$, 503) in at least one second frequency band higher than the first frequency band from the excitation signal upsampled and extended into the at least one second frequency band ($U_{HB1}(k)$) ;
 - scaling (507) the extended signal by a defined gain per subframe as a function of a ratio between the ratio between the energy per subframe and the energy per frame of the excitation signal of the low band and the ratio between the energy per subframe and the energy per frame of the extended signal;
 - a module (510) for filtering said scaled extended signal using a linear prediction filter whose coefficients are derived from the coefficients of the low band filter.
9. Audio frequency signal decoder, **characterized in that** it includes a frequency band extension device according to Claim 8.
10. Computer program including code instructions for implementing the steps of the frequency band extension method according to one of Claims 1 to 7 when these instructions are executed by a processor.
11. Storage medium able to be read by a frequency band extension device and on which there is recorded a computer program comprising code instructions for executing the steps of the frequency band extension method according to one of Claims 1 to 7.

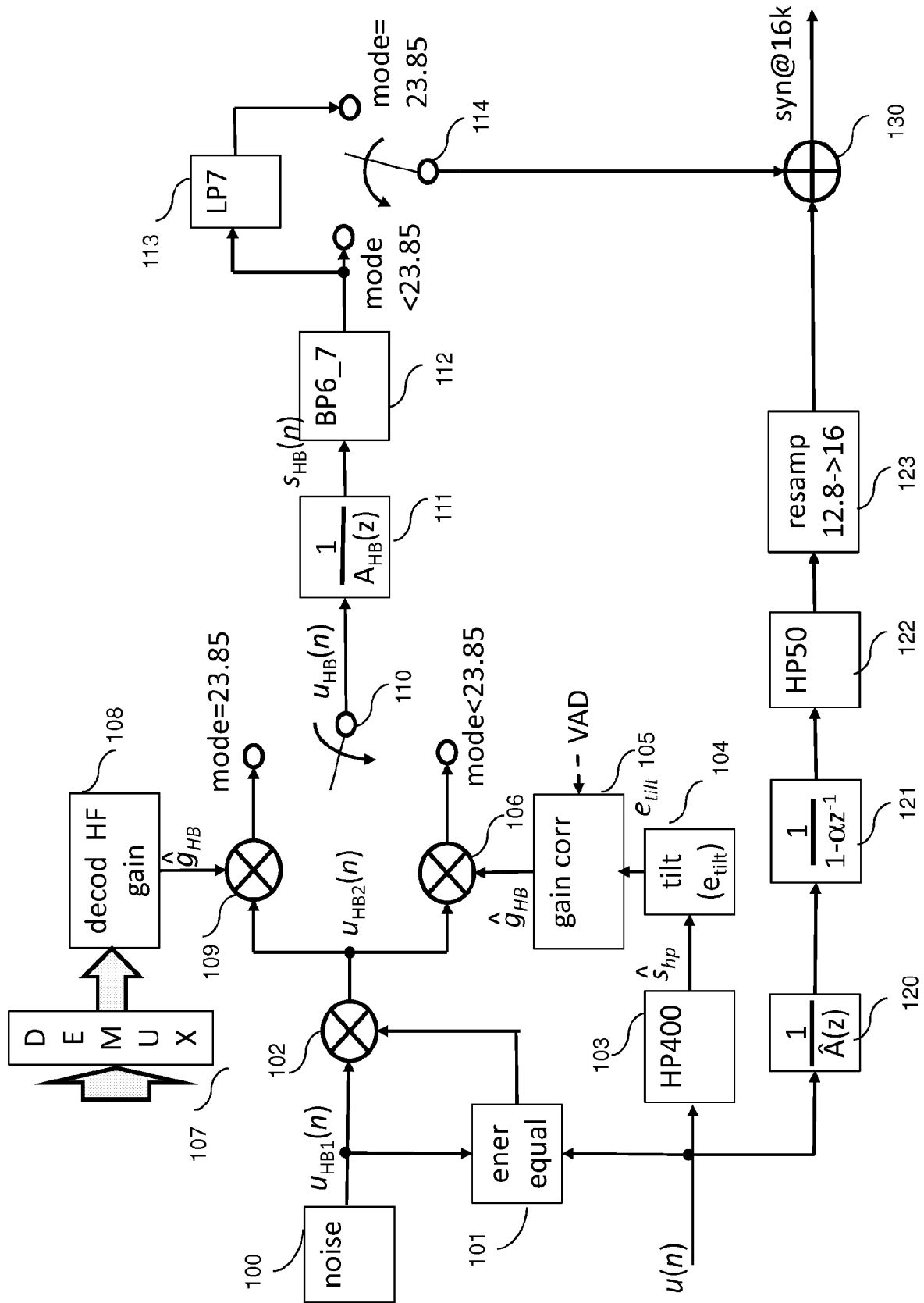


Fig.1

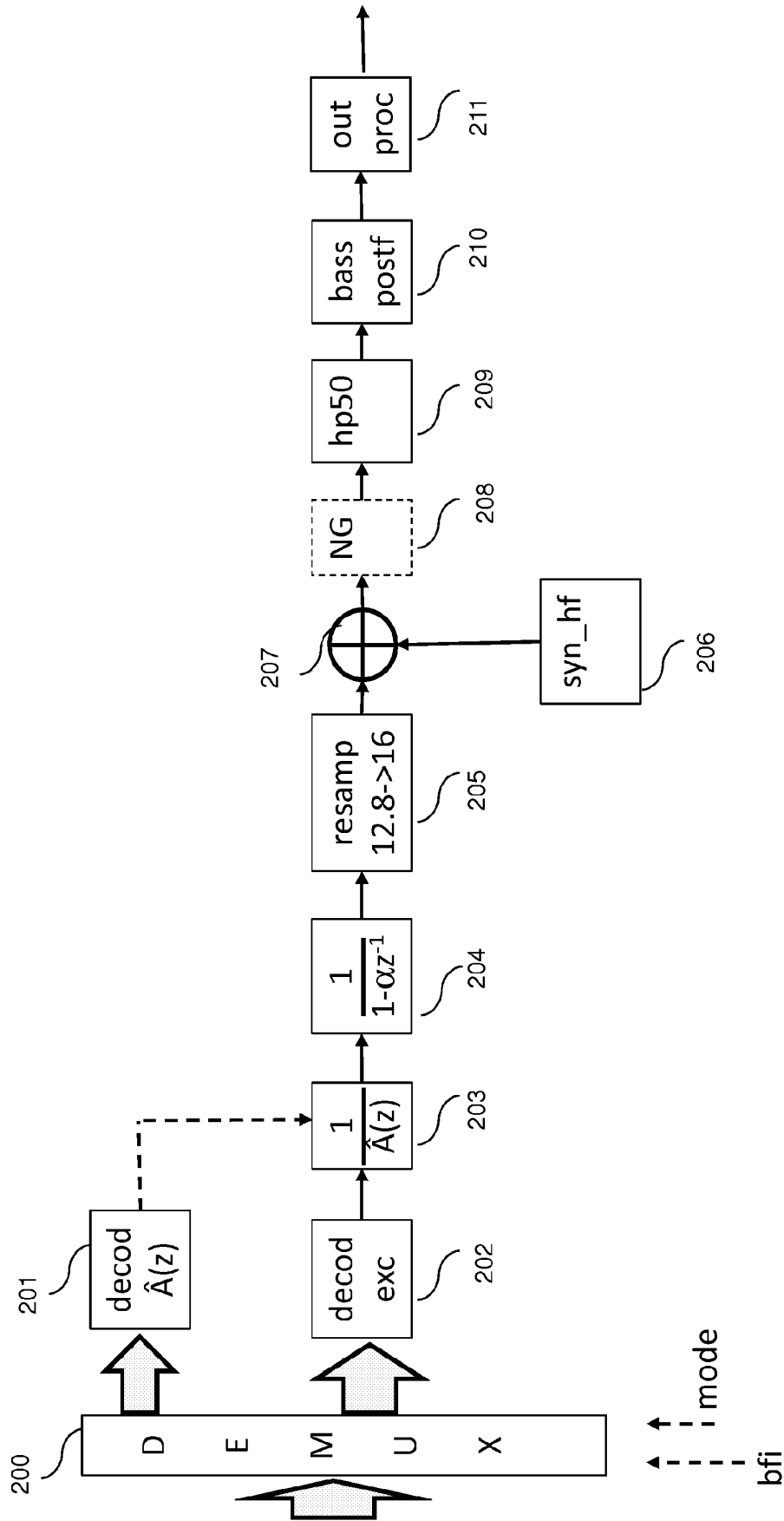


Fig.2

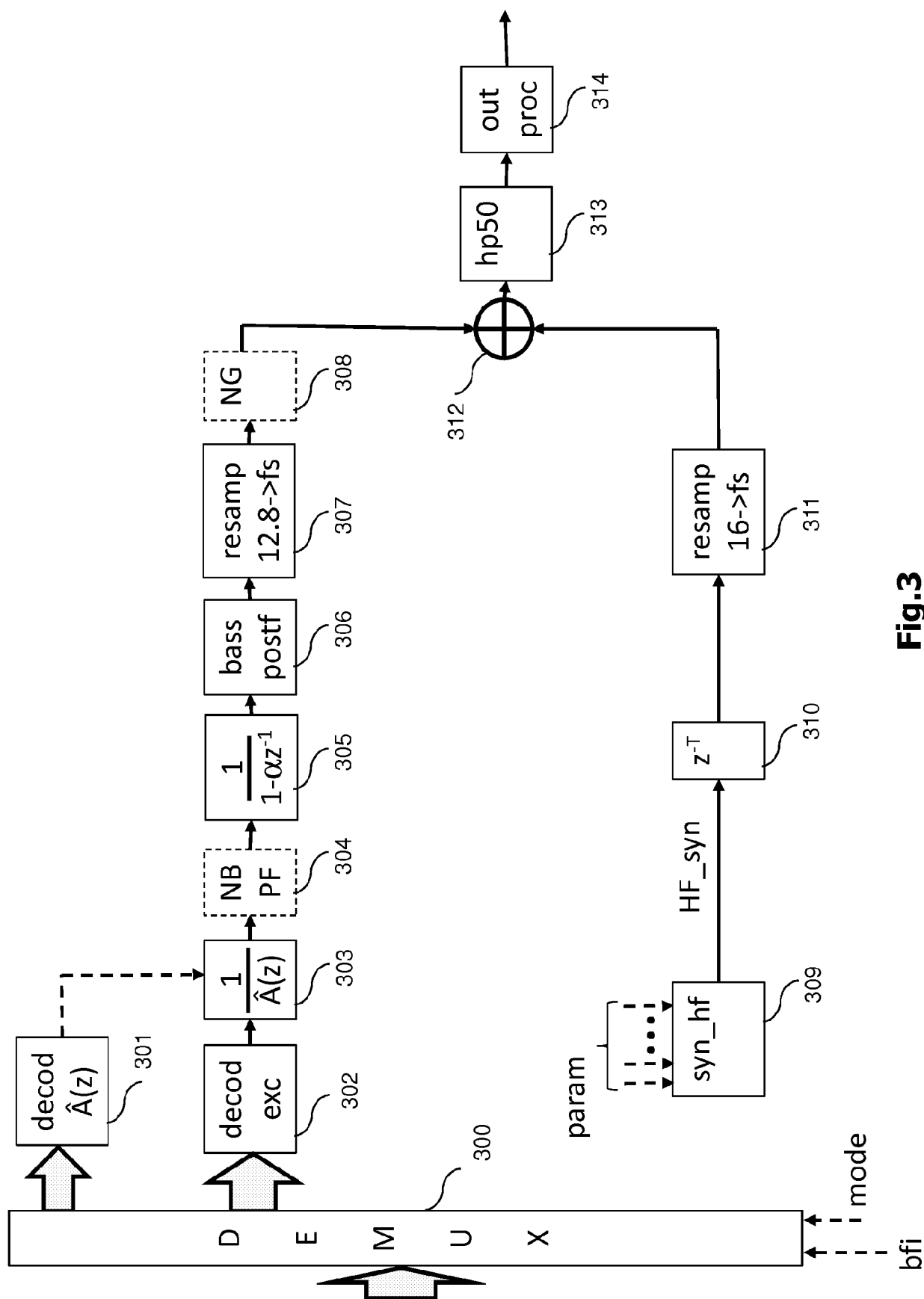


Fig.3

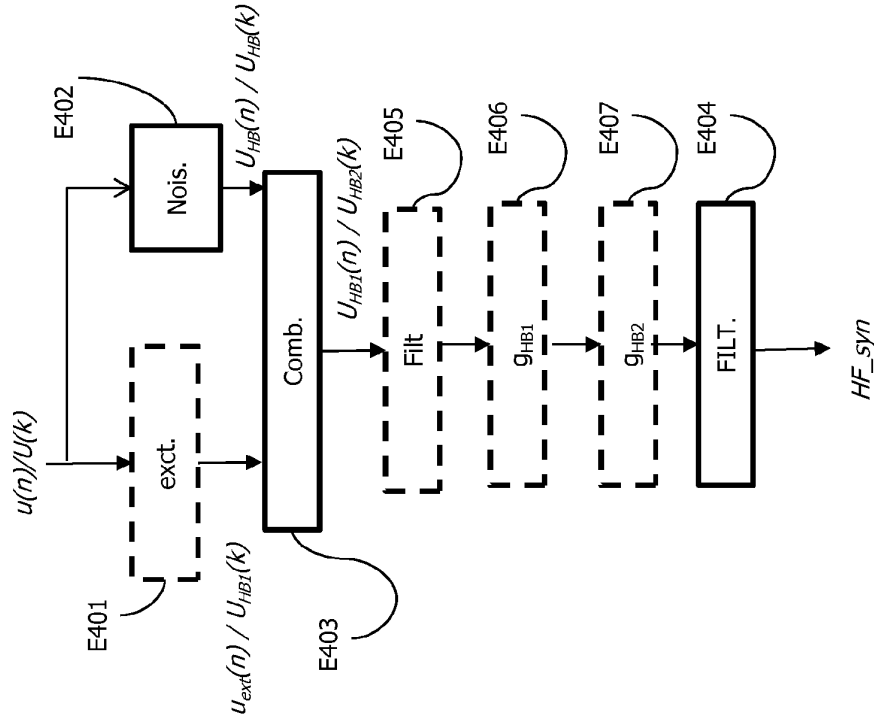


Fig.4

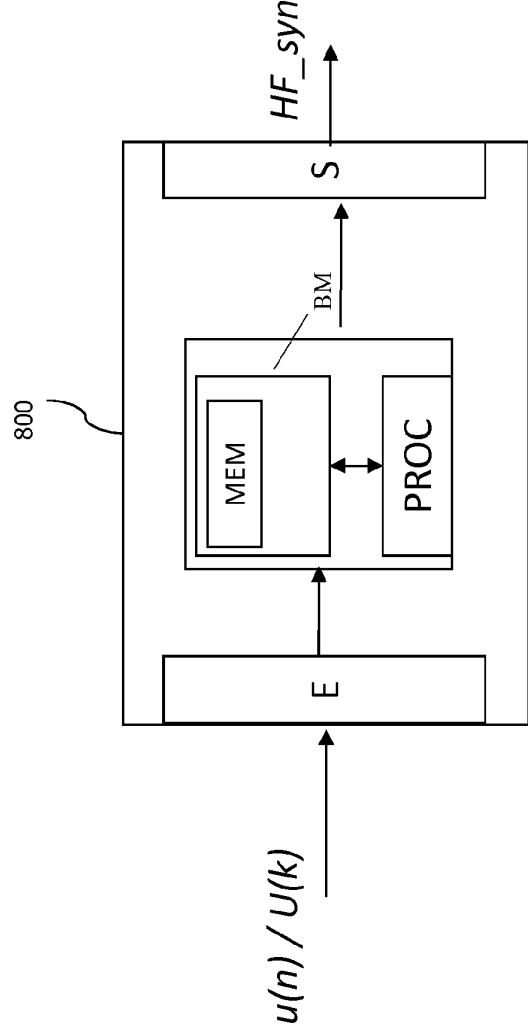


Fig.8

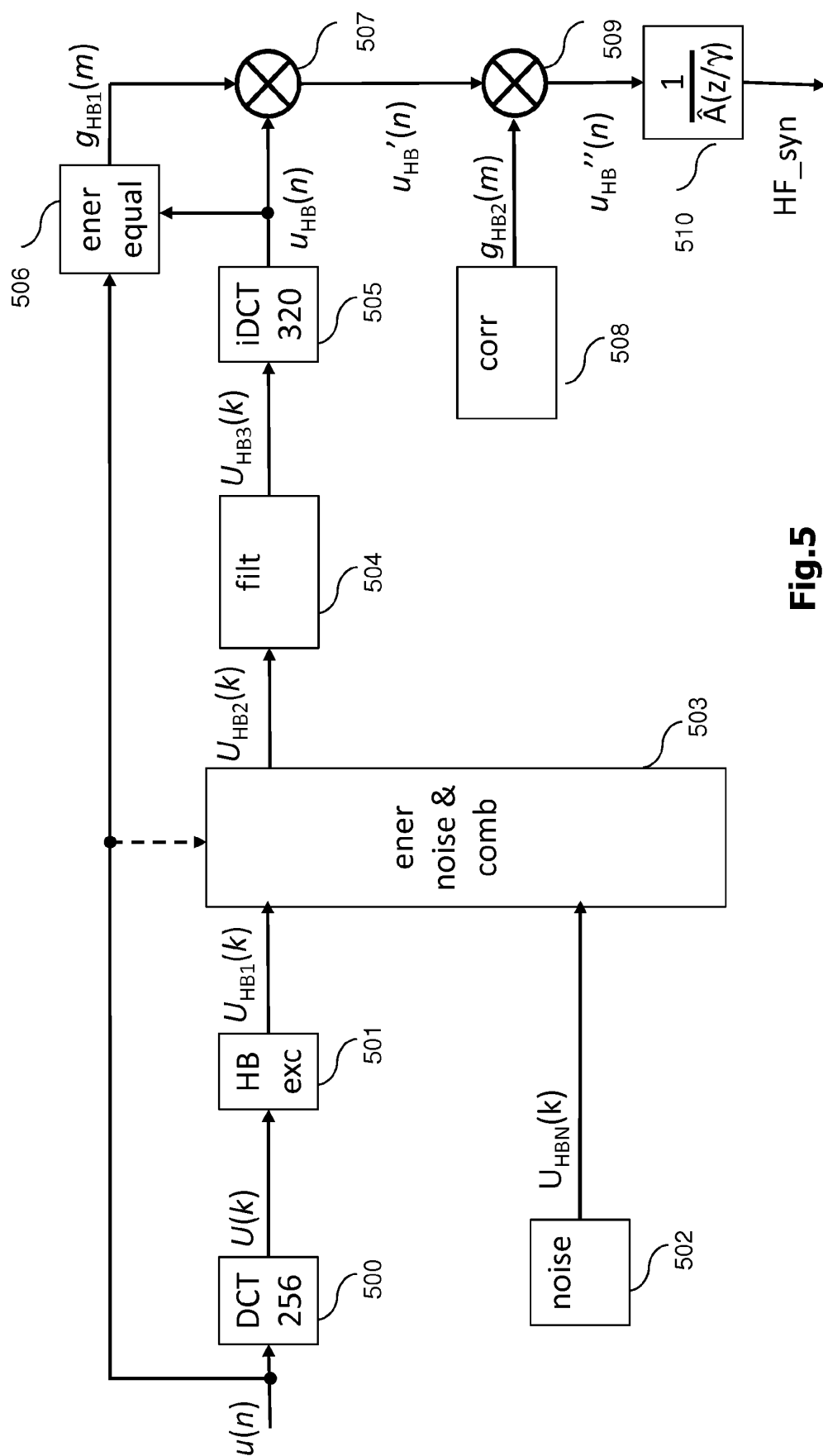


Fig.5

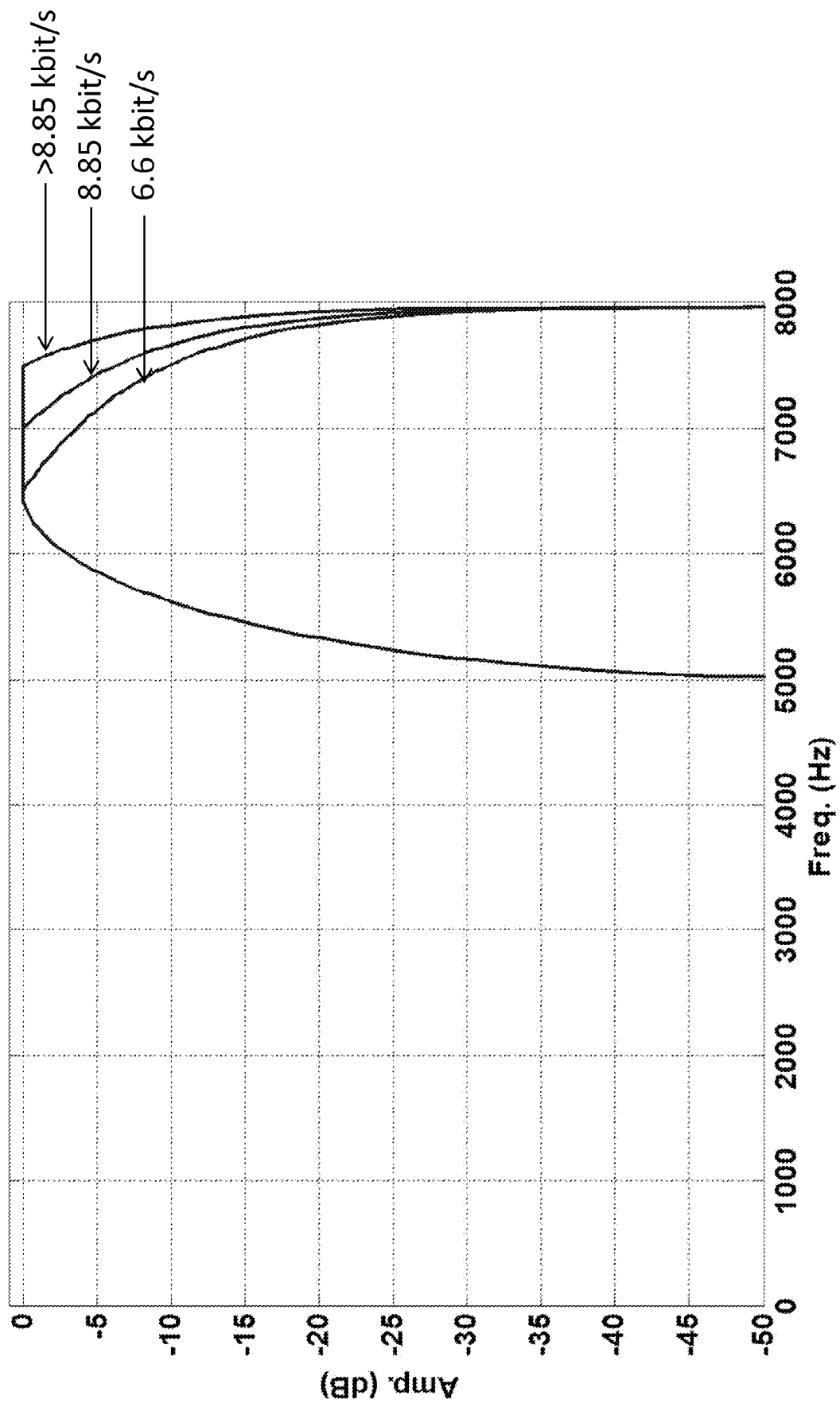


Fig.6

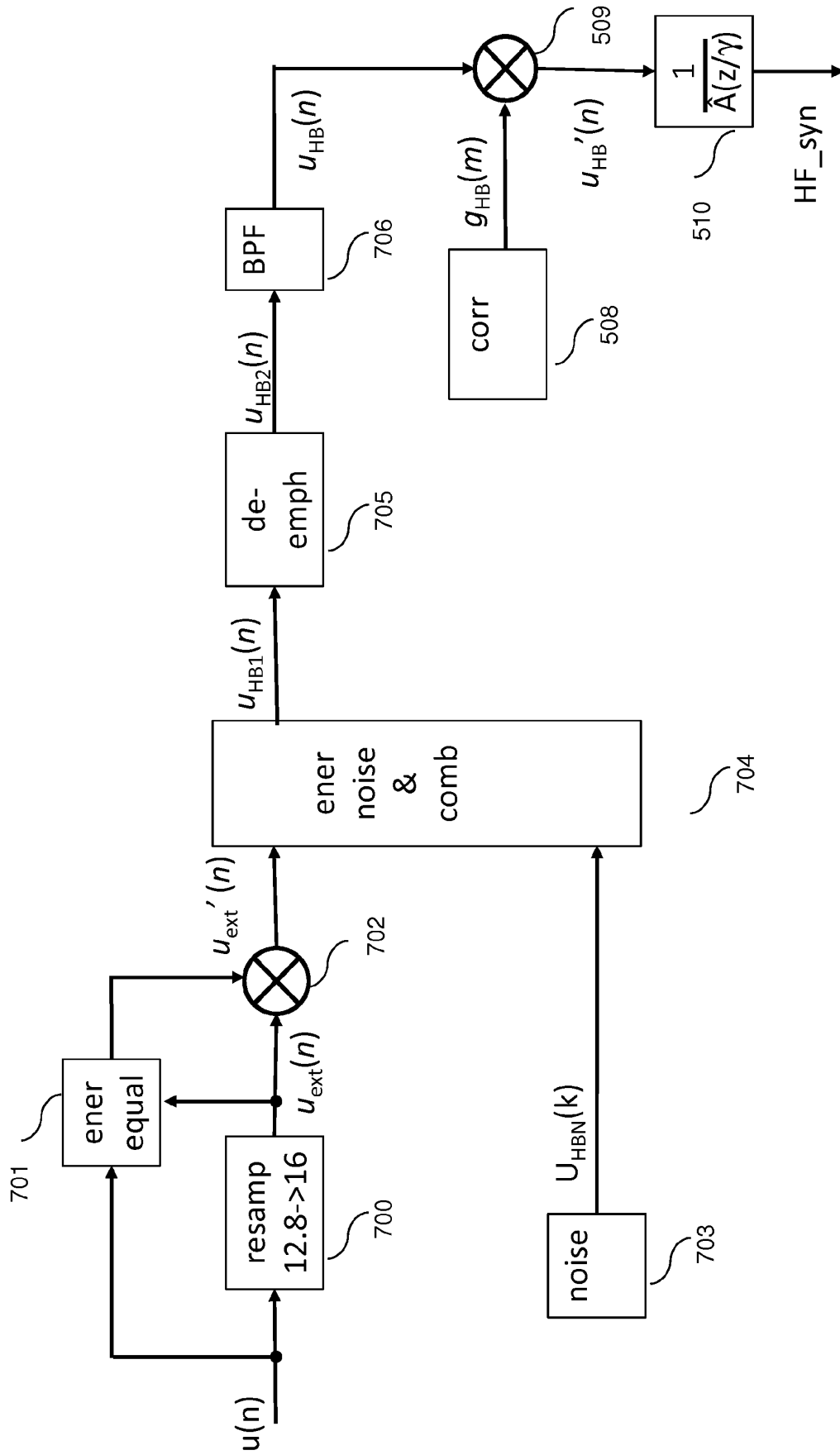


Fig.7

RÉFÉRENCES CITÉES DANS LA DESCRIPTION

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

Documents brevets cités dans la description

- US 2003050786 A1 [0022]

Littérature non-brevet citée dans la description

- Speech Coding and Synthesis. Elsevier, 1995 [0005]
- **M. BOSI ; R.E. GOLDBERG.** Introduction to Digital Audio Coding and Standards. Springer, 2002 [0005]
- Handbook of Speech Processing. Springer, 2008 [0005]
- **B. BESSETTE et al.** The adaptive multirate wide-band speech codec (AMR-WB). *IEEE Transactions on Speech and Audio Processing*, 2002, vol. 10 (8), 620-636 [0009]
- **DE D.M. ZHANG ; H.T. LI.** A Low Complexity Transform - Evolved DCT. *IEEE 14th International Conference on Computational Science and Engineering (CSE)*, Août 2011, 144-149 [0090]