



(11)

EP 3 020 043 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:

08.02.2017 Bulletin 2017/06

(51) Int Cl.:

G10L 21/038 ^(2013.01)

(86) Numéro de dépôt international:

PCT/FR2014/051720

(21) Numéro de dépôt: **14749907.3**

(22) Date de dépôt: **04.07.2014**

(87) Numéro de publication internationale:

WO 2015/004373 (15.01.2015 Gazette 2015/02)

(54) **FACTEUR D'ÉCHELLE OPTIMISÉ POUR L'EXTENSION DE BANDE DE FRÉQUENCE DANS UN DÉCODEUR DE SIGNAUX AUDIOFRÉQUENCES**

OPTIMIERTER SKALENFAKTOR FÜR FREQUENZBANDERWEITERUNG BEI EINEM
AUDIOFREQUENZSIGNALDECODIERER

OPTIMIZED SCALE FACTOR FOR FREQUENCY BAND EXTENSION IN AN AUDIOFREQUENCY
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: **12.07.2013 FR 1356909**

(43) Date de publication de la demande:

18.05.2016 Bulletin 2016/20

(73) Titulaire: **Orange**

75015 Paris (FR)

(72) Inventeurs:

- **KANIEWSKA, Magdalena**
3001 Leuven (BE)
- **RAGOT, Stéphane**
F-Lannion 22300 (FR)

(74) Mandataire: **Froger, Marie-Hélène**

Orange

IMTW/OLPS/IPL/PAT

Orange Gardens

44 Avenue de la République

CS 50010

92326 Châtillon Cedex (FR)

(56) Documents cités:

US-A1- 2009 201 983 US-A1- 2011 099 004

- **"Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions (3GPP TS 26.290 version 11.0.0 Release 11)", TECHNICAL SPECIFICATION, EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE (ETSI), 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS ; FRANCE, vol. 3GPP SA 4, no. V11.0.0, 1 octobre 2012 (2012-10-01), XP014075402, cité dans la demande**

EP 3 020 043 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 de détermination d'un facteur d'échelle optimisé servant à ajuster le niveau d'un signal d'excitation ou de façon équivalente d'un filtre lors d'une 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 Détection") et génération de bruit de confort (CNG pour "Confort Noise Génération") 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]. En effet, pour les signaux dont énergie croît quand la fréquence augmente (e_{tilt} proche de -1, g_{SP} proche de 2), le gain $\hat{g}_{HB}$ est d'habitude sous-estimé.

[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). 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 "Immittance 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].

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 122 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.

[0014] On peut identifier plusieurs inconvénients à la technique d'extension de bande du codec AMR-WB, en particulier :

- 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). 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.
- 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.

Un exemple d'extension de bande par approche temporelle est décrit dans la norme 3GPP TS 26.290 décrivant le codec AMR-WB+ (normalisé en 2005). Cet exemple est illustré dans les schémas-blocs des figures 2a (schéma global) et 2b (prédiction de gain par correction de niveau de réponses) qui correspondent respectivement aux figures 16 et 10 de la spécification 3GPP TS 26.290.

Dans le codec AMR-WB+, le signal d'entrée (mono) échantillonné à la fréquence F_s (en Hz) est divisé en deux bandes de fréquences disjointes, dans lesquelles deux filtres LPC sont calculés et codés séparément :

- un filtre LPC, noté $A(z)$, dans la bande basse ($0-F_s/4$) - sa version quantifiée est notée $\hat{A}(z)$
- un autre filtre LPC, noté $A_{HF}(z)$, dans la bande haute repliée spectralement ($F_s/4-F_s/2$) - sa version quantifiée est notée $\hat{A}_{HF}(z)$

L'extension de bande se fait dans le codec AMR-WB+ comme détaillé dans les sections 5.4 (codage HF) et 6.2 (décodage HF) de la spécification 3GPP TS 26.290. On en résume ici le principe : l'extension consiste à utiliser l'excitation décodée en basses fréquences (LF excit.) et à mettre en forme cette excitation par un gain temporel par sous-trame (bloc 205) et un filtrage LPC de synthèse (bloc 207) ; de plus, des traitements d'améliorations (post-traitement de l'excitation (bloc 206) et lissage de l'énergie du signal HF reconstruit (bloc 208) sont mis en oeuvre comme illustré à la figure 2a.

Il est important de remarquer que cette extension dans AMR-WB+ nécessite la transmission d'informations supplémentaires : les coefficients du filtre $\hat{A}_{HF}(z)$ en 204 et un gain de mise en forme temporelle par sous-trame (bloc 201). Une particularité de l'algorithme d'extension de bande dans AMR-WB+ est que le gain par sous-trame est quantifié par une approche prédictive ; autrement dit, on ne code pas les gains directement, mais plutôt des corrections de gain qui sont relatives à une estimation du gain notée g_{match} . Cette estimation, g_{match} , correspondant en fait à un facteur

d'égalisation du niveau entre les filtres $\hat{A}(z)$ et $\hat{A}_{HF}(z)$ à la fréquence de séparation entre bande basse et bande haute ($F_s/4$). Le calcul du facteur g_{match} (bloc 203) est détaillé à la figure 10 de la spécification 3GPP TS 26.290 reprise ici à la figure 2b. On ne détaillera pas plus ici cette figure. On retiendra pour résumer que les blocs 210 à 213 servent à

calculer l'énergie de la réponse impulsionnelle de $\frac{\hat{A}(z)}{(1 - 0.9z^{-1})\hat{A}_{HF}(z)}$,

en se rappelant que le filtre $\hat{A}_{HF}(z)$ modélise une bande haute repliée spectrale (à cause des propriétés spectrales du banc de filtre séparant les bandes basse et haute). Puisque les filtres sont interpolés par sous-trames, le gain g_{match} n'est calculé qu'une fois par trame, et il est interpolé par sous-trames.

La technique de codage des gains d'extension de bande dans AMR-WB+, et plus précisément la compensation de niveaux des filtres LPC en leur point de jonction, est une méthode adaptée dans le contexte d'une extension de bande par modèles LPC en bande basse et haute, et on peut remarquer qu'une telle compensation de niveau entre filtres LPC n'est pas présente dans l'extension de bande du codec AMR-WB. Cependant, on peut vérifier dans la pratique que l'égalisation directe du niveau entre les deux filtres LPC à la fréquence de séparation n'est pas une méthode optimale et peut provoquer une surestimation d'énergie en bande-haute et des artefacts audibles dans certains cas ; on rappelle qu'un filtre LPC représente une enveloppe spectrale, ainsi le principe de l'égalisation du niveau entre deux filtres LPC pour une fréquence donnée revient à ajuster le niveau relatif de deux enveloppes LPC. Or une telle égalisation réalisée en une fréquence précise n'assure pas une complète continuité et cohérence globale de l'énergie (en fréquence) au voisinage du point d'égalisation lorsque l'enveloppe fréquentielle du signal fluctue de façon significative dans ce voisinage. Une façon mathématique de poser le problème consiste à remarquer que la continuité entre deux courbes peut être assurée en les forçant à se rejoindre en un même point, mais rien ne garantit que les propriétés locales (dérivées successives) coïncident de façon à assurer une cohérence plus globale. Le risque en assurant une continuité ponctuelle entre des enveloppes LPC bandes basse et haute est de fixer l'enveloppe de LPC en bande haute à un niveau relatif trop fort ou trop faible, le cas d'un niveau trop fort étant plus dommageable car il résulte en des artefacts plus gênants. Par ailleurs, la compensation de gain dans AMR-WB+ est avant tout une prédiction du gain connue au codeur et au décodeur et qui sert à réduire le débit nécessaire à la transmission d'information de gain mettant à l'échelle le signal d'excitation bande haute. Or, dans le contexte d'une amélioration du codage/décodage AMR-WB de façon interopérable, il n'est pas possible de modifier le codage existant des gains par sous-trames (0.8 kbit/s) de l'extension de bande dans le mode 23.85 kbit/s d'AMR-WB. De plus, pour les débits strictement inférieurs à 23.85 kbit/s, la compensation de niveaux de filtres LPC en bandes basse et haute peut être appliquée dans l'extension de bande d'un décodage compatible avec AMR-WB, cependant l'expérience montre que cette seule technique dérivée du codage AMR-WB+, appliquée sans optimisation, peut engendrer des problèmes de surestimation d'énergie de la bande haute (>6 kHz).

Il existe donc un besoin pour améliorer la compensation de gains entre des filtres de prédiction linéaire de bande de fréquences différentes pour l'extension de bande de fréquence dans un codec de type AMR-WB ou une version interopérable de ce codec sans pour autant surestimer l'énergie dans une bande de fréquence et sans nécessiter d'informations supplémentaires du codeur.

[0015] La présente invention vient améliorer la situation.

[0016] A cet effet, l'invention vise un procédé de détermination d'un facteur d'échelle optimisé à appliquer à un signal d'excitation ou à un filtre lors d'un procédé d'extension de bande de fréquence d'un signal audiofréquence, le procédé d'extension de bande comportant une étape de décodage ou d'extraction, dans une première bande de fréquence, d'un signal d'excitation et de paramètres de la première bande de fréquence comprenant des coefficients d'un filtre de prédiction linéaire, une étape de génération d'un signal d'excitation étendu sur au moins une deuxième bande de fréquence et une étape de filtrage par un filtre de prédiction linéaire pour la deuxième bande de fréquence. Le procédé de détermination est tel qu'il comporte les étapes suivantes:

- détermination d'un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur au filtre de prédiction linéaire de la première bande de fréquence, les coefficients du filtre additionnel étant obtenus à partir des paramètres décodés ou extraits de la première bande de fréquence; et
- calcul du facteur d'échelle optimisé en fonction au moins des coefficients du filtre additionnel.

[0017] Ainsi, l'utilisation d'un filtre additionnel d'ordre inférieur au filtre de la première bande de fréquence à égaliser, permet d'éviter les surestimations d'énergie dans les hautes fréquences qui pourraient résulter de fluctuations locales de l'enveloppe et qui peuvent perturber l'égalisation des filtres de prédiction.

[0018] L'égalisation de gains entre les filtres de prédiction linéaire de la première et deuxième bande de fréquence est ainsi améliorée.

[0019] Dans une application avantageuse du facteur d'échelle optimisé ainsi obtenu, le procédé d'extension de bande comprend une étape d'application du facteur d'échelle optimisé au signal d'excitation étendu.

[0020] Dans un mode de réalisation adapté, l'application du facteur d'échelle optimisé est combinée à l'étape de filtrage dans la deuxième bande de fréquence.

[0021] Ainsi les étapes de filtrage et d'application du facteur d'échelle optimisé sont combinées à une seule étape de filtrage pour réduire la complexité de traitement.

5 **[0022]** Dans un mode de réalisation particulier, les coefficients du filtre additionnel sont obtenus par troncature de la fonction de transfert du filtre de prédiction linéaire de la première bande de fréquence pour obtenir un ordre inférieur.

[0023] Ce filtre additionnel d'ordre inférieur est donc obtenu de façon simple.

[0024] De plus, de façon à obtenir un filtre stable, les coefficients du filtre additionnel sont modifiés en fonction d'un critère de stabilité du filtre additionnel.

10 **[0025]** Dans un mode de réalisation particulier, le calcul du facteur d'échelle optimisé comporte les étapes suivantes:

- calcul des réponses en fréquence des filtres de prédiction linéaire des première et deuxième bandes de fréquence pour une fréquence commune;
- calcul de la réponse en fréquence du filtre additionnel pour cette fréquence commune;
- 15 - calcul du facteur d'échelle optimisé en fonction des réponses en fréquence ainsi calculées.

[0026] Ainsi, le facteur d'échelle optimisé est calculé de façon à éviter les artefacts gênants qui pourraient survenir dans le cas où la réponse en fréquence de filtre d'ordre supérieur de la première bande à proximité de la fréquence commune révélerait un pic ou une vallée du signal.

20 **[0027]** Dans un mode de réalisation particulier, le procédé comprend en outre les étapes suivantes, mises en oeuvre pour un débit de décodage prédéterminé:

- première mise à l'échelle du signal d'excitation étendu par un gain calculé par sous-trame fonction d'un rapport d'énergie entre le signal d'excitation décodé et le signal d'excitation étendu;
- 25 - deuxième mise à l'échelle du signal d'excitation issu de la première mise à l'échelle par un gain de correction décodé;
- ajustement de l'énergie de l'excitation pour la sous-trame courante par un facteur d'ajustement calculé en fonction de l'énergie du signal obtenu après la deuxième mise à l'échelle et en fonction du signal obtenu après application du facteur d'échelle optimisé.

30 **[0028]** Ainsi, des informations supplémentaires peuvent être utilisées pour améliorer la qualité du signal étendu pour un mode de fonctionnement prédéterminé.

[0029] L'invention vise également un dispositif de détermination d'un facteur d'échelle optimisé à appliquer à un signal d'excitation ou à un filtre dans un dispositif d'extension de bande de fréquence d'un signal audiofréquence, le dispositif d'extension de bande comprenant un module de décodage ou d'extraction, dans une première bande de fréquence, d'un signal d'excitation et de paramètres de la première bande de fréquence comprenant des coefficients d'un filtre de prédiction linéaire, un module de génération d'un signal d'excitation étendu sur au moins une deuxième bande de fréquence et un module de filtrage par un filtre de prédiction linéaire pour la deuxième bande de fréquence. Le dispositif de détermination est tel qu'il comprend:

- 40 - un module de détermination d'un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur au filtre de prédiction linéaire de la première bande de fréquence, les coefficients du filtre additionnel étant obtenus à partir des paramètres décodés ou extraits de la première bande de fréquence; et
- un module de calcul du facteur d'échelle optimisé en fonction au moins des coefficients du filtre additionnel.

45 **[0030]** L'invention vise un décodeur comportant un dispositif tel que décrit.

[0031] Elle vise un programme informatique comportant des instructions de code pour la mise en oeuvre des étapes du procédé de détermination d'un facteur d'échelle optimisé tel que décrit, lorsque ces instructions sont exécutées par un processeur.

50 **[0032]** Enfin l'invention se rapporte à un support de stockage, lisible par un processeur, intégré ou non au dispositif de détermination d'un facteur d'échelle optimisé, éventuellement amovible, mémorisant un programme informatique mettant en oeuvre un procédé de détermination d'un facteur d'échelle optimisé tel que décrit précédemment.

[0033] 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 :

- 55 - 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;
- les figures 2a et 2b présentent le codage de la bande haute dans le codec AMR-WB+ 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 utilisé selon un mode de réalisation de l'invention ;
- la figure 4 illustre un dispositif de détermination d'un facteur d'échelle optimisé par sous-trame en fonction du débit, selon un mode de réalisation de l'invention; et
- les figures 5a et 5b illustrent les réponses en fréquences des filtres utilisées pour le calcul du facteur d'échelle optimisé selon un mode de réalisation de l'invention;
- la figure 6 illustre sous forme d'organigramme, les étapes principales d'un procédé de détermination d'un facteur d'échelle optimisé selon un mode de réalisation de l'invention;
- la figure 7 illustre un mode de réalisation dans le domaine fréquentiel d'un dispositif de détermination de facteur d'échelle optimisé lors d'une extension de bande;
- la figure 8 illustre une réalisation matérielle d'un dispositif de détermination de facteur d'échelle optimisé lors d'une extension de bande selon l'invention.

[0034] La figure 3 illustre un exemple de décodeur, compatible avec la norme AMR-WB/G.722.2 dans lequel on retrouve une extension de bande comprenant une détermination d'un facteur d'échelle optimisé selon un mode de réalisation du procédé de l'invention, mis en oeuvre par le dispositif d'extension de bande illustré par le bloc 309.

[0035] Contrairement au décodage AMR-WB qui fonctionne avec une fréquence d'échantillonnage de sortie de 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 avec 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.

[0036] 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 l'extension de bande (HF pour hautes fréquences) utilisée pour 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 réalisation, la combinaison des bandes basse et haute pourra se faire à 16 kHz, après avoir ré-échantillonné la bande basse de 12.8 à 16 kHz, avant de ré-échantillonner le signal combiné à la fréquence f_s .

[0037] 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:

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

en suivant les notations de la clause 7.1.2.1 de la recommandation ITU-T G.718 d'un décodeur interopérable avec le codeur/décodeur AMR-WB, 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.

- 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 (dit « *bass postfilter* ») (bloc 306) atténuant le bruit inter-harmonique en basses fréquences 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 pour « améliorer » la qualité des silences par réduction du niveau.

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 de l'extension de bande.

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

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'extension de bande 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.

[0038] Dans le mode de réalisation décrit ici et en référence à la figure 7, , le décodeur 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 6400Hz et d'une deuxième bande de fréquence de 6400 à 8000Hz. 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.

[0039] A 23.85 kbit/s, l'information de correction de gain HF (0.8 kbit/s) transmise à 23.85 kbit/s est ici décodée. Son utilisation est détaillée plus loin, en référence à la figure 4. La partie synthèse bande haute est réalisée dans le bloc 309 représentant le dispositif d'extension de bande utilisé pour l'invention et qui est détaillé à la figure 7 dans un mode de réalisation.

[0040] Afin d'aligner les bandes basses et hautes décodées, un retard (bloc 310) est introduit 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). La valeur du retard T dépend de la façon de synthétiser le signal bande haute, de la fréquence f_s de même que du post-traitement des basses fréquences. Ainsi, de façon générale la valeur de T dans le bloc 310 devra être ajustée en fonction de la mise en oeuvre spécifique.

[0041] 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 "noisegate" de façon similaire à G.718 (bloc 314).

[0042] En référence à la figure 3, on décrit à présent un mode de réalisation d'un dispositif de détermination d'un facteur d'échelle optimisé à appliquer à un signal d'excitation lors d'un processus d'extension de bande de fréquence. Ce dispositif est inclus dans le bloc d'extension de bande 309 décrit précédemment.

[0043] Ainsi, le bloc 400, à partir d'un signal d'excitation décodé dans une première bande de fréquence $u(n)$, effectue une extension de bande pour obtenir un signal d'excitation étendu $u_{HB}(n)$ sur au moins une deuxième bande de fréquence.

[0044] On notera ici, que l'estimation de facteur d'échelle optimisé selon l'invention est indépendante de la façon d'obtenir le signal $u_{HB}(n)$. Une condition concernant son énergie est cependant importante. En effet, il faut que l'énergie de la bande haute de 6000 à 8000 Hz soit à un niveau similaire à l'énergie de la bande de 4000 à 6000 Hz du signal d'excitation décodé à la sortie du bloc 302. De plus, puisque le signal bande basse est désaccentué (bloc 305), il faut aussi appliquer la désaccentuation au signal d'excitation bande haute, soit en utilisant un filtre de désaccentuation propre, soit en multipliant par un facteur constant qui correspond à une atténuation moyenne du filtre mentionné. Cette condition ne s'applique pas au cas du débit 23.85 kbit/s qui utilise les informations supplémentaires transmises par le codeur. Dans ce cas l'énergie du signal d'excitation bande haute doit être cohérente avec l'énergie du signal correspondant au codeur, comme expliqué plus tard.

[0045] L'extension de bande de fréquence peut par exemple être mise en oeuvre de la même façon que pour le décodeur de type AMR-WB décrit en référence à la figure 1 dans les blocs 100 à 102, à partir d'un bruit blanc.

[0046] Dans un autre mode de réalisation, cette extension de bande peut s'effectuer à partir d'une combinaison d'un bruit blanc et d'un signal d'excitation décodé comme illustré et décrit ultérieurement pour les blocs 700 à 707 de la figure 7.

[0047] D'autres méthodes d'extension de bande de fréquence avec conservation du niveau d'énergie entre le signal d'excitation décodé et le signal d'excitation étendu comme décrit ci-dessous, peuvent bien sur être envisagées pour le bloc 400.

[0048] De plus, le module d'extension de bande peut également être indépendant du décodeur et peut effectuer une extension de bande d'un signal audio existant stocké ou transmis au module d'extension, avec une analyse du signal audio pour en extraire une excitation et un filtre LPC. Dans ce cas, le signal d'excitation en entrée du module d'extension, n'est plus un signal décodé mais un signal extrait après analyse, de même que les coefficients du filtre de prédiction linéaire de la première bande de fréquence utilisés dans le procédé de détermination du facteur d'échelle optimisé dans une mise en oeuvre de l'invention.

[0049] Dans l'exemple illustré à la figure 4, on considère tout d'abord le cas des débits <23.85kbit/s, pour lequel la détermination du facteur d'échelle optimisé se limite au bloc 401. On calcule dans ce cas un facteur d'échelle optimisé, noté $g_{HB2}(m)$. Dans un mode de réalisation, ce calcul est effectué de façon préférentielle par sous-trame et il consiste à égaliser les niveaux des réponses en fréquences des filtres LPC $1/\hat{A}(z)$ et $1/\hat{A}(z/\gamma)$ utilisés en basses et hautes fréquences, comme décrit ultérieurement en référence à la figure 7, avec des précautions supplémentaires pour éviter les cas de surestimations qui peuvent résulter en une énergie trop grande de la bande haute synthétisée et générer donc des artefacts audibles.

Dans un mode de réalisation alternatif, on pourra garder le filtre de synthèse HF extrapolé $1/\hat{A}^{ext}(z/\gamma)$ tel que mis en oeuvre dans le décodeur AMR-WB ou un décodeur interopérable avec le codeur/décodeur AMR-WB, par exemple selon la recommandation ITU-T G.718, au lieu du filtre $1/\hat{A}(z/\gamma)$. La compensation selon l'invention s'effectue alors à partir des filtres $1/\hat{A}(z)$ et $1/\hat{A}^{ext}(z/\gamma)$.

La détermination du facteur d'échelle optimisé s'effectue aussi par la détermination (en 401a) d'un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur au filtre de prédiction linéaire de la première bande de fréquence $1/\hat{A}(z)$, les coefficients du filtre additionnel étant obtenus à partir des paramètres décodés ou extraits de la première bande de fréquence. Le facteur d'échelle optimisé est ensuite calculé (en 401b) en fonction au moins de ces coefficients pour être appliqué au signal d'excitation étendu $u_{HB}(n)$.

[0050] Le principe de la détermination du facteur d'échelle optimisé, mis en oeuvre dans le bloc 401 est illustré aux **figures 5a et 5b** avec des exemples concrets obtenus à partir de signaux échantillonnés à 16 kHz ; les valeurs d'amplitude de réponses en fréquence, notées plus loin R , P , Q , de 3 filtres sont calculées à la fréquence commune de 6000 Hz (ligne verticale en pointillés) dans la sous-trame courante, dont l'indice m n'est pas rappelé ici dans la notations des filtres LPC interpolés par sous-trame pour alléger le texte. La valeur de 6000 Hz est choisie de sorte qu'elle soit proche de la fréquence de Nyquist de la bande basse, soit 6400 Hz. Il est préférable de ne pas prendre cette fréquence de Nyquist pour déterminer le facteur d'échelle optimisé. En effet l'énergie du signal décodé en basses fréquences est typiquement déjà atténuée à 6400 Hz. De plus, l'extension de bande décrite ici est effectuée sur une deuxième bande de fréquence dite bande haute qui va de 6000 à 8000 Hz. A noter que dans des variantes de l'invention, une autre fréquence que 6000 Hz pourra être choisie, sans perte de généralité pour déterminer le facteur d'échelle optimisé. On pourra aussi considérer le cas où les deux filtres LPC sont définis pour les bandes séparés (comme dans AMR-WB+). Dans ce cas R , P et Q seront calculées à la fréquence de séparation.

Les figures 5a et 5b illustrent comment sont définies les quantités R , P , Q .

La première étape consiste à calculer les réponses en fréquences R et P respectivement du filtre de prédiction linéaire de la première bande de fréquence (bande basse) et de la deuxième bande de fréquence (bande haute) à la fréquence de 6000 Hz. On calcule d'abord :

$$R = \frac{1}{|\hat{A}(e^{j\theta})|} = \frac{1}{\left| \sum_{i=0}^M \hat{a}_i e^{-ji\theta} \right|}$$

où $M = 16$ est l'ordre du filtre LPC décodé $1/\hat{A}(z)$, et θ correspond à la fréquence de 6000 Hz normalisée pour la fréquence d'échantillonnage de 12.8 kHz, soit :

$$\theta = 2\pi \frac{6000}{12800}.$$

Ensuite, de façon similaire, on calcule :

$$P = \frac{1}{\left| \hat{A}(e^{j\theta'} / \gamma) \right|} = \frac{1}{\left| \sum_{i=0}^M \hat{a}_i \gamma^i e^{-ji\theta'} \right|}$$

où

$$\theta' = 2\pi \frac{6000}{16000}.$$

Dans un mode de réalisation privilégié, les quantités P et R sont calculées selon le pseudo-code suivant:

```

15      px = py = 0
      rx = ry = 0
      for i=0 to 16
20          px = px + Ap[i]*exp_tab_p[i]
          py = py + Ap[i]*exp_tab_p[33-i]
          rx = rx + Aq[i]*exp_tab_q[i]
          ry = ry + Aq[i]*exp_tab_q[33-i]
25      end for
      P = 1/sqrt(px*px+py*py)
      R = 1/sqrt(rx*rx+ry*ry)
30

```

où $Aq[i] = \hat{a}_i$ correspond aux coefficients de $\hat{A}(z)$ (d'ordre 16), $Ap[i] = \gamma^i \hat{a}_i$ correspond aux coefficient de $\hat{A}(z/\gamma)$, sqrt() correspond à l'opération de racine carrée et les tableaux exp_tab_p et exp_tab_q de taille 34 contiennent les parties réelles et imaginaires des exponentielles complexes associée à la fréquence de 6000 Hz, avec

$$\text{exp_tab_p}[i] = \begin{cases} \cos\left(2\pi \frac{6000}{12800} i\right) & i = 0, \dots, 16 \\ -\sin\left(2\pi \frac{6000}{12800} (33-i)\right) & i = 17, \dots, 33 \end{cases}$$

$$\text{exp_tab_q}[i] = \begin{cases} \cos\left(2\pi \frac{6000}{16000} i\right) & i = 0, \dots, 16 \\ -\sin\left(2\pi \frac{6000}{16000} (33-i)\right) & i = 17, \dots, 33 \end{cases}$$

Le filtre de prédiction additionnel est obtenu par exemple en tronquant de façon adéquate le polynôme $\hat{A}(z)$ à l'ordre 2. En fait la troncature directe à l'ordre conduit au filtre $1 + \hat{a}_1 + \hat{a}_2$, ce qui peut poser problème car rien ne garantit en général que ce filtre d'ordre 2 est stable. Dans un mode de réalisation privilégiée, on détecte donc la stabilité du filtre $1 + \hat{a}_1 + \hat{a}_2$ et on utilise un filtre $1 + \hat{a}_1' + \hat{a}_2'$,

dont les coefficients sont tirés de $1 + \hat{a}_1 + \hat{a}_2$ en fonction de la détection d'instabilité. Plus précisément, on initialise :

$$\hat{a}_i' = \hat{a}_i, i=1, 2$$

La stabilité du filtre $1 + \hat{a}_1 + \hat{a}_2$ peut être vérifiée de différente façon, on utilise ici une conversion dans le domaine des coefficients PARCOR (ou coefficients de réflexion) en calculant :

$$k_1 = \hat{a}_1' / (1 + \hat{a}_2')$$

$$k_2 = \hat{a}_2'$$

La stabilité est vérifiée si $|k_i| < 1$, $i=1,2$. On modifie donc de façon conditionnelle la valeur de k_i avant d'assurer la stabilité du filtre, avec les étapes suivantes :

$$k_2 \leftarrow \begin{cases} \min(0.6, k_2) & k_2 > 0 \\ \max(-0.6, k_2) & k_2 < 0 \end{cases}$$

$$k_1 \leftarrow \begin{cases} \min(0.99, k_1) & k_1 > 0 \\ \max(-0.99, k_1) & k_1 < 0 \end{cases}$$

où $\min(.,.)$ et $\max(.,.)$ donnent respectivement le minimum et le maximum de 2 opérandes. On note que les valeurs de seuils, 0.99 pour k_1 et 0.6 pour k_2 , pourront être ajustées dans des variantes de l'invention. On rappelle que le premier coefficient de réflexion, k_1 , caractérise la pente spectrale (ou tilt) du signal modélisé à l'ordre 1 ; dans l'invention on sature la valeur de k_1 à une valeur proche de la limite de stabilité, afin de préserver cette pente et conserver un tilt similaire à celui de $1/\hat{A}(z)$. On rappelle aussi que le second coefficient de réflexion, k_2 , caractérise le niveau de résonance du modèle de signal à l'ordre 2 ; puisque l'utilisation d'un filtre d'ordre 2 vise à éliminer l'influence de telles résonances autour de la fréquence de 6000 Hz, on limite plus fortement la valeur de k_2 , cette limite est fixée à 0.6.

Les coefficients de $1 + \hat{a}_1' + \hat{a}_2'$ sont alors obtenus par:

$$\hat{a}_1' = (1 + k_2)k_1$$

$$\hat{a}_2' = k_2$$

On calcule donc finalement la réponse en fréquence du filtre additionnel:

$$Q = \frac{1}{\left| \sum_{k=0}^2 \hat{a}_k' e^{-jk\theta} \right|}$$

avec $\theta = 2\pi \frac{6000}{12800}$. Cette quantité est calculée de façon préférentielle selon le pseudo-code 12800 suivant :

```

 $qx = qy = 0$ 
for  $i=0$  to 2
     $qx = qx + As[i]*exp\_tab\_q[i];$ 
     $qy = qy + As[i]*exp\_tab\_q[33-i];$ 
end for
 $Q = 1/sqrt(qx*qx+qy*qy)$ 

```

où $As[i] = \hat{a}_i^*$.

Sans perte de généralité, on pourra calculer les coefficients du filtre d'ordre 2 autrement, par exemple en appliquant au filtre LPC $\hat{A}(z)$ d'ordre 16 la procédure de réduction de l'ordre LPC dite « STEP DOWN » décrite dans J.D. Markel and A.H. Gray, Linear Prediction of Speech, Springer Verlag, 1976 ou en effectuant deux itérations d'algorithme de Levinson-Durbin (ou STEP-UP) à partir des autocorrélations calculées sur le signal synthétisé (décodé) à 12.8 kHz et fenêtré.

Pour certains signaux, la quantité Q , calculée à partir des 3 premiers coefficients LPC décodés, prend mieux en compte l'influence de la pente spectrale (ou tilt) dans le spectre et évite l'influence de pics ou de vallées « parasites » proches de 6000 Hz qui peuvent biaiser ou élever la valeur de la quantité R , calculée à partir de tous les coefficients LPC.

Dans un mode de réalisation privilégié, le facteur d'échelle optimisé est déduit des quantités pré-calculées R , P , Q de façon conditionnelle, comme suit :

Si le tilt (calculé comme dans AMR-WB dans le bloc 104, par autocorrélation normalisée sous la forme $r(1)/r(0)$ où $r(i)$ est l'autocorrélation) est négatif (tilt < 0 comme représenté à la figure 5b), le calcul du facteur d'échelle se fait de la façon suivante:

Pour éviter des artefacts dus à des variations trop brusques d'énergie de la bande haute, on applique un lissage à la valeur de R . Dans un mode de réalisation privilégié, un lissage exponentiel est effectué avec un facteur fixe dans le temps (0.5) sous la forme :

$$R = 0.5R + 0.5R_{prev}$$

$$R_{prev} = R$$

où R_{prev} correspond à la valeur de R dans la sous-trame précédente et le facteur 0.5 est optimisé de façon empirique - bien entendu, le facteur 0.5 pourra être changé pour une autre valeur et d'autres méthodes de lissage sont également possibles. A noter que le lissage permet de réduire les variantes temporelles et évite donc des artefacts.

Le facteur d'échelle optimisé est alors donné par :

$$g_{HB2}(m) = \max(\min(R, Q), P) / P$$

Dans un mode de réalisation alternative, on pourra remplacer le lissage de R par un lissage de $g_{HB2}(m)$ tel que:

$$g_{HB2}(m) \leftarrow 0.5g_{HB2}(m) + 0.5g_{HB2}(m-1)$$

Si le tilt (calculé comme dans AMR-WB dans le bloc 104) est positif (tilt > 0 comme à la figure 5a), le calcul du facteur d'échelle se fait de la façon suivante:

La quantité R est lissée de façon adaptative dans le temps, avec un lissage plus fort quand R est faible - comme dans le cas précédent, ce lissage permet de réduire les variantes temporelles et évite donc des artefacts:

$$R = (1 - \alpha)R + \alpha R_{prev} \text{ avec } \alpha = 1 - R^2$$

5

$$R_{prev} = R$$

10

Ensuite, le facteur d'échelle optimisé est donné par :

$$g_{HB2}(m) = \min(R, P, Q) / P$$

15

Dans un mode de réalisation alternative, on pourra remplacer le lissage de R par un lissage de $g_{HB2}(m)$ tel que calculé ci-dessus.

20

$$g_{HB}(m) = (1 - \alpha)g_{HB}(m) + \alpha g_{HB}(m-1), m = 0, \dots, 3, \alpha = 1 - g_{HB}^2(m)$$

ou $g_{HB}(-1)$ est le facteur d'échelle ou gain calculé pour la dernière sous-trame de la trame précédente.

On prend ici le minimum de R, P, Q afin d'éviter de surestimer le facteur d'échelle.

Dans une variante, la condition ci-dessus dépendant uniquement du tilt pourra être étendue pour tenir compte non seulement du paramètre de tilt mais également d'autres paramètres afin d'affiner la décision. De plus, le calcul de $g_{HB2}(m)$ pourra être ajusté en fonction de ces dits paramètres supplémentaires.

25

Un exemple de paramètre supplémentaire est le nombre de passage par zéro (ZCR, zero crossing rate) qui peut être défini comme :

30

$$zcr_s = \frac{1}{2} \sum_{n=1}^{N-1} |\text{sgn}[s(n)] - \text{sgn}[s(n-1)]|$$

où

35

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$$

40

Le paramètre zcr donne généralement les résultats similaires au tilt. Un bon critère de classification est le ratio entre zcr_s calculé pour le signal synthétisé $s(n)$ et zcr_u calculé pour le signal d'excitation $u(n)$ à 12800 Hz. Ce ratio est entre 0 et 1, où 0 signifie que le signal a un spectre décroissant, 1 que le spectre est croissant (ce qui correspond à $(1 - \text{tilt})/2$). Dans ce cas, un ratio $zcr_s / zcr_u > 0.5$ correspond au cas $\text{tilt} < 0$, un ratio $zcr_s / zcr_u < 0.5$ correspond à $\text{tilt} > 0$.

45

Dans une variante, on pourra utiliser une fonction d'un paramètre tilt_{hp} où tilt_{hp} est le tilt calculé pour le signal synthétisé $s(n)$ filtré par un filtre passe haut avec une fréquence de coupure par exemple à 4800 Hz ; dans ce cas, la réponse $1/\hat{A}(z/\gamma)$ de 6 à 8 kHz (appliquée à 16 kHz) correspond à la réponse pondérée de $1/\hat{A}(z)$ de 4.8 à 6.4 kHz. Comme $1/\hat{A}(z/\gamma)$ a une réponse plus aplatie, il faut compenser ce changement de tilt. La fonction de facteur d'échelle selon tilt_{hp} est alors donnée dans un mode de réalisation par : $(1 - \text{tilt}_{hp})^2 + 0.6$. On multiplie donc Q et R par $\min(1, (1 - \text{tilt}_{hp})^2 + 0.6)$ quand $\text{tilt} > 0$

50

ou par $\max(1, (1 - \text{tilt}_{hp})^2 + 0.6)$ quand $\text{tilt} < 0$.

[0051] On considère maintenant le cas du débit de 23.85kbit/s, pour lequel une correction de gain est réalisée par les blocs 403 à 408. Cette correction de gain pourrait d'ailleurs faire l'objet d'une invention séparée. Dans ce mode particulier selon l'invention, l'information de correction de gain, notée $g_{HBcorr}(m)$, transmise par le codage (compatible) AMR-WB avec un débit de 0.8 kbit/s est utilisée pour améliorer la qualité à 23.85 kbit/s.

55

On suppose ici que le codage (compatible) AMR-WB a effectué une quantification de gain de correction sur 4 bits comme décrit dans la clause UIT-T G.722.2/5.11 ou de façon équivalente dans la clause 3GPP TS 26.190/5.11.

Dans le codeur AMR-WB, le gain de correction est calculé en comparant l'énergie du signal original échantillonné à 16

kHz et filtré par un filtre passe-bande 6-7 kHz, $s_{HB}(n)$ avec l'énergie du bruit blanc à 16 kHz filtré par un filtre de synthèse $1/\hat{A}(z/\gamma)$ et un filtre passe-bande 6-7 kHz (avant le filtrage l'énergie du bruit est mis à un niveau semblable à celui de l'excitation à 12.8 kHz), $s_{HB2}(n)$. Le gain est la racine du ratio d'énergie du signal original sur l'énergie du bruit divisé par deux. Dans un mode possible de réalisation on pourra changer le filtre passe-bande pour un filtre avec une bande plus large (par exemple de 6 à 7.6 kHz).

$$g_{HBcorr}(m) = \sqrt{\frac{\sum_{n=80m}^{80(m+1)-1} s_{HB}(n)^2}{\sum_{n=80m}^{80(m+1)-1} s_{HB2}(n)^2}}, m = 0, \dots, 3$$

Pour pouvoir appliquer l'information de gain reçue à 23.85 kbit/s (dans le bloc 407), il est important de ramener l'excitation à un niveau similaire à celui attendu au codage (compatible) AMR-WB. Ainsi, le bloc 404 effectue la mise à l'échelle du signal d'excitation selon l'équation suivante:

$$u_{HB1}(n) = g_{HB3}(m) u_{HB}(n), n = 80m, \dots, 80(m+1) - 1$$

où $g_{HB3}(m)$ est un gain par sous-trame calculé dans le bloc 403 sous la forme :

$$g_{HB3}(m) = \sqrt{\frac{\sum_{n=0}^{63} u(n)^2}{5 \cdot \sum_{n=0}^{79} u_{HB}(n)^2}}$$

où le facteur 5 au dénominateur sert à compenser la différence de largeur de bande entre le signal $u(n)$ et le signal $u_{HB}(n)$, sachant qu'au codage AMR-WB l'excitation HF est un bruit blanc sur la bande 0-8000 Hz. L'indice de 4 bits par sous-trame, noté $\text{index}_{HF_gain}(m)$, envoyé à 23.85 kbit/s est démultiplexé du train binaire (bloc 405) et décodé par le bloc 406 de la façon suivante :

$$g_{HBcorr}(m) = 2 \cdot HP_gain(\text{index}_{HF_gain}(m))$$

où $HP_gain(.)$ est le dictionnaire de quantification de gain HF défini dans le codage AMR-WB et rappelé ci-dessous :

Tableau 1 (dictionnaire de gain à 23.85 kbit/s)

i	$HP_gain(i)$	l	$HP_gain(l)$
0	0.110595703125000	8	0.342102050781250
1	0.142608642578125	9	0.372497558593750
2	0.170806884765625	10	0.408660888671875
3	0.197723388671875	11	0.453002929687500
4	0.226593017578125	12	0.511779785156250
5	0.255676269531250	13	0.599822998046875f
6	0.284545898437500	14	0.741241455078125
7	0.313232421875000	15	0.998779296875000

Le bloc 407 effectue la mise à l'échelle du signal d'excitation selon l'équation suivante:

$$u_{HB2}(n) = g_{HBcorr}(m)u_{HB1}(n), n = 80m, \dots, 80(m+1) - 1$$

Enfin, on ajuste l'énergie de l'excitation au niveau de la sous-trame courante avec les conditions suivantes (bloc 408).
On calcule :

$$fac(m) = \sqrt{\frac{\sum_{n=0}^{79} (g(m)g_{HB2}(m)u_{HB}(n))^2}{\sum_{n=0}^{79} u_{HB2}(n)^2}}$$

Le numérateur représente ici l'énergie de signal bande-haute qui serait obtenue dans le mode 23.05. Comme expliqué avant, pour les débits <23.85 kbit/s il faut conserver le niveau d'énergie entre le signal d'excitation décodé et le signal d'excitation étendu $u_{HB}(n)$, mais cette contrainte n'est pas nécessaire dans le cas du débit de 23.85 kbit/s, puisque $u_{HB}(n)$ est dans ce cas mis à l'échelle par le gain $g_{HB3}(m)$. Pour éviter les doubles multiplications certaines opérations de multiplications appliqués au signal dans le bloc 400 sont appliquées dans le bloc 402 en multipliant par $g(m)$. La valeur de $g(m)$ dépend de l'algorithme de synthèse de $u_{HB}(n)$ et doit être ajusté de telle sorte que le niveau d'énergie entre le signal d'excitation décodé en bande basse et le signal $g(m)u_{HB}(n)$ soit conservé.

Dans un mode de réalisation particulier, qui sera décrit en détail plus tard en référence à la figure 7, $g(m) = 0.6g_{HB1}(m)$, où $g_{HB1}(m)$ est un gain qui assure, pour le signal u_{HB} , le même ratio entre énergie par sous-trame et énergie par trame que pour le signal $u(n)$ et 0.6 correspond à la valeur moyenne d'amplitude de réponse en fréquence du filtre de désaccentuation de 5000 à 6400 Hz.

On suppose que dans le bloc 408 on dispose d'une information sur le tilt du signal bande basse - dans un mode de réalisation privilégié ce tilt est calculé comme dans le codec AMR-WB selon les blocs 103 et 104, cependant d'autres méthodes d'estimation du tilt sont possibles sans changer le principe de l'invention.

Si $fac(m) > 1$ ou tilt < 0, on prend :

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

Sinon :

$$u_{HB}'(n) = \max(\sqrt{1 - \text{tilt}}, fac(m)) \cdot u_{HB2}(n), n = 80m, \dots, 80(m+1) - 1$$

On notera que le calcul de facteur d'échelle optimisé présenté ici, notamment dans les blocs 401 et 402, se distingue de l'égalisation précitée de niveaux de filtres effectuée dans le codec AMR-WB+ par plusieurs aspects :

- Le facteur d'échelle optimisé est calculé directement à partir des fonctions de transfert des filtres LPC sans impliquer de filtrage temporel. Ceci simplifie le procédé.
- L'égalisation est faite de préférence à une fréquence différente de la fréquence de Nyquist (6400 Hz) associée à la bande basse. En effet, la modélisation LPC représente implicitement l'atténuation du signal typiquement causée par les opérations de ré-échantillonnage et donc la réponse en fréquence d'un filtre LPC peut subir à la fréquence de Nyquist une diminution qui ne se retrouve pas à la fréquence commune choisie.
- L'égalisation repose ici sur un filtre d'ordre moins élevé (ici d'ordre 2) en plus des 2 filtres à égaliser. Ce filtre additionnel permet d'éviter les effets de fluctuations spectrales locales (pic ou vallée) qui peuvent être présentes à la fréquence commune pour le calcul de la réponse en fréquence des filtres de prédiction.

Pour les blocs 403 à 408, l'avantage de l'invention est que la qualité du signal décodé à 23.85 kbit/s selon l'invention est améliorée par rapport à un signal décodé à 23.05 kbit/s, ce qui n'est pas le cas dans un décodeur AMR-WB. En fait, cet aspect de l'invention permet d'utiliser l'information supplémentaire (0.8 kbit/s) reçue à 23.85 kbit/s, mais de façon contrôlée (bloc 408), pour améliorer la qualité du signal d'excitation étendu au débit de 23.85.

Le dispositif de détermination du facteur d'échelle optimisé tel qu'illustré par les blocs 401 à 408 de la figure 4, met en oeuvre un procédé de détermination du facteur d'échelle optimisé décrit maintenant en référence à la figure 6.

[0052] Les étapes principales sont mises en oeuvre par le bloc 401.

[0053] Ainsi, un signal d'excitation étendu $u_{HB}(n)$ est obtenu lors d'un procédé d'extension de bande de fréquence E601 qui comporte une étape de décodage ou d'extraction, dans une première bande de fréquence dite bande basse, d'un signal d'excitation et de paramètres de la première bande de fréquence comme par exemple les coefficients du filtre de prédiction linéaire de la première bande de fréquence.

[0054] Une étape E602 détermine un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur à celui de la première bande de fréquence. Pour déterminer ce filtre, les paramètres de la première bande de fréquence décodés ou extraits sont utilisés.

[0055] Dans un mode de réalisation cette étape est effectuée par troncature de la fonction de transfert du filtre de prédiction linéaire de la bande basse pour obtenir un ordre de filtre inférieur, par exemple 2. Ces coefficients peuvent ensuite être modifiés en fonction d'un critère de stabilité comme expliqué précédemment en référence à la figure 4.

[0056] A partir des coefficients du filtre additionnel ainsi déterminé, une étape E603 est mise en oeuvre pour calculer le facteur d'échelle optimisé à appliquer au signal d'excitation étendu. Ce facteur d'échelle optimisé est par exemple calculé à partir de la réponse en fréquence du filtre additionnel à une fréquence commune entre la bande basse (première bande de fréquence) et la bande haute (deuxième bande de fréquence). Une valeur minimale pouvant être choisie entre la réponse en fréquence de ce filtre et celles des filtres bande basse et bande haute.

Ceci évite donc les surestimations d'énergie qui pouvaient exister dans les méthodes de l'état de l'art.

[0057] Cette étape de calcul du facteur d'échelle optimisé est par exemple décrite précédemment en référence à la figure 4 et aux figures 5a et 5b.

[0058] L'étape E604 réalisée par le bloc 402 ou 409 (selon le débit de décodage) pour l'extension de bande, applique le facteur d'échelle optimisé ainsi calculée au signal d'excitation étendu de façon à obtenir un signal d'excitation étendu optimisé $u_{HB}'(n)$.

[0059] Dans un mode de réalisation particulier, le dispositif de détermination du facteur d'échelle optimisé 708 est intégré dans un dispositif d'extension de bande décrit maintenant en référence à la **figure 7**. Ce dispositif de détermination du facteur d'échelle optimisé illustré par le bloc 708 met en oeuvre le procédé de détermination du facteur d'échelle optimisé décrit précédemment en référence à la figure 6.

[0060] Dans ce mode de réalisation, le bloc 400 d'extension de bande de la figure 4 comprend les blocs 700 à 707 de la figure 7 décrite maintenant.

[0061] Ainsi, à l'entrée du dispositif d'extension de bande, 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 de la figure 3.

[0062] On notera que dans ce mode de réalisation, la génération de l'excitation suréchantillonné 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).

[0063] Ainsi, la génération d'un 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.

[0064] 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.

[0065] 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 700) 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 $k = 0, \dots, 255$.

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.

[0066] 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.

[0067] 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.

[0068] Le spectre DCT, $U(k)$, de 256 échantillons couvrant la bande 0-6400 Hz (à 12.8 kHz), est ensuite étendu (bloc 701) 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 $start_band = 160$.

[0069] Le bloc 701 fonctionne comme module de génération d'un signal d'excitation sur-échantillonné et étendu et réalise un ré-échantillonnage de 12.8 à 16 kHz dans le domaine fréquentiel, en rajoutant $\frac{1}{4}$ d'échantillons ($k = 240, \dots, 319$) au spectre, le ratio entre 16 et 12.8 étant de 5/4.

[0070] De plus, le bloc 701 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 704 mais pourrait être réalisée séparément en dehors du bloc 704. 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.

[0071] Dans cet exemple de réalisation et selon la définition de $U_{HE1}(k)$, 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.

[0072] 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.

[0073] Dans une variante du mode de réalisation, la valeur de $start_band$ pourra être rendue adaptative autour de la valeur de 160. 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.

[0074] 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 le bloc 702 de génération de bruit, 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 703.

[0075] 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.

[0076] Le bloc 703 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 .

[0077] 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 associées à du bruit, soit (en appliquant la négation de la condition précédente):

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

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.

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

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

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. Dans un mode de réalisation privilégié, on prend

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

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

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:

$$\beta \leftarrow \beta \cdot f(\alpha)$$

$$\alpha \leftarrow \alpha \cdot f(\alpha)$$

où $f(\alpha)$ est une fonction décroissante de α , par exemple $f(\alpha) = b - a\sqrt{\alpha}$, $b = 1.1$, $a = 1.2$, $f(\alpha)$ 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 .

On remarque donc ici que le bloc 703 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.

[0078] Dans une variante simple, on peut considérer une réalisation du bloc 703, 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

[0079] Le bloc 704 réalise de façon optionnelle, une double opération d'application de réponse en fréquence de filtre passe-bande et de filtrage de désaccentuation (ou déemphasis) dans le domaine fréquentiel.

[0080] Dans une variante de l'invention, le filtrage de désaccentuation pourra être réalisé dans le domaine temporel, après le bloc 705 voire avant le bloc 700 ; cependant, dans ce cas, le filtrage passe-bande réalisé dans le bloc 704 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 :

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

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.

[0081] 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 de la figure 3. Ceci est important pour l'estimation et l'ajustement ultérieur de l'énergie de la synthèse HF.

[0082] 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.

[0083] 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.

[0084] 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).

[0085] Ce filtrage est effectué dans le domaine fréquentiel.

[0086] 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 :

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

La définition de $G_{hp}(k)$, $k = 0, \dots, 55$, est donnée par exemple au tableau 2 ci-dessous.

Tableau 2

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

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.

[0087] On notera aussi que le filtrage passe-bande pourra être adapté en définissant une seule étape de filtrage combinant les filtrages passe-haut et passe-bas.

[0088] 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. 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.

[0089] On notera aussi que dans le cas du débit de 23.85 kbit/s on n'effectue pas la désaccentuation de l'excitation $U_{HB2}(k)$ pour rester en accord avec la façon dont le gain de correction est calculé dans le codeur AMR-WB et pour éviter les doubles multiplications. Dans ce cas block 704 réalise seulement le filtrage passe-bas.

[0090] Le bloc 705 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 700, car la DCT-IV est orthonormée, sauf que la longueur de la transformée est de 320 au lieu de 256, et on obtient:

$$u_{HB0}(n) = \sum_{k=0}^{N_{16k}-1} U_{HB3}(k) \cos\left(\frac{\pi}{N_{16k}}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right)$$

où $N_{16k} = 320$ et $k = 0, \dots, 319$. Cette excitation échantillonnée à 16 kHz est ensuite de façon optionnelle mise à l'échelle

par des gains définis par sous-trame de 80 échantillons (bloc 707).

Dans un mode de réalisation privilégié, on calcule d'abord (bloc 706) 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

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

$$e_2(m) = \sum_{n=0}^{79} u_{HB0}(n+80m)^2 + \varepsilon$$

$$e_3(m) = e_1(m) \frac{\sum_{n=0}^{319} u_{HB0}(n)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon}$$

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

$$g_{HB1}(m) = \sqrt{\frac{\frac{\sum_{n=0}^{63} u(n+64m)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon}}{\frac{\sum_{n=0}^{79} u_{HB0}(n+80m)^2 + \varepsilon}{\sum_{n=0}^{319} u_{HB0}(n)^2 + \varepsilon}}}$$

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 $u(n)$.

Le bloc 707 effectue la mise à l'échelle du signal combiné selon l'équation suivante:

$$u_{HB}(n) = g_{HB1}(m)u_{HB0}(n), \quad n = 80m, \dots, 80(m+1) - 1$$

[0091] On notera que la réalisation du bloc 706 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.

[0092] 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.

[0093] On notera ici que dans le cas du débit 23.85 kbit/s les gains $g_{HB1}(m)$ sont calculés mais appliqués dans l'étape suivante, comme expliqué en référence à la figure 4, pour éviter les doubles multiplications. Dans ce cas $u_{HB}(n) = u_{HB0}(n)$.

[0094] Selon l'invention, le bloc 708 effectue ensuite un calcul de facteur d'échelle par sous-trame du signal (étapes

E602 à E 603 de la figure 6), comme décrit précédemment en référence à la figure 6 et détaillé en figure 4 et 5.

[0095] Enfin, l'excitation corrigée $u_{HB}'(n)$ est filtrée par le module de filtrage 710 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.

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 710.

[0096] Dans une variante de réalisation, l'étape de filtrage par un filtre 710 de prédiction linéaire pour la deuxième bande de fréquence est combinée à l'application du facteur d'échelle optimisé, ce qui permet de réduire la complexité de traitement. Ainsi les étapes de filtrage $1/\hat{A}(z/\gamma)$ et d'application du facteur d'échelle optimisé g_{HB2} sont combinées à une seule étape de filtrage $g_{HB2}/\hat{A}(z/\gamma)$ pour réduire la complexité de traitement.

[0097] 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.

[0098] 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.

[0099] 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.

[0100] Dans des variantes de l'extension de bande, 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.

[0101] 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.

[0102] L'extension de bande illustrée à la figure 7, s'applique dans ce cas à une bande basse qui n'est pas décodée mais analysée.

[0103] La figure 8 représente un exemple de réalisation matérielle d'un dispositif de détermination d'un facteur d'échelle optimisé 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.

[0104] 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 et optimisé ($u_{HB}'(n)$) par exemple à un module de filtrage comme le bloc 710 de la figure 7 ou à un module de ré-échantillonnage comme le module 311 de la figure 3.

[0105] Le bloc mémoire peut avantageusement comporter un programme informatique comportant des instructions de code pour la mise en oeuvre des étapes du procédé de détermination d'un facteur d'échelle optimisé à appliquer à un signal d'excitation ou à un filtre au sens de l'invention, lorsque ces instructions sont exécutées par le processeur PROC, et notamment les étapes de détermination (E602) d'un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur au filtre de prédiction linéaire de la première bande de fréquence, les coefficients du filtre additionnel étant obtenus à partir des paramètres décodés ou extraits de la première bande de fréquence, de calcul (E603) d'un facteur d'échelle optimisé en fonction au moins des coefficients du filtre additionnel.

[0106] Typiquement, la description de la figure 6 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.

[0107] La mémoire MEM enregistre de manière générale, toutes les données nécessaires à la mise en oeuvre du procédé.

[0108] Dans un mode possible de réalisation, le dispositif ainsi décrit peut également comporter les fonctions d'appli-

cation du facteur d'échelle optimisé au signal d'excitation étendu, d'extension de bande de fréquence, de décodage bande basse et autre fonctions de traitement décrites par exemple en figure 3 et 4 en plus des fonctions de détermination de facteur d'échelle optimisé selon l'invention.

5

Revendications

10

1. Procédé de détermination d'un facteur d'échelle optimisé à appliquer à un signal d'excitation ou à un filtre lors d'un procédé d'extension de bande de fréquence d'un signal audiofréquence, le procédé d'extension de bande (E601) comportant une étape de décodage ou d'extraction, dans une première bande de fréquence, d'un signal d'excitation et de paramètres de la première bande de fréquence comprenant des coefficients d'un filtre de prédiction linéaire, une étape de génération d'un signal d'excitation étendu sur au moins une deuxième bande de fréquence et une étape de filtrage par un filtre de prédiction linéaire pour la deuxième bande de fréquence, le procédé de détermination étant **caractérisé en ce qu'il** comporte les étapes suivantes:

15

- détermination (E602) d'un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur au filtre de prédiction linéaire de la première bande de fréquence, les coefficients du filtre additionnel étant obtenus à partir des paramètres décodés ou extraits de la première bande de fréquence; et
- calcul (E603) du facteur d'échelle optimisé en fonction au moins des coefficients du filtre additionnel.

20

2. Procédé selon la revendication 1, **caractérisé en ce que** le procédé d'extension de bande comprend une étape d'application (E604) du facteur d'échelle optimisé au signal d'excitation étendu.

25

3. Procédé selon la revendication 2, **caractérisé en ce que** l'application du facteur d'échelle optimisé est combinée à l'étape de filtrage dans la deuxième bande de fréquence.

30

4. Procédé selon la revendication 1, **caractérisé en ce que** les coefficients du filtre additionnel sont obtenus par troncature de la fonction de transfert du filtre de prédiction linéaire de la première bande de fréquence pour obtenir un ordre inférieur.

35

5. Procédé selon la revendication 4, **caractérisé en ce que** les coefficients du filtre additionnel sont modifiés en fonction d'un critère de stabilité du filtre additionnel.

40

6. Procédé selon la revendication 1, **caractérisé en ce que** le calcul du facteur d'échelle optimisé comporte les étapes suivantes:

- calcul des réponses en fréquence des filtres de prédiction linéaire des première et deuxième bandes de fréquence pour une fréquence commune;
- calcul de la réponse en fréquence du filtre additionnel pour cette fréquence commune;
- calcul du facteur d'échelle optimisé en fonction des réponses en fréquence ainsi calculées.

45

7. Procédé selon la revendication 1, **caractérisé en ce qu'il** comprend en outre les étapes suivantes, mises en oeuvre pour un débit de décodage prédéterminé:

50

- première mise à l'échelle du signal d'excitation étendu par un gain calculé par sous-trame fonction d'un rapport d'énergie entre le signal d'excitation décodé et le signal d'excitation étendu;
- deuxième mise à l'échelle du signal d'excitation issu de la première mise à l'échelle par un gain de correction décodé;
- ajustement de l'énergie de l'excitation pour la sous-trame courante par un facteur d'ajustement calculé en fonction de l'énergie du signal obtenu après la deuxième mise à l'échelle et en fonction du signal obtenu après application du facteur d'échelle optimisé.

55

8. Dispositif de détermination d'un facteur d'échelle optimisé à appliquer à un signal d'excitation ou à un filtre dans un dispositif d'extension de bande de fréquence d'un signal audiofréquence, le dispositif d'extension de bande (400) comprenant un module de décodage ou d'extraction, dans une première bande de fréquence, d'un signal d'excitation et de paramètres de la première bande de fréquence comprenant des coefficients d'un filtre de prédiction linéaire, un module de génération d'un signal d'excitation étendu sur au moins une deuxième bande de fréquence et un module de filtrage par un filtre de prédiction linéaire pour la deuxième bande de fréquence, le dispositif de déter-

mination étant **caractérisé en ce qu'il** comprend:

- un module de détermination (401a) d'un filtre de prédiction linéaire dit filtre additionnel, d'ordre inférieur au filtre de prédiction linéaire de la première bande de fréquence, les coefficients du filtre additionnel étant obtenus à partir des paramètres décodés ou extraits de la première bande de fréquence; et
- un module de calcul (401b) du facteur d'échelle optimisé en fonction au moins des coefficients du filtre additionnel.

9. Décodeur de signal audiofréquence **caractérisé en ce qu'il** comporte un dispositif de détermination d'un facteur d'échelle optimisé conforme à la revendication 8.

10. Programme informatique comportant des instructions de code pour la mise en oeuvre des étapes du procédé détermination d'un facteur d'échelle optimisé selon l'une des revendications 1 à 7, lorsque ces instructions sont exécutées par un processeur.

11. Support de stockage lisible par un dispositif de détermination d'un facteur d'échelle optimisé sur lequel est enregistré un programme informatique comprenant des instructions de code pour l'exécution des étapes du procédé de détermination d'un facteur d'échelle optimisé selon l'une des revendications 1 à 7.

Patentansprüche

1. Verfahren zum Bestimmen eines optimierten Skalenfaktors, der an ein Erregungssignal oder an ein Filter bei einem Frequenzband-Erweiterungsverfahren eines Audiofrequenzsignals anzulegen ist, wobei das Banderweiterungsverfahren (E601) einen Schritt des Decodierens oder der Extraktion aus einem ersten Frequenzband eines Erregungssignals und von Parametern des ersten Frequenzbands, das Koeffizienten eines linearen Vorhersagefilters umfasst, einen Schritt des Erzeugens eines Erregungssignals, das auf mindestens ein zweites Frequenzband erweitert ist, und einen Schritt des Filterns durch ein lineares Vorhersagefilter für das zweite Frequenzband umfasst, Bestimmungsverfahren **dadurch gekennzeichnet, dass** es die folgenden Schritte umfasst:

- Bestimmen (E602) eines linearen Vorhersagefilters, Zusatzfilter genannt, mit niedrigerer Rangordnung als das lineare Vorhersagefilter des ersten Frequenzbands, wobei die Koeffizienten des Zusatzfilters ausgehend von Parametern erhalten werden, die aus dem ersten Frequenzband decodiert oder extrahiert werden; und
- Berechnen (E603) des optimierten Skalenfaktors in Abhängigkeit von mindestens den Koeffizienten des Zusatzfilters.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das Frequenzband-Erweiterungsverfahren einen Schritt (E604) des Anwendens des optimierten Skalenfaktors auf das erweiterte Erregungssignal umfasst.

3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** das Anwenden des optimierten Skalenfaktors mit dem Filterschritt in dem zweiten Frequenzband kombiniert ist.

4. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Koeffizienten des Zusatzfilters durch Abkürzen der Transferfunktion des linearen Vorhersagefilters des ersten Frequenzbands erhalten werden, um einen unteren Rang zu erhalten.

5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** die Koeffizienten des Zusatzfilters in Abhängigkeit von einem Stabilitätskriterium des Zusatzfilters modifiziert werden.

6. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Berechnung des optimierten Skalenfaktors die folgenden Schritte umfasst:

- Berechnen der Frequenzgänge der linearen Vorhersagefilter des ersten und zweiten Frequenzbands für eine gemeinsame Frequenz;
- Berechnen des Frequenzgangs des Zusatzfilters für diese gemeinsame Frequenz;
- Berechnen des optimierten Skalenfaktors in Abhängigkeit von den Frequenzgängen, die so berechnet werden.

7. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es außerdem die folgenden Schritte umfasst, die für

einen vorbestimmten Dekodierungsdurchsatz umgesetzt werden:

- erstes Skalieren des erweiterten Erregungssignals durch eine Verstärkung, die durch Unterblock in Abhängigkeit von einem Energiezufuhrverhältnis zwischen dem decodierten Erregungssignal und dem erweiterten Erregungssignal berechnet wird;
- zweites Skalieren des Erregungssignals, das aus der ersten Skalierung durch eine decodierte Korrekturverstärkung hervorgegangen ist;
- Anpassen der Energie der Erregung für den aktuellen Unterblock durch einen Anpassfaktor, der in Abhängigkeit von der Energie des Signals, das nach der zweiten Skalierung erhalten wird, und in Abhängigkeit von dem Signal, das nach dem Anwenden des optimierten Skalierfaktors erhalten wird, berechnet wird.

8. Vorrichtung zum Bestimmen eines optimierten Skalenfaktors, der auf ein Erregungssignal oder ein Filter in einer Frequenzband-Erweiterungsvorrichtung eines Audiofrequenzsignals anzuwenden ist, wobei die Banderweiterungsvorrichtung (400) ein Decodier- oder Extraktionsmodul in einem ersten Frequenzband eines Erregungssignals und von Parametern des ersten Frequenzbands, die Koeffizienten eines linearen Vorhersagefilters umfassen, umfasst, ein Modul zum Erzeugen eines Erregungssignals, das auf mindestens ein zweites Frequenzband erweitert ist, und ein Modul zum Filtern durch ein lineares Vorhersagefilter für das zweite Frequenzband, Vorrichtung zum Bestimmen **dadurch gekennzeichnet, dass** sie Folgendes umfasst:

- ein Modul zum Bestimmen (401a) eines linearen Vorhersagefilters, Zusatzfilter genannt, mit niedrigerem Rang als das lineare Vorhersagefilter des ersten Frequenzbands, wobei die Koeffizienten des Zusatzfilters ausgehend von Parametern erhalten werden, die aus dem ersten Frequenzband decodiert oder extrahiert werden; und
- ein Modul zum Berechnen (401b) des Skalenfaktors, der in Abhängigkeit von mindestens Koeffizienten des Zusatzfilters optimiert ist.

9. Audiofrequenzsignaldecodierer, **dadurch gekennzeichnet, dass** er eine Vorrichtung zum Bestimmen eines optimierten Skalenfaktors gemäß Anspruch 8 umfasst.

10. Informatikprogramm, das Codeanweisungen zum Umsetzen der Schritte des Verfahrens zum Bestimmen eines optimierten Skalenfaktors nach einem der Ansprüche 1 bis 7 umfasst, wenn diese Anweisungen von einem Prozessor ausgeführt werden.

11. Speicherträger, der von einer Vorrichtung zum Bestimmen eines optimierten Skalenfaktors gelesen werden kann, auf dem ein Informatikprogramm aufgezeichnet ist, das Codeanweisungen zur Ausführung der Schritte des Verfahrens zum Bestimmen eines optimierten Skalenfaktors nach einem der Ansprüche 1 bis 7 umfasst.

Claims

1. Method for determining an optimized scale factor to be applied to an excitation signal or to a filter in an audio frequency signal frequency band extension method, the band extension method (E601) comprising a step of decoding or of extraction, in a first frequency band, of an excitation signal and of parameters of the first frequency band comprising coefficients of a linear prediction filter, a step of generation of an extended excitation signal on at least one second frequency band and a step of filtering, by a linear prediction filter, for the second frequency band, the determination method being **characterized in that** it comprises the following steps:

- determination (E602) of a linear prediction filter called additional filter, of lower order than the linear prediction filter of the first frequency band, the coefficients of the additional filter being obtained from the parameters decoded or extracted from the first frequency band; and
- computation (E603) of the optimized scale factor as a function at least of the coefficients of the additional filter.

2. Method according to Claim 1, **characterized in that** the band extension method comprises a step of application (E604) of the optimized scale factor to the extended excitation signal.

3. Method according to Claim 2, **characterized in that** the application of the optimized scale factor is combined with the step of filtering in the second frequency band.

4. Method according to Claim 1, **characterized in that** the coefficients of the additional filter are obtained by truncation

of the transfer function of the linear prediction filter of the first frequency band to obtain a lower order.

- 5 5. Method according to Claim 4, **characterized in that** the coefficients of the additional filter are modified as a function of a stability criterion of the additional filter.

- 10 6. Method according to Claim 1, **characterized in that** the computation of the optimized scale factor comprises the following steps:

- 10 - computation of the frequency responses of the linear prediction filters of the first and second frequency bands for a common frequency;
 - computation of the frequency response of the additional filter for this common frequency;
 - computation of the optimized scale factor as a function of the duly computed frequency responses.

- 15 7. Method according to Claim 1, **characterized in that** it further comprises the following steps, implemented for a predetermined decoding bit rate:

- 20 - first scaling of the extended excitation signal by a gain computed for each subframe as a function of an energy ratio between the decoded excitation signal and the extended excitation signal;
 - second scaling of the excitation signal obtained from the first scaling by a decoded correction gain;
 - adjustment of the energy of the excitation for the current subframe by an adjustment factor computed as a function of the energy of the signal obtained after the second scaling and as a function of the signal obtained after application of the optimized scale factor.

- 25 8. Device for determining an optimized scale factor to be applied to an excitation signal or to a filter in an audio frequency signal frequency band extension device, the band extension device (400) comprising a module for decoding or extracting, in a first frequency band, an excitation signal and parameters of the first frequency band comprising coefficients of a linear prediction filter, a module for generating an extended excitation signal on at least one second frequency band and a module for filtering, by a linear prediction filter, for the second frequency band, the determination device being **characterized in that** it comprises:

- 30 - a module (401a) for determining a linear prediction filter called additional filter, of lower order than the linear prediction filter of the first frequency band, the coefficients of the additional filter being obtained from the parameters decoded or extracted from the first frequency band; and
 - a module (401b) for computing the optimized scale factor as a function at least of the coefficients of the additional filter.

- 35 9. Audio frequency signal decoder, **characterized in that** it comprises a device for determining an optimized scale factor according to Claim 8.

- 40 10. Computer program comprising code instructions for implementing the steps of the method for determining an optimized scale factor according to one of Claims 1 to 7, when these instructions are executed by a processor.

- 45 11. Storage medium that can be read by a device for determining an optimized scale factor on which is stored a computer program comprising code instructions for the execution of steps of the method for determining an optimized scale factor according to one of Claims 1 to 7.

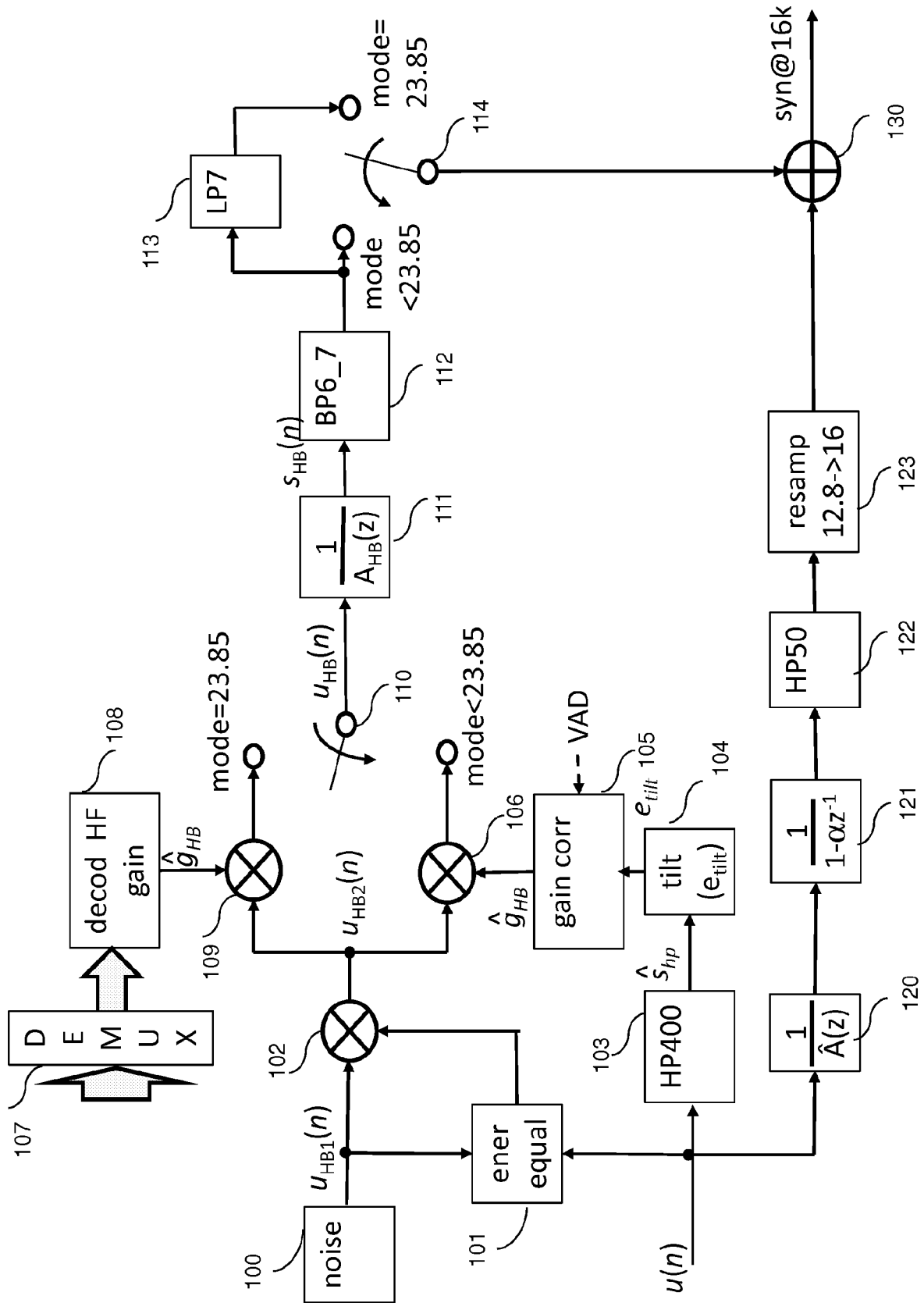


Fig.1

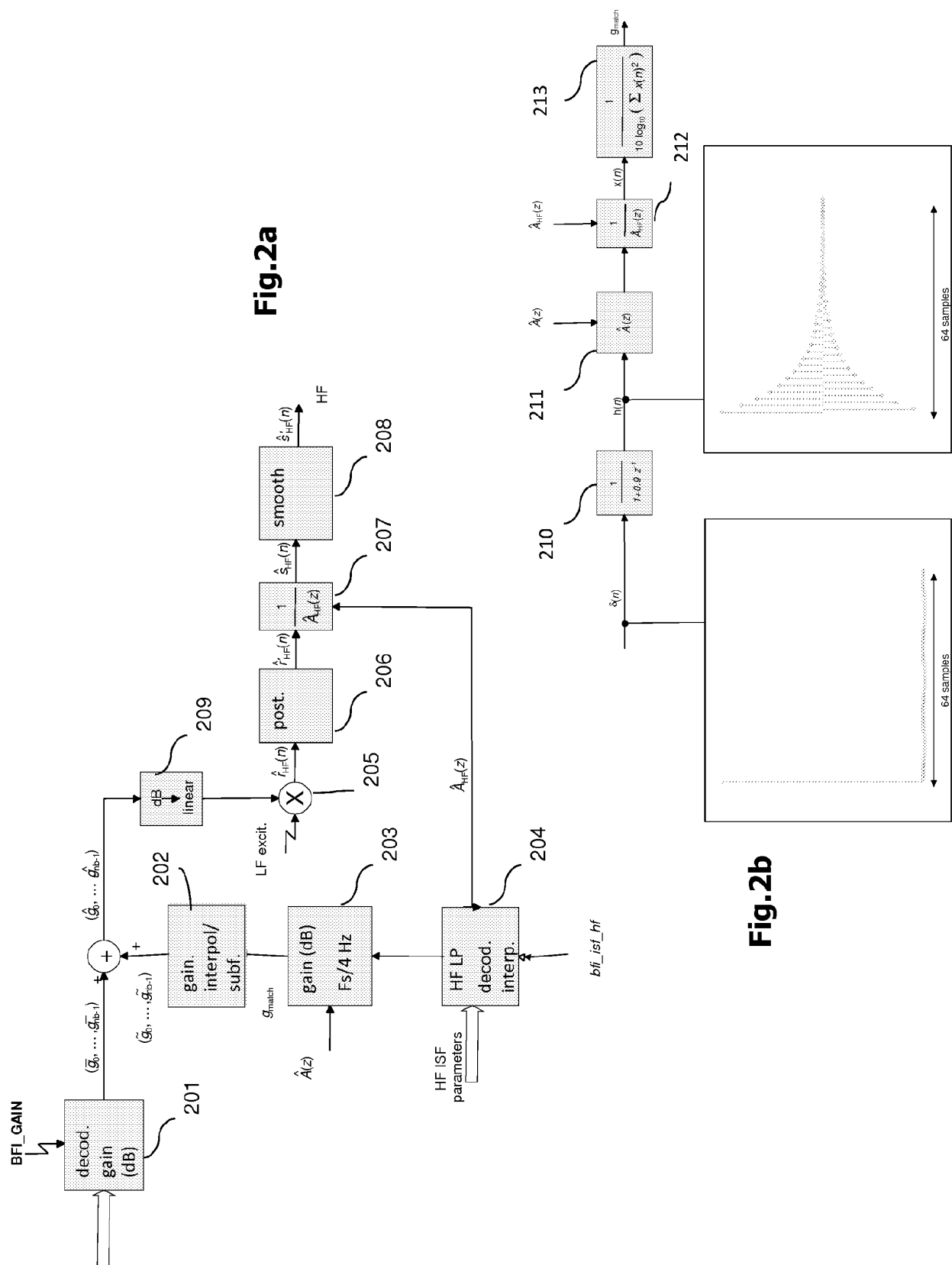


Fig. 2b

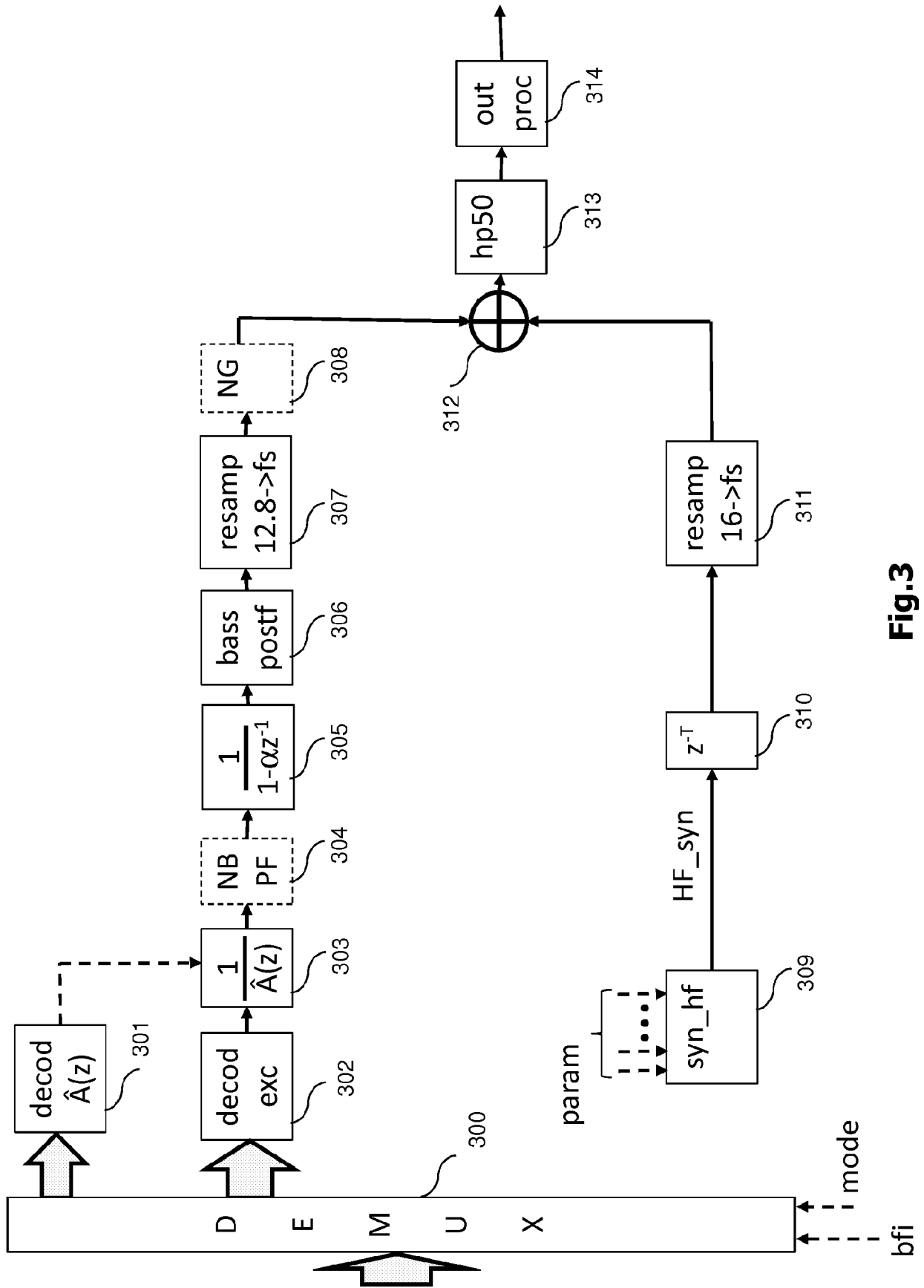


Fig.3

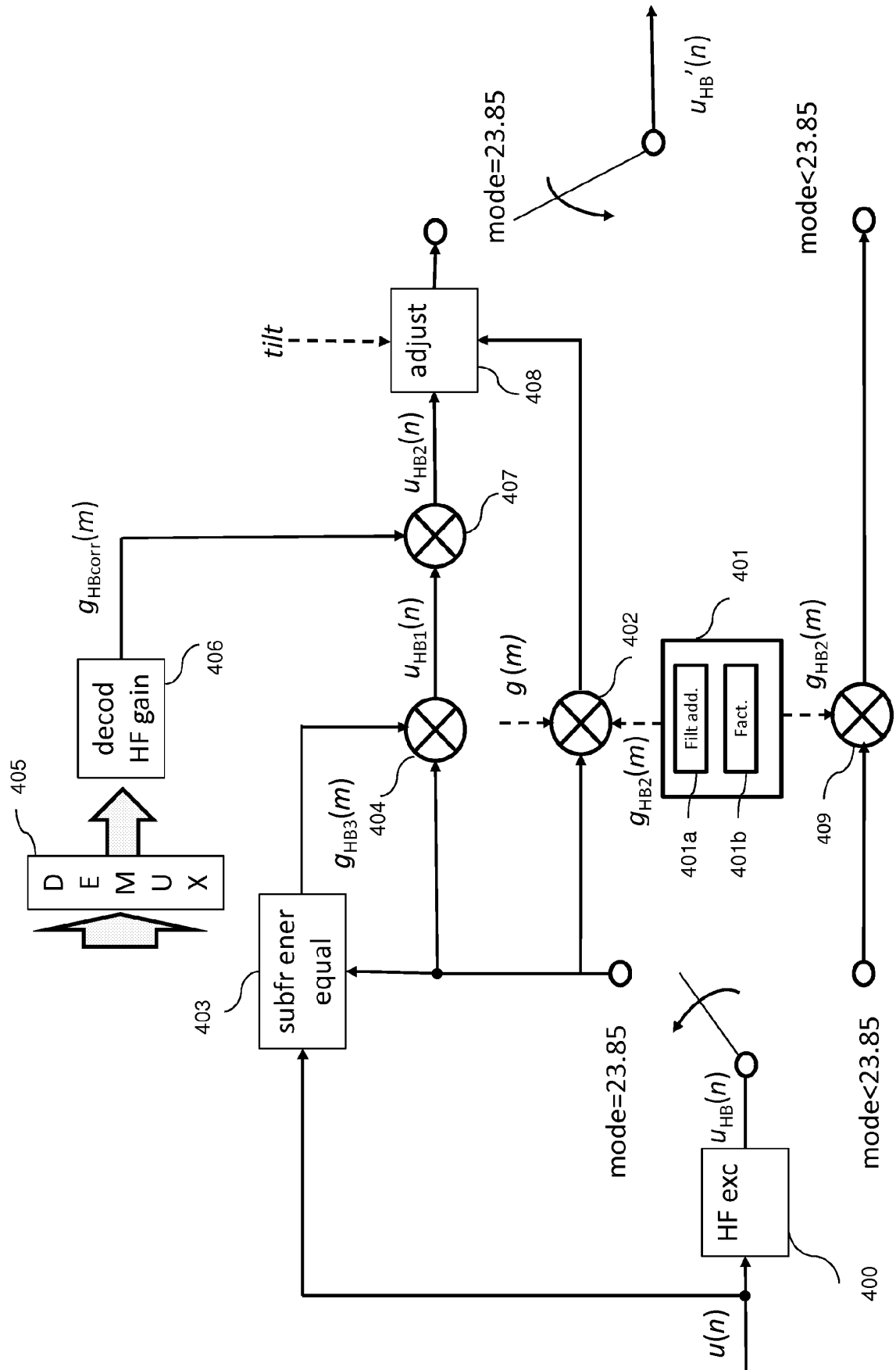


Fig.4

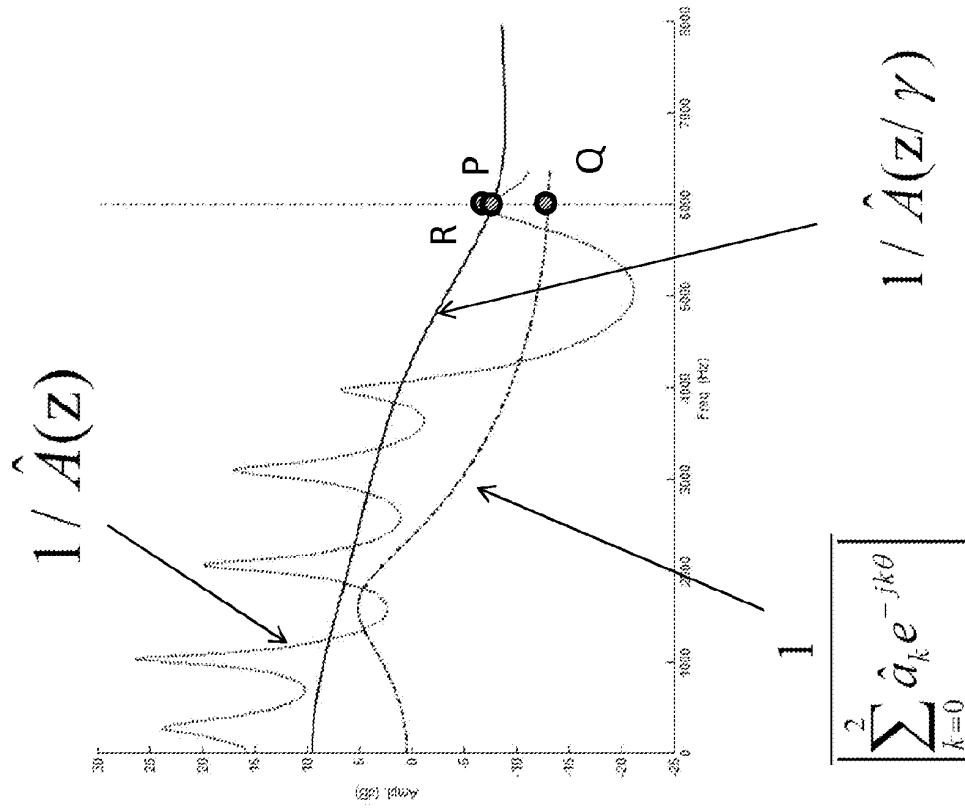


Fig.5a

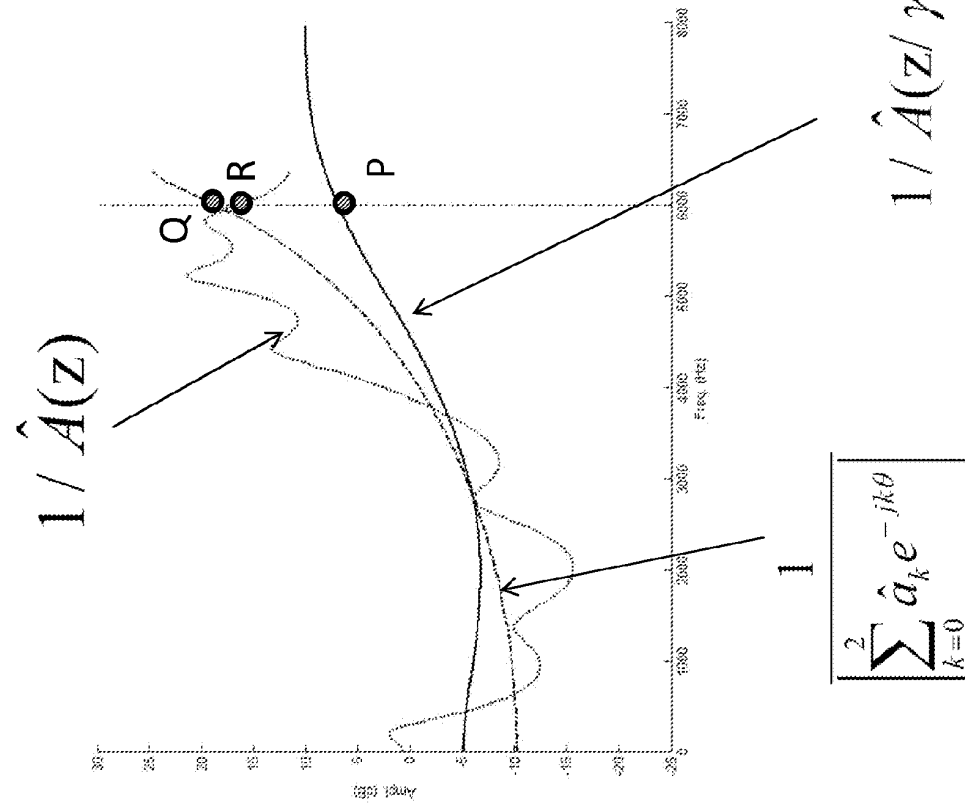


Fig.5b

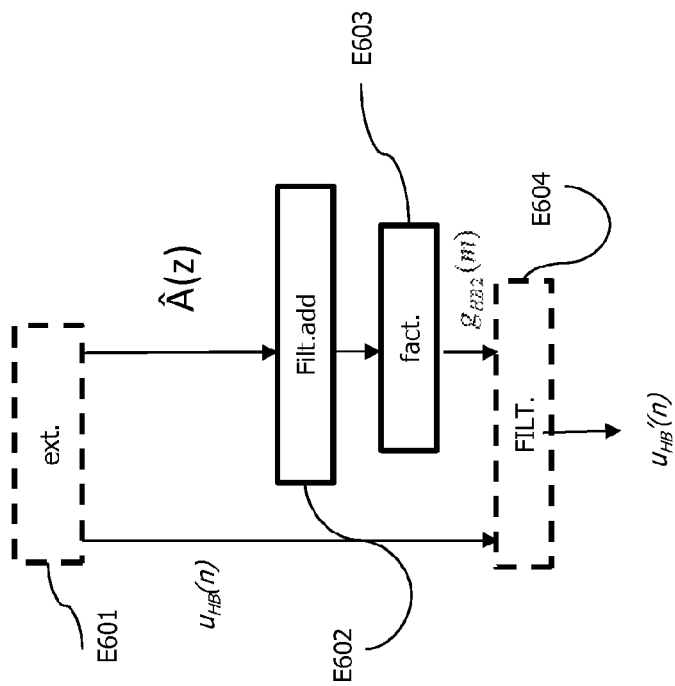


Fig.6

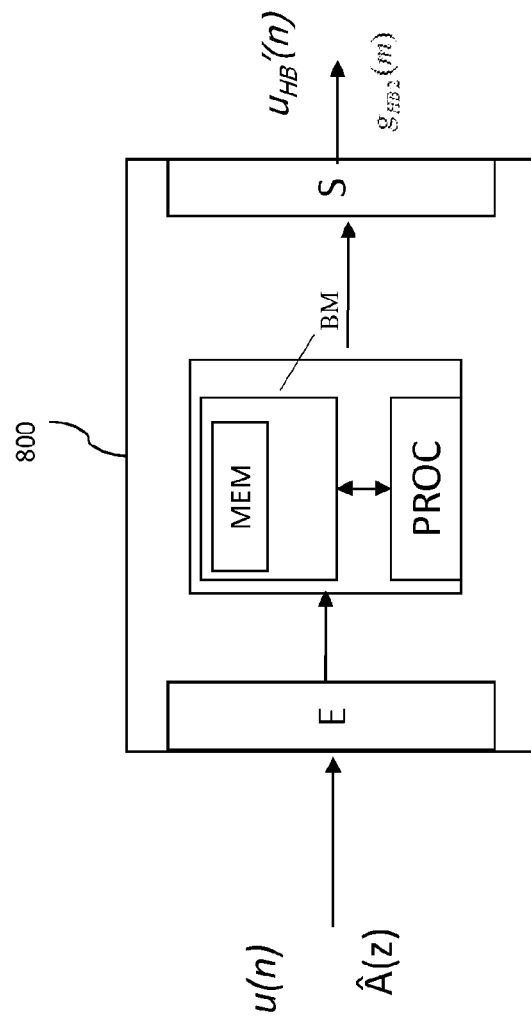


Fig.8

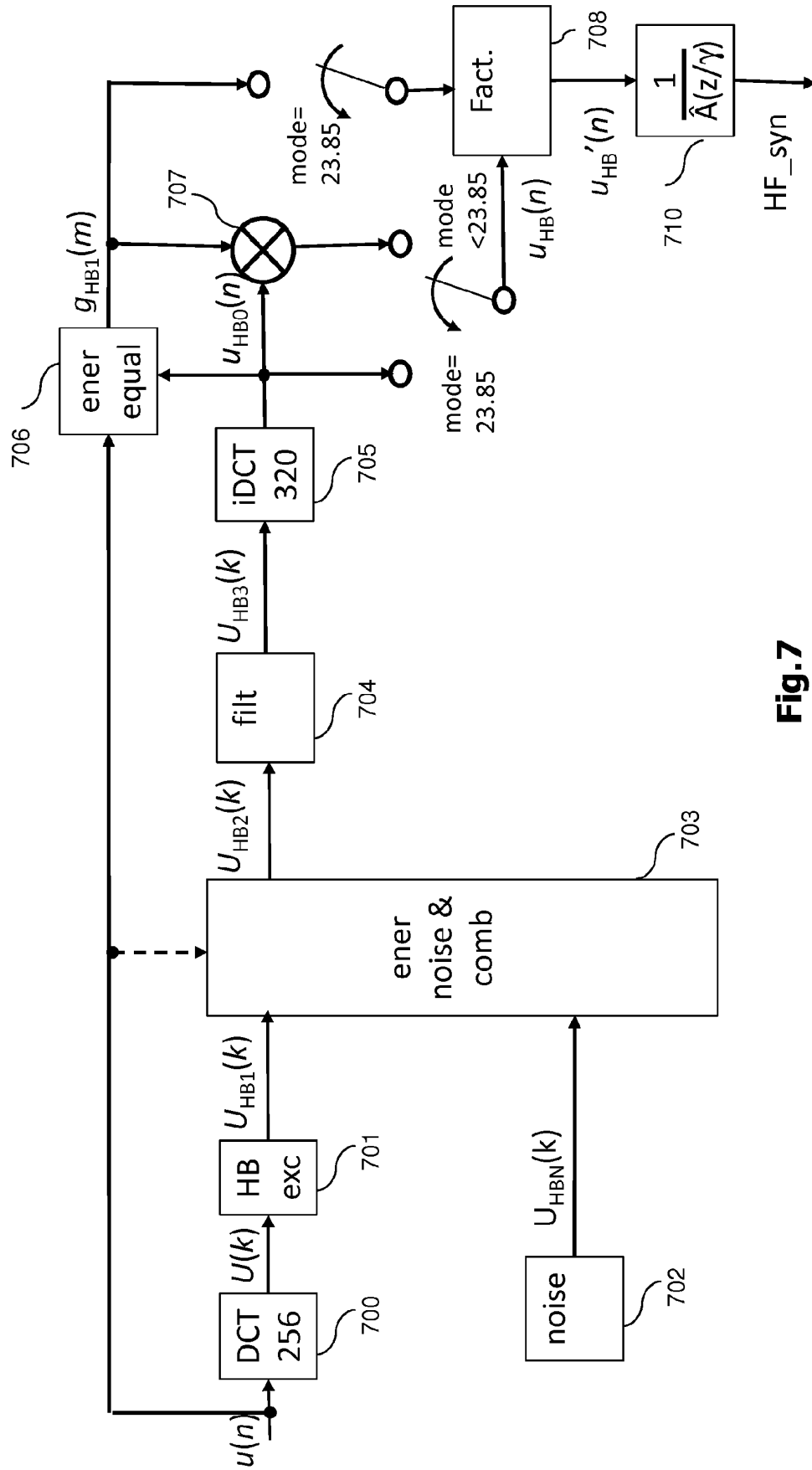


Fig.7

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.

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]
- STEP DOWN. **J.D. MARKEL ; A.H. GRAY.** Linear Prediction of Speech. Springer Verlag, 1976 [0050]
- **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 [0066]