



(12) DEMANDE DE BREVET EUROPEEN

(43) Date de publication:
02.07.1997 Bulletin 1997/27

(51) Int Cl. 6: G10L 9/14

(21) Numéro de dépôt: 96402715.5

(22) Date de dépôt: 12.12.1996

(84) Etats contractants désignés:
DE GB IT

• Le Guyader, Alain
22300 Lannion (FR)

(30) Priorité: 15.12.1995 FR 9514925

(74) Mandataire: Loisel, Bertrand et al
Cabinet Plasseraud,
84, rue d'Amsterdam
75440 Paris Cédex 09 (FR)

(71) Demandeur: FRANCE TELECOM
75015 Paris (FR)

(72) Inventeurs:
• Quinquis, Catherine
22300 Lannion (FR)

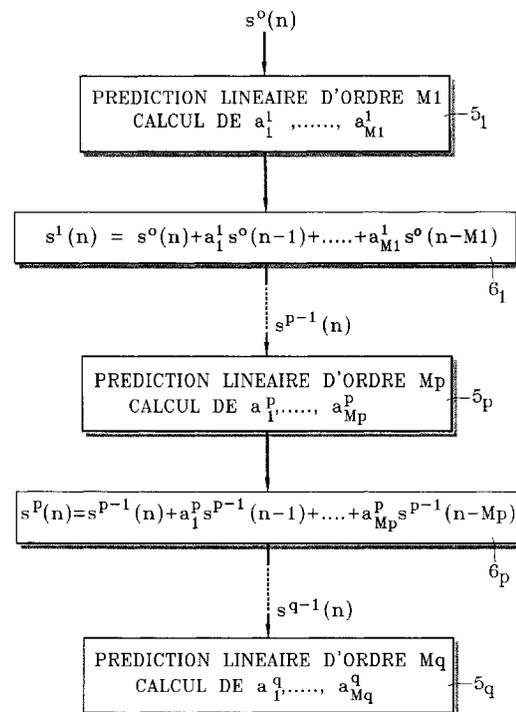
(54) Procédé d'analyse par prédiction linéaire d'un signal audiofréquence, et procédés de codage et de décodage d'un signal audiofréquence en comportant application

(57) Le procédé d'analyse par prédiction linéaire est utilisé pour déterminer des paramètres spectraux représentatifs de l'enveloppe spectrale du signal audiofréquence. Ce procédé comprend q étages de prédiction successifs (5_p), q étant un entier supérieur à 1. A chaque étage de prédiction p ($1 \leq p \leq q$), on détermine des paramètres représentant un nombre prédéfini M_p de coefficients $a_1^p, \dots, a_{M_p}^p$ de prédiction linéaire d'un signal d'entrée dudit étage. Le signal audiofréquence à analyser constitue le signal d'entrée du premier étage. Le signal d'entrée ($s^p(n)$) d'un étage p+1 est constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p filtré par un filtre de fonction de transfert

$$A^p(z) = 1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i}$$

Utilisation notamment dans des codeurs audio en bande élargie.

FIG. 1



Description

La présente invention concerne un procédé d'analyse par prédiction linéaire d'un signal audiofréquence. Ce procédé trouve une application particulière, mais non exclusive, dans des codeurs audio à prédiction, notamment dans des codeurs à analyse par synthèse, dont le type le plus répandu est le codeur CELP ("Code-Excited Linear Prediction").

Les techniques de codage prédictif à analyse par synthèse sont actuellement très répandues pour le codage de la parole en bande téléphonique (300-3400 Hz) à des débits pouvant descendre jusqu'à 8 kbit/s, tout en conservant une qualité téléphonique. Pour la bande audio (de l'ordre de 20 kHz), les techniques de codage par transformée sont utilisées pour des applications de diffusion et de stockage de signaux vocaux et musicaux. Cependant, ces techniques impliquent des retards de codage relativement importants (plus grands que 100 ms), ce qui produit en particulier des difficultés de participation dans les communications de groupe où l'interactivité est très importante. Les techniques prédictives produisent un retard plus faible, dépendant essentiellement de la longueur des trames d'analyse par prédiction linéaire (typiquement 10 à 20 ms), et trouvent pour cette raison des applications même pour le codage de signaux vocaux et/ou musicaux ayant une largeur de bande supérieure à la bande téléphonique.

Les codeurs prédictifs utilisés pour la compression de débit réalisent une modélisation de l'enveloppe spectrale du signal. Cette modélisation résulte d'une analyse par prédiction linéaire d'ordre M (M=10 typiquement en bande étroite), consistant à déterminer M coefficients a_i de prédiction linéaire du signal d'entrée. Ces coefficients caractérisent un filtre de synthèse utilisé au décodeur, dont la fonction de transfert est de la forme $1/A(z)$ avec

$$A(z) = 1 + \sum_{i=1}^M a_i z^{-i} \quad (1)$$

L'analyse par prédiction linéaire a un domaine d'application général plus large que celui du codage de la parole. Dans certaines applications, l'ordre M de la prédiction constitue l'une des variables que l'analyse par prédiction linéaire vise à obtenir, cette variable étant influencée par le nombre de pics présents dans le spectre du signal analysé (voir US-A-5 142 581).

Le filtre calculé par l'analyse par prédiction linéaire peut avoir diverses structures, conduisant à différents choix de paramètres pour la représentation des coefficients (les coefficients a_i eux-mêmes, les paramètres LAR, LSF, LSP, les coefficients de réflexion ou PARCOR...). Avant l'avènement des processeurs de signal numérique (DSP), il était courant d'employer des structures récursives pour le filtre calculé, par exemple des structures faisant appel aux coefficients PARCOR du type décrit dans l'article de F. ITAKURA et S. SAITO "Digital Filtering Techniques for Speech Analysis and Synthesis", Proc. of the 7th International Congress on Acoustics, Budapest 1971, pages 261-264 (voir FR-A-2 284 946 ou US-A-3 975 587).

Dans les codeurs à analyse par synthèse, les coefficients a_i servent également à construire un filtre de pondération perceptuelle utilisé par le codeur pour déterminer le signal d'excitation à appliquer au filtre de synthèse à court terme pour obtenir un signal synthétique représentatif du signal de parole. Cette pondération perceptuelle accentue les portions du spectre où les erreurs de codage sont les plus perceptibles, c'est-à-dire les zones interformantiques. La fonction de transfert $W(z)$ du filtre de pondération perceptuelle est habituellement de la forme

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)} \quad (2)$$

où γ_1 et γ_2 sont deux coefficients d'expansion spectrale tels que $0 \leq \gamma_2 \leq \gamma_1 \leq 1$. Une amélioration du masquage du bruit a été apportée par E. Ordentlich et Y. Shoham, dans leur article "Low-Delay Code-Excited Linear Predictive Coding of Wideband Speech at 32 kbps", Proc. ICASSP, Toronto, Mai 1991, pages 9-12. Cette amélioration consiste à combiner pour la pondération perceptuelle le filtre $W(z)$ avec un autre filtre modélisant la pente du spectre. Cette amélioration est particulièrement appréciable dans le cas de codage de signaux à forte dynamique spectrale (bande élargie ou bande audio) pour lesquels les auteurs ont montré une importante amélioration de la qualité subjective du signal reconstruit.

Dans la plupart des décodeurs CELP actuels, les coefficients de prédiction linéaire a_i sont également utilisés pour définir un post-filtre servant à atténuer les zones fréquentielles entre les formants et les harmoniques du signal de parole, sans modifier la pente du spectre du signal. Une forme habituelle de la fonction de transfert de ce post-filtre est :

$$H_{PF}(z) = G_P \frac{A(z/\beta_1)}{A(z/\beta_2)} (1 - \mu r_1 z^{-1}) \quad (3)$$

5 où G_P est un facteur de gain compensant l'atténuation des filtres, β_1 et β_2 sont des coefficients tels que $0 \leq \beta_1 \leq \beta_2 \leq 1$, μ est une constante positive et r_1 désigne le premier coefficient de réflexion dépendant des coefficients a_i .

La modélisation de l'enveloppe spectrale du signal par les coefficients a_i constitue donc un élément essentiel du processus de codage et de décodage, en ce sens qu'elle doit représenter le contenu spectral du signal à reconstituer au décodeur et qu'elle pilote aussi bien le masquage du bruit de quantification que le post-filtrage au décodeur.

10 Pour des signaux à forte dynamique spectrale, l'analyse par prédiction linéaire habituellement pratiquée ne parvient pas à modéliser fidèlement l'enveloppe du spectre. Souvent, les signaux de parole sont sensiblement plus énergétiques aux basses fréquences qu'aux fréquences élevées, de sorte que l'analyse par prédiction linéaire conduit certes à une modélisation précise aux basses fréquences, mais au détriment de la modélisation du spectre aux fréquences plus élevées. Cet inconvénient devient particulièrement gênant dans le cas du codage en bande élargie.

15 Un but de la présente invention est d'améliorer la modélisation du spectre d'un signal audiofréquence dans un système faisant appel à un procédé d'analyse par prédiction linéaire. Un autre but est de rendre les performances d'un tel système plus homogènes pour des signaux d'entrée différents (parole, musique, sinusoïdes, signaux DTMF...), des largeurs de bande différentes (bande téléphonique, bande élargie, bande hifi...), des conditions différentes d'enregistrement (microphone directif, antenne acoustique...) et de filtrage.

20 L'invention propose ainsi un procédé d'analyse par prédiction linéaire d'un signal audiofréquence, pour déterminer des paramètres spectraux dépendant d'un spectre à court terme du signal audiofréquence, comprenant q étages de prédiction successifs, q étant un entier supérieur à 1. A chaque étage de prédiction p ($1 \leq p \leq q$), on détermine des paramètres représentant un nombre prédéfini M_p de coefficients $a_1^p, \dots, a_{M_p}^p$ de prédiction linéaire d'un signal d'entrée dudit étage, le signal audiofréquence analysé constituant le signal d'entrée du premier étage, et le signal d'entrée d'un

25 étage $p+1$ étant constitué par le signal d'entrée de l'étage p filtré par un filtre de fonction de transfert

$$A^D(z) = 1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i} \quad (4)$$

30

Le nombre M_p de coefficients de prédiction linéaire peut notamment augmenter d'un étage au suivant. Ainsi, le premier étage pourra rendre compte assez fidèlement de la pente générale du spectre ou du signal, tandis que les

35 étages suivants affineront la représentation des formants du signal. On évite ainsi, dans le cas de signaux à forte dynamique, de trop privilégier les zones les plus énergétiques au risque d'une modélisation médiocre des autres zones fréquentielles pouvant être perceptuellement importantes.

35

Un second aspect de l'invention concerne une application de ce procédé d'analyse par prédiction linéaire dans un codeur audiofréquence à analyse par synthèse à adaptation "forward". L'invention propose ainsi un procédé de

40 codage d'un signal audiofréquence comprenant les étapes suivantes :

40

- analyse par prédiction linéaire d'un signal audiofréquence numérisé en trames successives pour déterminer des paramètres définissant un filtre de synthèse à court terme ;
- détermination de paramètres d'excitation définissant un signal d'excitation à appliquer au filtre de synthèse à court

45 terme pour produire un signal synthétique représentatif du signal audiofréquence ; et

- production de valeurs de quantification des paramètres définissant le filtre de synthèse à court terme et des paramètres d'excitation,

45

dans lequel l'analyse par prédiction linéaire est un processus à q étages successifs tel que défini ci-dessus, et dans lequel le filtre de prédiction à court terme a une fonction de transfert de la forme $1/A(z)$ avec

50

$$A(z) = \prod_{p=1}^q A^p(z) \quad (5)$$

55

La fonction de transfert $A(z)$ ainsi obtenue peut également être utilisée pour définir selon la formule (2) la fonction

de transfert du filtre de pondération perceptuelle lorsque le codeur est un codeur à analyse par synthèse avec détermination en boucle fermée du signal d'excitation. Une autre possibilité intéressante est d'adopter des coefficients d'expansion spectrale γ_1 et γ_2 pouvant varier d'un étage au suivant, c'est-à-dire de donner au filtre de pondération perceptuelle une fonction de transfert de la forme

5

$$W(z) = \prod_{p=1}^q [A^p(z/\gamma_1^p) / A^p(z/\gamma_2^p)] , \quad (6)$$

10

où γ_1^p, γ_2^p désignent des paires de coefficients d'expansion spectrale tels que $0 \leq \gamma_2^p \leq \gamma_1^p \leq 1$ pour $1 \leq p \leq q$.

L'invention est également applicable au niveau d'un décodeur associé. Le procédé de décodage ainsi mis en oeuvre selon l'invention comprend les étapes suivantes :

15

- on reçoit des valeurs de quantification de paramètres définissant un filtre de synthèse à court terme et des paramètres d'excitation, les paramètres définissant le filtre de synthèse à court terme comprenant un nombre $q > 1$ de jeux de coefficients de prédiction linéaire, chaque jeu comportant un nombre prédéfini de coefficients ;
- on produit un signal d'excitation sur la base des valeurs de quantification des paramètres d'excitation ;
- on produit un signal audiofréquence synthétique en filtrant le signal d'excitation par un filtre de synthèse ayant une fonction de transfert de la forme $1/A(z)$ avec

20

$$A(z) = \prod_{p=1}^q \left(1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i} \right) \quad (7)$$

25

où les coefficients $a_1^p, \dots, a_{M_p}^p$ correspondent au p -ième jeu de coefficients de prédiction linéaire pour $1 \leq p \leq q$.

30

Cette fonction de transfert $A(z)$ peut également être utilisée pour définir un post-filtre dont la fonction de transfert comporte, comme dans la formule (3) ci-dessus, un terme de la forme $A(z/\beta_1)/A(z/\beta_2)$, où β_1 et β_2 désignent des coefficients tels que $0 \leq \beta_1 \leq \beta_2 \leq 1$.

Une variante intéressante consiste à remplacer ce terme de la fonction de transfert du post-filtre par :

35

$$\prod_{p=1}^q [A^p(z/\beta_1^p) / A^p(z/\beta_2^p)] , \quad (8)$$

40

où β_1^p, β_2^p désignent des paires de coefficients tels que $0 \leq \beta_1^p \leq \beta_2^p \leq 1$ pour $1 \leq p \leq q$.

L'invention s'applique également à des codeurs audiofréquence à adaptation "backward". L'invention propose ainsi un procédé de codage d'un premier signal audiofréquence numérisé en trames successives, comprenant les étapes suivantes :

45

- analyse par prédiction linéaire d'un second signal audiofréquence pour déterminer des paramètres définissant un filtre de synthèse à court terme ;
- détermination de paramètres d'excitation définissant un signal d'excitation à appliquer au filtre de synthèse à court terme pour produire un signal synthétique représentatif du premier signal audiofréquence, ce signal synthétique constituant ledit second signal audiofréquence pour au moins une trame suivante ; et
- production de valeurs de quantification des paramètres d'excitation,

50

dans lequel l'analyse par prédiction linéaire est un processus à q étages successifs tel que défini ci-dessus, et dans lequel le filtre de prédiction à court terme a une fonction de transfert de la forme $1/A(z)$ avec

55

$$A(z) = \prod_{p=1}^q A^p(z)$$

5 Pour une mise en oeuvre dans un décodeur associé, l'invention propose un procédé de décodage d'un flux binaire pour construire en trames successives un signal audiofréquence codé par ledit flux binaire, comprenant les étapes suivantes :

- 10
- on reçoit des valeurs de quantification de paramètres d'excitation ;
 - on produit un signal d'excitation sur la base des valeurs de quantification des paramètres d'excitation ;
 - on produit un signal audiofréquence synthétique en filtrant le signal d'excitation par un filtre de synthèse à court terme ;
 - on effectue une analyse par prédiction linéaire du signal synthétique pour obtenir des coefficients du filtre de
- 15 synthèse à court terme pour au moins une trame suivante,

dans lequel l'analyse par prédiction linéaire est un processus à q étages successifs tel que défini ci-dessus, et dans lequel le filtre de prédiction à court terme a une fonction de transfert de la forme $1/A(z)$ avec

20

$$A(z) = \prod_{p=1}^q A^p(z)$$

25 L'invention permet encore de réaliser des codeurs/décodeurs audiofréquence mixtes, c'est-à-dire faisant appel à la fois à des schémas d'adaptation "forward" et "backward", le ou les premiers étages de prédiction linéaire correspondant à une analyse "forward" et le ou les derniers étages à une analyse "backward". L'invention propose ainsi un procédé de codage d'un premier signal audiofréquence numérisé en trames successives, comprenant les étapes

30 suivantes :

- analyse par prédiction linéaire du premier signal audiofréquence pour déterminer des paramètres définissant une première composante d'un filtre de synthèse à court terme ;
 - détermination de paramètres d'excitation définissant un signal d'excitation à appliquer au filtre de synthèse à court
- 35 terme pour produire un signal synthétique représentatif du premier signal audio-fréquence ;
- production de valeurs de quantification des paramètres définissant la première composante du filtre de synthèse à court terme et des paramètres d'excitation ;
 - filtrage du signal synthétique par un filtre de fonction de transfert correspondant à l'inverse de la fonction de transfert de la première composante du filtre de synthèse à court terme ; et
 - analyse par prédiction linéaire du signal synthétique filtré pour obtenir des coefficients d'une seconde composante
- 40 du filtre de synthèse à court terme pour au moins une trame suivante,

dans lequel l'analyse par prédiction linéaire du premier signal audiofréquence est un processus à q_F étages successifs, q_F étant un entier au moins égal à 1, ledit processus à q_F étages comportant, à chaque étage de prédiction

45 p ($1 \leq p \leq q_F$), la détermination de paramètres représentant un nombre prédéfini MFP de coefficients $a_1^{F,p}, \dots, a_{MFP}^{F,p}$ de prédiction linéaire d'un signal d'entrée dudit étage, le premier signal audiofréquence constituant le signal d'entrée du premier étage, et le signal d'entrée d'un étage $p+1$ étant constitué par le signal d'entrée de l'étage p filtré par un filtre de fonction de transfert

50

$$A^{F,p}(z) = 1 + \sum_{i=1}^{MFP} a_i^{F,p} \cdot z^{-i},$$

55 la première composante du filtre de synthèse à court terme ayant une fonction de transfert de la forme $1/A^F(z)$ avec

$$A^F(z) = \prod_{p=1}^{q_F} A^{F,p}(z) ,$$

5 et dans lequel l'analyse par prédiction linéaire du signal synthétique filtré est un processus à q_B étages successifs, q_B étant un entier au moins égal à 1, ledit processus à q_B étages comportant, à chaque étage de prédiction p ($1 \leq p \leq q_B$), la détermination de paramètres représentant un nombre prédéfini MBp de coefficients $a_1^{B,p}, \dots, a_{MBp}^{B,p}$ de prédiction linéaire d'un signal d'entrée dudit étage, le signal synthétique filtré constituant le signal d'entrée du premier étage, et le signal d'entrée d'un étage $p+1$ étant constitué par le signal d'entrée de l'étage p filtré par un filtre de fonction de transfert

$$A^{B,p}(z) = 1 + \sum_{i=1}^{MBp} a_i^{B,p} \cdot z^{-i} ,$$

15 la seconde composante du filtre de synthèse à court terme ayant une fonction de transfert de la forme $1/A^B(z)$ avec

$$A^B(z) = \prod_{p=1}^{q_B} A^{B,p}(z) ,$$

25 et le filtre de synthèse à court terme ayant une fonction de transfert de la forme $1/A(z)$ avec $A(z) = A^F(z) \cdot A^B(z)$.

Pour une mise en oeuvre dans un décodeur mixte associé, l'invention propose un procédé de décodage d'un flux binaire pour construire en trames successives un signal audiofréquence codé par ledit flux binaire, comprenant les étapes suivantes :

- on reçoit des valeurs de quantification de paramètres définissant une première composante d'un filtre de synthèse à court terme et de paramètres d'excitation, les paramètres définissant la première composante du filtre de synthèse à court terme représentant un nombre q_F au moins égal à 1 de jeux de coefficients de prédiction linéaire $a_1^{F,p}, \dots, a_{MFp}^{F,p}$ pour $1 \leq p \leq q_F$, chaque jeu p comportant un nombre prédéfini MFp de coefficients, la première composante du filtre de synthèse à court terme ayant une fonction de transfert de la forme $1/A^F(z)$ avec

$$A^F(z) = \prod_{p=1}^{q_F} A^{F,p}(z) = \prod_{p=1}^{q_F} \left(1 + \sum_{i=1}^{MFp} a_i^{F,p} \cdot z^{-i} \right) ;$$

- on produit un signal d'excitation sur la base des valeurs de quantification des paramètres d'excitation ;
- on produit un signal audiofréquence synthétique en filtrant le signal d'excitation par un filtre de synthèse à court terme de fonction de transfert $1/A(z)$ avec $A(z) = A^F(z) \cdot A^B(z)$, $1/A^B(z)$ représentant la fonction de transfert d'une seconde composante du filtre de synthèse à court terme ;
- on filtre le signal synthétique par un filtre de fonction de transfert $A^F(z)$; et
- on effectue une analyse par prédiction linéaire du signal synthétique filtré pour obtenir des coefficients de la seconde composante du filtre de synthèse à court terme pour au moins une trame suivante,

dans lequel l'analyse par prédiction linéaire du signal synthétique filtré est un processus à q_B étages tel que défini ci-dessus, et dans lequel le filtre de synthèse à court terme a une fonction de transfert de la forme $1/A(z) = 1/[A^F(z) \cdot A^B(z)]$ avec

$$A^B(z) = \prod_{p=1}^{q_B} A^{B,p}(z)$$

5

Bien qu'on accorde une importance particulière aux applications de l'invention dans le domaine du codage/ décodage à analyse par synthèse, il convient d'observer que le procédé d'analyse par prédiction linéaire à étages multiples proposé selon l'invention comporte de nombreuses autres applications dans le traitement de signaux audio, par exemple dans les codeurs prédictifs par transformée, dans des systèmes de reconnaissance de parole, dans des systèmes d'accentuation de parole (speech enhancement) ...

10

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

- 15 - la figure 1 est un organigramme d'un procédé d'analyse par prédiction linéaire selon l'invention ;
- la figure 2 est un diagramme spectral comparant les résultats d'un procédé selon l'invention avec ceux d'un procédé conventionnel d'analyse par prédiction linéaire ;
- les figures 3 et 4 sont des schémas synoptiques d'un décodeur et d'un codeur CELP pouvant mettre en oeuvre l'invention ;
- 20 - les figures 5 et 6 sont des schémas synoptiques de variantes de décodeur et de codeur CELP pouvant mettre en oeuvre l'invention ; et
- les figures 7 et 8 sont des schémas synoptiques d'autres variantes de décodeur et de codeur CELP pouvant mettre en oeuvre d'invention.

25 Le signal audiofréquence à analyser dans le procédé illustré par la figure 1 est noté $s^0(n)$. Il est supposé disponible sous forme d'échantillons numériques, l'entier n désignant les instants d'échantillonnage successifs. Le procédé d'analyse par prédiction linéaire comprend q étages successifs $5_1, \dots, 5_p, \dots, 5_q$. A chaque étage de prédiction 5_p ($1 \leq p \leq q$), on effectue une prédiction linéaire d'ordre M_p d'un signal d'entrée $s^{p-1}(n)$. Le signal d'entrée du premier étage 5_1 est constitué par le signal audiofréquence à analyser $s^0(n)$, tandis que le signal d'entrée d'un étage 5_{p+1} ($1 \leq p < q$) est constitué par le signal $s^p(n)$, obtenu à une étape notée 6_p en appliquant au signal d'entrée $s^{p-1}(n)$ du p-ième étage 5_p un filtrage au moyen d'un filtre de fonction de transfert

30

$$A^p(z) = 1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i} \tag{4}$$

35

40 où les coefficients a_i^p ($1 \leq i \leq M_p$) sont les coefficients de prédiction linéaire obtenus à l'étage 5_p . Les méthodes d'analyse par prédiction linéaire pouvant être mises en oeuvre dans les différents étages $5_1, \dots, 5_q$ sont bien connues dans la technique.

On pourra par exemple se reporter aux ouvrages "Digital Processing of Speech Signals" de L.R. Rabiner et R.W. Shafer, Prentice-Hall Int., 1978 et "Linear Prediction of Speech" de J.D. Markel et A.H. Gray, Springer Verlag Berlin Heidelberg, 1976. On peut notamment utiliser l'algorithme de Levinson-Durbin, qui comporte les étapes suivantes (pour chaque étage 5_p) :

45

- évaluation de M_p autocorrélations $R(i)$ ($0 \leq i \leq M_p$) du signal d'entrée $s^{p-1}(n)$ de l'étage sur une fenêtre d'analyse de Q échantillons :

50

$$R(i) = \sum_{n=i}^{Q-1} s^*(n) \cdot s^*(n-i)$$

55

avec $s^*(n) = s^{p-1}(n) \cdot f(n)$, $f(n)$ désignant une fonction de fenêtrage de longueur Q, par exemple une fonction rectangulaire ou une fonction de Hamming ;

- évaluation récursive des coefficients a_i^p :

$$E(0) = R(0)$$

5

Pour i allant de 1 à M_p , faire

10

$$r_i^p = \left[R(i) + \sum_{j=1}^{i-1} a_j^p \cdot r_{i-j}^p \right] / E(i-1)$$

$$a_i^p = -r_i^p$$

15

$$E(i) = [1 - (r_i^p)^2] \cdot E(i-1)$$

Pour j allant de 1 à $i-1$, faire

20

$$a_j^{p,i} = a_j^{p,i-1} - r_i^p \cdot a_{i-j}^{p,i-1}$$

25

Les coefficients a_i^p ($i=1, \dots, M_p$) sont pris égaux aux a_i^{p, M_p} obtenus à la dernière itération. La quantité $E(M_p)$ est l'énergie de l'erreur résiduelle de prédiction de l'étage p . Les coefficients r_i^p , compris entre -1 et 1, sont appelés coefficients de réflexion. Ils peuvent être représentés par les rapports logarithmiques (log-area-ratios) $LAR_i^p = LAR(r_i^p)$, la fonction LAR étant définie par $LAR(r) = \log_{10}[(1-r)/(1+r)]$.

30

Dans un certain nombre d'applications, on a besoin d'opérer une quantification des coefficients de prédiction obtenus. La quantification peut être effectuée sur les coefficients a_i^p directement, sur les coefficients de réflexion associés r_i^p ou sur les rapports logarithmiques LAR_i^p . Une autre possibilité est de quantifier des paramètres de raie spectrale (LSP pour "line spectrum pairs", ou LSF pour "line spectrum frequencies"). Les M_p fréquences de raie spectrale ω_i^p ($1 \leq i \leq M_p$), normalisées entre 0 et π , sont telles que les nombres complexes $1, \exp(j\omega_2^p), \exp(j\omega_4^p), \dots, \exp(j\omega_{M_p}^p)$, soient les racines du polynôme $PP(z) = AP(z) - z^{-(M_p+1)} AP(z^{-1})$ et que les nombres complexes $\exp(j\omega_1^p), \exp(j\omega_3^p), \dots, \exp(j\omega_{M_p-1}^p)$, et -1 soient les racines du polynôme $QP(z) = AP(z) + z^{-(M_p+1)} AP(z^{-1})$. La quantification peut porter sur les fréquences normalisées ω_i^p ou sur leurs cosinus.

35

L'analyse peut être effectuée à chaque étage de prédiction 5_p selon l'algorithme classique de Levinson-Durbin ci-dessus rappelé. D'autres algorithmes fournissant les mêmes résultats, développés plus récemment, peuvent être utilisés avantageusement, notamment l'algorithme de Levinson éclaté (voir "A new Efficient Algorithm to Compute the LSP Parameters for Speech Coding", par S. Saoudi, J.M. Boucher et A. Le Guyader, Signal Processing, Vol.28, 1992, pages 201-212), ou l'utilisation des polynômes de Chebyshev (voir "The Computation of Line Spectrum Frequencies Using Chebyshev Polynomials", par P. Kabal et R.P. Ramachandran, IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, n°6, pages 1419-1426, décembre 1986).

40

Lorsque l'analyse multi-étages représentée sur la figure 1 est réalisée pour définir un filtre de prédiction à court terme du signal audiofréquence $s^0(n)$, on donne à la fonction de transfert $A(z)$ de ce filtre, la forme

45

$$A(z) = \prod_{p=1}^q A^p(z) \tag{5}$$

50

On note que cette fonction de transfert obéit à la forme générale classique donnée par la formule (1), avec $M = M_1 + \dots + M_q$. Toutefois, les coefficients a_i de la fonction $A(z)$ obtenus avec le processus de prédiction multi-étages diffèrent en général de ceux que procure le processus classique de prédiction en un seul étage.

55

Les ordres M_p des prédictions linéaires effectuées augmentent de préférence d'un étage au suivant : $M_1 < M_2 < \dots < M_q$. Ainsi, l'allure de l'enveloppe spectrale du signal analysé est modélisée relativement grossièrement au premier étage 5_1 ($M_1=2$ par exemple), et cette modélisation s'affine d'étage en étage sans perdre l'information globale fournie par le premier étage. On évite ainsi que soient insuffisamment pris en compte des paramètres comme la pente générale du spectre qui sont perceptuellement importants, particulièrement dans le cas de signaux en bande élargie et/ou à

forte dynamique spectrale.

Dans une réalisation typique, le nombre q d'étages de prédiction successifs est égal à 2. Si on a pour objectif un filtre de synthèse d'ordre M, on peut alors prendre M1=2 et M2=M-2, les coefficients a_i du filtre (équation (1)) étant donnés par :

$$a_1 = a_1^1 + a_1^2 \quad (9)$$

$$a_2 = a_2^1 + a_1^1 a_1^2 + a_2^2 \quad (10)$$

$$a_k = a_2^1 a_{k-2}^2 + a_1^1 a_{k-1}^2 + a_k^2 \text{ pour } 2 < k \leq M-2 \quad (11)$$

$$a_{M-1} = a_2^1 a_{M-3}^2 + a_1^1 a_{M-2}^2 \quad (12)$$

$$a_M = a_2^1 a_{M-2}^2 \quad (13)$$

Pour la représentation et éventuellement la quantification du spectre à court terme, il est possible d'adopter l'un des jeux de paramètres spectraux précédemment évoqués (a_i^p, r_i^p, LAR_i^p, ω_i^p ou cos ω_i^p pour 1 ≤ i ≤ M_p) pour chacun des étages (1 ≤ p ≤ q), ou encore les mêmes paramètres spectraux mais pour le filtre composé calculé selon les relations (9) à (13) (a_i, r_i, LAR_i, ω_i ou cos ω_i pour 1 ≤ i ≤ M). Le choix entre ces paramètres de représentation, ou d'autres encore, dépend des contraintes de chaque application particulière.

Le graphique de la figure 2 montre une comparaison des enveloppes spectrales d'une portion voisée de 30 ms d'un signal de parole, modélisées par un processus classique de prédiction linéaire à un étage avec M=15 (courbe II) et par un processus selon l'invention de prédiction linéaire en q=2 étages avec M1=2 et M2=13 (courbe III). La fréquence d'échantillonnage F_e du signal était de 16 kHz. Le spectre du signal (module de sa transformée de Fourier) est représenté par la courbe I. Ce spectre est représentatif des signaux audiofréquence qui ont, en moyenne, plus d'énergie aux basses fréquences qu'aux hautes fréquences. La dynamique spectrale est parfois supérieure à celle de la figure 2 (60 dB). Les courbes (II) et (III) correspondent aux enveloppes spectrales modélisées |1/A(e^{2jπf/F_e})|. On voit que le procédé d'analyse selon l'invention améliore sensiblement la modélisation du spectre, particulièrement aux hautes fréquences (f > 4 kHz). La pente générale du spectre et ses formants en haute fréquence sont mieux respectés par le processus d'analyse en plusieurs étages.

L'invention est décrite ci-après dans son application à un codeur de parole de type CELP.

Le processus de synthèse de parole mis en oeuvre dans un codeur et un décodeur CELP est illustré sur la figure 3. Un générateur d'excitation 10 délivre un code d'excitation c_k appartenant à un répertoire prédéterminé en réponse à un index k. Un amplificateur 12 multiplie ce code d'excitation par un gain d'excitation β, et le signal résultant est soumis à un filtre 14 de synthèse à long terme. Le signal de sortie u du filtre 14 est à son tour soumis à un filtre 16 de synthèse à court terme, dont la sortie s constitue ce qu'on considère ici comme le signal de parole synthétique. Ce signal synthétique est appliqué à un post-filtre 17 destiné à améliorer la qualité subjective de la parole reconstruite. Les techniques de post-filtrage sont bien connues dans le domaine du codage de parole (voir J.H. Chen et A. Gersho : "Adaptive postfiltering for quality enhancement of coded speech", IEEE Trans. on Speech and Audio Processing, Vol. 3-1, pages 59-71, janvier 1995). Dans l'exemple représenté, les coefficients du post-filtre 17 sont obtenus à partir des paramètres LPC caractérisant le filtre de synthèse à court terme 16. On comprendra que, comme dans certains décodeurs CELP actuels, le post-filtre 17 pourrait également comporter une composante de post-filtrage à long terme.

Les signaux précités sont des signaux numériques représentés par exemple par des mots de 16 bits à une cadence d'échantillonnage F_e égale par exemple à 16 kHz pour un codeur en bande élargie (50-7000 Hz). Les filtres de synthèse 14, 16 sont en général des filtres purement récurrents. Le filtre 14 de synthèse à long terme a typiquement une fonction de transfert de la forme 1/B(z) avec B(z)=1-Gz^{-T}. Le retard T et le gain G constituent des paramètres de prédiction à long terme (LTP) qui sont déterminés d'une manière adaptative par le codeur. Les paramètres LPC définissant le filtre 16 de synthèse à court terme sont déterminés au codeur par un procédé d'analyse par prédiction linéaire du signal de parole. Dans les codeurs et décodeurs CELP habituels, la fonction de transfert du filtre 16 est généralement de la forme 1/A(z) avec A(z) de la forme (1). La présente invention propose d'adopter une forme semblable de la fonction de transfert, dans laquelle A(z) est décomposée selon (7) comme indiqué précédemment. A titre d'exemple, les paramètres des différents étages peuvent être q=2, M1=2, M2=13 (M=M1+M2=15).

On désigne ici par "signal d'excitation" le signal $u(n)$ appliqué au filtre de synthèse à court terme 14. Ce signal d'excitation comporte une composante LTP $G \cdot u(n-T)$ et une composante résiduelle, ou séquence d'innovation, $\beta c_k(n)$. Dans un codeur à analyse par synthèse, les paramètres caractérisant la composante résiduelle et, optionnellement, la composante LTP sont évalués en boucle fermée, en utilisant un filtre de pondération perceptuelle.

La figure 4 montre le schéma d'un codeur CELP. Le signal de parole $s(n)$ est un signal numérique, par exemple fourni par un convertisseur analogique-numérique 20 traitant le signal de sortie amplifié et filtré d'un microphone 22. Le signal $s(n)$ est numérisé en trames successives de Λ échantillons elles-mêmes divisées en sous-trames, ou trames d'excitation, de L échantillons (par exemple $\Lambda=160$, $L=32$).

Les paramètres LPC, LTP et EXC (index k et gain d'excitation β) sont obtenus au niveau du codeur par trois modules d'analyse respectifs 24, 26, 28. Ces paramètres sont ensuite quantifiés de façon connue en vue d'une transmission numérique efficace, puis soumis à un multiplexeur 30 qui forme le signal de sortie du codeur. Ces paramètres sont également fournis à un module 32 de calcul d'états initiaux de certains filtres du codeur. Ce module 32 comprend essentiellement une chaîne de décodage telle que celle représentée sur la figure 3. Comme le décodeur, le module 32 opère sur la base des paramètres LPC, LTP et EXC quantifiés. Si une interpolation des paramètres LPC est effectuée au décodeur, comme il est courant, la même interpolation est effectuée par le module 32. Le module 32 permet de connaître au niveau du codeur les états antérieurs des filtres de synthèse 14, 16 du décodeur, déterminés en fonction des paramètres de synthèse et d'excitation antérieurs à la sous-trame considérée.

Dans une première étape du processus de codage, le module 24 d'analyse à court terme détermine les paramètres LPC définissant le filtre de synthèse à court terme, en analysant les corrélations à court terme du signal de parole $s(n)$. Cette détermination est effectuée par exemple une fois par trame de Λ échantillons, de manière à s'adapter à l'évolution du contenu spectral du signal de parole. Elle consiste selon l'invention à mettre en oeuvre le procédé d'analyse illustré par la figure 1 avec $s^0(n)=s(n)$.

L'étape suivante du codage consiste en la détermination des paramètres LTP de prédiction à long terme. Ceux-ci sont par exemple déterminés une fois par sous-trame de L échantillons. Un soustracteur 34 soustrait du signal de parole $s(n)$ la réponse à un signal d'entrée nul du filtre de synthèse à court terme 16. Cette réponse est déterminée par un filtre 36 de fonction de transfert $1/A(z)$ dont les coefficients sont donnés par les paramètres LPC qui ont été déterminés par le module 24, et dont les états initiaux \hat{s} sont fournis par le module 32 de façon à correspondre aux $M=M_1 + \dots + M_q$ derniers échantillons du signal synthétique. Le signal de sortie du soustracteur 34 est soumis à un filtre 38 de pondération perceptuelle dont le rôle est d'accentuer les portions du spectre où les erreurs sont les plus perceptibles, c'est-à-dire les zones inter-formantiques.

La fonction de transfert $W(z)$ du filtre de pondération perceptuelle 38 est de la forme $W(z)=AN(z)/AP(z)$ où $AN(z)$ et $AP(z)$ sont des fonctions de transfert d'ordre M de type RIF (réponse impulsionnelle finie). Les coefficients respectifs b_i et c_i ($1 \leq i \leq M$) des fonctions $AN(z)$ et $AP(z)$ sont calculés pour chaque trame par un module 39 d'évaluation de la pondération perceptuelle qui les fournit au filtre 38. Une première possibilité est de prendre $AN(z)=A(z/\gamma_1)$ et $AP(z)=A(z/\gamma_2)$ avec $0 \leq \gamma_2 \leq \gamma_1 \leq 1$, ce qui revient à la forme habituelle (2) avec $A(z)$ de la forme (7). Dans le cas d'un signal en bande élargie avec $q=2$, $M_1=2$ et $M_2=13$, on a trouvé que le choix $\gamma_1=0,92$ et $\gamma_2=0,6$ fournissait de bons résultats.

L'invention permet toutefois, avec une surcharge de calculs très faible, d'avoir une plus grande souplesse quant à la mise en forme du bruit de quantification, en adoptant la forme (6) pour $W(z)$, soit :

$$AN(z) = \prod_{p=1}^q A^p(z/\gamma_1^p)$$

$$AP(z) = \prod_{p=1}^q A^p(z/\gamma_2^p)$$

Dans le cas d'un signal en bande élargie avec $q=2$, $M_1=2$ et $M_2=13$, on a trouvé que le choix $\gamma_1^1=0,9$, $\gamma_2^1=0,65$, $\gamma_1^2=0,95$ et $\gamma_2^2=0,75$ fournissait de bons résultats. Le terme $A^1(z/\gamma_1^1)/A^1(z/\gamma_2^1)$ permet de régler la pente générale du filtre 38, tandis que le terme $A^2(z/\gamma_1^2)/A^2(z/\gamma_2^2)$ permet de régler le masquage au niveau des formants.

L'analyse LTP en boucle fermée effectuée par le module 26 consiste, de façon classique, à sélectionner pour chaque sous-trame le retard T qui maximise la corrélation normalisée:

$$\left[\sum_{n=0}^{L-1} x'(n) \cdot y_T(n) \right]^2 / \left[\sum_{n=0}^{L-1} [y_T(n)]^2 \right]$$

5 où $x'(n)$ désigne le signal de sortie du filtre 38 pendant la sous-trame considérée, et $y_T(n)$ désigne le produit de convolution $u(n-T) \cdot h'(n)$. Dans l'expression ci-dessus, $h'(0), h'(1), \dots, h'(L-1)$ désigne la réponse impulsionnelle du filtre de synthèse pondéré, de fonction de transfert $W(z)/A(z)$. Cette réponse impulsionnelle h' est obtenue par un module 10 de calcul de réponses impulsionnelles, en fonction des coefficients b_i et c_i fournis par le module 39 et des paramètres LPC qui ont été déterminés pour la sous-trame, le cas échéant après quantification et interpolation. Les échantillons $u(n-T)$ sont les états antérieurs du filtre 14 de synthèse à long terme, fournis par le module 32. Pour les retards T inférieurs à la longueur d'une sous-trame, les échantillons manquants $u(n-T)$ sont obtenus par interpolation sur la base des échantillons antérieurs, ou à partir du signal de parole. Les retards T , entiers ou fractionnaires, sont sélectionnés 15 dans une fenêtre déterminée. Pour réduire la plage de recherche en boucle fermée, et donc pour réduire le nombre de convolutions $y_T(n)$ à calculer, on peut d'abord déterminer un retard T' en boucle ouverte par exemple une fois par trame, puis sélectionner les retards en boucle fermée pour chaque sous-trame dans un intervalle réduit autour de T' . La recherche en boucle ouverte consiste plus simplement à déterminer le retard T' qui maximise l'autocorrélation du signal de parole $s(n)$ éventuellement filtré par le filtre inverse de fonction de transfert $A(z)$. Une fois que le retard T a 20 été déterminé, le gain G de prédiction à long terme est obtenu par :

$$G = \left[\sum_{n=0}^{L-1} x'(n) \cdot y_T(n) \right] / \left[\sum_{n=0}^{L-1} [y_T(n)]^2 \right]$$

25 Pour rechercher l'excitation CELP relative à une sous-trame, le signal $Gy_T(n)$, qui a été calculé par le module 26 pour le retard optimal T , est d'abord soustrait du signal $x'(n)$ par le soustracteur 42. Le signal résultant $x(n)$ est soumis à un filtre à rebours 44 qui fournit un signal $D(n)$ donné par :

$$D(n) = \sum_{i=n}^{L-1} x(i) \cdot h(i-n)$$

35 où $h(0), h(1), \dots, h(L-1)$ désigne la réponse impulsionnelle du filtre composé des filtres de synthèse et du filtre de pondération perceptuelle, calculée par le module 40. En d'autres termes, le filtre composé a pour fonction de transfert $W(z)/[A(z) \cdot B(z)]$. En notation matricielle, on a donc :

$$D = (D(0), D(1), \dots, D(L-1)) = x \cdot H$$

45 avec

$$x = (x(0), x(1), \dots, x(L-1))$$

50 et

55

5

$$H = \begin{pmatrix} h(0) & 0 & \cdot & \cdot & \cdot & 0 \\ h(1) & h(0) & 0 & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ h(L-2) & \cdot & & & h(0) & 0 \\ h(L-1) & h(L-2) & \cdot & \cdot & h(1) & h(0) \end{pmatrix}$$

10

Le vecteur D constitue un vecteur-cible pour le module 28 de recherche de l'excitation. Ce module 28 détermine un mot de code du répertoire qui maximise la corrélation normalisée P_k^2/α_k^2 dans laquelle :

15

$$P_k = D \cdot c_k^T$$

20

$$\alpha_k^2 = c_k \cdot H^T \cdot H \cdot c_k^T = c_k \cdot U \cdot c_k^T$$

L'indice k optimal ayant été déterminé, le gain d'excitation β est pris égal à $\beta = P_k/\alpha_k^2$.

En référence à la figure 3, le décodeur CELP comprend un démultiplexeur 8 recevant le flux binaire issu du codeur. Les valeurs quantifiées des paramètres d'excitation EXC et des paramètres de synthèse LTP et LPC sont fournies au générateur 10, à l'amplificateur 12 et aux filtres 14, 16 pour reconstituer le signal synthétique \hat{s} , qui est soumis au post-filtre 17 puis converti en analogique par le convertisseur 18 avant d'être amplifié puis appliqué à un haut-parleur 19 pour restituer la parole originale.

25

Dans le cas du décodeur de la figure 3, les paramètres LPC sont par exemple constitués par des index de quantification des coefficients de réflexion r_i^p (également appelés coefficients de corrélation partielle ou PARCOR) relatifs aux différents étages de prédiction linéaire. Un module 15 récupère les valeurs quantifiées des r_i^p à partir des index de quantification, et les convertit pour fournir les q jeux de coefficients de prédiction linéaire. Cette conversion est par exemple effectuée par la même méthode récursive que dans l'algorithme de Levinson-Durbin.

30

Les jeux de coefficients a_i^p sont fournis au filtre 16 de synthèse à court terme constitué par une succession de q filtres/étages de fonctions de transfert $1/A^1(z), \dots, 1/A^q(z)$ données par la relation (4). Le filtre 16 pourrait également être en un seul étage de fonction de transfert $1/A(z)$ donnée par la relation (1) dans laquelle les coefficients a_i ont été calculés selon les relations (9) à (13).

35

Les jeux de coefficients a_i^p sont également fournis au post-filtre 17 qui, dans l'exemple considéré, a une fonction de transfert de la forme

40

$$H_{PF}(z) = G_p \frac{APN(z)}{APP(z)} (1 - \mu r_1 z^{-1})$$

où APN(z) et APP(z) sont des fonctions de transfert d'ordre M de type RIF, G_p est un facteur de gain constant, μ est une constante positive et r_1 désigne le premier coefficient de réflexion. Le coefficient de réflexion r_1 peut être celui associé aux coefficients a_i du filtre de synthèse composé, qu'il est alors nécessaire de calculer. On peut également prendre pour r_1 le premier coefficient de réflexion du premier étage de prédiction ($r_1 = r_1^1$) moyennant un éventuel ajustement de la constante μ . Pour le terme APN(z)/APP(z), une première possibilité est de prendre APN(z)=A(z/ β_1) et APP(z)=A(z/ β_2) avec $0 \leq \beta_1 \leq \beta_2 \leq 1$, ce qui revient à la forme habituelle (3) avec A(z) de la forme (7).

45

Comme dans le cas du filtre de pondération perceptuelle du codeur, l'invention permet d'adopter des coefficients β_1 et β_2 différents d'un étage au suivant (formule (8)), soit :

50

$$APP(z) = \prod_{p=1}^q A^p(z/\beta_2^p)$$

55

$$APN(z) = \prod_{p=1}^q A^P(z/\beta_1^p)$$

5

Dans le cas d'un signal en bande élargie avec $q=2$, $M_1=2$ et $M_2=13$, on a trouvé que le choix $\beta_1^1=0,7$, $\beta_2^1=0,9$, $\beta_1^2=0,95$ et $\beta_2^2=0,97$ fournissait de bons résultats.

L'invention a été décrite ci-dessus dans son application à un codeur prédictif à adaptation "forward", c'est-à-dire dans lequel le signal audiofréquence faisant l'objet de l'analyse par prédiction linéaire est le signal d'entrée du codeur. L'invention s'applique également à des codeurs/décodeurs prédictifs à adaptation "backward", dans lesquels le signal synthétique fait l'objet de l'analyse par prédiction linéaire au codeur et au décodeur (voir J.H. Chen et al: "A Low-Delay CELP Coder for the CCITT 16 kbit/s Speech Coding Standard", IEEE J.SAC, Vol.10, n°5, pages 830-848, juin 1992). Les figures 5 et 6 montrent respectivement un décodeur CELP et un codeur CELP à adaptation "backward" mettant en oeuvre la présente invention. Des références numériques identiques à celles des figures 3 et 4 ont été utilisées pour désigner des éléments analogues.

Le décodeur à adaptation "backward" reçoit seulement les valeurs de quantification des paramètres définissant le signal d'excitation $u(n)$ à appliquer au filtre de synthèse à court terme 16. Dans l'exemple considéré, ces paramètres sont l'index k et le gain associé β ainsi que les paramètres LTP. Le signal synthétique $\hat{s}(n)$ est traité par un module 124 d'analyse par prédiction linéaire multi-étages identique au module 24 de la figure 3. Le module 124 fournit les paramètres LPC au filtre 16 pour une ou plusieurs trames suivantes du signal d'excitation, et au post-filtre 17 dont les coefficients sont obtenus comme décrit précédemment.

Le codeur correspondant, représenté sur la figure 6, effectue l'analyse par prédiction linéaire multi-étages sur le signal synthétique généré localement et non sur le signal audio $s(n)$. Il comprend ainsi un décodeur local 132 consistant essentiellement en les éléments notés 10, 12, 14, 16 et 124 du décodeur de la figure 5. Outre les échantillons u du dictionnaire adaptatif et les états initiaux \hat{s} du filtre 36, le décodeur local 132 fournit les paramètres LPC obtenus par analyse du signal synthétique, qui sont utilisés par le module 39 d'évaluation de la pondération perceptuelle et le module 40 de calcul des réponses impulsionnelles h et h' . Pour le reste, le fonctionnement du codeur est identique à celui du codeur décrit en référence à la figure 4, sauf que le module d'analyse LPC 24 n'est plus nécessaire. Seuls les paramètres EXC et LTP sont envoyés vers le décodeur.

Les figures 7 et 8 sont des schémas synoptiques d'un décodeur CELP et d'un codeur CELP à adaptation mixte. Les coefficients de prédiction linéaire du ou des premiers étages résultent d'une analyse "forward" du signal audiofréquence effectuée par le codeur, tandis que les coefficients de prédiction linéaire du ou des derniers étages résultent d'une analyse "backward" du signal synthétique effectuée par le décodeur (et par un décodeur local prévu dans le codeur). Des références numériques identiques à celles des figures 3 à 6 ont été utilisée pour désigner des éléments analogues.

Le décodeur mixte illustré sur la figure 7 reçoit les valeurs de quantification des paramètres EXC, LTP définissant le signal d'excitation $u(n)$ à appliquer au filtre de synthèse à court terme 16, et les valeurs de quantification des paramètres LPC/F déterminés par l'analyse "forward" effectuée par le codeur. Ces paramètres LPC/F représentent q_F jeux de coefficients de prédiction linéaire $a_1^{F,p}, \dots, a_{M_F p}^{F,p}$ pour $1 \leq p \leq q_F$, et définissent une première composante $1/A^F(z)$ de la fonction de transfert $1/A(z)$ du filtre 16 :

45

$$A^F(z) = \prod_{p=1}^{q_F} A^{F,p}(z) = \prod_{p=1}^{q_F} \left[1 + \sum_{i=1}^{M_F p} a_i^{F,p} \cdot z^{-i} \right] ;$$

Pour l'obtention de ces paramètres LPC/F, le codeur mixte représenté sur la figure 8 comporte un module 224/F qui analyse le signal audiofréquence à coder $s(n)$ de la manière décrite en référence à la figure 1 si $q_F > 1$, ou en un seul étage si $q_F = 1$.

L'autre composante $1/A^B(z)$ du filtre de synthèse à court terme 16 de fonction de transfert $1/A(z) = 1/[A^F(z) \cdot A^B(z)]$ est donnée par

55

$$A^B(z) = \prod_{p=1}^{q_B} A^{B,p}(z) = \prod_{p=1}^{q_B} \left(1 + \sum_{i=1}^{MBp} a_i^{B,p} \cdot z^{-i} \right)$$

5
10
15
Pour déterminer les coefficients $a_i^{B,p}$, le décodeur mixte comporte un filtre inverse 200 de fonction de transfert $A^F(z)$ qui filtre le signal synthétique $\hat{s}^0(n)$ produit par le filtre de synthèse à court terme 16 pour produire un signal synthétique filtré $\hat{s}^0(n)$. Un module 224/B effectue l'analyse par prédiction linéaire de ce signal $\hat{s}^0(n)$ de la manière décrite en référence à la figure 1 si $q_B > 1$, ou en un seul étage si $q_B = 1$. Les coefficients LPC/B ainsi obtenus sont fournis au filtre de synthèse 16 pour définir sa seconde composante pour la trame suivante. Ils sont également fournis, de même que les coefficients LPC/F au post-filtre 17, dont les composantes APN(z) et APP(z) sont soit de la forme $APN(z) = A(z/\beta_1)$, $APP(z) = A(z/\beta_2)$, soit de la forme :

$$APN(z) = \left[\prod_{p=1}^{q_F} A^{F,p}(z/\beta_1^{F,p}) \right] \cdot \left[\prod_{p=1}^{q_B} A^{B,p}(z/\beta_1^{B,p}) \right]$$

$$APP(z) = \left[\prod_{p=1}^{q_F} A^{F,p}(z/\beta_2^{F,p}) \right] \cdot \left[\prod_{p=1}^{q_B} A^{B,p}(z/\beta_2^{B,p}) \right]$$

les paires de coefficient $\beta_1^{F,p}$, $\beta_2^{F,p}$ et $\beta_1^{B,p}$, $\beta_2^{B,p}$ étant optimisables séparément avec $0 \leq \beta_1^{F,p} \leq \beta_2^{F,p} \leq 1$ et $0 \leq \beta_1^{B,p} \leq \beta_2^{B,p} \leq 1$.

Le décodeur local 232 prévu dans le codeur mixte consiste essentiellement en les éléments notés 10, 12, 14, 16, 200 et 224/B du décodeur de la figure 7. Outre les échantillons u du dictionnaire adaptatif et les états initiaux \hat{s} du filtre 36, le décodeur local 232 fournit les paramètres LPC/B qui sont utilisés, avec les paramètres LPC/F fournis par le module d'analyse 224/F, par le module 39 d'évaluation de la pondération perceptuelle et le module 40 de calcul des réponses impulsionnelles h et h' .

La fonction de transfert du filtre de pondération perceptuelle 38 évaluée par le module 39 est soit de la forme $W(z) = A(z/\gamma_1)/A(z/\gamma_2)$, soit de la forme

$$W(z) = \left[\prod_{p=1}^{q_F} \left[\frac{A^{F,p}(z/\gamma_1^{F,p})}{A^{F,p}(z/\gamma_2^{F,p})} \right] \right] \cdot \left[\prod_{p=1}^{q_B} \left[\frac{A^{B,p}(z/\gamma_1^{B,p})}{A^{B,p}(z/\gamma_2^{B,p})} \right] \right]$$

les paires de coefficients $\gamma_1^{F,p}$, $\gamma_2^{F,p}$, et $\gamma_1^{B,p}$, $\gamma_2^{B,p}$ étant optimisables séparément avec $0 \leq \gamma_2^{F,p} \leq \gamma_1^{F,p} \leq 1$ et $0 \leq \gamma_2^{B,p} \leq \gamma_1^{B,p} \leq 1$.

Pour le reste, le fonctionnement du codeur mixte est identique à celui du codeur décrit en référence à la figure 4. Seuls les paramètres EXC, LTP et LPC/F sont envoyés vers le décodeur.

Revendications

1. Procédé d'analyse par prédiction linéaire d'un signal audiofréquence ($s^0(n)$), pour déterminer des paramètres spectraux dépendant d'un spectre à court terme du signal audiofréquence, le procédé comprenant q étages de prédiction successifs (5_p), q étant un entier supérieur à 1, caractérisé en ce qu'à chaque étage de prédiction p ($1 \leq p \leq q$), on détermine des paramètres représentant un nombre M_p , prédéfini pour chaque étage p , de coefficients a_1^p, \dots ,

a_{Mp}^p de prédiction linéaire d'un signal d'entrée dudit étage, le signal audiofréquence à analyser constituant le signal d'entrée ($s^0(n)$) du premier étage, et le signal d'entrée ($s^p(n)$) d'un étage p+1 étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p filtré par un filtre de fonction de transfert

5

$$A^p(z) = 1 + \sum_{i=1}^{Mp} a_i^p \cdot z^{-i}$$

10

2. Procédé d'analyse selon la revendication 1, caractérisé en ce que le nombre Mp de coefficients de prédiction linéaire augmente d'un étage au suivant.

15

3. Procédé de codage d'un signal audiofréquence, comprenant les étapes suivantes :

20

- analyse par prédiction linéaire du signal audiofréquence ($s(n)$) numérisé en trames successives pour déterminer des paramètres (LPC) définissant un filtre de synthèse à court terme (16) ;
- détermination de paramètres d'excitation (k, β, LTP) définissant un signal d'excitation ($u(n)$) à appliquer au filtre de synthèse à court terme (16) pour produire un signal synthétique ($\hat{s}(n)$) représentatif du signal audiofréquence ; et
- production de valeurs de quantification des paramètres définissant le filtre de synthèse à court terme et des paramètres d'excitation,

25

caractérisé en ce que l'analyse par prédiction linéaire est un processus à q étages successifs (5_p), q étant un entier supérieur à 1, ledit processus comportant, à chaque étage de prédiction p ($1 \leq p \leq q$), la détermination de paramètres représentant un nombre Mp , prédéfini pour chaque étage p , de coefficients a_1^p, \dots, a_{Mp}^p de prédiction linéaire d'un signal d'entrée dudit étage, le signal audiofréquence à coder ($s(n)$) constituant le signal d'entrée ($s^0(n)$) du premier étage, et le signal d'entrée ($s^p(n)$) d'un étage p+1 étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p filtré par un filtre de fonction de transfert

30

$$A^p(z) = 1 + \sum_{i=1}^{Mp} a_i^p \cdot z^{-i}$$

35

le filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A(z)$ avec

40

$$A(z) = \prod_{p=1}^q A^p(z)$$

45

4. Procédé de codage selon la revendication 3, caractérisé en ce que le nombre Mp de coefficients de prédiction linéaire augmente d'un étage au suivant.

50

5. Procédé de codage selon la revendication 3 ou 4, caractérisé en ce que certains au moins des paramètres d'excitation sont déterminés en minimisant l'énergie d'un signal d'erreur résultant du filtrage de la différence entre le signal audiofréquence ($s(n)$) et le signal synthétique ($\hat{s}(n)$) par au moins un filtre de pondération perceptuelle (38) dont la fonction de transfert est de la forme $W(z) = A(z/\gamma_1)/A(z/\gamma_2)$ où γ_1 et γ_2 désignent des coefficients d'expansion spectrale tels que $0 \leq \gamma_2 \leq \gamma_1 \leq 1$.

55

6. Procédé de codage selon la revendication 3 ou 4, caractérisé en ce que certains au moins des paramètres d'excitation sont déterminés en minimisant l'énergie d'un signal d'erreur résultant du filtrage de la différence entre le signal audiofréquence ($s(n)$) et le signal synthétique ($\hat{s}(n)$) par au moins un filtre de pondération perceptuelle (38) dont la fonction de transfert est de la forme

$$W(z) = \prod_{p=1}^q [A^p(z/\gamma_1^p) / A^p(z/\gamma_2^p)] ,$$

5 où γ_1^p, γ_2^p désignent des paires de coefficients d'expansion spectrale tels que $0 \leq \gamma_2^p \leq \gamma_1^p \leq 1$ pour $1 \leq p \leq q$.

7. Procédé de décodage d'un flux binaire pour construire un signal audiofréquence codé par ledit flux binaire, caractérisé en ce que :

- 10
- on reçoit des valeurs de quantification de paramètres (LPC) définissant un filtre de synthèse à court terme (16) et de paramètres d'excitation (k, β, LTP), les paramètres définissant le filtre de synthèse représentant un nombre q plus grand que 1 de jeux de coefficients de prédiction linéaire (a_i^p), chaque jeu p comportant un nombre prédéfini M_p de coefficients ;
 - 15 - on produit un signal d'excitation ($u(n)$) sur la base des valeurs de quantification des paramètres d'excitation ; et
 - on produit un signal audiofréquence synthétique ($\hat{s}(n)$) en filtrant le signal d'excitation par un filtre de synthèse (16) ayant une fonction de transfert de la forme $1/A(z)$ avec

20

$$A(z) = \prod_{p=1}^q \left(1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i} \right)$$

25 où les coefficients $a_1^p, \dots, a_{M_p}^p$ correspondent au p -ième jeu de coefficients de prédiction linéaire pour $1 \leq p \leq q$.

8. Procédé de décodage selon la revendication 7, caractérisé en ce que ledit signal audiofréquence synthétique ($\hat{s}(n)$) est appliqué à un post-filtre (17) dont la fonction de transfert ($H_{PF}(z)$) comporte un terme de la forme $A(z/\beta_1) / A(z/\beta_2)$, où β_1 et β_2 désignent des coefficients tels que $0 \leq \beta_1 \leq \beta_2 \leq 1$.

30 9. Procédé de décodage selon la revendication 7, caractérisé en ce que ledit signal audiofréquence synthétique ($\hat{s}(n)$) est appliqué à un post-filtre (17) dont la fonction de transfert ($H_{PF}(z)$) comporte un terme de la forme

35

$$\prod_{p=1}^q [A^p(z/\beta_1^p) / A^p(z/\beta_2^p)] ,$$

40 où β_1^p, β_2^p désignent des paires de coefficients tels que $0 \leq \beta_1^p \leq \beta_2^p \leq 1$ pour $1 \leq p \leq q$, et $A^p(z)$ représente, pour le p -ième jeu de coefficients de prédiction linéaire, la fonction

45

$$A^p(z) = 1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i}$$

10. Procédé de codage d'un premier signal audiofréquence numérisé en trames successives, comprenant les étapes suivantes :

- 50
- analyse par prédiction linéaire d'un second signal audiofréquence ($\hat{s}(n)$) pour déterminer des paramètres (LPC) définissant un filtre de synthèse à court terme (16) ;
 - détermination de paramètres d'excitation (k, β, LTP) définissant un signal d'excitation ($u(n)$) à appliquer au filtre de synthèse à court terme (16) pour produire un signal synthétique ($\hat{s}(n)$) représentatif du premier signal audiofréquence, ce signal synthétique constituant ledit second signal audiofréquence pour au moins une trame
 - 55 - suivante ; et
 - production de valeurs de quantification des paramètres d'excitation,

caractérisé en ce que l'analyse par prédiction linéaire est un processus à q étages successifs (5_p), q étant un entier supérieur à 1, ledit processus comportant, à chaque étage de prédiction p ($1 \leq p \leq q$), la détermination de paramètres représentant un nombre M_p , prédéfini pour chaque étage p, de coefficients $a_1^p, \dots, a_{M_p}^p$ de prédiction linéaire d'un signal d'entrée dudit étage, le second signal audiofréquence ($\hat{s}(n)$) constituant le signal d'entrée ($s^0(n)$) du premier étage, et le signal d'entrée ($s^p(n)$) d'un étage p+1 étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p filtré par un filtre de fonction de transfert

$$A^p(z) = 1 + \sum_{i=1}^{M_p} a_i^p \cdot z^{-i},$$

le filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A(z)$ avec

$$A(z) = \prod_{p=1}^q A^p(z)$$

11. Procédé de codage selon la revendication 10, caractérisé en ce que le nombre M_p de coefficients de prédiction linéaire augmente d'un étage au suivant.

12. Procédé de codage selon la revendication 10 ou 11, caractérisé en ce que certains au moins des paramètres d'excitation sont déterminés en minimisant l'énergie d'un signal d'erreur résultant du filtrage de la différence entre le premier signal audiofréquence ($s(n)$) et le signal synthétique ($\hat{s}(n)$) par au moins un filtre de pondération perceptuelle (38) dont la fonction de transfert est de la forme $W(z) = A(z/\gamma_1)/A(z/\gamma_2)$ où γ_1 et γ_2 désignent des coefficients d'expansion spectrale tels que $0 \leq \gamma_2 \leq \gamma_1 \leq 1$.

13. Procédé de codage selon la revendication 10 ou 11, caractérisé en ce que certains au moins des paramètres d'excitation sont déterminés en minimisant l'énergie d'un signal d'erreur résultant du filtrage de la différence entre le premier signal audiofréquence ($s(n)$) et le signal synthétique ($\hat{s}(n)$) par au moins un filtre de pondération perceptuelle (38) dont la fonction de transfert est de la forme

$$W(z) = \prod_{p=1}^q [A^p(z/\gamma_1^p) / A^p(z/\gamma_2^p)],$$

où γ_1^p, γ_2^p désignent des paires de coefficients d'expansion spectrale tels que $0 \leq \gamma_2^p \leq \gamma_1^p \leq 1$ pour $1 \leq p \leq q$.

14. Procédé de décodage d'un flux binaire pour construire en trames successives un signal audiofréquence codé par ledit flux binaire, caractérisé en ce que :

- on reçoit des valeurs de quantification de paramètres d'excitation (k, β, LTP) ;
- on produit un signal d'excitation ($u(n)$) sur la base des valeurs de quantification des paramètres d'excitation ;
- on produit un signal audiofréquence synthétique ($\hat{s}(n)$) en filtrant le signal d'excitation par un filtre de synthèse à court terme (16) ;
- on effectue une analyse par prédiction linéaire du signal synthétique ($\hat{s}(n)$) pour obtenir des coefficients du filtre de synthèse à court terme (16) pour au moins une trame suivante,

et en ce que l'analyse par prédiction linéaire est un processus à q étages successifs (5_p), q étant un entier supérieur à 1, ledit processus comportant, à chaque étage de prédiction p ($1 \leq p \leq q$), la détermination de paramètres représentant un nombre M_p , prédéfini pour chaque étage p, de coefficients $a_1^p, \dots, a_{M_p}^p$ de prédiction linéaire d'un signal d'entrée dudit étage, le signal synthétique ($\hat{s}(n)$) constituant le signal d'entrée ($s^0(n)$) du premier étage, et le signal d'entrée ($s^p(n)$) d'un étage p+1 étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p filtré par un filtre de fonction de transfert

$$A^P(z) = 1 + \sum_{i=1}^{MP} a_i^P \cdot z^{-i},$$

5

le filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A(z)$ avec

10

$$A(z) = \prod_{p=1}^Q A^P(z)$$

15 15. Procédé de décodage selon la revendication 14, caractérisé en ce que ledit signal audiofréquence synthétique $\hat{s}(n)$ est appliqué à un post-filtre (17) dont la fonction de transfert ($H_{PF}(z)$) comporte un terme de la forme $A(z/\beta_1)/A(z/\beta_2)$, où β_1 et β_2 désignent des coefficients tels que $0 \leq \beta_1 \leq \beta_2 \leq 1$.

20 16. Procédé de décodage selon la revendication 14, caractérisé en ce que ledit signal audiofréquence synthétique $\hat{s}(n)$ est appliqué à un post-filtre (17) dont la fonction de transfert ($H_{PF}(z)$) comporte un terme de la forme

25

$$\prod_{p=1}^Q [A^P(z/\beta_1^P) / A^P(z/\beta_2^P)],$$

où β_1^P, β_2^P désignent des paires de coefficients tels que $0 \leq \beta_1^P \leq \beta_2^P \leq 1$ pour $1 \leq p \leq Q$.

30 17. Procédé de codage d'un premier signal audiofréquence numérisé en trames successives, caractérisé en ce qu'il comprend les étapes suivantes :

- analyse par prédiction linéaire du premier signal audiofréquence ($s(n)$) pour déterminer des paramètres (LPC/F) définissant une première composante d'un filtre de synthèse à court terme (16) ;
- détermination de paramètres d'excitation (k, β, LTP) définissant un signal d'excitation ($u(n)$) à appliquer au filtre de synthèse à court terme (16) pour produire un signal synthétique ($\hat{s}(n)$) représentatif du premier signal audiofréquence ;
- production de valeurs de quantification des paramètres définissant la première composante du filtre de synthèse à court terme et des paramètres d'excitation ;
- filtrage du signal synthétique ($\hat{s}(n)$) par un filtre de fonction de transfert correspondant à l'inverse de la fonction de transfert de la première composante du filtre de synthèse à court terme ; et
- analyse par prédiction linéaire du signal synthétique filtré ($\hat{s}^0(n)$) pour obtenir des coefficients d'une seconde composante du filtre de synthèse à court terme pour au moins une trame suivante,

45 en ce que l'analyse par prédiction linéaire du premier signal audiofréquence ($s(n)$) est un processus à q_F étages successifs (5_p), q_F étant un entier au moins égal à 1, ledit processus à q_F étages comportant, à chaque étage de prédiction p ($1 \leq p \leq q_F$), la détermination de paramètres représentant un nombre MF_p , prédéfini pour chaque étage p , de coefficients $a_1^{F,p}, \dots, a_{MF_p}^{F,p}$ de prédiction linéaire d'un signal d'entrée dudit étage, le premier signal audiofréquence ($s(n)$) constituant le signal d'entrée ($s^0(n)$) du premier étage du processus à q_F étages, et le signal d'entrée ($s^p(n)$) d'un étage $p+1$ du processus à q_F étages étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p du processus à q_F étages filtré par un filtre de fonction de transfert

55

$$A^{F,p}(z) = 1 + \sum_{i=1}^{MF_p} a_i^{F,p} \cdot z^{-i},$$

la première composante du filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A^F$

(z) avec

5

$$A^F(z) = \prod_{p=1}^{Q_F} A^{F,p}(z) ,$$

10

et en ce que l'analyse par prédiction linéaire du signal synthétique filtré est un processus à q_B étages successifs (5_p), q_B étant un entier au moins égal à 1, ledit processus à q_B étages comportant, à chaque étage de prédiction p ($1 \leq p \leq q_B$), la détermination de paramètres représentant un nombre MB_p , prédéfini pour chaque étage p , de coefficients $a_1^{B,p}, \dots, a_{MB_p}^{B,p}$ de prédiction linéaire d'un signal d'entrée dudit étage, le signal synthétique filtré ($\hat{s}^0(n)$) constituant le signal d'entrée ($s^0(n)$) du premier étage du processus à q_B étages, et le signal d'entrée ($s^p(n)$) d'un étage $p+1$ du processus à q_B étages étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p du processus à q_B étages filtré par un filtre de fonction de transfert

15

20

$$A^{B,p}(z) = 1 + \sum_{i=1}^{MB_p} a_i^{B,p} \cdot z^{-i} ,$$

25

la seconde composante du filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A^B(z)$ avec

30

$$A^B(z) = \prod_{p=1}^{q_B} A^{B,p}(z) ,$$

et le filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A(z)$ avec $A(z) = A^F(z) \cdot A^B(z)$.

35

18. Procédé de codage selon la revendication 17, caractérisé en ce que certains au moins des paramètres d'excitation sont déterminés en minimisant l'énergie d'un signal d'erreur résultant du filtrage de la différence entre le premier signal audiofréquence ($s(n)$) et le signal synthétique ($\hat{s}(n)$) par au moins un filtre de pondération perceptuelle (38) dont la fonction de transfert est de la forme $W(z) = A(z/\gamma_1)/A(z/\gamma_2)$ où γ_1 et γ_2 désignent des coefficients d'expansion spectrale tels que $0 \leq \gamma_2 \leq \gamma_1 \leq 1$.

40

19. Procédé de codage selon la revendication 17, caractérisé en ce que certains au moins des paramètres d'excitation sont déterminés en minimisant l'énergie d'un signal d'erreur résultant du filtrage de la différence entre le premier signal audiofréquence ($s(n)$) et le signal synthétique ($\hat{s}(n)$) par au moins un filtre de pondération perceptuelle (38) dont la fonction de transfert est de la forme

45

50

$$W(z) = \left[\prod_{p=1}^{Q_F} \left[\frac{A^{F,p}(z/\gamma_1^{F,p})}{A^{F,p}(z/\gamma_2^{F,p})} \right] \right] \cdot \left[\prod_{p=1}^{q_B} \left[\frac{A^{B,p}(z/\gamma_1^{B,p})}{A^{B,p}(z/\gamma_2^{B,p})} \right] \right]$$

55

où $\gamma_1^{F,p}, \gamma_2^{F,p}$ désignent des paires de coefficients d'expansion spectrale tels que $0 \leq \gamma_2^{F,p} \leq \gamma_1^{F,p} \leq 1$ pour $1 \leq p \leq Q_F$, et $\gamma_1^{B,p}, \gamma_2^{B,p}$ désignent des paires de coefficients d'expansion spectrale tels que $0 \leq \gamma_2^{B,p} \leq \gamma_1^{B,p} \leq 1$ pour $1 \leq p \leq q_B$.

20. Procédé de décodage d'un flux binaire pour construire en trames successives un signal audiofréquence codé par ledit flux binaire, caractérisé en ce que :

- on reçoit des valeurs de quantification de paramètres (LPC/F) définissant une première composante d'un filtre de synthèse à court terme (16) et de paramètres d'excitation (k,β,LTP), les paramètres définissant la première composante du filtre de synthèse à court terme représentant un nombre q_F au moins égal à 1 de jeux de coefficients de prédiction linéaire $a_1^{F,p}, \dots, a_{MFp}^{F,p}$ pour $1 \leq p \leq q_F$, chaque jeu p comportant un nombre prédéfini MFp de coefficients, la première composante du filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A^F(z)$ avec

$$A^F(z) = \prod_{p=1}^{q_F} A^{F,p}(z) = \prod_{p=1}^{q_F} \left(1 + \sum_{i=1}^{MFp} a_i^{F,p} \cdot z^{-i} \right) ;$$

- on produit un signal d'excitation (u(n)) sur la base des valeurs de quantification des paramètres d'excitation ;
- on produit un signal audiofréquence synthétique ($\hat{s}(n)$) en filtrant le signal d'excitation par un filtre de synthèse à court terme (16) de fonction de transfert $1/A(z)$ avec $A(z)=A^F(z) \cdot A^B(z)$, $1/A^B(z)$ représentant la fonction de transfert d'une seconde composante du filtre de synthèse à court terme (16) ;
- on filtre le signal synthétique ($\hat{s}(n)$) par un filtre de fonction de transfert $A^F(z)$; et
- on effectue une analyse par prédiction linéaire du signal synthétique filtré ($\hat{s}^0(n)$) pour obtenir des coefficients de la seconde composante du filtre de synthèse à court terme (16) pour au moins une trame suivante,

et en ce que l'analyse par prédiction linéaire du signal synthétique filtré est un processus à q_B étages successifs (5_p), q_B étant un entier au moins égal à 1, ledit processus comportant, à chaque étage de prédiction p ($1 \leq p \leq q_B$), la détermination de paramètres représentant un nombre MBp, prédéfini pour chaque étage p, de coefficients $a_1^{B,p}, \dots, a_{MBp}^{B,p}$ de prédiction linéaire d'un signal d'entrée dudit étage, le signal synthétique filtré ($\hat{s}^0(n)$) constituant le signal d'entrée ($s^0(n)$) du premier étage, et le signal d'entrée ($s^p(n)$) d'un étage p+1 étant constitué par le signal d'entrée ($s^{p-1}(n)$) de l'étage p filtré par un filtre de fonction de transfert

$$A^{B,p}(z) = 1 + \sum_{i=1}^{MBp} a_i^{B,p} \cdot z^{-i} ,$$

la seconde composante du filtre de synthèse à court terme (16) ayant une fonction de transfert de la forme $1/A^B(z)$ avec

$$A^B(z) = \prod_{p=1}^{q_B} A^{B,p}(z)$$

21. Procédé de décodage selon la revendication 20, caractérisé en ce que ledit signal audiofréquence synthétique ($\hat{s}(n)$) est appliqué à un post-filtre (17) dont la fonction de transfert ($H_{PF}(z)$) comporte un terme de la forme $A(z/\beta_1) / A(z/\beta_2)$, où β_1 et β_2 désignent des coefficients tels que $0 \leq \beta_1 \leq \beta_2 \leq 1$.

22. Procédé de décodage selon la revendication 20, caractérisé en ce que ledit signal audiofréquence synthétique ($\hat{s}(n)$) est appliqué à un post-filtre (17) dont la fonction de transfert ($H_{PF}(z)$) comporte un terme de la forme

5

$$\left[\prod_{p=1}^{q_F} \left[\frac{A^{F,p}(z/\beta_1^{F,p})}{A^{F,p}(z/\beta_2^{F,p})} \right] \right] \cdot \left[\prod_{p=1}^{q_B} \left[\frac{A^{B,p}(z/\beta_1^{B,p})}{A^{B,p}(z/\beta_2^{B,p})} \right] \right]$$

10

où $\beta_1^{F,p}, \beta_2^{F,p}$ désignent des paires de coefficients tels que $0 \leq \beta_1^{F,p} \leq \beta_2^{F,p} \leq 1$ pour $1 \leq p \leq q_F$, et $\beta_1^{B,p}, \beta_2^{B,p}$ désignent des paires de coefficients tels que $0 \leq \beta_1^{B,p} \leq \beta_2^{B,p} \leq 1$ pour $1 \leq p \leq q_B$.

15

20

25

30

35

40

45

50

55

FIG. 1

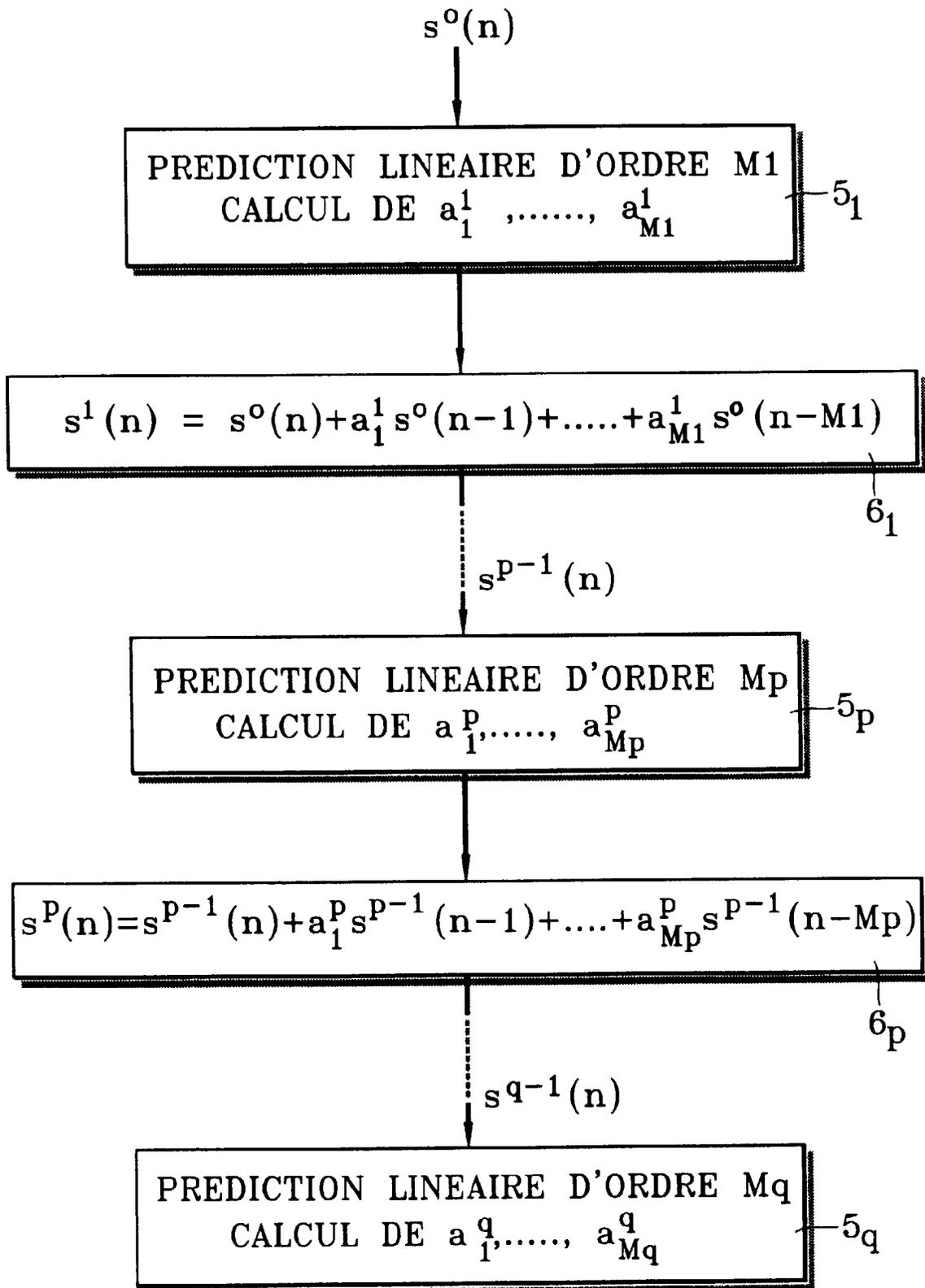
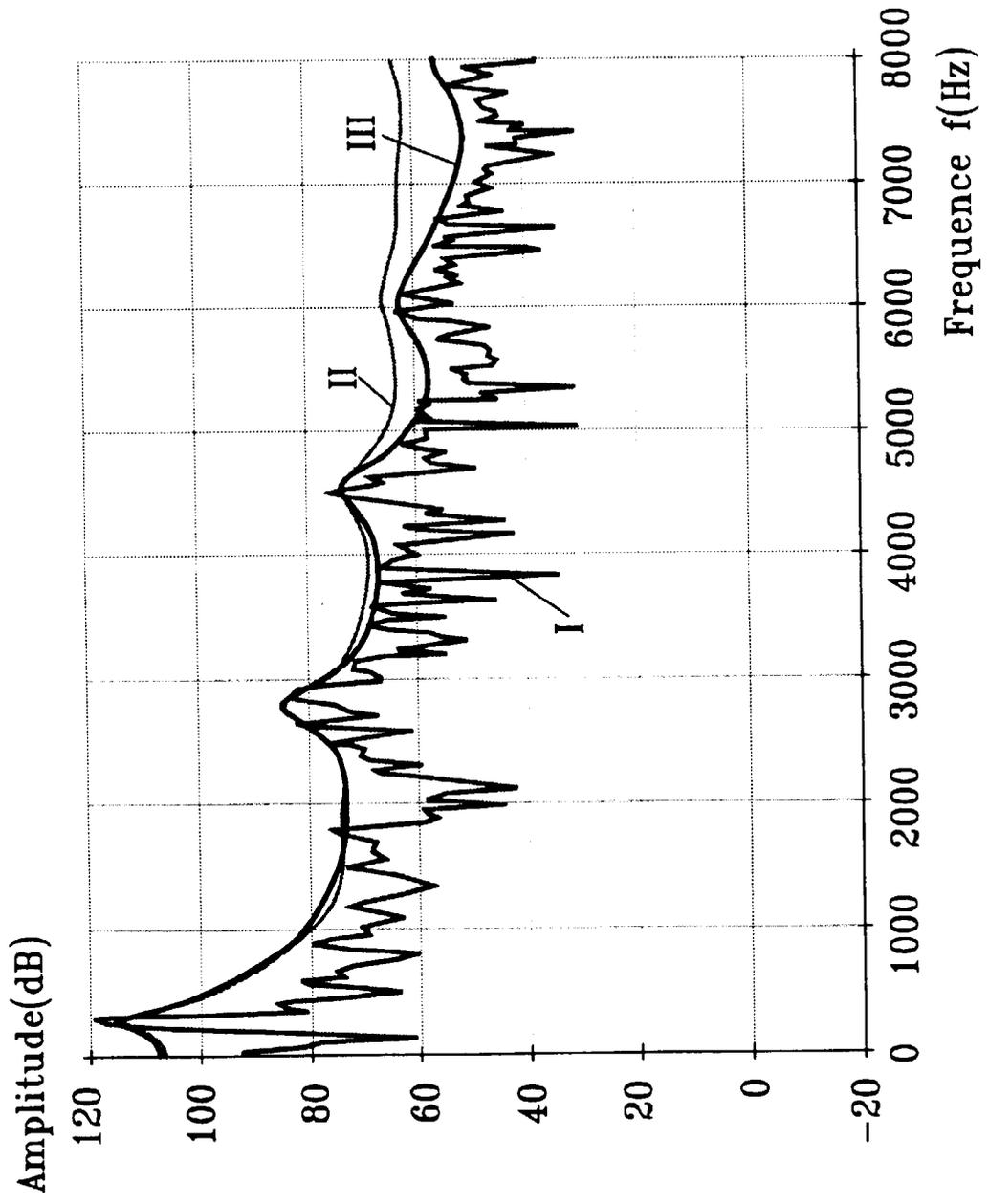
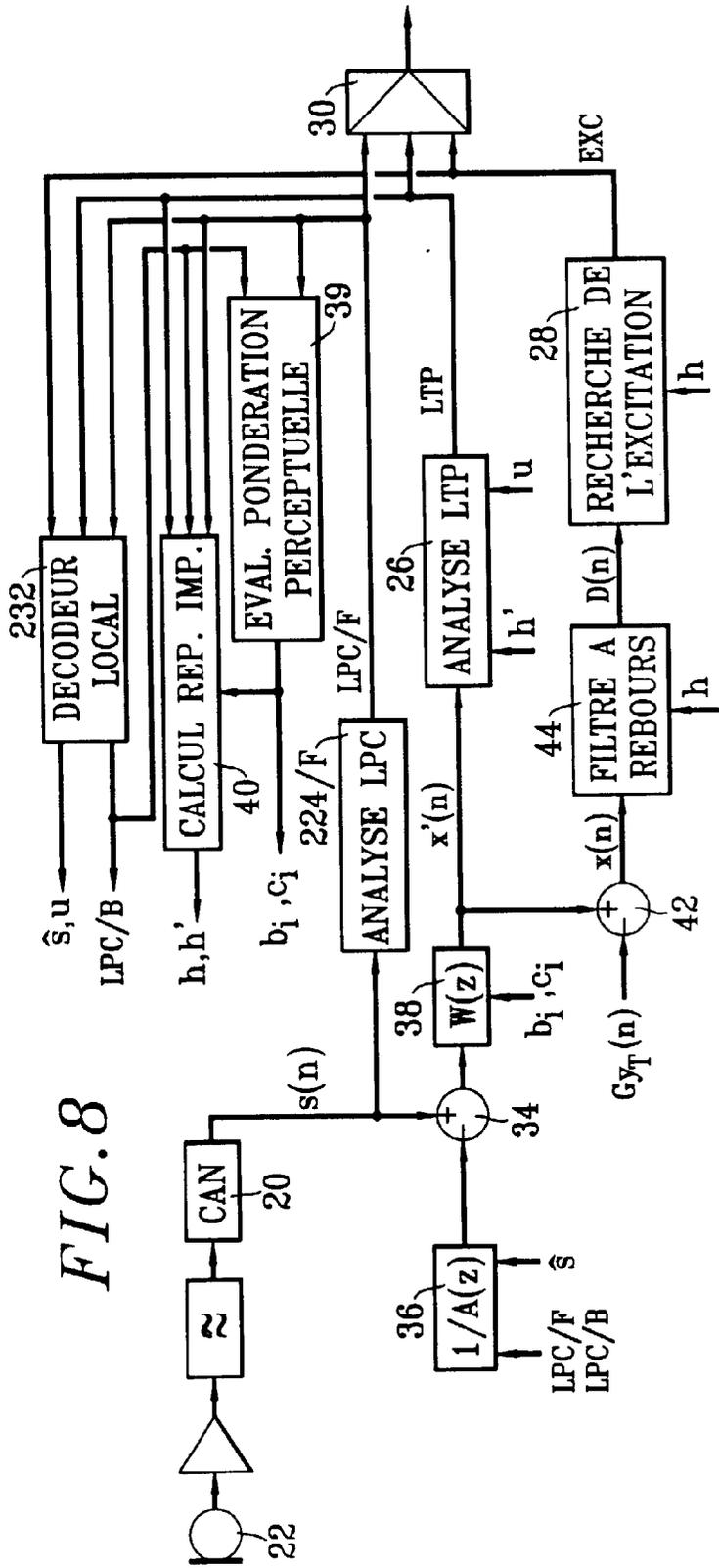
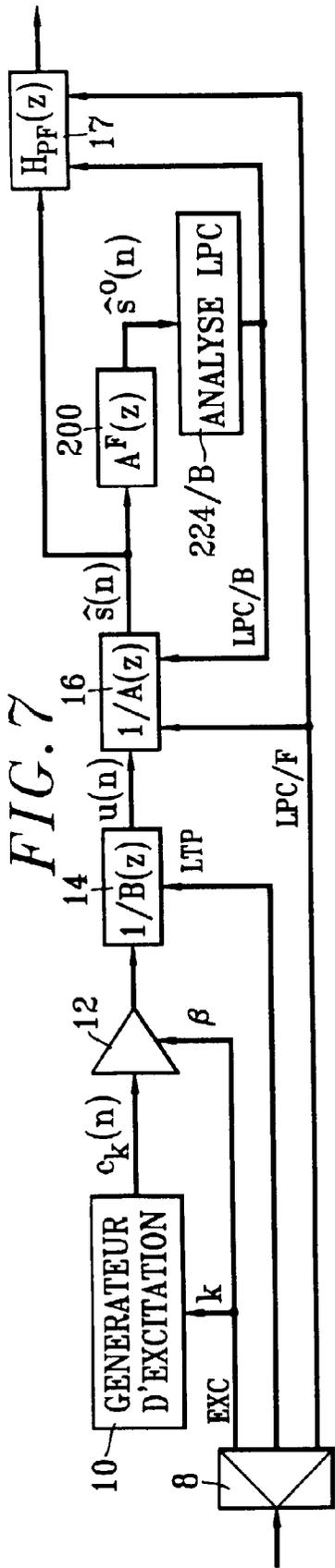


FIG. 2







Office européen
des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numero de la demande
EP 96 40 2715

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.6)
Y	FR 2 284 946 A (INT STANDARD ELECTRIC CORP) 9 Avril 1976 * revendications 1-6; figures 4,11 * * page 10, ligne 32 - page 11, ligne 10 * ---	1-4,7, 10,11, 14,17,20	G10L9/14
Y	US 5 142 581 A (TOKUDA KIYOHITO ET AL) 25 Août 1992 * abrégé; revendication 1; figure 1 * ---	1-4,7, 10,11, 14,17,20	DOMAINES TECHNIQUES RECHERCHES (Int.Cl.6) G10L
A	ICASSP-94. 1994 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (CAT. NO.94CH3387-8), PROCEEDINGS OF ICASSP '94. IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, ADELAIDE, SA, AUSTRALIA, 19-22 APRIL 1, ISBN 0-7803-1775-0, 1994, NEW YORK, NY, USA, IEEE, USA, pages I/493-6 vol.1, XP002013349 KWOK-WAH LAW ET AL: "A novel split residual vector quantization scheme for low bit rate speech coding" * abrégé; figure 1 * ---	1-4,7, 10,11, 14,17,20	
A	WO 83 02346 A (MOTOROLA INC) 7 Juillet 1983 * abrégé * --- -/--	1,2,4,11	
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche LA HAYE		Date d'achèvement de la recherche 24 Février 1997	Examineur Van Doremalen, J
CATEGORIE DES DOCUMENTS CITES X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire		T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	

EPO FORM 1502 03.82 (P/MC02)



Office européen
des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numero de la demande
EP 96 40 2715

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.6)
D,A	SPEECH PROCESSING 1, TORONTO, MAY 14 - 17, 1991, vol. 1, 14 Mai 1991, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, pages 9-12, XP000245155 ORDENTLICH E ET AL: "LOW-DELAY CODE-EXCITED LINEAR-PREDICTIVE CODING OF WIDEBAND SPEECH AT 32 KBPS" * abrégé * * page 10, colonne de gauche, ligne 10 - ligne 20 * -----	5,6,8,9, 12,13, 15,16, 18,19, 21,22	
			DOMAINES TECHNIQUES RECHERCHES (Int.Cl.6)
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche LA HAYE		Date d'achèvement de la recherche 24 Février 1997	Examineur Van Doremalen, J
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>			

EPO FORM 1503 01.82 (P04C02)