

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 481 895 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:
10.12.1997 Bulletin 1997/50

(51) Int Cl.⁶: **G10L 9/14**

(21) Numéro de dépôt: **91402774.3**

(22) Date de dépôt: **17.10.1991**

(54) **Procédé de transmission, à bas débit, par codage CELP d'un signal de parole et système correspondant**

Verfahren und Einrichtung zur Übertragung mit niedriger Bitrate eines Sprachsignals mittels CELP-Codierung

Method and apparatus for low bit rate transmission of a speech signal using CELP coding

(84) Etats contractants désignés:
DE GB

(30) Priorité: **19.10.1990 FR 9012980**

(43) Date de publication de la demande:
22.04.1992 Bulletin 1992/17

(73) Titulaire: **FRANCE TELECOM**
75015 Paris (FR)

(72) Inventeur: **Di Francesco, Renaud**
F-94250 Gentilly (FR)

(74) Mandataire: **Fréchède, Michel et al**
Cabinet Plasseraud
84, rue d'Amsterdam
F-75009 Paris (FR)

(56) Documents cités:
EP-A- 0 379 296 **WO-A-91/13432**

- **ADVANCED IN SPEECH CODING (IEEE WORKSHOP ON SPEECH CODING FOR TELECOMMUNICATIONS, Vancouver, 5 - 8 septembre 1989), pages 145-156, Kluwer**
- **Academic Publishers, Dordrecht, NL; R.A. SALAMI: "Binary pulse excitation: a novel approach to low complexity CELP coding"**

EP 0 481 895 B1

Il est rappelé que: Dans un délai de neuf mois à compter de la date de publication de la mention de la délivrance du brevet européen, toute personne peut faire opposition au brevet européen délivré, auprès de l'Office européen des brevets. L'opposition doit être formée par écrit et motivée. Elle n'est réputée formée qu'après paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

Description

L'invention est relative à un procédé de transmission, à bas débit, par codage CELP d'un signal de parole et au système correspondant.

5 La technique de codage de signaux de parole selon le processus de codage CELP ("Code Excited Linear Prediction" : prédiction linéaire excitée par codes) est actuellement utilisée et a fait l'objet de nombreux travaux. Cette technique de codage d'échantillons numériques représentant le signal de parole est une technique de codage hybride dans laquelle le signal de parole est modélisé par des filtres de prédiction linéaire et des résidus de cette prédiction. De manière générale, les codeurs CELP, ainsi que représenté de manière schématique en figure 1a et 1b, essaient de façon exhaustive tous les éléments d'une liste de formes d'onde. La forme d'onde réalisant la meilleure synthèse du signal est retenue, et son indice, ou adresse caractéristique, est transmis au décodeur. Cette méthode est appelée analyse par synthèse. La liste des formes d'onde, mémorisée au niveau du codeur et du décodeur est appelée dictionnaire.

10 La qualité d'un codeur CELP dépend étroitement du dictionnaire choisi, de la méthode de détermination-modélisation des filtres de prédiction linéaire utilisés, ces deux paramètres constituant deux degrés de liberté, non indépendants, permettant d'adapter un codeur CELP particulier aux besoins d'une application déterminée.

15 Une telle technique de codage CELP convient aux applications de codage à bas débit (compris entre 4 à 24 kbits/s). Pour une description plus détaillée de ce type de codage, on pourra utilement se reporter à l'article intitulé "A robust and fast CELP coder at 16 Kbits/s" publié par A. LE GUYADER, D. MASSALOUX et F. ZURCHER, CNET Lannion France, dans la revue Speech Communication n°7, 1988.

20 De manière générale, dans ce type de codeur, décodeur, le signal numérique à analyser, transmettre et reconstituer, est découpé en blocs, ou trames. Chaque bloc comprenant L valeurs est considéré comme un vecteur d'un espace vectoriel de dimension L. Le signal d'excitation courant constitué par un vecteur v, lu dans le dictionnaire des formes d'onde, doit minimaliser un critère de distorsion perceptuel de la forme :

25 $\min \|\chi - H \cdot v\|^2$, dans lequel χ désigne un signal cible issu du signal original O à transmettre après pondération perceptuelle et H désigne une matrice de dimension LxL de réponse impulsionnelle issue du produit des fonctions de transfert du filtre de synthèse et de la pondération perceptuelle. On rappellera que la pondération perceptuelle a pour objet, relativement au bruit de codage, analogue à un bruit blanc, de relativiser dans le domaine fréquentiel la contribution de ce dernier au signal perçu effectivement. La matrice H est une matrice triangulaire de la forme :

30

$$H(L) = \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & 0 \\ h_1 & h_0 & 0 & 0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 \\ h_3 & h_2 & h_1 & h_0 & 0 & 0 \\ h_4 & h_3 & h_2 & h_1 & h_0 & 0 \\ - & - & - & - & - & - \\ h_{L-1} & h_{L-2} & h_3 & h_2 & h_1 & h_0 \end{bmatrix}$$

35 Au cours du processus de codage, à chaque vecteur de référence v_i est associée une valeur de gain adaptatif g_k prise parmi un dictionnaire de valeurs de gain G, ce qui permet, suite à application du gain g_k au vecteur v_i pour former un vecteur $v_{k,i}$, de satisfaire au critère de distorsion minimale précité.

40 Afin de diminuer la complexité des calculs, très nombreux en fonction de la dimension L des vecteurs et du débit du signal de parole, on a, dans certains travaux, proposé d'utiliser comme vecteur de référence, afin de réaliser le signal d'excitation, des vecteurs dont les composantes n'ont pour valeur que les valeurs +1, 0 ou -1, le dictionnaire des vecteurs étant alors constitué sous forme de dictionnaire de vecteurs ternaires. Une telle utilisation dans un processus de codage de type CELP de vecteurs ternaires de ce type a été mentionnée dans la demande de brevet européen EP 0 347 307 publiée le 20.12.1989.

45 Toutefois, dans un tel processus de codage, on notera que tous les vecteurs de référence contiennent nécessairement la même énergie. En outre, la recherche de la séquence ou vecteur de référence optimum ne peut être réduite au calcul de produits purement scalaires que dans le cas où l'auto-corrélation est elle-même normalisée et présente des termes nuls dont l'écartement correspond aux composantes non nulles des séquences ou vecteurs de référence.

Un tel mode opératoire ne permet donc pas de prendre en compte, comme vecteur de référence, la totalité des possibilités des combinaisons des valeurs ternaires des composantes des vecteurs de référence, la minimalisation du critère de distorsion ne pouvant dans tous les cas être optimale.

Enfin, le document EP-A-0 379 296 décrit un procédé de transmission d'un signal de parole à bas débit, de type CELP. Ce procédé met en oeuvre un processus de calcul du minimum d'écart quadratique d'un vecteur initial et d'un vecteur de référence à partir d'un dictionnaire de vecteurs factorisé en produit de deux dictionnaires de vecteurs de base et de gains. Le gain est séparé en amplitude et en signe et le processus de calcul consiste à trouver trois indices satisfaisant le critère d'erreur quadratique minimum.

La présente invention a pour objet de remédier aux inconvénients précités, afin, notamment, d'apporter une simplification des calculs par l'introduction comme vecteur de référence, dans le dictionnaire des vecteurs de référence, ou directions, de la quasi-totalité des combinaisons des valeurs n-aires des composantes des vecteurs, n étant un nombre impair.

Un autre objet de la présente invention est la mise en oeuvre, préalablement au processus classique d'application d'un gain adaptatif à chacun des vecteurs de référence, d'un processus de correction par application d'un facteur d'échelle, introduisant la répartition de l'énergie du signal d'excitation en fonction du spectre fréquentiel de ce dernier, afin de tenir compte de la non-uniformité de la distribution d'énergie du signal dans le domaine des fréquences.

Un autre objet de la présente invention est enfin la mise en oeuvre d'un procédé de transmission à bas débit d'un signal de parole dans lequel, chaque vecteur de référence, constituant le signal d'excitation, peut être régénéré au niveau d'un décodeur à partir des seules valeurs d'indice ou d'adresse du vecteur de référence optimal satisfaisant au critère de minimum de distorsion au niveau du codeur, ce qui a pour effet de simplifier considérablement et de réduire les coûts de fabrication des décodeurs précités.

Le procédé de transmission d'un signal de parole à bas débit selon la présente invention comporte un processus de codage d'échantillons numériques de parole par prédiction linéaire excitée par codes pour engendrer un signal de code, un processus de transmission du signal de code et un processus de décodage du signal de code reçu. Le processus de codage correspond à un processus dans lequel une forme d'onde est représentée par un bloc d'échantillons comportant L valeurs d'échantillons et constituant un vecteur initial (o) de dimension L est représenté, à partir d'un filtre de synthèse, par une forme d'onde de référence sélectionnée parmi un dictionnaire de formes d'onde de référence formant chacune un vecteur de référence (v) sur critère d'écart quadratique minimum du vecteur initial (o) vis-à-vis de la forme d'onde ou vecteur de référence (v), $\min \|\chi - H.v\|^2$, où χ représente un vecteur cible obtenu par pondération perceptuelle du vecteur initial (o) et H une matrice de dimension LxL de réponse impulsionnelle issue du produit du filtre de synthèse et de la pondération perceptuelle linéaire. Ce processus est remarquable en ce que ce critère de sélection consiste à établir un dictionnaire factorisé en produit d'un premier dictionnaire Y_n , $n=2m+1$ entier impair, m étant un nombre entier non nul, de vecteurs de base y_i , de forme n-aire, de dimension L, dont chaque valeur des composantes, $y_{i(j)}$, peut prendre une valeur appartenant à l'ensemble I_m des entiers relatifs consécutifs compris entre -m et m, soit $I_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ constituant un ensemble à $n=2m+1$ éléments, chacun de ces vecteurs de base étant multiplié par un facteur d'échelle v_i tenant compte de la distribution de l'énergie d'excitation dans le domaine de fréquence du signal, et d'un deuxième dictionnaire $G(y)$ de gains g_k , de façon à représenter ainsi le dictionnaire de formes d'ondes ou vecteurs de référence, chaque vecteur de référence vérifiant la relation $v_{k,i} = g_k \cdot v_i \cdot y_i$. On notera que la valeur $n/2$ correspond à la division entière de n par 2. La valeur minimum de l'écart quadratique $\|\chi - g_k \cdot H.v_i \cdot y_i\|^2$ est alors établie par le calcul du maximum de $C(g_k, v_i, y_i) = 2 g_k \langle \chi | H.v_i \cdot y_i \rangle - g_k^2 \|H.v_i \cdot y_i\|^2$ par le calcul de tous les produits scalaires $\langle \chi | H.v_i \cdot y_i \rangle$ et de toutes les énergies perceptuelles $\|H.v_i \cdot y_i\|^2$, ce qui permet d'attribuer au vecteur initial (o) le vecteur de référence optimal correspondant v_{k^*, i^*} , avec $v_{k^*, i^*} = g_{k^*} \cdot v_{i^*} \cdot y_{i^*}$, ce vecteur de référence optimal étant représenté par les seules valeurs d'indice k^*, i^* satisfaisant au critère $\min \|\chi - g_k \cdot H.v_i \cdot y_i\|^2$.

Le processus de transmission à bas débit d'un signal de parole, selon la présente invention, consiste à transmettre, comme signal de code, les seules valeurs des indices k^*, i^* représentatifs de chaque vecteur de référence optimal v_{k^*, i^*} .

Le processus de décodage d'un signal de parole codé transmis à bas débit selon un signal de code, conformément à l'objet de la présente invention, est remarquable en ce que, afin d'assurer le décodage du signal de code, ce processus consiste à discriminer les valeurs des indices k^*, i^* constituant le signal de code, à décomposer la valeur de l'indice i^* , représentatif du vecteur de référence optimal, en base n pour régénérer le vecteur de base correspondant y_{i^*} , à effectuer, à partir de la valeur de l'indice i^* , du facteur d'échelle v_{i^*} et du gain adaptatif g_{k^*} correspondants, une correction du vecteur de base correspondant régénéré pour constituer le vecteur de référence régénéré v_{k^*, i^*} . Une opération de filtrage de synthèse est effectuée sur le vecteur de référence régénéré v_{k^*, i^*} pour engendrer le signal de parole reconstitué.

Le procédé objet de la présente invention, les processus de codage, de transmission et de décodage et le système et circuits de codage, de transmission et de décodage permettant la mise en oeuvre de ce procédé, trouvent avantageusement application à la transmission de signaux de parole à bas débit, notamment entre mobiles par exemple.

L'invention sera mieux comprise à la lecture de la description ci-après et à l'observation des dessins dans lesquels, outre les figures la et 1b relatives à l'art antérieur,

- la figure 2 représente au point a) d'une part les étapes de traitement dans un processus de codage conforme à l'objet de la présente invention, et au point b) d'autre part les opérations effectuées, sur les vecteurs de base, dans les étapes représentées au point a), pour des vecteurs n-aires,
- la figure 3a représente aux points 1, 2 et 3 les modules de traitement de vecteurs impulsion constituant des vecteurs de base privilégiés, dans un traitement de type récursif permettant d'engendrer un premier dictionnaire de vecteurs de base,
- la figure 3b représente successivement les opérations effectuées sur les vecteurs de base pour engendrer, de manière itérative, le premier dictionnaire de vecteurs de base précité, dans un cas particulier où $n=3$, les vecteurs de base étant des vecteurs ternaires,
- la figure 4 représente de manière analogue à la figure 3a, 3b un processus de calcul de la réponse impulsionnelle pour tous les vecteurs ternaires y_i excitant le filtre de synthèse et le filtre de pondération perceptuelle en cascade présentant la fonction de transfert H ,
- la figure 5 représente en ses différents points a), b), c) et d) des graphes représentatifs des processus de calcul des énergies perceptuelles des vecteurs ternaires à partir des réponses partielles impulsionnelles de la fonction de transfert H ,
- la figure 6 représente des graphes représentatifs des processus de calcul des produits scalaires,
- la figure 7 représente un organigramme des étapes de traitement des valeurs d'indice optimal k^*, i^* reçus au cours du processus de décodage,
- la figure 8 représente un schéma synoptique d'un circuit de codage dans un système de transmission de parole à bas débit conforme à l'objet de la présente invention,
- la figure 9 représente un schéma synoptique d'un circuit de décodage dans un système de transmission de parole à bas débit conforme à l'objet de la présente invention.

Le procédé de transmission d'un signal de parole à bas débit, objet de la présente invention, sera tout d'abord décrit en liaison avec les figures 2 a et b.

Selon la figure 2 précitée, le procédé objet de l'invention comporte un processus de codage d'échantillons numériques de parole par prédiction linéaire excitée par codes. Ce processus permet d'engendrer un signal de code. Le procédé comporte en outre un processus de transmission du signal de code et un processus de décodage du signal de code reçu.

Selon la figure 2 précitée, le processus de codage correspond à un processus dans lequel une forme d'onde représentée par un bloc d'échantillons comportant L valeurs d'échantillons, ou trames, constitue un vecteur initial noté o de dimension L , ce vecteur étant représenté, et la forme d'onde correspondante, à partir d'un filtre de synthèse par une forme d'onde de référence, notée v , sélectionnée parmi un dictionnaire de formes d'onde de référence formant chacune un vecteur de référence précité. La sélection est effectuée sur critère d'écart quadratique minimum du vecteur initial o vis-à-vis de la forme d'onde ou vecteur de référence v , ce critère s'écrivant :

$$\min \|\chi - H \cdot v\|^2.$$

Dans cette relation, χ représente un vecteur cible obtenu par pondération perceptuelle du vecteur initial o et H représente une matrice de dimension $L \times L$ de réponse impulsionnelle issue du produit du filtre de synthèse et de la pondération perceptuelle linéaire précitée.

Selon le procédé objet de la présente invention, le processus de codage est tel que le critère de sélection consiste à établir un dictionnaire factorisé en produit d'un premier dictionnaire Y_n de vecteurs de base notés y_i . Chaque vecteur de base est un vecteur de base de forme n -aire, c'est-à-dire que les composantes $y_{i(j)}$ de ces vecteurs de base, avec $j \in [0, L-1]$, peuvent prendre n valeurs discrètes différentes. D'une manière générale, chaque valeur des composantes $y_{i(j)}$ peut prendre une valeur comprise dans l'ensemble I_m des entiers relatifs consécutifs compris entre $-m$ et m , soit $I_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ constituant un ensemble à $n=2m+1$ éléments.

Selon une caractéristique avantageuse du procédé objet de la présente invention, chaque vecteur de base y_i est corrigé par un facteur d'échelle v_i tenant compte de la distribution de l'énergie d'excitation dans le domaine de fréquence du signal. On notera que, de la manière la plus générale, les facteurs d'échelle v_i sont déterminés à partir d'une base de données, de manière expérimentale, la base de données étant constituée par un enregistrement d'échantillons de parole significatifs sur plusieurs heures par exemple et pour plusieurs locuteurs d'une même langue d'expression ou de plusieurs langues distinctes, l'expérience montrant que la diversité des langues d'expression n'intervient qu'au second degré dans la détermination des facteurs d'échelle v_i précités. Une description plus détaillée d'une table des facteurs d'échelles v_i pour des vecteurs ternaires de dimension $L=5$ sera donnée ultérieurement dans la description.

On notera simplement que selon ce principe les facteurs d'échelles v_i sont déterminés pour chaque vecteur de base y_i correspondant par un processus d'identification de chaque vecteur de base y_i sur une séquence délocalisée de L échantillons de parole successifs récursifs de la base de données, tri des coefficients d'adaptation les plus faibles et moyennisation d'un nombre u de coefficients d'identification ou d'adaptation pour obtenir le facteur d'échelle v_i correspondant associé au vecteur de base y_i précité.

Le dictionnaire factorisé précédemment mentionné est également constitué par un deuxième dictionnaire constituant le produit précité, ce deuxième dictionnaire étant noté $G(y)$ et étant formé par un dictionnaire de gains g_k . Le dictionnaire factorisé constitue ainsi un dictionnaire de formes d'onde ou de vecteurs de référence. Chaque vecteur de référence vérifie ainsi la relation $v_{k,i} = g_k \cdot v_i \cdot y_i$.

5 On notera bien entendu, ainsi que représenté sur la figure 2a, que l'opération de correction effectuée par l'application du facteur d'échelle v_i ne constitue pas une simple pondération des composantes $y_{i(j)}$ de chaque vecteur de base y_i car chaque coefficient de facteur d'échelle v_i est représentatif de la distribution de l'énergie d'excitation dans le domaine de fréquence d'un signal de parole.

10 Ainsi qu'on l'a représenté au point a) de la figure 2, le procédé objet de l'invention consiste alors à établir la valeur minimum de l'écart quadratique $\|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$ par le calcul d'une fonction notée :

$$C(g_k, v_i, y_i) = 2 g_k \langle \chi | H \cdot v_i \cdot y_i \rangle - g_k^2 \|H \cdot v_i \cdot y_i\|^2$$

15 par calcul de tous les produits scalaires $\langle \chi | H \cdot v_i \cdot y_i \rangle$ et de toutes les énergies perceptuelles $\|H \cdot y_i\|^2$.

Le calcul précité permet alors d'attribuer au vecteur initial ou le vecteur de référence optimal correspondant noté v_{k^*, i^*} , avec $V_{k^*, i^*} = g_{k^*} \cdot v_{i^*} \cdot y_{i^*}$. Bien entendu, conformément à un objet particulièrement intéressant de la présente invention, ce vecteur de référence optimal est représenté par les seules valeurs des paramètres d'indice k^*, i^* satisfaisant au critère précité :

20 $\min \| \chi - g_k \cdot H \cdot v_i \cdot y_i \|^2$.

Une description plus détaillée des opérations effectuées au niveau de chaque vecteur de base y_i , ces vecteurs de base étant des vecteurs n -aires de dimension L dont les composantes $y_{i(j)}$ ont pour valeur au plus la valeur $n/2$ ou éventuellement $-n/2$, par valeurs entières et par incrément de 1, sera donnée en liaison avec le point b) de la figure 2.

25 Au point b) précité, on a représenté successivement les vecteurs de base notés y_0, y_1, y_i, y_K avec $K = (n-3)/2$, chaque composante ayant pour valeur l'une des valeurs de la forme n -aire. On a ensuite représenté la correction par application du facteur d'échelle v_i , lequel, pour les raisons précédemment mentionnées, ne constitue pas une simple pondération analogue à l'application adaptative du gain g_k , à chaque valeur des composantes $y_{i(j)}$ des vecteurs de base y_i étant appliqué le facteur d'échelle v_i correspondant déterminé dans les conditions précédemment mentionnées. Sur le même point b) on a enfin représenté l'application du gain adaptatif g_k , chaque composante $y_{i(j)}$ des vecteurs de base y_i étant alors multipliée par le produit $g_k \cdot v_i$.

30 On comprendra bien sûr que, dans la mise en oeuvre du processus de codage tel que représenté aux points a) et b) de la figure 2 précédemment mentionnée, la valeur minimum de l'écart quadratique $\min \| \chi - g_k \cdot H \cdot v_i \cdot y_i \|^2$ est évaluée par sélection de l'élément de gain g_k correspondant du deuxième dictionnaire $G(y)$ permettant de rendre minimum la différence $|g - g_k|$ où g vérifie la relation :

$$g = \frac{\langle \chi | H \cdot v_i \cdot y_i \rangle}{\|H \cdot v_i \cdot y_i\|^2}$$

40 Une description plus détaillée de l'arrangement des vecteurs de base y_i pour constituer le dictionnaire ou premier dictionnaire Y_n de dimension L de vecteurs de base y_i sera maintenant donnée en liaison avec les figures 3a et 3b.

De manière générale, on comprendra que le dictionnaire Y_n de vecteurs de base y_i de forme n -aire $[-n/2, \dots, 0, \dots, n/2]$ de dimension L comprend tous les vecteurs de base dont les L composantes ont pour valeur les valeurs n -aires précitées, à l'exception du vecteur nul. De manière générale, l'indice i des vecteurs de base est pris égal à la valeur en base n de chaque vecteur de base après transcodage des valeurs $[-n/2, \dots, 0, \dots, n/2]$ en valeurs correspondantes (0, 1, 2, ..., n). On comprendra ainsi que les vecteurs de base y_i de forme n -aire sont arrangés en fonction de leur indice i , cet indice i ayant pour valeur la valeur en base n de chaque vecteur.

45 On comprendra également que l'ensemble des vecteurs de base y_i constituant le dictionnaire Y_n est défini à partir des $n/2 \cdot L$ vecteurs impulsion dont une seule composante $y_{i(j)}$ d'ordre j , avec $j \in [0, L-1]$, est égale à $-1, -2, \dots, -n/2$. A chaque vecteur impulsion sont associés les vecteurs de base apparentés présentant des valeurs de composantes d'ordre $q \leq j$ identique, chaque vecteur apparenté à un vecteur impulsion de rang q , avec $q = j$ pour $y_{i(j)}$ différent de 0, étant obtenu par combinaison linéaire du vecteur impulsion de rang $j=q$ et des vecteurs impulsion ou apparentés de rang $j=q'$ supérieur.

50 Une description plus détaillée de la mise en oeuvre du dictionnaire des vecteurs de base y_i dans le cas de vecteurs ternaires et de la manière d'engendrer ces vecteurs de base sera donnée en liaison avec les figures 3a et 3b, des vecteurs de base de dimension L et de forme n -aire pouvant être engendrés selon le même principe sans sortir du cadre de l'objet de la présente invention.

55 Sur les figures 3a et 3b on a représenté respectivement des cellules opérateurs permettant d'engendrer, à partir

des vecteurs impulsion précédemment définis et de sous-dictionnaires constitués par le vecteur impulsion considéré et les vecteurs apparentés correspondant à chaque vecteur impulsion, le dictionnaire complet comprenant la réunion de l'ensemble de tous les sous-dictionnaires.

5 Chaque opérateur tel que représenté en figure 3a comprend un opérateur dit opérateur de retard R dont la fonction de transfert est notée Z^{-1} , selon la notation classique de transformée en Z, un opérateur symétriseur noté Sy lequel a pour fonction de multiplier les composantes de tous les vecteurs présentés à son entrée par la valeur +1, par la valeur 0 puis par la valeur -1 et un sommateur, noté S, recevant la sortie de l'opérateur de retard R et du symétriseur Sy. Le sommateur S reçoit la sortie de l'opérateur de retard R par l'intermédiaire d'un commutateur I, en position F, ou le vecteur nul $[0,0,0,0]$ de dimension L en position O. Les opérateurs représentés en figure 3a sont constitués par un seul et même opérateur représenté en 1), 2) et 3) à des étapes différentes de processus de traitement pour engendrer les vecteurs de base y_i du dictionnaire Y_n précité.

En début de processus de génération des vecteurs de base y_i , tel que représenté au point 1) de la figure 3a, l'impulsion initiale ou vecteur d'impulsion δ_{L-1} est présente à l'entrée de l'opérateur de retard R. Le symétriseur Sy est alors alimenté par un sous-dictionnaire noté $\overline{D_0}$ lequel est constitué initialement par le vecteur impulsion δ_{L-1} précité. Le symétriseur Sy délivre un sous-dictionnaire symétrique noté $\overline{D_0}$, tel que représenté en figure 3b, et le sommateur S, lequel reçoit le vecteur impulsion δ_{L-2} délivré par l'opérateur de retard R, vecteur impulsion de rang $q = L-2$, ou le vecteur nul, et le sous-dictionnaire symétrique $\overline{D_0}$, délivre en sortie le dictionnaire D1 constitué par les vecteurs de base y_0, y_1, y_2 et y_3 . On notera bien sûr, ainsi que représenté en figure 3b, qu'au vecteur impulsion δ_{L-2} est associé le sous-dictionnaire D1 formé par les vecteurs apparentés y_1, y_2, y_3 au vecteur impulsion δ_{L-2} et par le vecteur impulsion à L-1 initial formant le vecteur de base y_0 , ainsi que le vecteur nul. Bien entendu, et de manière réursive ainsi que représenté au point 2) de la figure 3a, l'opérateur permettant d'engendrer les vecteurs de base y_i est tel que celui-ci reçoit au niveau de l'opérateur de retard R le vecteur impulsion à L-m, au niveau du symétriseur Sy, le dictionnaire noté D m-1 formé récurivement comme le dictionnaire D1, le sommateur S tel que représenté au point 2 de la figure 3a délivrant alors à partir du vecteur impulsion δ_{L-m-1} précité délivré par l'opérateur de retard R ou du vecteur nul et par le sous-dictionnaire $\overline{D_{m-1}}$ le sous dictionnaire Dm.

Par itération il est ainsi possible et de manière réursive d'engendrer à partir de l'ensemble des vecteurs impulsion, tel que précédemment décrit, les vecteurs apparentés et les sous-dictionnaires correspondants puis finalement le dictionnaire complet. On notera que, sur la figure 3b, les * représentés au niveau des composantes $y_{i(j)}$ pour le processus de traitement de niveau m correspondent à des valeurs 0, -1 ou +1 lorsque les vecteurs sont des vecteurs ternaires. Bien entendu, dans le cas de vecteurs n-aires les * représentent des valeurs comprises entre $-n/2$ et $+n/2$, dans les conditions précédemment mentionnées.

On notera que le dictionnaire ternaire total, somme ou réunion de tous les sous-dictionnaires de niveau intermédiaire m, jusqu'à L peut être obtenu pour les seules valeurs positives ou négatives des composantes $y_{i(j)}$, le dictionnaire total pouvant alors être obtenu par symétrisation par l'intermédiaire d'un opérateur de symétrisation tel que Sy.

35 De la même manière, le calcul de la réponse partielle à un instant $t = L-1$, c'est-à-dire à un instant relatif correspondant à l'occurrence du vecteur impulsion δ_{L-1} , du système H constitué par le filtre de synthèse et par le filtre de pondération perceptuelle excité par les vecteurs de base ternaire y_i peut être décrit à l'aide des opérateurs précédents. La réponse partielle à l'instant $t = L-1$ est notée $SL-1(y_i)$.

40 Au niveau du premier opérateur de calcul, noté 1 sur la figure 4, cet opérateur est tel que les réponses impulsionnelles du système H au temps relatif 0, 1, 2, L-1 c'est-à-dire les valeurs $h_0, h_1, h_{L-2}, h_{L-1}$ sont appliquées à l'opérateur précité.

On rappelle que l'opérateur SL-1 représente ici aussi l'addition à chaque élément h_{L-m-1} ou à la valeur zéro de toutes les réponses partielles à $t = L-1$ des vecteurs du dictionnaire symétrisé délivré par le symétriseur Sy de niveau m.

On obtient ainsi $S_{L-1(D_m)}$ ensemble des réponses $t = L-1$ des vecteurs de Dm.

45 L'opérateur de symétrisation Sy multiplie les éléments de $S_{L-1(D_{m-1})}$ par +1, 0, -1 et réalise, ainsi que décrit précédemment, la réunion des éléments distincts obtenus. Enfin, le dernier opérateur représenté en 3 à la figure 4 fournit la réponse à $t = L-1$ des vecteurs ternaires y_i dont la première coordonnée est -1.

50 On notera que la réponse du système linéaire de la matrice H aux vecteurs ternaires qui lui sont appliqués peut donc être réalisée selon la même architecture que précédemment en appliquant la transformation linéaire H à chaque noeud de cette architecture.

Les énergies perceptuelles des vecteurs ternaires peuvent alors se déduire des seules réponses partielles précédemment décrites à $t = L-1$.

En effet, : la réponse de la matrice H à l'excitation par un vecteur y_i s'écrit :

$$\begin{array}{r}
 5 \\
 10 \\
 15 \\
 20 \\
 25 \\
 30
 \end{array}
 \begin{array}{c}
 \text{H} \qquad \qquad \text{y}_i \qquad \qquad \text{H} \cdot \text{y}_i \\
 \\
 \text{Hy}_i = \begin{bmatrix} h_0 & 0 \\ h_1 & h_0 \\ h_2 & h_1 \\ h_3 & h_2 \\ h_4 & h_3 \\ h_{L-1} & \dots h_1 h_0 \end{bmatrix} \begin{bmatrix} y_i(0) \\ y_i(1) \\ - \\ - \\ - \\ y_i(L-1) \end{bmatrix} = \begin{bmatrix} h_0 y_i(0) \\ h_1 y_i(1) \\ - \\ - \\ - \\ h_{L-2} y_i(0) + \dots + h_0 y_i(L-2) \\ h_{L-1} y_i(0) + \dots + h_0 y_i(L-1) \end{bmatrix} \\
 \\
 = \begin{bmatrix} 0 \\ - \\ - \\ - \\ \text{SL-1}(y_i) \end{bmatrix} + \begin{bmatrix} h_0 y_i(0) \\ h_1 y_i(0) + h_0 y_i(1) \\ - \\ - \\ - \\ h_{L-2} y_i(0) + \dots + h_0 y_i(L-2) \\ 0 \end{bmatrix}
 \end{array}$$

En effet, par définition la réponse à l'instant relatif $t = L-1$, notée $\text{SL-1}(y_i)$ est la coordonnée d'ordre $L-1$ de Hy_i .
Cependant, on peut écrire :

$$\begin{array}{r}
 35 \\
 40 \\
 45
 \end{array}
 \begin{array}{c}
 \text{y}'_i = \begin{bmatrix} h_0 y_i(0) \\ h_1 y_i(0) + h_0 y_i(1) \\ - \\ - \\ - \\ h_{L-2} y_i(0) + h_0 y_i(L-2) \\ 0 \end{bmatrix}
 \end{array}$$

et

50

55

$$\begin{matrix} 5 \\ 10 \\ 15 \end{matrix}
 \quad
 \mathbf{y}^i = \begin{bmatrix} 0 \\ h_0 y_i(0) \\ h_1 y_i(0) + h_0 y_i(1) \\ - \\ - \\ - \\ h_{L-2} y_i(0) + \dots + h_0 y_i(L-2) \end{bmatrix} = H \begin{bmatrix} 0 \\ y_i(0) \\ y_i(1) \\ - \\ - \\ - \\ y_i(L-2) \end{bmatrix}$$

On notera que y^i et y^i ont même norme et l'on peut démontrer, en notant z^{-1} l'opérateur de retard élémentaire, la relation ci-après :

$$\begin{matrix} 20 \end{matrix}
 \quad
 \|y^i\|^2 = \|y^i\|^2 = \|H.z^{-1}y_i\|^2$$

$$\|H.y_i\|^2 = S_{L-1}(y_i)^2 + \|H.z^{-1}y_i\|^2$$

25 Cependant, si y_i appartient à D_m , $z^{-1}.y_i$ appartient à D_{m-1} .

Un processus itératif permet donc de calculer les énergies perceptuelles pour D_0 puis D_1 puis D_{L-1} . La valeur initiale est pour $D_0 = \delta L-1$, c'est-à-dire le vecteur impulsion précédemment représenté en figure 3, h_0^2 .

Un schéma de principe du processus de numérotation et de calcul des différentes entités mises en oeuvre par le critère de sélection conformément à l'objet de la présente invention sera décrit en liaison avec les figures 5a et 5b.

30 De manière générale, ainsi que représenté en figure 5a, les vecteurs de base y_i ainsi que déjà décrit précédemment peuvent être engendrés selon le graphe de génération globale à raison de $3^0 = 1$ vecteur est engendré au niveau 0, le vecteur y_0 , 3^1 sont engendrés au niveau 1, les vecteurs y_1 , y_2 , y_3 et ainsi de suite, 3^{L-1} vecteurs de base au niveau $L-1$.

35 La cellule élémentaire de détriplement est représentée en figure 5b à partir des vecteurs impulsion notés $\theta-1$, θ_0 et θ_1 . On notera que la sommation des vecteurs d'impulsion θ_1 , θ_0 et $\theta-1$ revient à remplacer la dernière coordonnée du vecteur de base incident par les valeurs de composantes +1, 0 ou -1.

On notera que l'architecture telle que représentée en figure 5a et 5b est celle d'une structure linéaire de graphes ternaires. Pour une structure n-aire on obtient un graphe n-aire.

40 Il est également possible d'obtenir une réalisation pratique de calcul de l'expression $\|H.y_i\|^2 = S_{L-1}(y_i)^2 + \|H.z^{-1}y_i\|^2$ grâce à l'architecture analogue ci-après. Cette architecture sera décrite en liaison avec les figures 5c et 5d.

On appelle $E(i)$ l'expression $E(i) = \|H.y_i\|^2$.

Ainsi qu'on l'a représenté en figure 5c le graphe global pour l'obtention des énergies est parcouru de droite à gauche, l'énergie initiale $E(0)$ étant égale à $S_{L-1}(0)^2$.

La cellule élémentaire constitutive du graphe représenté en figure 5c est représentée en figure 5d.

45 On notera que la numérotation des vecteurs, c'est-à-dire l'affectation de leur indice i de vecteurs de base, peut correspondre soit à une numérotation rétrograde, tout indice p de la numérotation directe d'un vecteur ternaire vérifiant la relation correspondante en numérotation p' rétrograde $p' = 3^{L-p}-1$. On comprendra bien entendu que l'ensemble des calculs peut être effectué soit avec la numérotation directe soit avec la numérotation rétrograde cette dernière étant préférée. Il est ensuite possible de transmettre les valeurs d'indice rétrograde par exemple ou les valeurs d'indice direct sur la ligne de transmission ainsi qu'il sera décrit ultérieurement dans la description.

50 On notera en outre, conformément aux pratiques antérieures en matière de codage de type CELP, que préalablement au filtrage de synthèse chaque vecteur de référence v_{k^*,j^*} peut avantageusement être pondéré par un facteur de niveau prédit, noté a . Ce facteur de niveau prédit a est représentatif de l'énergie moyenne du signal d'excitation estimé sur au moins trois vecteurs excitation antérieurs successifs. Une telle opération sur les composantes $y_{i(j)}$ de chaque vecteur de référence ne sera pas décrite car elle correspond à une opération connue de l'homme de métier.

55 Une description plus détaillée d'un processus de calcul des produits scalaires de la forme $\langle 2\chi | H.y_i \rangle$ où $\chi = x/\sigma$ pour tous les vecteurs de base y_i sera maintenant décrite en liaison avec la figure 6.

On notera qu'en fait, compte tenu du facteur de niveau prédit σ introduit effectivement dans le processus de codage

EP 0 481 895 B1

objet de la présente invention, il s'agit en fait de calculer l'expression $\langle 2x/\sigma | H.y_i \rangle$ pour tous les vecteurs ternaires y_i .
L'expression précédente est alors calculée en filtrant l'expression $2x/\sigma$ par la matrice transposée de la matrice H , soit H^t .

Cette expression s'écrit :

5

$$\langle 2x/\sigma | H.y_i \rangle = \langle H^t.(2x/\sigma) | y_i \rangle$$

En posant $x' =$

10

$$x' = H^t(2x/\sigma) = \begin{bmatrix} 2/\sigma \cdot (h_0x(0) + h_1x(1) + \dots + h_{L-1} x(L-1)) \\ 2/\sigma \cdot (h_0x(1) + h_2x(2) + \dots + h_{L-2} x(L-1)) \\ \cdot \\ \cdot \\ \cdot \\ 2/\sigma \cdot h_0x(L-1) \end{bmatrix}$$

15

20

On peut obtenir l'expression $\langle x' | y_i \rangle$ pour les vecteurs y_i de base ternaire de la façon ci-après : on calcule l'expression :

25

$$- \langle H^t(2x/\sigma) | y_i \rangle = F(y_i).$$

30

Le processus de calcul tel que représenté grâce à l'opérateur en figure 6 permet, de manière analogue au calcul des réponses partielles $S_{L-1}(y_i)$ précédemment décrit, d'obtenir les quantités x'_0 , x'_{L-m-1} , x'_{L-2} et donc les produits scalaires précités, le vecteur nul étant remplacé par la valeur nulle.

En ce qui concerne la détermination et l'attribution du facteur d'échelle v_i à chacun des vecteurs de base y_i on rappelle que chaque facteur d'échelle v_i peut être déterminé à partir d'une pluralité N de trames, à partir d'une base de données de signal de parole, le facteur d'échelle v_i pour chaque vecteur de base y_i étant choisi de façon à rendre minimum pour la trame considérée le résidu de filtrage des trames précitées. On rappellera que plusieurs processus de détermination de chaque facteur d'échelle v_i peuvent être envisagés.

35

A titre d'exemple non limitatif, dans le cas de vecteurs de base de type ternaire et de dimension $L=5$ le tableau des facteurs d'échelle v_i est donné ci-dessous par la table des 121 valeurs des facteurs d'échelle. La première valeur multiplie $(-1, -1, -1, -1, -1) \dots$, la dernière $(0, 0, 0, 0, -1)$.

40

1.50, 1.66, 1.77, 1.28, 1.46, 1.36, 0.86, 2.47, 1.68, 1.51, 1.12, 1.04, 1.38, 1.86, 1.51, 4.23, 3.47, 1.96, 1.25, 2.28, 0.77, 2.50, 3.51, 0.87, 1.11, 1.16, 0.95, 1.29, 1.23, 1.85, 1.34, 1.55, 1.60, 1.51, 1.44, 1.21, 1.45, 1.95, 1.45, 1.73, 4.06, 1.73, 1.32, 1.39, 2.43, 1.38, 4.62, 1.35, 1.92, 2.15, 1.44, 2.20, 1.95, 1.07, 0.88, 1.56, 1.48, 1.33, 1.64, 1.70, 1.44, 3.33, 1.10, 1.89, 0.80, 2.07, 1.27, 1.57, 3.82, 1.28, 1.31, 1.34, 1.94, 1.86, 1.25, 1.06, 2.15, 1.39, 0.89, 1.24, 1.32, 1.17, 1.45, 0.57, 1.28, 2.00, 4.88, 2.14, 2.98, 2.24, 1.23, 1.66, 1.41, 1.82, 3.44, 1.14, 3.15, 3.91, 1.60, 0.95, 1.74, 1.50, 1.12, 2.98, 1.16, 1.23, 1.34, 1.00, 2.06, 2.52, 4.52, 1.93, 2.89, 3.21, 1.39, 2.44, 2.38, 4.55, 3.00, 2.49, 3.17

45

Les valeurs optimales des indices k^* et i^* ayant été déterminées et numérotées de façon directe ou de façon rétrograde ainsi que décrit précédemment dans la description, en ce qui concerne notamment la valeur des indices i , la transmission de parole à bas débit est effectuée par la seule transmission, comme signal de code, des valeurs des indices k^* et i^* représentatifs de chaque vecteur de référence v_{k^*, i^*} .

50

En ce qui concerne la transmission des indices k^* et i^* précités on notera que la transmission peut être effectuée à l'aide de protocoles de transmission classique dans lesquels une redondance de l'information transmise est introduite afin d'assurer une transmission à taux d'erreur sensiblement nul. On comprendra bien sur que la valeur i^* peut être transmise soit en numérotation directe soit en numérotation rétrograde, soit selon une numérotation traduite dont la table de traduction est connue du codeur comme du décodeur.

55

Une description plus détaillée du processus de décodage des informations transmises, c'est-à-dire du signal de code ainsi transmis conformément au procédé objet de l'invention, sera maintenant donnée en liaison avec la figure 7.

Conformément à la figure 7 précitée, le processus de décodage consiste à discriminer en 1000 les valeurs des

indices k^* et i^* constituant le signal de code puis à décomposer en 1001 la valeur de l'indice i^* représentatif du vecteur de référence optimal en base n afin de régénérer le vecteur de base correspondant y_{i^*} .

La régénération du vecteur de base y_{i^*} est effectuée en 1002 à partir de la valeur de l'indice i^* et du facteur d'échelle v_{i^*} correspondant, une correction du vecteur de base correspondant régénéré étant effectuée pour constituer le vecteur de référence $v_{k^*,i^*} = v_{i^*} \cdot y_{i^*}$.

Suite à l'opération précitée, le processus de décodage consiste à effectuer une opération de filtrage de synthèse 1003 du vecteur de référence pour engendrer le signal de parole reconstruit.

On notera bien sûr que, de même que dans le cas du processus de codage, dans le processus de codage du procédé objet de la présente invention, chaque vecteur de référence v_{k^*,i^*} préalablement au filtrage de synthèse est pondéré par un facteur de niveau prédit σ lequel est estimé sur au moins trois vecteurs excitation antérieurs successifs. La détermination du niveau prédit σ ne sera pas décrite en détail car elle correspond au niveau du processus de décodage des opérations normalement connues de l'homme du métier.

Une description plus détaillée d'un système de transmission d'un signal de parole à bas débit conforme à l'objet de la présente invention sera décrit en liaison avec les figures 8 et 9.

Selon la figure 8, le circuit de codage comporte un générateur 1 d'un premier dictionnaire Y_n de vecteurs de base y_i de forme n -aire de dimension L , les composantes de ces vecteurs, ainsi que mentionné précédemment, pouvant prendre les valeurs comprises entre $-n/2$ à $n/2$. On notera bien entendu que le générateur du dictionnaire Y_n peut avantageusement être constitué par des moyens calculateurs comportant les opérateurs tels que décrits en figures 3a, 3b par exemple et/ou un circuit de mémorisation lequel peut être constitué par une mémoire vive associée à ce circuit calculateur ou par une mémoire morte. Dans ce cas, la mémoire morte est associée à un séquenceur rapide lequel permet d'effectuer une lecture successive des vecteurs de base y_i selon les indices en numérotation directe ou rétrograde ainsi que décrit précédemment.

En outre, le circuit de codage tel que représenté en figure 8 comprend un circuit 2 correcteur des vecteurs de base y_i par un facteur d'échelle v_i . Le circuit correcteur peut être constitué par une table de valeurs mémorisées en mémoire morte, ce circuit correcteur permettant d'engendrer un vecteur de base corrigé noté $\bar{y}_i = v_i \cdot y_i$ pour chaque vecteur de base y_i . Un multiplexeur rapide noté MUX permet successivement de lire les valeurs correspondantes du vecteur de base corrigé \bar{y}_i et de délivrer cette valeur correspondante à un circuit 3 générateur d'un deuxième dictionnaire de gain adaptatif g_k . De manière classique, le circuit 3 générateur du deuxième dictionnaire $G(y)$ peut comporter avantageusement un circuit amplificateur, noté 30, relié à une table des valeurs g_k constituant le deuxième dictionnaire précité. Ainsi, le circuit générateur 3 du deuxième dictionnaire $G(y)$ délivre les vecteurs de référence $v_{k,i} = g_k \cdot v_i \cdot y_i$.

On notera bien sûr que le circuit de codage objet de la présente invention comporte également un circuit amplificateur 4 lequel permet d'appliquer à chaque vecteur de référence $v_{k,i}$ le coefficient de prédiction de niveau a tel que celui-ci a été défini précédemment dans la description.

En outre, et de manière classique, le circuit de codage objet de la présente invention comporte ensuite disposés en cascade le filtre de synthèse noté 5 et le filtre de pondération perceptuelle noté 6 de transmission H ainsi que décrit précédemment dans la description. Un sommateur 7 permet de recevoir d'une part le signal d'origine par l'intermédiaire d'un même filtre de pondération perceptuelle 6 après inversion de la différence des signaux délivrés par le sommateur 7 sommateur algébrique, permettant l'application sur le signal ainsi obtenu du critère de distorsion minimum.

Dans ce but, le circuit de codage objet de la présente invention comprend un circuit de calcul du minimum de distorsion 8 lequel comprend un premier circuit calculateur 80 du produit $2 \cdot g_k \cdot \langle x / \sigma | H \cdot v_i \cdot y_i \rangle$ dans lequel l'expression $\langle x / \sigma | H \cdot v_i \cdot y_i \rangle$ désigne le produit scalaire du vecteur cible x et du vecteur reconstitué et pondéré perceptuellement obtenu par le produit de la matrice H et du vecteur de base corrigé $v_i \cdot y_i$. Le premier circuit calculateur 80 délivre un premier résultat de calcul r_1 .

Un deuxième circuit calculateur 81 permet d'effectuer le calcul de l'énergie du vecteur reconstitué et pondéré perceptuellement, cette énergie étant de la forme $g_k^2 \| H \cdot v_i \cdot y_i \|^2$.

On notera que les circuits calculateurs 80 et 81 peuvent être constitués par des modules de programme dont les graphes de calcul ont été explicités en figures 4 et 5 a) à d) respectivement. Le deuxième circuit de calcul 81 délivre un deuxième résultat de calcul noté r_2 . Un comparateur 83 permet de comparer la valeur des résultats de calcul r_1 et r_2 ce qui permet de déterminer par discrimination des valeurs des indices i et k , les indices i^* et k^* pour lesquels le critère de minimum de l'écart quadratique est satisfait. La discrimination des indices i^* et k^* est effectuée par exemple par un programme de tri noté 84 sur la figure 8. Les valeurs des indices k^* et i^* sont alors délivrées, ces indices étant représentatifs du vecteur de référence correspondant v_{k^*,i^*} .

Sur la figure 8, on a également représenté le circuit de transmission conforme à l'objet de la présente invention, ce circuit de transmission permettant de délivrer en tant que signal de code représentatif du signal de parole les seules valeurs des indices k^* et i^* . Ce circuit de transmission ne présente pas de caractéristique particulière dans la mesure où il peut en fait être constitué par un système de transmission de type classique utilisé dans les dispositifs de transmission de signaux de parole par codage de type CELP de l'art antérieur.

Une description plus détaillée d'un circuit de décodage permettant la mise en oeuvre du procédé objet de l'invention

est représentée en figure 9.

Conformément à la figure précitée, le circuit de décodage comporte un module 10 de discrimination des valeurs des indices i^* , k^* du signal de code reçu, le signal de code étant transmis bien entendu selon un protocole particulier lequel n'entre pas dans l'objet de la présente invention. En outre, le circuit de discrimination 10 effectuant ainsi une transformation série parallèle des informations relatives aux indices i^* , k^* , le circuit de décodage comporte un circuit de décomposition en base n de la valeur de l'indice i^* .

On comprendra bien sûr que parallèlement l'indice k^* est traité. Dans ce but, le circuit de décodage tel que représenté en figure 9 comprend une table des valeurs de gain adaptatif g_k notée 11, laquelle, sur réception de la valeur de l'indice k^* , permet de délivrer la valeur de gain adaptatif correspondant g_{k^*} . Ce circuit 11 peut être avantageusement constitué par une mémoire morte dans laquelle les valeurs de gain adaptatif g_k sont mémorisées.

En outre, un circuit générateur 12 du facteur d'échelle vi^* est prévu. Ce circuit peut consister en une mémoire morte formant une table de consultation, laquelle à la valeur i^* fait correspondre la valeur vi^* . Un circuit multiplicateur 12a permet d'engendrer un coefficient produit $A = \sigma \cdot g_{k^*} \cdot vi^*$ à partir des valeurs vi^* , g_{k^*} et du coefficient de niveau prédit σ .

Ainsi qu'on l'a représenté également en figure 9, le circuit de décodage comprend un circuit 13 générateur du vecteur de base régénéré \widehat{y}_{i^*} par décomposition en base n de la valeur de l'indice i^* . Dans ce but, un circuit 14 fait correspondre à la valeur i^* par transcodage des composantes en base n de la valeur de l'indice i^* , la valeur $[-n/2, \dots, 0, \dots, n/2]$, ce qui permet d'engendrer un vecteur de référence régénéré v_{k^*, i^*} du produit du vecteur \widehat{y}_{i^*} de base régénéré et du produit A .

Un filtre de synthèse 15 permet à partir du vecteur de référence régénéré \widehat{v}_{k^*, i^*} d'engendrer le signal de parole reconstruit.

Le fonctionnement du circuit de décodage tel que représenté en figure 9 peut être résumé de la façon ci-après selon un fonctionnement préférentiel.

La double multiplication réalisée au niveau du multiplicateur 12 donne un facteur d'amplitude noté $A = \sigma \cdot g_{k^*} \cdot vi^*$.

Si l'indice i^* du vecteur ternaire transmis correspond à la numérotation rétrograde on désigne par :

$i' = (3^L - 3) / 2 - i^*$ et on effectue la synthèse du vecteur d'excitation ou vecteur de référence reconstitué v_{k^*, i^*} comme suit :

- étape courante (j, t) ,
- si j modulo 3 vaut 0 alors $v_{k^*, i^*}(L-1-t) = -A$
- si j modulo 3 vaut 1 alors $v_{k^*, i^*}(L-1-t) = 0$
- si j modulo 3 vaut 2 alors $v_{k^*, i^*}(L-1-t) = A$

où $v_{k^*, i^*}(L-1-t)$ représente la composante de v_{k^*, i^*} à l'ordre $L-1-t$.

On notera que j est divisé par 3, division entière, et t est augmenté de 1, addition de 1 à un nombre entier. On initialise la première étape par $j=i'$ et $t=0$.

Bien entendu, on répète l'étape courante jusqu'à ce que $t=L-1$ inclus.

Si au contraire i^* provient d'une numérotation directe, ainsi que décrit précédemment, alors $i'=i$ et les opérations sur j modulo 3 sont effectuées comme précédemment mentionné.

On a ainsi décrit un procédé et un système de transmission de la parole à bas débit particulièrement performant dans la mesure où un avantage important réside dans le fait que le dictionnaire Y_n n'a pas à être mémorisé au niveau du décodeur. Ainsi seuls les indices du vecteur de référence sont transmis au décodeur, un calcul permettant en temps réel de reconstituer le vecteur de référence correspondant ce qui permet un gain de ressource de mémoire au niveau de chaque décodeur utilisé. En outre et en raison des processus de génération des vecteurs de base, des processus de calcul des produits scalaires et des énergies perceptuelles, il n'est pas non plus nécessaire de mémoriser les vecteurs de base au niveau du codeur, ce qui permet un gain substantiel en matériel de mise en oeuvre.

On comprendra également que les algorithmes de calcul décrits dans la description de l'objet de la présente invention permettent d'obtenir une très grande rapidité de calcul de par la rationalisation des opérations de calcul utilisés, et une simplification des matériels nécessaires à leur mise en oeuvre.

On notera enfin que le procédé et le système de transmission d'un signal de parole codé à bas débit objet de la présente invention ont été décrits dans le cas où le codage de type CELP met en oeuvre des vecteurs de base de type n -aire, le nombre n n'étant en principe pas limité. Bien entendu, un mode de réalisation préférentiel a été donné dans le cas où $n=3$, les vecteurs de base étant ainsi des vecteurs ternaires.

Toutefois un mode de réalisation basé sur le même principe a pu être réalisé pour des vecteurs pour lesquels $n=5$. Le dictionnaire Y_n est alors réalisé à partir d'un alphabet à cinq symboles, les valeurs obtenues étant par exemple, de manière non limitative, le symbole 0, le symbole 0,5 et le symbole 1 plus les symboles symétriques -0,5 et -1, ce qui peut être ramené à des valeurs entières quelconques par changement d'échelle.

Dans la mise en oeuvre d'un dictionnaire à cinq symboles, on a pu ainsi réaliser un procédé et un système de transmission à débit variable pouvant atteindre jusqu'à 24 Kbits par seconde.

Revendications

1. Procédé de transmission d'un signal de parole à bas débit comportant un processus de codage d'échantillons numériques de parole par prédiction linéaire excitée par codes, pour engendrer un signal de code, un processus de transmission du signal de code et un processus de décodage de signal de code reçu, le processus de codage correspondant à un processus dans lequel une forme d'onde représentée par un code d'échantillons comportant L valeurs d'échantillons et constituant un vecteur initial (o) de dimension L est représenté, à partir d'un filtre de synthèse, par une forme d'onde de référence sélectionnée parmi un dictionnaire de formes d'onde de référence formant chacune un vecteur de référence (v) sur critère d'écart quadratique minimum dudit vecteur initial (o) vis-à-vis de ladite forme d'onde ou vecteur de référence (v), $\min \|\chi - H.v\|^2$, où χ représente un vecteur cible obtenu par pondération perceptuelle dudit vecteur initial (o) et H représente une matrice de dimension LxL de réponse impulsionnelle issue du produit du filtre de synthèse et de la pondération perceptuelle linéaire, ledit critère de sélection consistant :

- à établir un dictionnaire factorisé en produit d'un premier dictionnaire Y_n , $n=2m+1$ entier impair, m étant un nombre entier non nul, de vecteurs de base y_i , de dimension L, dont chaque valeur des composantes, $y_{i(j)}$, peut prendre une valeur appartenant à l'ensemble I_m des entiers relatifs consécutifs compris entre -m et m, soit $I_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ constituant un ensemble à $n=2m+1$ éléments, chacun de ces vecteurs de base étant multiplié par un facteur d'échelle v_i associé tenant compte de la distribution de l'énergie d'excitation dans le domaine de fréquence du signal, et d'un deuxième dictionnaire $G(y)$ de gains g_k adaptatifs, de façon à représenter ainsi le dictionnaire de formes d'ondes ou vecteurs de référence, chaque vecteur de référence vérifiant la relation $v_{k,i} = g_k.v_i.y_i$,
- à établir la valeur minimum de l'écart quadratique $\|\chi - g_k.H.v_i.y_i\|^2$ par le calcul du maximum de $C(g_k, v_i, y_i) = 2 g_k \langle \chi | H.v_i.y_i \rangle - g_k^2 \|H.v_i.y_i\|^2$ par calcul de tous les produits scalaires $\langle \chi | H.v_i.y_i \rangle$ et de toutes les énergies perceptuelles $\|H.v_i.y_i\|^2$, ce qui permet d'attribuer au vecteur initial (o) le vecteur de référence optimal correspondant v_{k^*,i^*} avec $v_{k^*,i^*} = g_{k^*}.v_{i^*}.y_{i^*}$, ce vecteur de référence optimal étant représenté par les seules valeurs des paramètres d'indice (k^*, i^*) satisfaisant au critère $\min \|\chi - g_k.H.v_i.y_i\|^2$.

2. Procédé selon la revendication 1, caractérisé en ce que ladite valeur minimum de l'écart quadratique $\min \|\chi - g_k.H.v_i.y_i\|^2$ est évaluée par sélection de l'élément de gain g_k correspondant du deuxième dictionnaire $G(y)$ permettant de rendre minimum la différence $lg - g^*l$ où g vérifie la relation :

$$g = \frac{\langle \chi | H.v_i.y_i \rangle}{\|H.v_i.y_i\|^2}$$

3. Procédé selon l'une des revendications 1 ou 2, caractérisé en ce que ledit premier dictionnaire Y_n , $n=2m+1$ de vecteurs de base y_i , de forme n-aire $[-n/2, \dots, 0, \dots, n/2]$ de dimension L comprend tous les vecteurs de base dont les L composantes ont pour valeur l'une des valeurs $[-n/2, \dots, 0, \dots, n/2]$ à l'exception du vecteur nul, l'indice i des vecteurs de base étant pris égal à la valeur en base n de chaque vecteur de base après transcodage des valeurs $[-n/2, \dots, 0, \dots, n/2]$ en valeur correspondante $[0, 1, 2, \dots, n]$.

4. Procédé selon la revendication 3, caractérisé en ce que l'ensemble des vecteurs de base y_i constituant ledit dictionnaire Y_n est défini à partir des $n/2.L$ vecteurs impulsion, dont une seule composante $y_{i(j)}$ d'ordre j avec $j \in [0, L-1]$ est égal à -1, -2, ..., -n/2, à chaque vecteur impulsion étant associés les vecteurs de base apparentés présentant des valeurs de composante d'ordre $q \leq j$ identiques, chaque vecteur apparenté à un vecteur impulsion de rang q avec $q = j$ pour $y_{i(j)} \neq 0$ étant obtenu par combinaison linéaire dudit vecteur impulsion de rang q et des vecteurs impulsion ou apparentés de rang q supérieur.

5. Procédé selon l'une des revendications 1 à 4, caractérisé en ce que pour chaque vecteur de base y_i , le facteur d'échelle v_i qui lui est associé est déterminé expérimentalement, à partir d'une pluralité N de trames constituées par L valeurs de signal de parole et formant une base de données, le facteur d'échelle v_i pour chaque vecteur de base y_i étant choisi de façon à rendre minimum, pour la trame considérée, le résidu de filtrage desdites trames.

6. Procédé selon la revendication 1, caractérisé en ce que, afin d'assurer la transmission du signal de parole à bas débit, le processus de transmission consiste à transmettre comme signal de code les seules valeurs des indices (k^*, i^*) représentatives de chaque vecteur de référence v_{k^*,i^*} .

7. Procédé selon la revendication 1 et 2, caractérisé en ce que, afin d'assurer le décodage du signal de code, ce procédé consiste :

- à discriminer (1000) les valeurs des indices k^* , i^* constituant le signal de code,
- à décomposer (1001) la valeur de l'indice i^* , représentatif du vecteur de référence optimal en base n pour régénérer le vecteur de base correspondant y_{i^*} ,
- à effectuer (1002), à partir de la valeur de l'indice i^* et du facteur d'échelle v_{i^*} correspondants, une correction du vecteur de base correspondant régénéré pour constituer le vecteur de référence $v_{k^*,i^*} = v_{i^*} \cdot y_{i^*}$,
- à effectuer une opération de filtrage de synthèse (1003) du vecteur de référence pour engendrer le signal de parole reconstruit.

8. Procédé selon l'une des revendications précédentes, caractérisé en ce que préalablement au filtrage de synthèse, chaque vecteur de référence v_{k^*,i^*} est pondéré (1004) par un facteur de niveau prédit σ représentatif de l'énergie moyenne du signal d'excitation estimé sur au moins trois vecteurs excitation antérieurs successifs.

9. Système de transmission d'un signal de parole à bas débit comportant un circuit de codage d'échantillons numériques de parole par prédiction linéaire excitée par codes pour engendrer un signal de code, un circuit de transmission de ce signal de code, et un circuit de décodage du signal de code transmis, le circuit de codage comprenant un filtre de synthèse permettant de représenter une forme d'onde constituée par un bloc de L échantillons constituant un vecteur initial (o) de dimension L , par une forme d'onde de référence sélectionnée parmi un dictionnaire de formes d'onde de référence formant chacune un vecteur de référence (v) sur critère d'écart quadratique minimum dudit vecteur initial (o) vis-à-vis de la forme d'onde ou vecteur de référence (v), des moyens de pondération perceptuelle dudit vecteur initial (o) pour engendrer un vecteur cible χ à partir dudit vecteur initial (o), ledit critère d'écart quadratique minimum du vecteur initial (o) vis-à-vis de ladite forme d'onde ou vecteur de référence (v) étant de la forme $\min \|\chi - H \cdot v\|^2$ où H représente une matrice de dimension $L \times L$ de réponse impulsionnelle issue du produit dudit filtre de synthèse et de ladite pondération perceptuelle linéaire, ledit circuit de codage, afin de mettre en oeuvre ledit critère de sélection, comportant :

- des moyens (1) générateurs d'un premier dictionnaire Y_n , $n=2m+1$ entier impair, m étant un nombre entier non nul, de vecteurs de base y_i de dimension L , dont chaque valeur des composantes $y_{i(j)}$ peut prendre une valeur appartenant à l'ensemble I_m des entiers relatifs consécutifs compris entre $-m$ et m , soit $I_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ constituant un ensemble à $n=2m+1$ éléments,
- des moyens (2) multiplicateurs de chacun desdits vecteurs de base y_i par un facteur d'échelle v_i associé, ce facteur d'échelle tenant compte de la distribution de l'énergie d'excitation dans le domaine de fréquence du signal, lesdits moyens correcteurs permettant d'engendrer un vecteur de base corrigé $\bar{y}_i = v_i \cdot y_i$ pour chaque valeur de base y_i ,
- des moyens (3) générateurs d'un deuxième dictionnaire $G(y)$ de gains adaptatifs g_k , comprenant des moyens (30) multiplicateurs permettant à partir desdits vecteurs de base corrigés \bar{y}_i et des valeurs de gain g_k d'engendrer lesdits vecteurs de référence $v_{k,i} = g_k \cdot v_i \cdot y_i$,
- des premiers moyens (80) de calcul du produit $2g_k \langle \chi | H \cdot v_i \cdot y_i \rangle$ où $\langle \chi | H \cdot v_i \cdot y_i \rangle$ désigne le produit scalaire dudit vecteur cible χ et du vecteur reconstitué pondéré perceptuellement obtenu par le produit de la matrice H et du vecteur de base corrigé $v_i \cdot y_i$, lesdits premiers moyens de calcul délivrant un premier résultat de calcul (r_1),
- des deuxièmes moyens (81) de calcul de l'énergie du vecteur reconstitué pondéré perceptuellement $g_k^2 \|H \cdot v_i \cdot y_i\|^2$, lesdits deuxièmes moyens de calcul délivrant un deuxième résultat de calcul (r_2),
- des moyens (82) de comparaison desdits premiers et deuxièmes résultats de calcul, ce qui permet de déterminer, par discrimination des valeurs des indices i, k , les indices i^* et k^* pour lesquels le critère de minimum de l'écart quadratique est satisfait, le vecteur de référence correspondant v_{k^*,i^*} avec $v_{k^*,i^*} = g_{k^*} \cdot v_{i^*} \cdot y_{i^*}$ étant représenté par les seules valeurs des indices k^* , i^* satisfaisant au critère $\min \|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$.

10. Système selon la revendication 9, caractérisé en ce que le circuit de transmission permet la transmission, en tant que signal de code représentatif du signal de parole, des seules valeurs des indices k^* et i^* .

11. Système selon la revendication 9, caractérisé en ce que le circuit de décodage comprend :

- des moyens (10) de discrimination des valeurs des indices i^*, k^* du signal de code reçu,
- des moyens (11) générateurs d'un dictionnaire $G(y)$ de gains adaptatifs g_{k^*} à partir des valeurs k^* discriminées,
- des moyens (12) générateurs du facteur d'échelle v_{i^*} correspondant,
- des moyens (12a) multiplicateurs pour engendrer un coefficient produit $\sigma \cdot g_{k^*} \cdot v_{i^*}$ à partir des valeurs i^*, g_{k^*} et

d'un coefficient de niveau prédit σ ,

- des moyens (13) de décomposition en base n de la valeur d'indice i^* ,
- des moyens (14) générateurs du vecteur de base régénéré $\widehat{y}i^*$ correspondant à la valeur i^* par transcodage des composantes en base n de la valeur d'indice i^* , à chaque valeur $n, \dots, 2, 1, 0$ d'expression en base n de la valeur d'indice i^* étant associée respectivement la valeur $[-n/2, \dots, 0, \dots, n/2]$ ce qui permet d'engendrer un vecteur de référence régénéré \widehat{v}_{k^*, i^*} , un filtre de synthèse permettant à partir du vecteur de référence régénéré \widehat{v}_{k^*, i^*} d'engendrer le signal de parole reconstruit.

12. Système selon l'une des revendications 9 à 11, caractérisé en ce que ledit circuit de codage, respectivement du décodage, comporte en outre, en amont du filtre de synthèse, un circuit de correction du vecteur de référence v_{k^*, i^*} respectivement vecteur de référence régénéré \widehat{v}_{k^*, i^*} par un facteur de niveau prédit représentatif de l'énergie moyenne du signal d'excitation estimé sur au moins trois vecteurs excitation antérieurs successifs.

Patentansprüche

1. Verfahren zur Übertragung eines Sprachsignals mit niedriger Rate, umfassend einen Codiervorgang zur Codierung von digitalen Sprachmustern durch codegesteuerte lineare Prädiktion, um ein Codesignal zu erzeugen, einen Übertragungsvorgang zur Übertragung des Codesignals und einen Decodiervorgang zur Decodierung des empfangenen Codesignals, wobei der Codiervorgang einem Vorgang entspricht, bei dem eine durch einen Mustercode, der L Musterwerte umfaßt und einen Anfangsvektor (o) der Dimension L bildet, dargestellte Wellenform anhand eines Synthesefilters durch eine Referenzwellenform dargestellt wird, die aus einem Verzeichnis von jeweils einen Referenzvektor (v) bildenden Referenzwellenformen nach dem Kriterium der kleinsten quadratischen Abweichung $\min \|\chi - H \cdot v\|^2$ des Anfangsvektors (o) von der Referenzwellenform oder dem Referenzvektor (v) ausgewählt wird, wobei χ einen durch Wahrnehmungswichtung erhaltenen Zielvektor des Anfangsvektors (o) darstellt und H eine sich aus dem Produkt des Synthesefilters und der linearen Wahrnehmungswichtung ergebende Impulsantwortmatrix der Dimension L x L darstellt, wobei das Auswahlkriterium darin besteht:

- ein Verzeichnis aufzustellen, das zerlegt ist in das Produkt eines ersten Verzeichnisses Y_n von Basisvektoren y_i der Dimension L, wobei $n = 2m + 1$ ganzzahlig und ungerade ist, wobei m eine ganze Zahl ungleich null ist, wobei der Wert jeder Komponente $y_{i(j)}$ der Basisvektoren y_i einen Wert annehmen kann, der der Menge I_m von aufeinanderfolgenden ganzen Zahlen angehört, die zwischen -m und m enthalten sind, d.h. einer Menge $I_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ mit $n = 2m + 1$ Elementen, wobei jeder der Basisvektoren mit einem zugehörigen Skalierungsfaktor v_i multipliziert wird, der die Verteilung der Anregungsenergie im Frequenzbereich des Signals berücksichtigt, und eines zweiten Verzeichnisses $G(y)$ von adaptiven Verstärkungen g_k , um so das Verzeichnis der Referenzwellenformen oder Referenzvektoren darzustellen, wobei jeder Referenzvektor der Beziehung $v_{k,i} = g_k \cdot v_i \cdot y_i$ genügt,
- den minimalen Wert der quadratischen Abweichung $\|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$ durch Berechnung des Maximums von $C(g_k, v_i \cdot y_i) = 2 g_k \langle \chi | H \cdot v_i \cdot y_i \rangle - g_k^2 \|H \cdot v_i \cdot y_i\|^2$ zu ermitteln, und zwar durch Berechnung aller Skalarprodukte $\langle \chi | H \cdot v_i \cdot y_i \rangle$ sowie aller Wahrnehmungsenergien $\|H \cdot y_i\|^2$, was es erlaubt, dem Anfangsvektor (o) den entsprechenden optimalen Referenzvektor v_{k^*, i^*} mit $v_{k^*, i^*} = g_{k^*} \cdot v_{i^*} \cdot y_{i^*}$ zuzuordnen, wobei dieser optimale Referenzvektor durch die alleinigen Werte der Indexparameter (k^*, i^*) dargestellt wird, die dem Kriterium $\min \|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$ genügen.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß der minimale Wert der quadratischen Abweichung $\|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$ durch Wahl des entsprechenden Verstärkungselements g_k des zweiten Verzeichnisses $G(y)$ berechnet wird, das es erlaubt, die Differenz $I_g - g_k \cdot I$ minimal zu machen, wobei g der Beziehung genügt:

$$g = \frac{\langle \chi | H \cdot v_i \cdot y_i \rangle}{\|H \cdot v_i \cdot y_i\|^2}$$

3. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß das erste Verzeichnis Y_n , $n = 2m + 1$, der Basisvektoren y_i n-wertiger Form $[-n/2, \dots, 0, \dots, n/2]$ der Dimension L alle die Basisvektoren mit Ausnahme des Nullvektors umfaßt, deren L Komponenten als Wert einen der Werte $[-n/2, \dots, 0, \dots, n/2]$ besitzen, wobei der Index i der Basisvektoren gleich dem Wert auf Basis n jedes Basisvektors nach Umcodierung der Werte $[-n/2, \dots, 0, \dots, n/2]$ in einen entsprechenden Wert $[0, 1,$

2, ... n] genommen wird.

4. Verfahren nach Anspruch 3,

dadurch gekennzeichnet, daß die das Verzeichnis Y_n bildende Menge der Basisvektoren y_i anhand von $n/2$ -L Impulsvektoren definiert wird, von denen eine einzige Komponente $y_{i(j)}$ der Ordnung j mit $j \in [0, L-1]$ gleich $-1, -2, \dots, -n/2$ ist, wobei jedem Impulsvektor die verwandten Basisvektoren zugeordnet werden, die identische Werte der Komponente der Ordnung $q \leq j$ aufweisen, wobei jeder Vektor, der einem Impulsvektor vom Rang q mit $q = j$ für $y_{i(j)} \neq 0$ verwandt ist, durch lineare Kombination des Impulsvektors vom Rang q und von Impulsvektoren oder verwandten Vektoren vom Rang größer als q erhalten wird.

5. Verfahren nach einem der Ansprüche 1 bis 4,

dadurch gekennzeichnet, daß für jeden Basisvektor y_i der diesem zugeordnete Skalierungsfaktor v_i anhand einer Mehrzahl N von Blöcken, die von L Sprachsignalwerten gebildet sind und eine Datenbasis bilden, experimentell bestimmt wird, wobei der Skalierungsfaktor v_i für jeden Basisvektor y_i so gewählt wird, daß für den betrachteten Block der Rest der Filterung der Blöcke minimal wird.

6. Verfahren nach Anspruch 1,

dadurch gekennzeichnet, daß zur Gewährleistung der Übertragung des Sprachsignals mit niedriger Rate der Übertragungsvorgang darin besteht, als Codesignal allein die Werte der Indizes (k^*, i^*) zu übertragen, die für die jeweiligen Referenzvektoren v_{k^*, i^*} repräsentativ sind.

7. Verfahren nach Anspruch 1 und 2,

dadurch gekennzeichnet, daß zur Gewährleistung der Decodierung des Codesignals das Verfahren darin besteht:

- die Werte der das Codesignal bildenden Indizes k^*, i^* zu diskriminieren (1000),
- den Wert des Index i^* , der für den optimalen Referenzvektor auf Basis n repräsentativ ist, zu zerlegen (1001), um den entsprechenden Basisvektor y_{i^*} zu regenerieren,
- anhand des Werts des Index i^* und des entsprechenden Skalierungsfaktors v_{i^*} eine Korrektur des entsprechenden regenerierten Basisvektors durchzuführen (1002), um den Referenzvektor $v_{k^*, i^*} = v_{i^*} \cdot y_{i^*}$ zu bilden,
- eine Synthesefilteroperation an dem Referenzvektor durchzuführen (1003), um das rekonstruierte Sprachsignal zu erzeugen.

8. Verfahren nach einem der vorhergehenden Ansprüche,

dadurch gekennzeichnet, daß vor der Synthesefilterung jeder Referenzvektor v_{k^*, i^*} mit einem vorhergesagten Niveaufaktor σ gewichtet wird (1004), der für die mittlere Energie des Anregungssignals repräsentativ ist und anhand wenigstens dreier vorhergehender aufeinanderfolgender Anregungsvektoren geschätzt wird.

9. System zur Übertragung eines Sprachsignals mit niedriger Rate, umfassend einen Codierkreis zur Codierung digitaler Sprachmuster

durch codegesteuerte lineare Prädiktion, um ein Codesignal zu erzeugen, einen Übertragungskreis zur Übertragung dieses Codesignals und einen Decodierkreis zur Decodierung des übertragenen Codesignals, wobei der Codierkreis ein Synthesefilter umfaßt, das es erlaubt, eine Wellenform, die von einem einen Anfangsvektor (o) der Dimension L bildenden Block von L Mustern gebildet ist, durch eine Referenzwellenform darzustellen, die aus einem Verzeichnis von jeweils einen Referenzvektor (v) bildenden Referenzwellenformen nach dem Kriterium der minimalen quadratischen Abweichung des Anfangsvektors (o) von der Referenzwellenform oder dem Referenzvektor (v) ausgewählt wird, und Wichtungsmittel zur Wahrnehmungswichtung des Anfangsvektors (o) umfaßt, um anhand des Anfangsvektors (o) einen Zielvektor χ zu erzeugen, wobei das Kriterium der minimalen quadratischen Abweichung des Anfangsvektors (o) von der Referenzwellenform oder dem Referenzvektor (v) die Form $\min \|\chi - H \cdot v\|^2$ besitzt, wobei H eine sich aus dem Produkt des Synthesefilters und der linearen Wahrnehmungswichtung ergebende Impulsantwortmatrix der Dimension $L \times L$ darstellt, wobei der Codierkreis zur Erfüllung des Auswahlkriteriums umfaßt:

- Erzeugungsmittel (1) zur Erzeugung eines ersten Verzeichnisses Y_n von Basisvektoren y_i der Dimension L , wobei $n = 2m + 1$ ganzzahlig und ungerade ist, wobei m eine ganze Zahl ungleich null ist, wobei der Wert jeder Komponente $y_{i(j)}$ der Basisvektoren y_i einen Wert annehmen kann, der der Menge l_m von aufeinanderfolgenden ganzen Zahlen angehört, die zwischen $-m$ und m enthalten sind, d.h. einer Menge $l_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ mit $n = 2m + 1$ Elementen,
- Multiplikationsmittel (2) zur Multiplikation jedes der Basisvektoren y_i mit einem zugehörigen Skalierungsfaktor

v_i , wobei dieser Skalierungsfaktor die Verteilung der Anregungsenergie im Frequenzbereich des Signals berücksichtigt, wobei die Korrekturmittel für jeden Basiswert y_i die Erzeugung eines korrigierten Basisvektors $\bar{y}_i = v_i \cdot y_i$ ermöglichen,

- Erzeugungsmittel (3) zur Erzeugung eines zweiten Verzeichnisses $G(y)$ von adaptiven Verstärkungen g_k , wobei diese Erzeugungsmittel (3) Multiplikatormittel (30) umfassen, die ausgehend von den korrigierten Basisvektoren \bar{y}_i und den Verstärkungswerten g_k die Erzeugung der Referenzvektoren $v_{k,i} = g_k \cdot v_i \cdot y_i$ ermöglichen,
- erste Berechnungsmittel (80) zur Berechnung des Produkts $2g_k \langle \chi | H \cdot v_i \cdot y_i \rangle$, wobei $\langle \chi | H \cdot v_i \cdot y_i \rangle$ das Skalarprodukt des Zielvektors χ und des durch das Produkt der Matrix H und des korrigierten Basisvektors $v_i \cdot y_i$ erhaltenen, wahrnehmungsmäßig gewichteten nachgebildeten Vektors bezeichnet, wobei die ersten Berechnungsmittel ein erstes Berechnungsergebnis (r_1) bereitstellen,
- zweite Berechnungsmittel (81) zur Berechnung der Energie $g_k^2 \|H \cdot v_i \cdot y_i\|^2$ des wahrnehmungsmäßig gewichteten nachgebildeten Vektors, wobei die zweiten Berechnungsmittel ein zweites Berechnungsergebnis (r_2) bereitstellen,
- Vergleichsmittel (82) zum Vergleichen des ersten und des zweiten Berechnungsergebnisses, was es erlaubt, durch Diskriminierung der Werte der Indizes i, k die Indizes i^* und k^* zu bestimmen, für die das Kriterium der minimalen quadratischen Abweichung erfüllt ist, wobei der entsprechende Referenzvektor v_{k^*,i^*} mit $v_{k^*,i^*} = g_{k^*} \cdot v_{i^*} \cdot y_{i^*}$ allein durch die Werte der Indizes k^*, i^* dargestellt wird, die dem Kriterium $\min \| \chi - g_k \cdot H \cdot v_i \cdot y_i \|^2$ genügen.

10. System nach Anspruch 9,

dadurch gekennzeichnet, daß der Übertragungskreis die Übertragung der bloßen Werte der Indizes k^* und i^* als für das Sprachsignal repräsentatives Codesignal erlaubt.

11. System nach Anspruch 9,

dadurch gekennzeichnet, daß der Decodierkreis umfaßt:

- Diskriminierungsmittel (10) zur Diskriminierung der Werte der Indizes i^*, k^* des empfangenen Codesignals,
- Erzeugungsmittel (11) zur Erzeugung eines Verzeichnisses $G(y)$ von adaptiven Verstärkungen g_{k^*} anhand der diskriminierten Werte k^* ,
- Erzeugungsmittel (12) zur Erzeugung eines entsprechenden Skalierungsfaktors v_{i^*} ,
- Multiplikatormittel (12a) zur Erzeugung eines Produktkoeffizienten $\sigma \cdot g_{k^*} \cdot v_{i^*}$ anhand der Werte i^*, k^* und eines vorhergesagten Niveaueffizienten σ ,
- Zerlegungsmittel (13) zur Zerlegung des Werts des Index i^* auf Basis n ,
- Erzeugungsmittel (14) zur Erzeugung des dem Wert i^* entsprechenden regenerierten Basisvektors \widehat{v}_{i^*} durch Umcodierung der Komponenten des Werts des Index i^* auf Basis n , wobei jedem Wert $n-2, 1, 0$ des Ausdrucks auf Basis n des Werts des Index i^* jeweils der Wert $[-n/2, \dots, 0, \dots, n/2]$ zugeordnet wird, was die Erzeugung eines regenerierten Referenzvektors \widehat{v}_{k^*,i^*} ermöglicht, wobei ein Synthesefilter es erlaubt, anhand des regenerierten Referenzvektors \widehat{v}_{k^*,i^*} das nachgebildete Sprachsignal zu erzeugen.

12. System nach einem der Ansprüche 9 bis 11,

dadurch gekennzeichnet, daß der Codierkreis bzw. der Decodierkreis darüber hinaus dem Synthesefilter vorgeschaltet einen Korrekturkreis zur Korrektur des Referenzvektors v_{k^*,i^*} bzw. des regenerierten Referenzvektors \widehat{v}_{k^*,i^*} mit einem vorhergesagten Niveaufaktor umfassen, der für die mittlere Energie des Anregungssignals repräsentativ ist und anhand wenigstens dreier vorhergehender aufeinanderfolgender Anregungsvektoren geschätzt wird.

Claims

1. A method of transmitting a speech signal at low throughput comprising a procedure for coding digital samples of speech by code excited linear prediction, in order to generate a code signal, a procedure for transmitting the code signal and a received code signal decoding procedure, the coding procedure corresponding to a procedure in which a waveform represented by a sample block comprising L sample values and constituting an initial vector (o) of dimension L is represented, on the basis of a synthesizing filter, by a reference waveform chosen from a dictionary of reference waveforms each forming a reference vector (v) relating to a criterion of minimum square deviation of the said initial vector (o) in relation to the said waveform or reference vector (v), $\min \| \chi - H \cdot v \|^2$, where χ represents a target vector obtained by perceptual weighting of the said initial vector (o) and H represents a pulse-response matrix of dimension $L \times L$ resulting from the product of the synthesizing filter and of the linear perceptual weighting, wherein the said selection criterion consists:

- in establishing a dictionary factorized as a product of a first dictionary Y_n , with $n=2m+1$ being an odd integer, and m being an integer different from zero, of basis vectors y_i , of dimension L , each component $y_{i(j)}$ of said basis vectors having a given value belonging to the set of successive related integer values l_m comprised between $-m$ and m , with $l_m = \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ forming a set of values comprising $n=2m+1$ elements, these basis vectors each being weighted by an associated scale factor v_i which takes account of the distribution of excitation energy in the frequency domain of the signal and of a second dictionary $G(y)$ of adaptive gains g_k , in such a way as to thus represent the dictionary of waveforms or reference vectors, each reference vector satisfying the relation $v_{k,i} = g_k \cdot v_i \cdot y_i$,
- in establishing the minimum value of the square deviation $\min \|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$ by calculating the maximum of $C(g_k, v_i, y_i) = 2g_k \langle \chi | H \cdot v_i \cdot y_i \rangle - g_k^2 \|H \cdot v_i \cdot y_i\|^2$ by calculating all the scalar products $\langle \chi | H \cdot v_i \cdot y_i \rangle$ and all the perceptual energies $\|H \cdot y_i\|^2$, this making it possible to assign to the initial vector (o) the corresponding optimal reference vector v_{k^*, i^*} with $v_{k^*, i^*} = g_{k^*} \cdot v_i \cdot y_{i^*}$, this optimal reference vector being represented by the sole values of the index parameters (k^*, i^*) satisfying the criterion $\min \|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$.

2. The method as claimed in claim 1, characterized by said minimum value of the square deviation $\min \|\chi - g_k \cdot H \cdot v_i \cdot y_i\|^2$ being evaluated by selecting the corresponding gain element g_k of the second dictionary $G(y)$ making it possible to minimize the difference $|g - g_{k^*}|$ where g satisfies the relation:

$$g = \frac{\langle \chi | H \cdot v_i \cdot y_i \rangle}{\|H \cdot v_i \cdot y_i\|^2}.$$

3. The method as claimed in one of the claims 1 or 2, characterized by said first dictionary Y_n , $n=2m+1$, of basis vectors y_i , of n -ary form $[-n/2, \dots, 0, \dots, n/2]$ of dimension L comprising all the basis vectors whose L components have the value of one of the values $[-n/2, \dots, 0, \dots, n/2]$ with the exception of the null vector, the index i of the basis vectors being made equal to the base n value of each basis vector after transcoding of the values $[-n/2, 0, \dots, n/2]$ into a corresponding value $[0, 1, 2, \dots, n]$.

4. The method as claimed in claim 3, characterized by the set of basis vectors y_i constituting the said dictionary Y_n being defined from the $n/2 \cdot L$ pulse vectors, of which a single component $y_{i(j)}$ of order j with $j \in [0, L-1]$ is equal to $-1, -2, \dots, -n/2$, each pulse vector being associated with the allied basis vectors having identical component values of order $q < j$, each vector allied to a pulse vector of rank q with $q = j$ for $y_{i(j)} \neq 0$ being obtained by linear combination of the said pulse vector of rank q and of the pulse or allied vectors of higher rank q .

5. The method as claimed in one of the claims 1 to 4, characterized by for each basis vector y_i , the scale factor v_i associated with it is determined experimentally, from a plurality N of frames consisting of L speech-signal values and forming a database, the scale factor v_i for each basis vector y_i being selected in such a way as to minimize, from the relevant frame, the filtering residue from the said frames.

6. The method as claimed in claim 1, characterized by in order to ensure the transmission of the speech signal at low throughput, the transmission procedure consisting in transmitting as code signal the sole values of the indices (k^*, i^*) representing each reference vector v_{k^*, i^*} .

7. The method as claimed in claims 1 and 2, characterized by in order to ensure the decoding of the code signal, this method consisting:

- in distinguishing (1000) the values of the indices k^*, i^* constituting the code signal,
- in decomposing (1001) the value of the indice i^* , representing the optimal reference vector to base n in order to generate the corresponding basis vector y_{i^*} ,
- in performing (1002), from the corresponding value of the index i^* and of the corresponding scale factor v_{i^*} , a correction of the corresponding regenerated basis vector in order to build up the reference vector $v_{k^*, i^*} = v_{i^*} \cdot y_{i^*}$,
- in performing a synthesizing filtering operation (1003) of the reference vector in order to generate the reconstructed speech signal.

8. The method according to one of the preceding claims characterized by prior to the synthesizing filtering, each reference vector v_{k^*, i^*} being weighted (1004) by a predicted level factor a representing the average energy of the excitation signal estimated over at least three successive earlier excitation vectors.

9. A system for transmitting a speech signal at low throughput comprising a circuit for coding digital samples of speech by code excited linear prediction in order to generate a code signal, a circuit for transmitting this code signal, and a circuit for decoding the transmitted code signal, the coding circuit comprising a synthesizing filter making it possible to represent a waveform consisting of a sample block constituting an initial vector (o) of dimension L, by a reference waveform chosen from a dictionary of reference waveforms each forming a reference vector (v) relating to a criterion of minimum square deviation of the said initial vector (o) in relation to the waveform or reference vector (v), means of perceptual weighting of the said initial vector (o) in order to generate a target vector χ from the said initial vector (o), the said criterion of minimum square deviation of the initial vector (o) in relation to the said waveform or reference vector (v) being of the form $\min \|\chi - H.v\|^2$ where H represents a pulse-response matrix of the dimension LxL resulting from the product of the said synthesizing filter and of the said linear perceptual weighting, characterized by, in order to implement the said selection criterion, the said coding circuit comprising:

- means (1) for generating a first dictionary Y_n , with $n=2m+1$ being an odd integer, and m being an integer different from zero, of basis vectors y_i of dimension L, each component $y_{i(j)}$ of said basis vectors having a given value belonging to the set of successive relative integer values l_m comprised between -m and m, with $l_m = \{m, m-1, \dots, -1, 0, 1, \dots, m-1, m\}$ forming a set of values comprising $n=2m+1$ elements,
- means (2) for multiplying the said basis vectors y_i by a scale factor v_i , this scale factor taking into account of the distribution of the excitation energy in the frequency domain of the signal, the said correcting means making it possible to generate a corrected basis vector $\hat{y}_i = v_i.y_i$ for each basis vector y_i ,
- means (3) for generating a second dictionary G(y) of adaptive gains g_k , comprising multiplier means making it possible, from the said corrected basis vector y_i and from the gain values g_k , to generate the said reference vectors $v_{k,i} = g_k.v_i.y_i$,
- first means (80) of calculating the product $2g_k \langle \chi | H.v_i.y_i \rangle$ where $\langle \chi | H.v_i.y_i \rangle$ designates the scalar product of the said target vector χ and of the perceptually weighted reconstituted vector obtained by the product of the matrix H and of the corrected basis vector $v_i.y_i$, the said first calculation means delivering a first calculation result (r1),
- second means (81) of calculating the energy of the perceptually weighted reconstituted vector $g_k^2 \|H.v_i.y_i\|^2$, the said second calculation means delivering a second calculation result (r2),
- means (82) of comparing the said first and second calculation results, this making it possible to determine, by distinguishing the values of the indices i,k, the indices i^* and k^* for which the criterion of minimum square deviation is satisfied, the corresponding reference vector v_{k^*,i^*} with $v_{k^*,i^*} = g_{k^*}.v_{i^*}.y_{i^*}$ being represented by the sole values of the indices k^* , i^* satisfying the criterion $\min \|\chi - g_k.H.v_i.y_i\|^2$.

10. The system as claimed in claim 9, characterized by the transmission circuit making it possible to transmit, in the guise of a code signal representing the speech signal, the sole values of the indices k^* and i^* .

11. The system as claimed in claim 9, characterized by the decoding circuit comprising :

- means (10) for distinguishing the values of the indices i^* , k^* of the code signal received,
- means (11) generating a dictionary G(y) for adaptive gains g_{k^*} from the distinguished values k^* ,
- means (12) generating the corresponding scale factor v_{i^*} ,
- multiplying means (12a) for generating a product coefficient $\sigma.g_{k^*}.v_{i^*}$ from the values i^*, g_{k^*} and from a predicted level coefficient σ ,
- means (13) of decomposing to base n the index value i^* ,
- means (14) generating the regenerated basis vector \hat{y}_{i^*} corresponding to the value i^* by transcoding of the components to base n of the index value i^* , each value $n, \dots, 2, 1, 0$ of the expression to base n of the index value i^* being associated with respectively the value $[n/2, \dots, 0, \dots, n/2]$, this making it possible to generate a regenerated reference vector \widehat{v}_{k^*,i^*} , a synthesizing filter making it possible on the basis of the regenerated reference vector \widehat{v}_{k^*,i^*} to generate the reconstructed speech signal.

12. The system as claimed in one of claims 9 to 11, characterized by the said coding, respectively decoding circuit further comprising, upstream of the synthesizing filter, a circuit for correcting the reference vector v_{k^*,i^*} , respectively regenerated reference vector \widehat{v}_{k^*,i^*} by a predicted level factor representing the average energy of the excitation signal estimated over at least three successive earlier excitation vectors.

FIG.1a ART ANTERIEUR

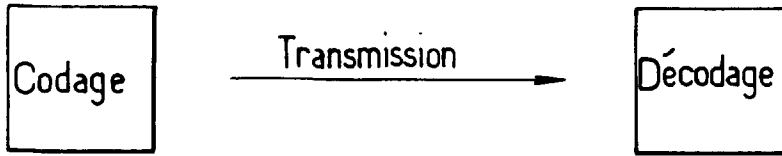
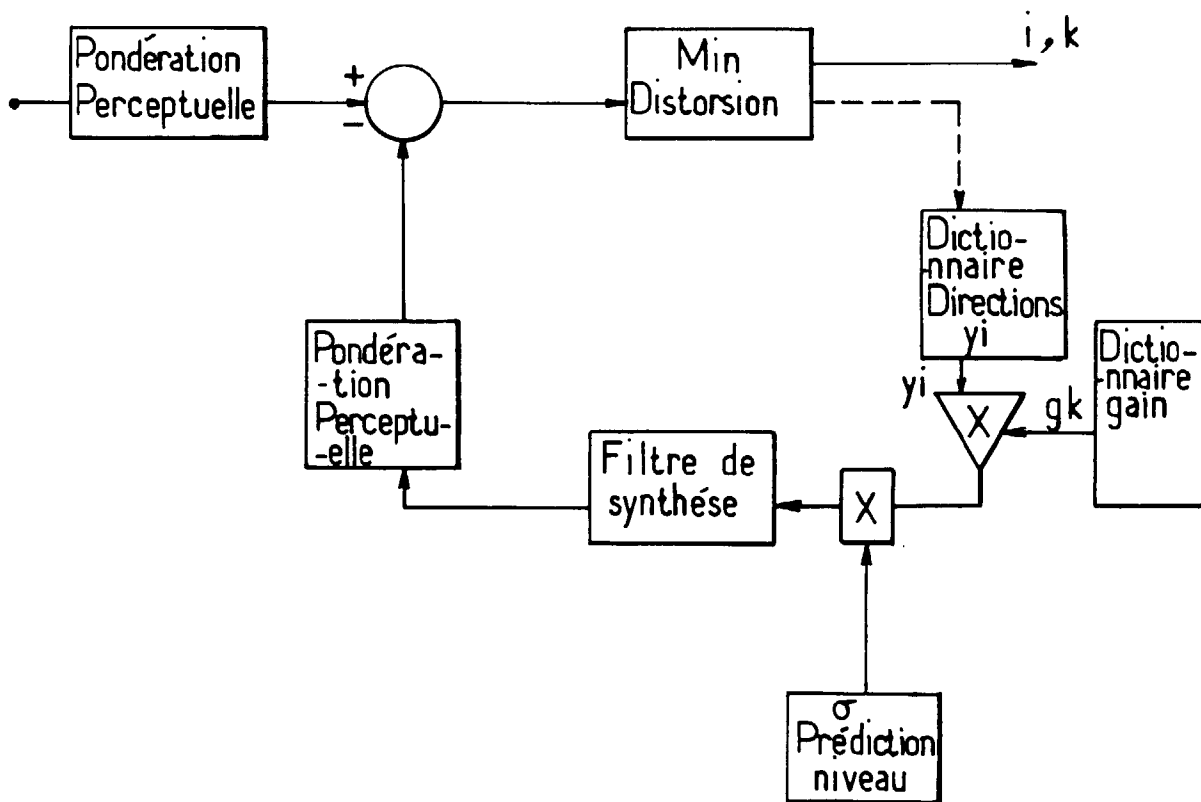


FIG.1b . ART ANTERIEUR



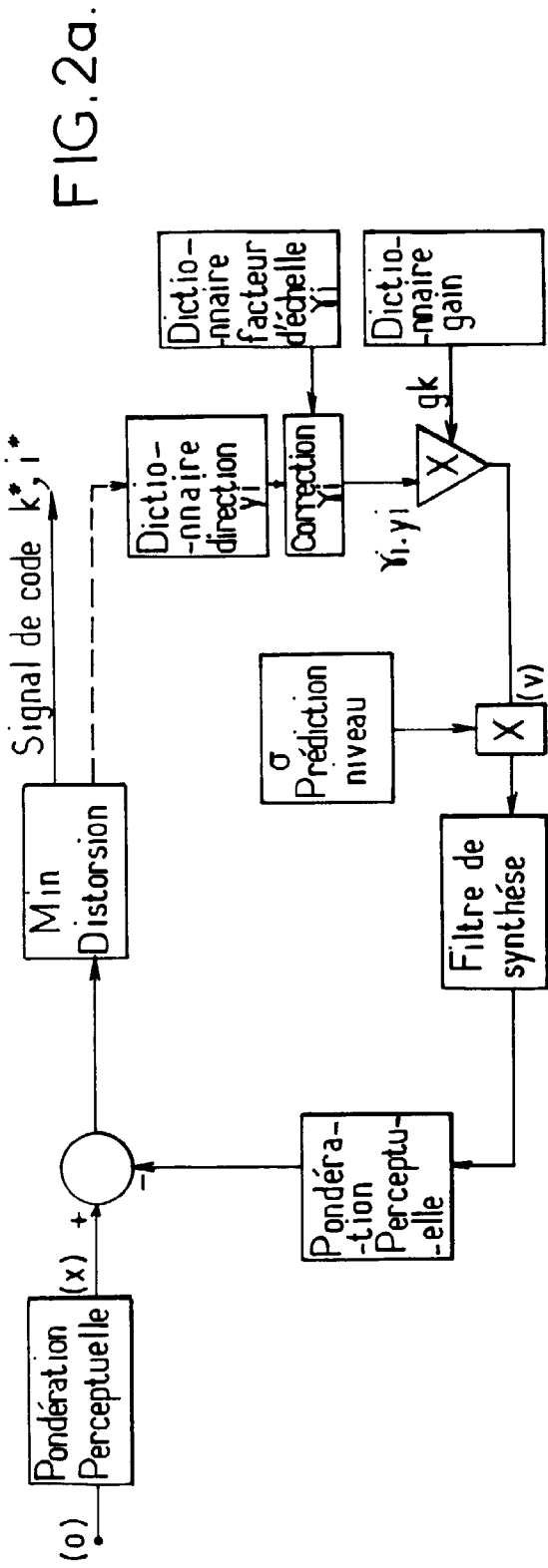


FIG. 2b.

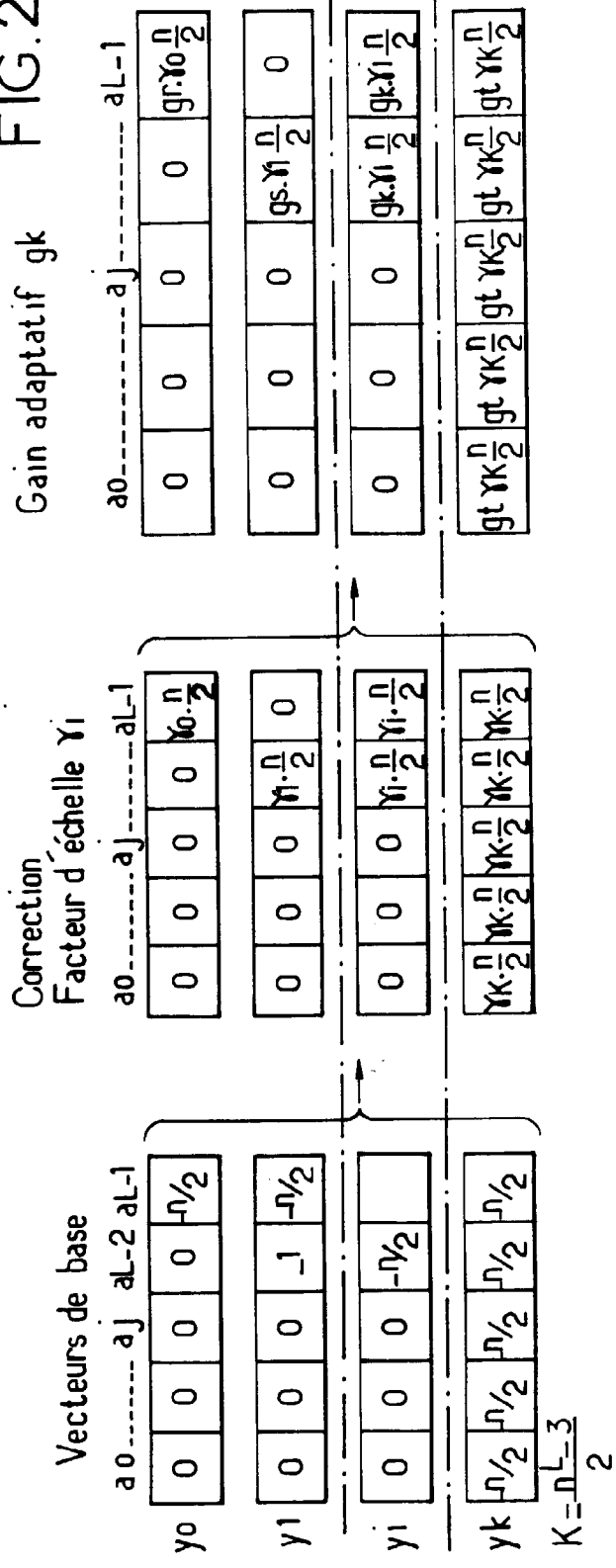


FIG.3a.

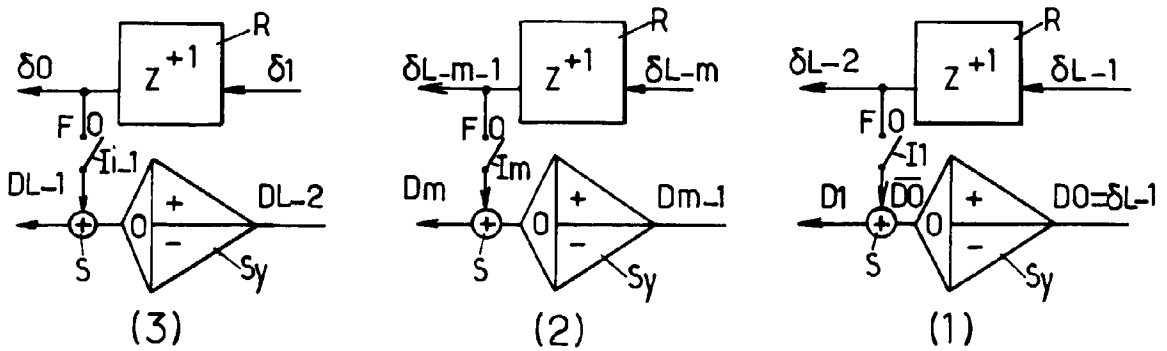


FIG.3b.

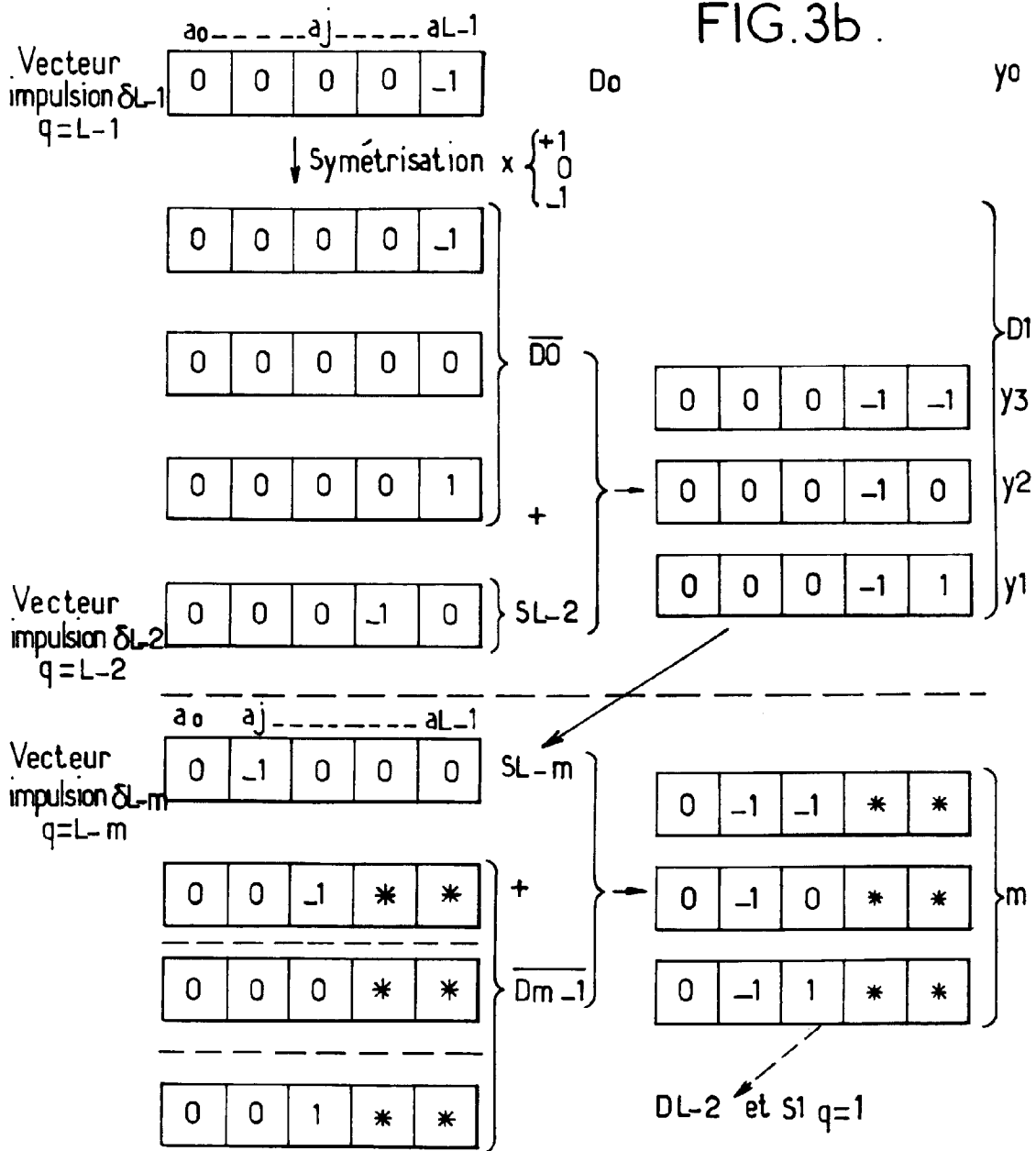


FIG.4.

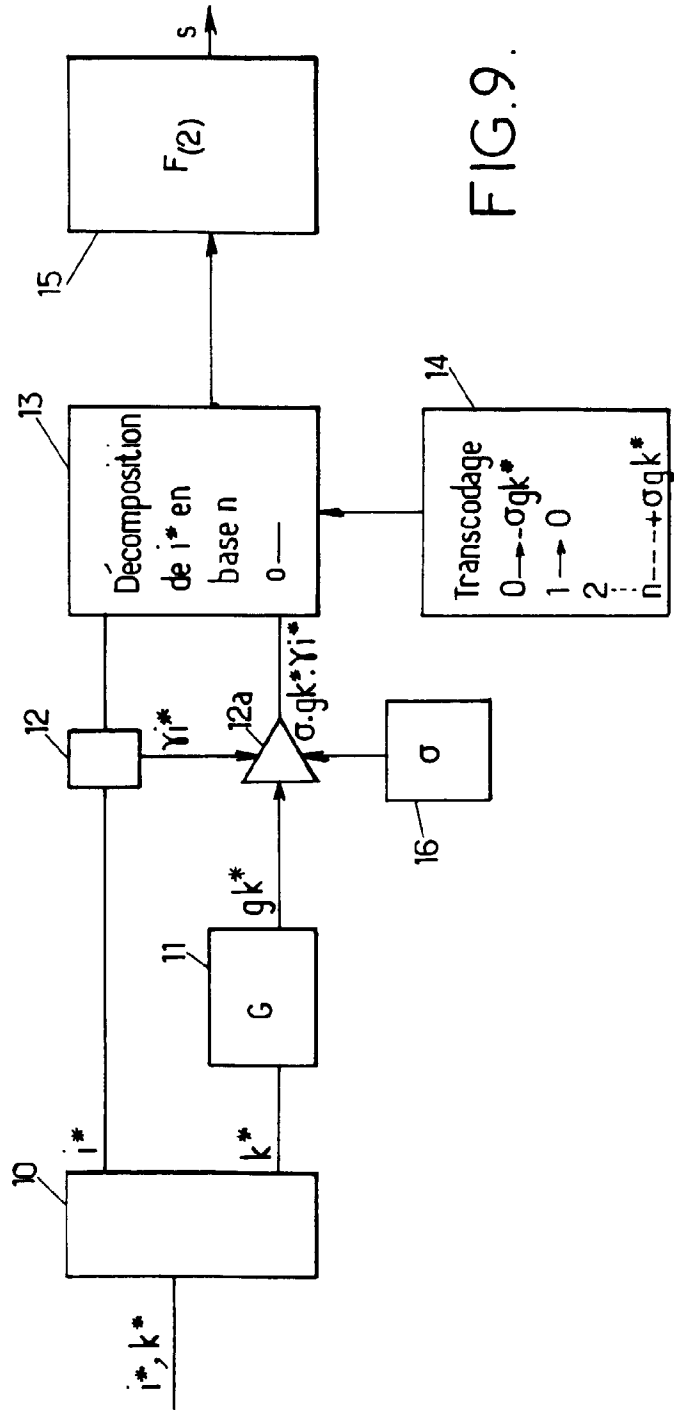
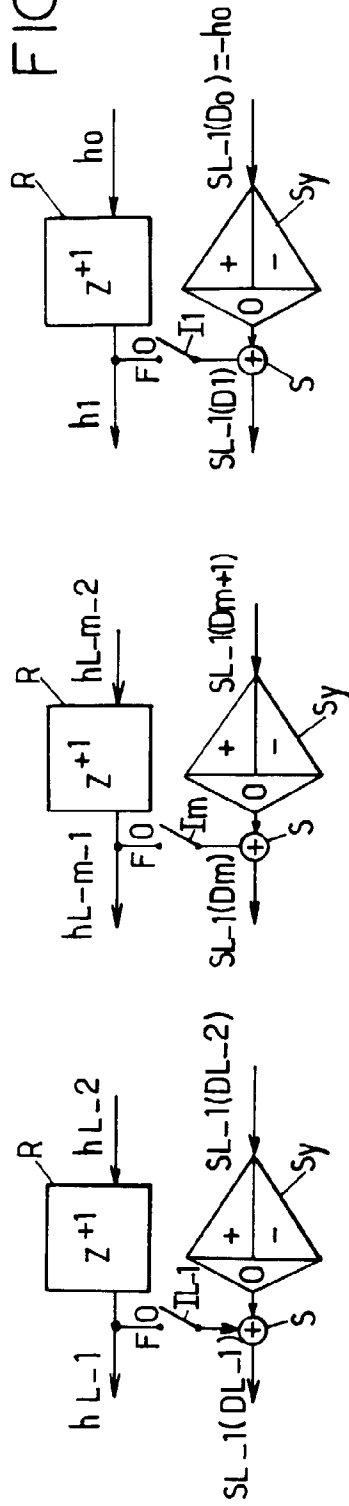


FIG.9.

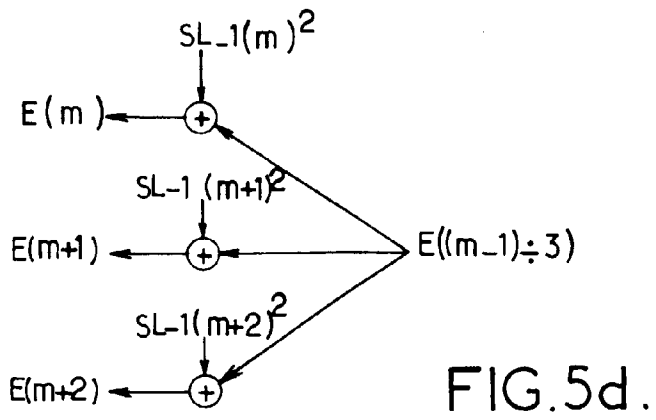
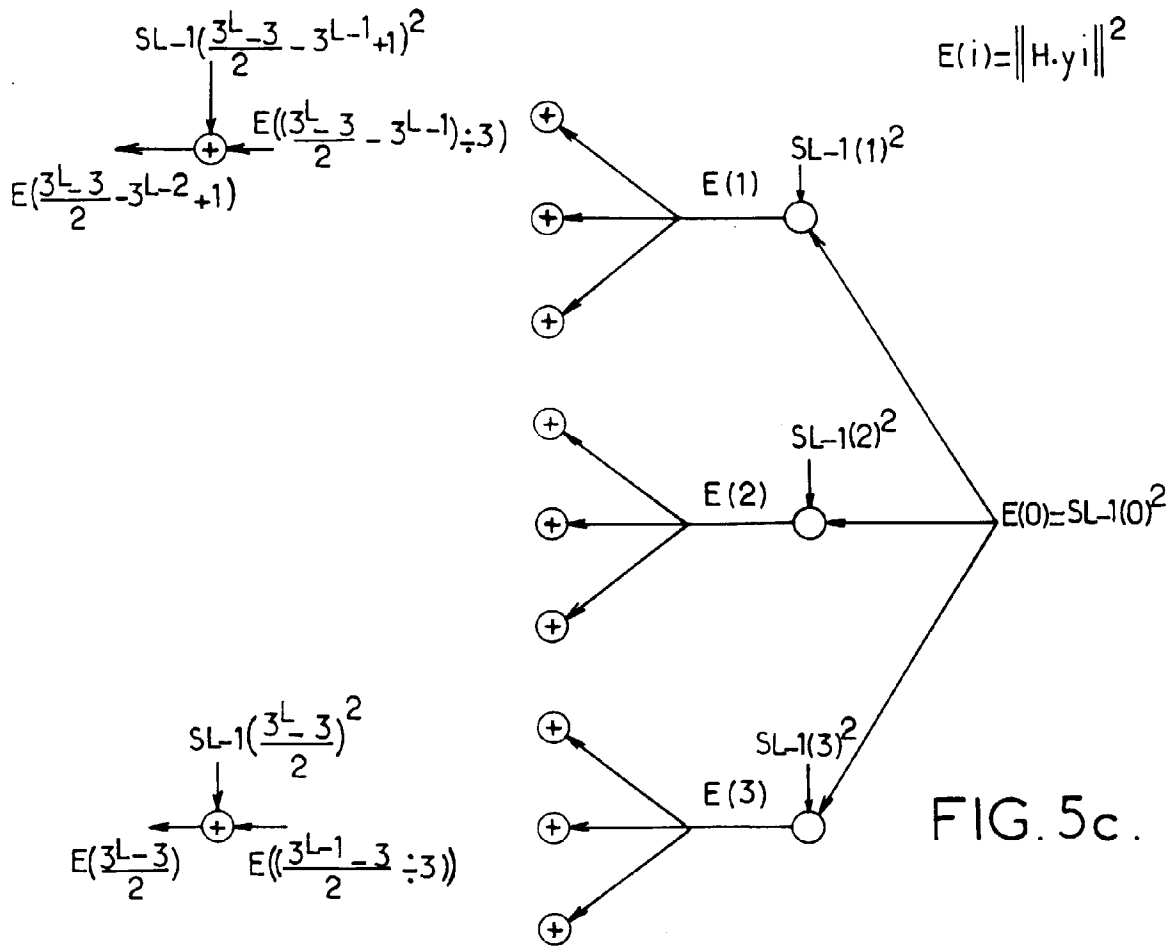


FIG. 6.

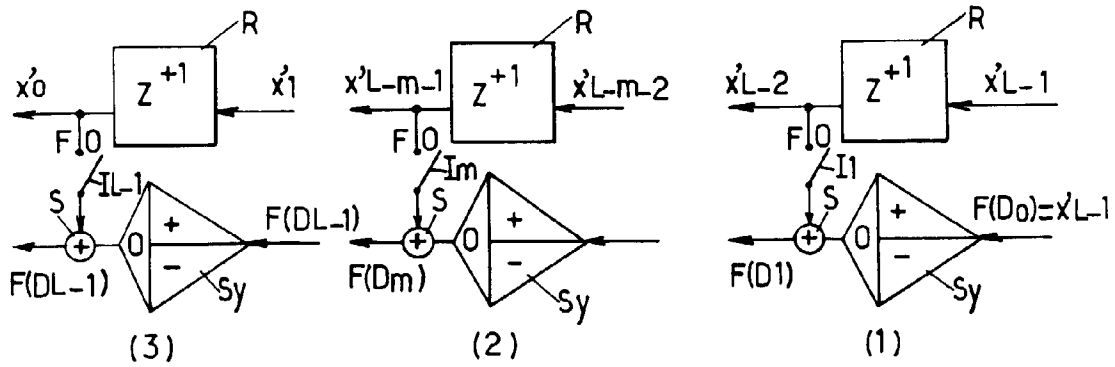
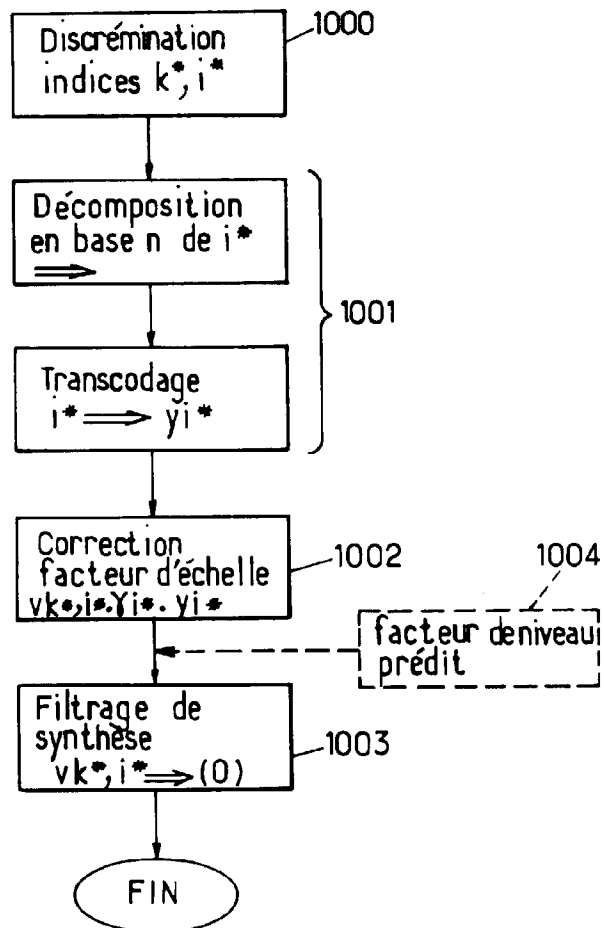


FIG. 7.



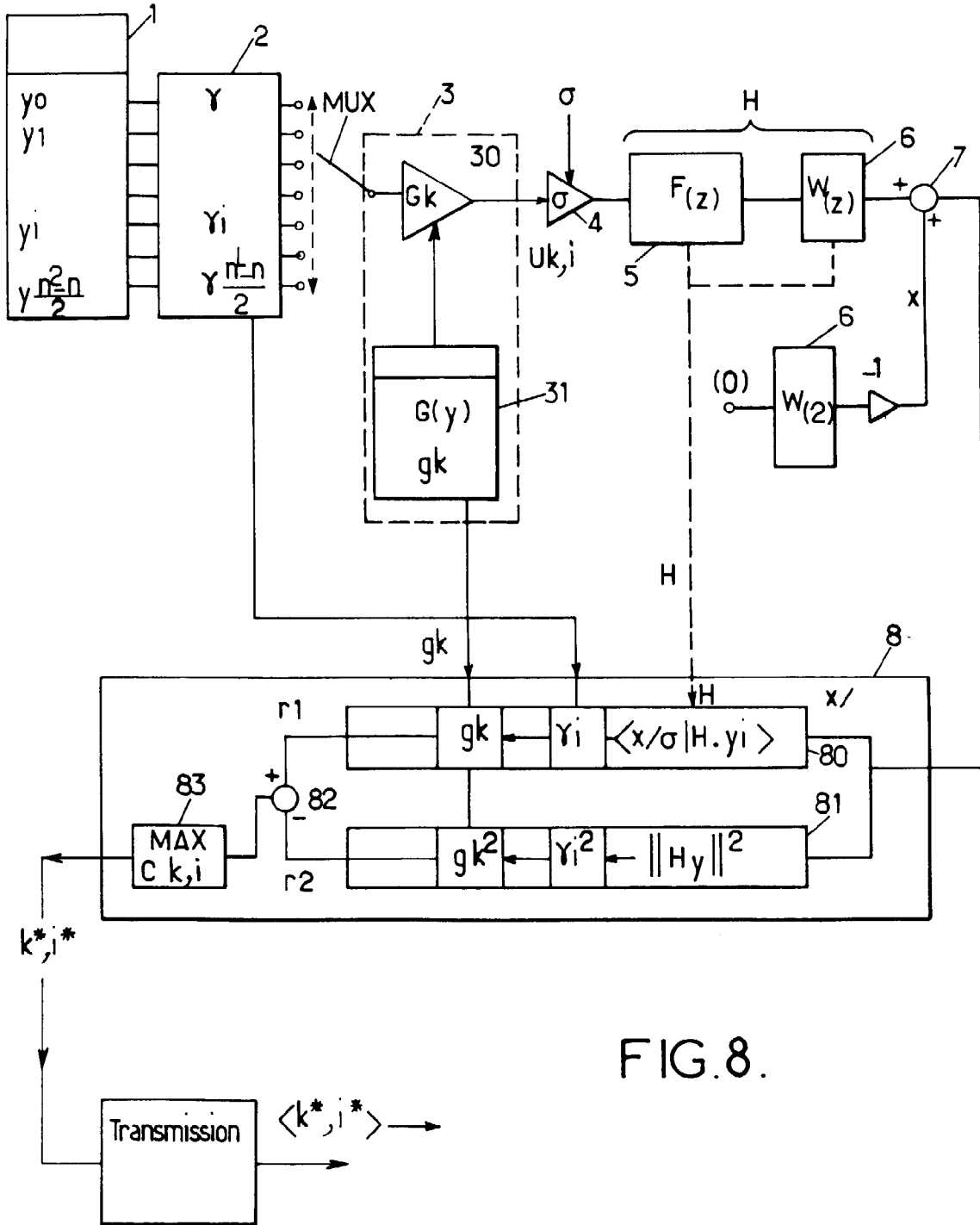


FIG. 8.