(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Numéro de publication : **0 542 585 A2**

(12) DEMANDE DE BREVET EUROPEEN

(21) Numéro de dépôt : **92402716.2**

(51) Int. Cl.^{5} : **G10L 9/14**

(22) Date de dépôt : **06.10.92**

(30) Priorité : **15.10.91 FR 9112669**

(43) Date de publication de la demande :
19.05.93 Bulletin 93/20

(84) Etats contractants désignés :
DE FR GB IT

(71) Demandeur : **THOMSON-CSF**
51, Esplanade du Général de Gaulle
F-92800 Puteaux (FR)

(72) Inventeur : **Laurent, Pierre-André,**
Thomson-CSF
SCPI, Cédex 67
F-92045 Paris la Défense (FR)

(74) Mandataire : **Lincot, Georges et al**
THOMSON-CSF, SCPI, B.P. 329, 50, rue
Jean-Pierre Timbaud
F-92402 Courbevoie Cédex (FR)

(54) Procédé de quantification d'un filtre prédicteur pour vocodeur à très faible débit.

(57) Le procédé consiste à partager le signal de parole en paquets d'un nombre déterminé de trames de durée constante en affectant (4) à chaque trame un poids fonction de la puissance moyenne du signal de parole dans la trame, à déterminer pour chaque trame les coefficients correspondants du filtre prédicteur en prenant ceux déjà déterminés (5) dans les trames voisines si son poids est similaire à au moins une des trames voisines ou en calculant ceux-ci isolement (6) ou par interpolation (7) à partir des coefficients des filtres voisins dans les autres cas.

Application : Vocodeurs.

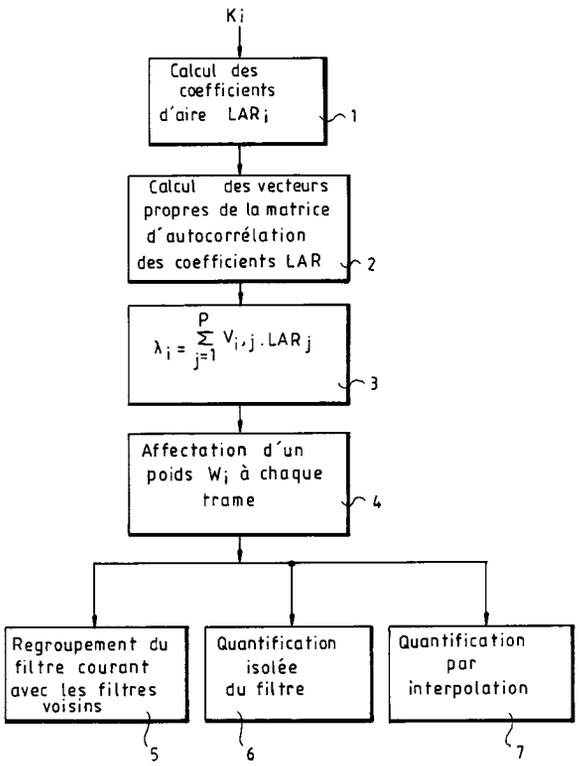

FIG. 1

EP 0 542 585 A2

La présente invention concerne un procédé de quantification d'un filtre prédicteur pour vocodeur à très faible débit.

Elle s'applique notamment aux vocodeurs à prédiction linéaire, similaires à ceux décrits par exemple dans la Revue Technique THOMSON-CSF, volume 14 n°3, septembre 1982, pages 715 à 731, suivant lesquels le signal de parole est identifié à la sortie d'un filtre numérique dont l'entrée reçoit soit une forme d'onde périodique correspondant à celles des sons voisés comme le sont les voyelles, soit une forme d'onde aléatoire correspondant à celles des sons non voisés comme le sont la plupart des consonnes.

Il est connu que la qualité auditive des vocodeurs à prédiction linéaire dépend en grande partie de la précision avec laquelle leur filtre prédicteur est quantifié et que cette qualité diminue généralement lorsque le débit numérique entre vocodeurs diminue car la précision de quantification du filtre devient alors insuffisante. D'une manière générale le signal de parole est segmenté en trames indépendantes de durées constantes et le filtre est renouvelé à chaque trame. Ainsi pour arriver à un débit d'environ 1820 bits par seconde, il faut, selon une réalisation standard normalisée, représenter le filtre par un paquet de 41 bits transmis toutes les 22,5 millisecondes. Pour des liaisons non standard à plus faible débit, de l'ordre de 800 bits par seconde, moins de 800 bits par seconde doivent être transmis pour représenter le filtre ce qui constitue approximativement un rapport de 3 en débit par rapport aux réalisations standards. Pour obtenir malgré tout une précision de quantification suffisante du filtre prédicteur l'approche classique consiste à mettre en oeuvre un schéma de quantification vectorielle intrinsèquement plus efficace que celui utilisé dans les systèmes standards où les 41 bits mis en oeuvre servent à quantifier scalairement les P = 10 coefficients de leur filtre de prédiction. La méthode repose sur l'utilisation d'un dictionnaire contenant un nombre déterminé de filtres standards obtenus par apprentissage. Elle consiste à transmettre uniquement la page ou l'index où se trouve le filtre standard le plus proche du filtre idéal. L'avantage est dans la réduction du débit binaire qui est obtenu, seulement 10 à 15 bits par filtre étant transmis au lieu des 41 bits nécessaires en mode de quantification scalaire, mais cette réduction de débit est obtenue au prix d'une très forte augmentation de la taille de mémoire nécessaire pour stocker le dictionnaire et d'une charge de calcul importante imputable à la complexité de l'algorithme de recherche des filtres dans le dictionnaire. Malheureusement le dictionnaire qui est ainsi créé n'est jamais universel et permet en fait de ne quantifier correctement que les filtres de prédiction qui sont proches de ceux de la base d'apprentissage. De ce fait, il apparaît que le dictionnaire ne peut à la fois avoir une taille raisonnable et permettre de quantifier correctement des filtres de prédiction résultant de l'analyse de parole pour tous les locuteurs, pour toutes les langues et dans toutes les conditions de prises de son.

Enfin les schémas de quantification standards fussent-ils de type vectoriel, cherchent avant tout à rendre minimum la distance spectrale entre le filtre d'origine et le filtre quantifié transmis et il n'est pas garanti que cette méthode soit la meilleure du fait des propriétés psycho-acoustiques de l'oreille qui ne peuvent être réduites simplement à celles connues des analyseurs de spectre.

Le but de l'invention est de pallier les inconvénients précités.

A cet effet l'invention a pour objet un procédé de quantification d'un filtre prédicteur pour vocodeur à très faible débit caractérisé en ce qu'il consiste à partager le signal de parole en paquets d'un nombre déterminé de trames de longueur de temps constante en affectant à chaque trame un poids fonction de la puissance moyenne du signal de parole dans la trame, à déterminer pour chaque trame les coefficients correspondants du filtre prédicteur en prenant ceux déjà déterminés dans les trames voisines si son poids est similaire à au moins une des trames voisines ou en calculant ceux-ci isolément ou par interpolation à partir des coefficients des filtres voisins dans les autres cas.

Le procédé selon l'invention a pour principal avantage qu'il ne demande pas d'apprentissage préalable pour former un dictionnaire et qu'il est de ce fait indifférent au type de locuteur, à la langue utilisée ou à la réponse en fréquence des parties analogiques du vocodeur. Il a également pour avantage de présenter, pour une complexité de réalisation raisonnable une qualité de restitution du signal de parole acceptable ne dépendant que de la qualité des algorithmes d'analyse de la parole utilisés.

D'autres caractéristiques et avantages de l'invention apparaîtront dans la description qui suit faite en regard des dessins annexés qui représentent :

- La figure 1 des premières étapes du procédé selon l'invention sous la forme d'un organigramme.
- La figure 2 un espace vectoriel à 2 dimensions figurant une répartition de coefficients d'aire dérivés des coefficients de réflexion modélisant le conduit vocal dans les vocodeurs.
- La figure 3 un exemple de regroupement des coefficients de filtre prédicteur suivant un nombre déterminé de trames du signal de parole permettant une simplification du processus de quantification des coefficients du filtre prédicteur des vocodeurs.
- La figure 4 un tableau illustrant le nombre de configurations possibles obtenues par regroupement de coefficients de filtres pour 1, 2 ou 3 trames et les configurations pour lesquelles les coefficients du filtre prédicteur pour une trame courante sont obtenus par interpolation.

- La figure 5 les dernières étapes du procédé selon l'invention sous la forme d'un organigramme.

Le procédé selon l'invention qui est représenté par l'organigramme de la figure 1 repose sur le principe qu'il n'est pas utile de transmettre les coefficients du filtre de prédiction trop souvent et qu'il faut plutôt adapter la transmission à ce que l'oreille peut percevoir. Suivant ce principe la cadence de renouvellement des coefficients du filtre est réduite pour transmettre les coefficients toutes les 30 millisecondes par exemple au lieu de toutes les 22,5 millisecondes comme cela est habituellement réalisé dans les solutions standards. D'autre part, le procédé selon l'invention tient compte du fait que le spectre du signal de parole est généralement corrélé d'une trame à l'autre en regroupant ensemble plusieurs trames avant tout codage. Dans les cas où le signal de parole est stable c'est-à-dire que son spectre en fréquence change peu au cours du temps ou dans le cas où le spectre en fréquence présente de fortes résonances une quantification fine est effectuée, par contre si ce signal est instable ou peu résonant la quantification effectuée est plus fréquente mais plus grossière, car l'oreille dans ce cas ne perçoit pas de différence. Enfin pour représenter le filtre de prédiction la base de coefficients utilisés contient un jeu de p coefficients facile à quantifier par une quantification scalaire efficace.

Comme dans les procédés standards le filtre de prédiction est représenté sous la forme d'un jeu de p coefficients obtenus à partir du signal de parole échantillonné original éventuellement préaccentué. Ces coefficients sont les coefficients de réflexion notés K_i qui modélisent au mieux le conduit vocal. Leur valeur absolue est choisie inférieure à l'unité pour que la condition de stabilité du filtre de prédiction soit toujours respectée. Lorsque ces coefficients ont une valeur absolue proche de 1 ceux-ci sont quantifiés finement pour tenir compte du fait que la réponse en fréquence du filtre devient alors très sensible à la moindre erreur. Comme représenté par les étapes 1 à 7 sur l'organigramme de la figure 1, le procédé consiste d'abord par distordre à l'étape 1 de façon non linéaire les coefficients de réflexion en les transformant en coefficients d'aire notés LAR_i de l'abréviation anglo-saxonne LOG AREA RATIO, par la relation :

$$LAR_i = cste.\log\left(\frac{1+K_i}{1-K_i}\right) \quad i = 1...P \qquad (1)$$

L'avantage d'utiliser les coefficients LAR est qu'ils sont plus commodes à traiter que les coefficients K_i puisque leur valeur est toujours comprise entre $-\infty$ et $+\infty$ et qu'en les quantifiant de façon linéaire les mêmes résultats peuvent être obtenus qu'en utilisant une quantification non linéaire des coefficients K_i. D'autre part, l'analyse en composantes principales du nuage de points ayant les coefficients LAR_i comme coordonnées dans un espace à P dimensions montre, comme représenté de façon simplifiée dans l'espace à deux dimensions de la figure 2, des directions privilégiées dont il est tenu compte dans la quantification pour la rendre plus efficace. Ainsi si V_1, V_2 etc V_P sont des vecteurs propres de la matrice d'autocorrélation des coefficients LAR, une quantification efficace est obtenue en considérant les projections des jeux des coefficients LAR sur les vecteurs propres. Selon ce principe la quantification a lieu aux étapes 2 et 3 sur des quantités λ_i telles que :

$$\lambda_i = \sum_{j=1}^{P} V_{i,j}\, LAR_j \quad i = 1...P \qquad (2)$$

Pour chacun des λ_i il est alors effectué une quantification uniforme entre une valeur minimale λ_{imini} et une valeur maximale λ_{imax} avec un nombre de bits n_i qui est calculé par les moyens classiques en fonction du nombre total N de bits utilisés pour quantifier le filtre et les pourcentages d'inertie correspondant aux vecteurs propres V_i.

Pour tirer profit de la non indépendance des spectres en fréquence d'une trame à la suivante, un nombre déterminé de trames sont regroupées avant quantification, et pour apporter davantage de soin à la quantification du filtre dans les trames qui sont les plus perçues par l'oreille, chaque trame est affectée à l'étape 4 d'un poids W_t (t étant compris entre 1 et L), qui est une fonction croissante de l'importance acoustique de chaque trame t considérée. La règle de pondération tient compte, du niveau sonore de la trame concernée car plus le niveau sonore d'une trame est élevé par rapport aux trames avoisinantes et plus celle-ci attire l'attention, ainsi que de l'état résonant ou non des filtres, seuls les filtres résonants étant convenablement quantifiés.

Une bonne mesure du poids W_t de chaque trame est obtenue en appliquant la relation

$$W_t = F\left(\frac{P_t}{\prod_{i=1}^{P}(1 - K_{t,i}^2)}\right) \qquad (3)$$

Dans la relation (3), P_t désigne la puissance moyenne du signal de parole dans chaque trame d'indice t et $K_{t,i}$ désigne les coefficients de réflexion du filtre prédicteur correspondant. Le dénominateur de l'expression précédente entre parenthèses représente l'inverse du gain du filtre de prédiction, ce gain étant élevé lorsque le filtre est résonant. La fonction F est une fonction monotone croissante incorporant un mécanisme de régulation pour à éviter que certaines trames aient un poids trop faible ou trop élevé par rapport à leurs voisines. Ainsi par exemple une règle de détermination des poids W_t peut être d'adopter pour la trame d'indice t que la quantité F soit supérieure à deux fois le poids W_{t-1} de la trame t-1, dans ce cas le poids W_t est déterminé pour être égal à deux fois le poids W_{t-1}. Par contre, si pour la trame d'indice t la quantité F est inférieure à une demi fois la quantité W_{t-1} du poids de la trame t-1 le poids W_t peut être pris égal à la moitié du poids W_{t-1}. Enfin pour les autres cas, le poids W_t peut être rendu égal à F.

Compte-tenu du fait de la quantification directe des L filtres d'un paquet de trames courant ne peut être envisagée du fait qu'elle conduirait à quantifier chaque filtre avec un nombre de bits largement insuffisant pour obtenir une qualité acceptable et du fait que les filtres de prédiction des trames voisines ne sont pas indépendants, il est considéré aux étapes 5, 6 et 7, que pour un filtre donné trois cas peuvent se présenter suivant que le signal dans la trame est auditivement très important et que le filtre courant peut être regroupé avec son ou ses filtres voisins, que l'ensemble peut être quantifié en une seule fois ou enfin que le filtre courant peut être approximé par interpolation à partir des filtres voisins.

L'ensemble de ces règles conduit par exemple, pour un nombre de filtres L=6 d'un bloc de trames à ne quantifier que trois filtres s'il est possible de regrouper trois filtres avant quantification ce qui conduit à prendre deux schémas de quantifications possibles. Un exemple de regroupement est représenté à la figure 3. Pour l'ensemble des six trames représentées il apparaît que les trames 1 et 2 sont regroupées et quantifiées ensemble, que les filtres des trames 4 et 6 sont quantifiés isolément et que les filtres des trames 3 et 5 sont obtenus par interpolation. Sur ce schéma les rectangles grisés représentent les filtres quantifiés, les cercles représentent les filtres vrais et les tiretés les interpolations. Le nombre de configurations possibles est représenté par le tableau de la figure 4. Sur ce tableau les chiffres 1, 2 ou 3 placés dans la colonne configuration indiquent des groupements respectifs de 1, 2 ou 3 filtres successifs et le chiffre 0 indique que le filtre courant est obtenu par interpolation.

Cette répartition permet d'affecter au mieux le nombre de bits nécessaires à appliquer à chaque filtre effectivement quantifié. Par exemple, dans le cas où seulement n=84 bits de quantification des filtres sont disponibles dans un paquet de six trames correspondant à 14 bits en moyenne par trame et si n_1, n_2 et n_3 désignent les nombres de bits alloués aux trois filtres quantifiés, ces nombres peuvent être choisis parmi les valeurs 24, 28, 32 et 36 de façon que leur somme soit égale à 84 ce qui donne dix possibilités au total. La manière de choisir les nombres n_1, n_2 et n_3 est alors considérée comme un sous schéma de quantification pour reprendre l'exemple ci-dessus de la figure 3. L'application des règles précédentes conduit par exemple à regrouper et quantifier les filtres 1 et 2 ensemble sur n_1 = 28 bits, à quantifier isolément respectivement sur n_2 = 32 et n_3 = 24 bits les filtres 4 et 6 et à obtenir les filtres 3 et 5 par interpolation.

De façon à obtenir la meilleure quantification pour l'ensemble des six filtres sachant qu'il y a 32 schémas de base, chacun offrant dix sous schémas correspondant à 320 possibilités sans explorer exhaustivement chacune des possibilités offertes le choix a lieu en appliquant les méthodes connues de calcul de distance entre filtres et en calculant pour chaque filtre quantifié l'erreur de quantification et l'erreur d'interpolation. Sachant que les coefficients λ_i sont quantifiés de façon simple la distance entre filtres peut être mesurée selon l'invention par le calcul d'une distance euclidienne pondérée de la forme

$$D(F_1, F_2) = \sum_{i=)1}^{P} \gamma i\left(\lambda_{1,i} - \lambda_{2,i}\right)^2 \qquad (4)$$

où les coefficients γ_i sont fonctions simples des pourcentages d'inertie associés aux vecteurs propres V_i et F_1 et F_2 sont les deux filtres dont on mesure la distance. Ainsi pour remplacer les filtres de trame T_{t+1}, etc, T_{t+k-1}

par un filtre unique il suffit de rendre minimum l'erreur totale en utilisant un filtre dont les coefficients sont donnés par la relation

$$\lambda_j = \frac{\sum_{i=0}^{k-1} W_{t+i}\,\lambda_{t+i,j}}{\sum_{i=0}^{k-1} W_{t+i}}, \quad j = 1 \ldots P \qquad (5)$$

où $\lambda_{t+i,j}$ représente le $j^{ème}$ coefficient du filtre de prédiction de la trame t+i. Le poids à affecter au filtre est alors simplement la somme des poids des filtres originaux qu'il approxime. L'erreur de quantification est alors obtenue en appliquant la relation

$$E(N_j) = \frac{1}{12}\sum_{i=1}^{P} \gamma_i \left(\frac{\lambda_{i\,max} - \lambda_{i\,min}}{2^{n_{i,j}}} \right)^2 \quad \text{avec} \sum_{i=1}^{P} n_{j,i} = N_j \qquad (6)$$

Comme il n'existe qu'un nombre fini de valeurs de N_j les quantités E_{Nj} sont calculées de préférence une fois pour toutes et qui permet de les stocker par exemple dans une mémoire morte. De la sorte la contribution d'un filtre donné de rang t à l'erreur totale de quantification est obtenue en tenant compte de trois facteurs qui sont, le poids W_t qui joue en tant que facteur multiplicatif, l'erreur déterministe commise éventuellement en le remplaçant par un filtre moyen partagé avec son ou ses voisins, et l'erreur de quantification théorique E_{Ng} calculée précédemment qui dépend du nombre de bits de quantification utilisés. Ainsi si F est le filtre par lequel le filtre F_t de la trame t est remplacé, la contribution du filtre de la trame t à l'erreur totale de quantification vérifie une expression de la forme :

$$E_t = W_t\{E(N_j) + D(F,F_t)\} \quad (7)$$

Les coefficients λ_i des filtres interpolés entre deux filtres F_1 et F_2 sont obtenus en effectuant la somme pondérée des coefficients de même rang des filtres F_1 et F_2 suivant une relation de la forme :

$$\lambda_i = \alpha\lambda_{1,i} + (1 + \alpha)\lambda_{2,i} \text{ pour } i = 1 \quad (8)$$

Par voie de conséquence, l'erreur de quantification associée à ces filtres est, en omettant les coefficients W_t qui leur sont associés, la somme de l'erreur d'interpolation, c'est-à-dire de la distance entre chaque filtre interpolé et le filtre de la trame T, $D(F_I,F_t)$ et de la somme pondérée des erreurs de quantification des deux filtres F_1 et F_2, à partir desquels est faite l'interpolation, à savoir :

$$\alpha^2 E(N_1) + (1 - \alpha)^2 E(N_2) \quad (9)$$

si les deux filtres sont quantifiés avec respectivement N_1 et N_2 bits.

Cette façon de calculer permet d'obtenir l'erreur totale de quantification à partir des seuls filtres quantifiés en faisant pour chaque filtre K quantifié la somme de l'erreur de quantification due à l'utilisation de N_k bits pondérés par le poids du filtre K, (ce poids pouvant être la somme de poids des filtres qu'il moyenne si c'est le cas), de l'erreur de quantification induite sur le ou les filtres qu'il sert à interpoler pondérés par une fonction du ou des coefficients α et le ou les poids du ou des filtres en question et de l'erreur déterministe faite délibérément en remplaçant certains filtres par leur moyenne pondérée et en interpolant d'autres.

A titre d'exemple, en retournant au groupement de la figure 3, un schéma de quantification correspondant peut être obtenu en quantifiant :

- les filtres F_1 et F_2 regroupés sur N_1 bits en considérant un filtre moyen F défini symboliquement par la relation :

$$F = (W_1F_1 + W_2F_2)/(W_1 + W_2) \quad (10)$$

- le filtre F_4 sur N_2 bits
- le filtre F_6 sur N_3 bits

et les filtres F_3 et F_5 par interpolation.

L'erreur déterministe qui est indépendante des quantifications est alors la somme des termes :

- $W_1\,D(F,F_1)$: distance pondérée entre F et F_1
- $W_2\,D(F,F_2)$: distance pondérée entre F et F_2
- $W_3\,D(F_3, (1/2\,F + 1/2\,F_4))$ pour le filtre 3, interpolé
- $W_5\,D(F_5, (1/2\,F + 1/2\,F_6))$ pour le filtre 4, interpolé

- 0 pour le filtre 4, quantifié directement
- 0 pour le filtre 6, quantifié directement.

L'erreur de quantification est, quant à elle, la somme des termes :

- $(W_1 + W_2)\ E(N_1)$ pour le filtre composite moyen F
- $W_4\ E(N_2)$ pour le filtre 4, quantifié tel quel sur N_2 bits
- $W_6\ E(N_3)$ pour le filtre 6, quantifié tel quel sur N_3 bits
- $W_3\ (1/4\ E(N_1) + 1/4\ E(N_2)$ pour le filtre 3, obtenu par interpolation
- $W_5\ (1/4\ E(N_1) + 1/4\ E(N_3)$ pour le filtre 5, obtenu par interpolation

soit encore la somme des termes :

- $E(N_1)$ pondéré par un poids $\omega_1 = W_1 + W_2 + 1/4\ W_3$
- $E(N_2)$ pondéré par $\omega_2 = 1/4\ W_3 + W_4 + 1/4 W_5$
- $E(N_3)$ pondéré par $\omega_3 = 1/4\ W_5 + W_6$

L'algorithme de quantification complet qui est représenté à la figure 5 comporte trois passes conçues de telle sorte qu'à chaque passe seuls les schémas de quantification les plus probables soient conservés.

- La première passe représentée en 8 sur la figure 5 s'effectue au fur et à mesure que les trames de parole arrivent. Elle consiste, dans chaque trame à effectuer tous les calculs d'erreurs déterministes faisables dans la trame t et à modifier en conséquence l'erreur totale à affecter à tous les schémas de quantification concernés. Par exemple, pour la trame 3 de la figure 3, les deux filtres moyens seront calculés en groupant les trames 1, 2 et 3 ou 2 et 3 qui se terminent dans la trame 3, ainsi que les erreurs correspondantes ; puis l'erreur d'interpolation est calculée pour tous les schémas de quantification où la trame 2 est obtenue par interpolation à partir des trames 1 et 3.

A la fin de la trame L, toutes les erreurs déterministes obtenues sont affectées aux différents schémas de quantification.

Une pile peut alors être crée qui ne contient plus que les schémas de quantification donnant les plus faibles erreurs et qui seuls sont susceptibles de donner de bons résultats. Typiquement, il peut être retenu environ le tiers des schémas de quantification d'origine.

La deuxième passe qui est représentée en 9 sur la figurez 5 vise à sélectionner les sous-schémas de quantification (répartitions des nombres de bits alloués aux différents filtres à quantifier) qui donnent les meilleurs résultats, pour les seuls schémas de quantification retenus. Cette sélection passe par le calcul de poids fictifs ω_i qui sont calculés pour les seuls filtres qui sont à quantifier (filtres éventuellement composites) en tenant compte des filtres voisins obtenus par interpolation. Une fois ces poids fictifs calculés, une deuxième pile de taille plus réduite est créée qui ne contient plus que les paires (schéma de quantification + sous-schémas), pour lesquelles la somme de l'erreur déterministe et de l'erreur de quantification (pondérée par les poids fictifs) est minimale.

Enfin la dernière phase qui est représentée en 10 sur la figure 5 consiste à effectuer la quantification complète selon les seuls schémas (+ sous-schémas) finalement sélectionnés dans la deuxième pile, et, naturellement, à retenir celui qui donnera l'erreur totale minimale.

Afin d'obtenir la meilleure quantification possible, il est encore possible d'envisager, si la charge de calcul le permet, d'utiliser une mesure de distance plus élaborée, à savoir celle connue d'Itakura-Saito qui est une mesure de la distorsion spectrale totale, ou, autrement dit, de l'"erreur de prédiction". Dans ce cas en désignant par $R_{t0}, R_{t1}, \ldots, R_{tP}$ les P + 1 premiers coefficients d'autocorrélation du signal dans une trame t, donnés par :

$$R_{t,k} = \frac{1}{N} \sum_{n=n_0}^{n=n_0+N-1} S_n\ S_{n-k} \qquad (11)$$

où N est la durée d'analyse utilisée dans la trame t, et n_0 la première position d'analyse du signal S échantillonné. Le filtre prédicteur est alors entièrement décrit par une transformée en z, $P(_z)$ telle que :

$$P(z) = \frac{1}{\sum_{i=0}^{P} a_i\ z^{-i}} \quad \text{avec } a_0 = 1 \qquad (12)$$

dans laquelle les coefficients a_i sont calculés itérativement à partir des coefficients de réflexion K_i déduits des coefficients LAR, eux-mêmes déduits des coefficients λ en inversant les relations (1) et (2) décrites

précédemment.

A l'initialisation des calculs

$a^0_0 = 1$ et $a^0_1 = 1$

et à l'itération p(p = 1...P) : les coefficients a_i sont définis par

$a^p_i = a^{p-1}_i + K_p\, a^{p-1}_{p-i}$ pour i = 0...p et $a^p_{p+1} = 0$

L'erreur de prédiction vérifie alors la relation :

$$E_t = R_{t0}\, B_{t,0} + 2 \sum_{i=1}^{P} R_{t,i}\, B_{t,i} \qquad (13)$$

avec

$$B_{t,i} = \sum_{i=0}^{P-i} \tilde{a}_j \tilde{a}_{j+1} \qquad (14)$$

Dans le relations (13) et (14), le signe "~" signifie que les valeurs sont obtenues à partir des coefficients λ quantifiés. Par définition, cette erreur est minimale s'il n'y a pas quantification car les K_i sont justement calculés pour qu'elle le soit.

L'intérêt de procéder ainsi est que l'ensemble de l'algorithme de quantification obtenu est peu coûteux en termes de puissance de calcul puisqu'en fin de compte, en retournant à l'exemple de la figure 3, sur les 320 possibilités de codage, seules quatre ou cinq possibilités sont sélectionnées et examinées en détail. Ceci permet de conserver des algorithmes d'analyse performants, ce qui est essentiel pour un vocodeur.

Revendications

1. Procédé de quantification d'un filtre prédicteur pour vocodeur à très faible débit caractérisé en ce qu'il consiste à partager le signal de parole en paquets d'un nombre déterminé de trames de durée constante en affectant (4) à chaque trame un poids fonction de la puissance moyenne du signal de parole dans la trame, à déterminer pour chaque trame les coefficients correspondants du filtre prédicteur en prenant ceux déjà déterminés (5) dans les trames voisines si son poids est similaire à au moins une des trames voisines ou en calculant ceux-ci isolément (6) ou par interpolation (7) à partir des coefficients des filtres voisins dans les autres cas.

2. Procédé selon la revendication 1, caractérisé en ce qu'il consiste à quantifier dans chaque paquet de trames le filtre avec des nombres de bits différents en fonction des regroupements entre trames réalisés pour calculer les coefficients du filtre, en gardant constant la somme du nombre de bits de quantification disponible dans chaque paquet.

3. Procédé selon la revendication 2, caractérisé en ce que le nombre de bits de quantification du filtre dans chaque trame est déterminé en effectuant une mesure de distance entre filtres pour ne quantifier le filtre qu'avec les coefficients donnant une erreur totale de quantification minimum.

4. Procédé selon la revendication 3, caractérisé en ce que la mesure de distance est de type euclidien.

5. Procédé selon la revendication 3, caractérisé en ce que la mesure de distance est celle d'ITKURA-SAITO.

6. Procédé selon la revendication 2, caractérisé en ce qu'il consiste à sélectionner (8) dans chaque trame un nombre déterminé de sous-schémas de quantification d'erreurs les plus faibles, à calculer (9) dans chaque sous-schéma sélectionné un poids de trame fictif en tenant compte des filtres voisins pour ne retenir (10) que le sous-schéma dont l'erreur de quantification pondérée par le poids fictif est minimum.

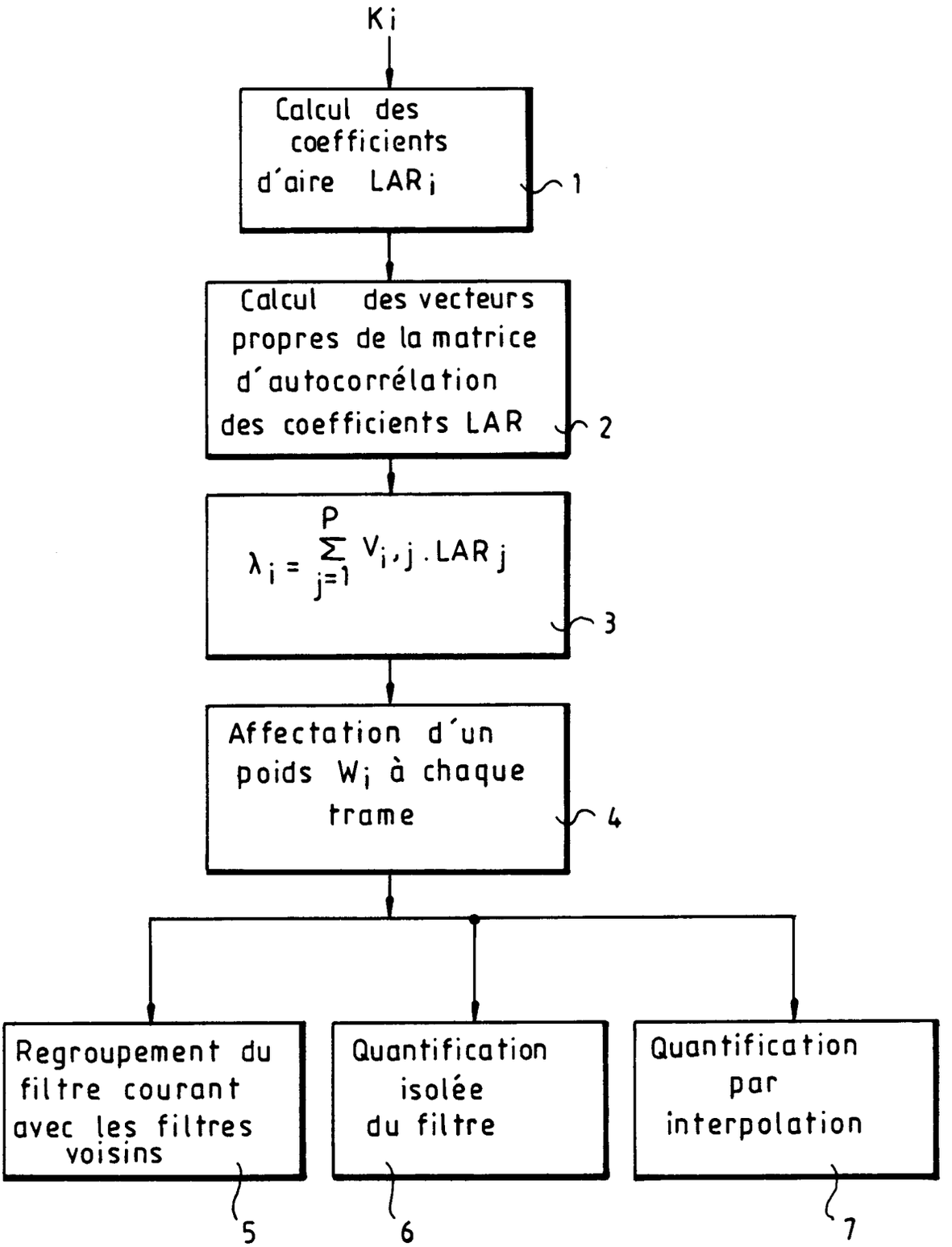
Ki
Calcul des coefficients d'aire LARi
1
Calcul des vecteurs propres de la matrice d'autocorrélation des coefficients LAR
2
$\lambda_i = \sum_{j=1}^{P} V_{i,j} . LAR_j$
3
Affectation d'un poids Wi à chaque trame
4
Regroupement du filtre courant avec les filtres voisins
5
Quantification isolée du filtre
6
Quantification par interpolation
7

FIG. 1

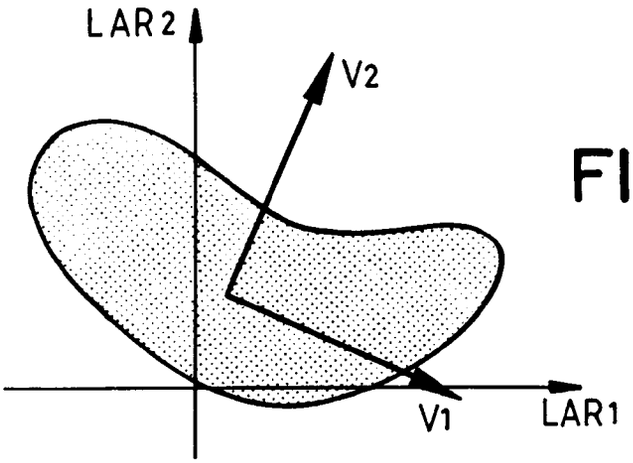

FIG. 2

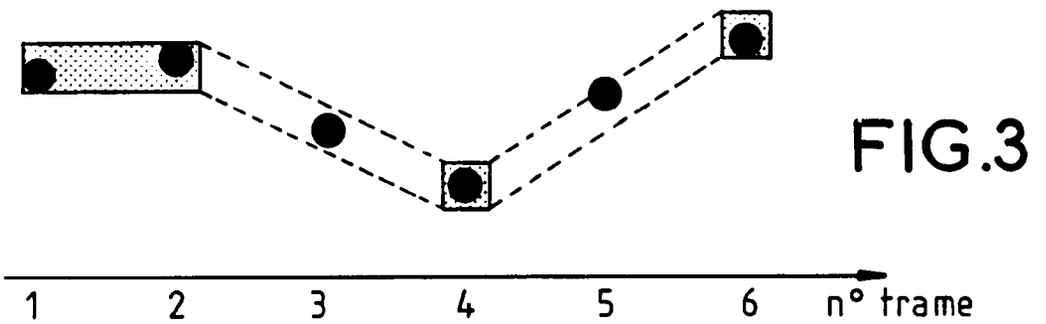

FIG.3

N°	CONFIGURATION	N°	CONFIGURATION
1:	3,2,1	17:	1,3,0,1
2:	3,1,2	18:	2,2,0,1
3:	1,3,2	19:	3,1,0,1
4:	2,3,1	20:	1,0,0,2,1
5:	2,1,3	21:	1,0,0,1,2
6:	1,2,3	22:	1,0,1,0,2
7:	2,2,2	23:	1,1,0,0,2
8:	1,0,3,1	24:	1,0,2,0,1
9:	1,0,2,2	25:	2,0,0,1,1
10:	1,0,1,3	26:	2,0,1,0,1
11:	2,0,2,1	27:	2,1,0,0,1
12:	2,0,1,2	28:	1,2,0,0,1
13:	1,1,0,3	29:	1,0,0,0,1,1
14:	1,2,0,2	30:	1,1,0,0,0,1
15:	2,1,0,2	31:	1,0,1,0,0,1
16:	3,0,1,1	32:	1,0,0,1,0,1

FIG.4

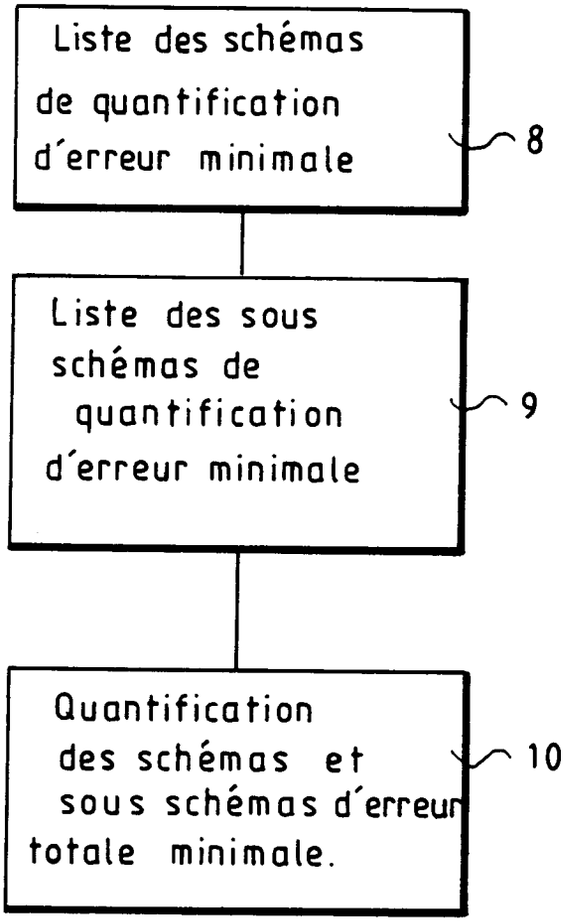
Liste des schémas de quantification d'erreur minimale
8
Liste des sous schémas de quantification d'erreur minimale
9
Quantification des schémas et sous schémas d'erreur totale minimale.
10

FIG.5