

CONCEPTION D'UN SYSTÈME D'ÉNUMÉRATION EN TREILLIS POUR LE QUANTIFICATEUR SCALAIRE DU FS1016

Bachir Boudraa*, Malika Boudraa*, Mouloud Djamah*, Aissa Boukhari*, et Bernard Guerin**

* Université des sciences et de la technologie Houari Boumédiène (USTHB),

Laboratoire de communication parlée et de traitement du signal, Faculté d'Electronique et d'Informatique,
BP 32, El alia, Alger, ALGERIE. E-mail : bboudraa@usthb.dz, mboudraa@usthb.dz

** Institut de la communication parlée, INPG, 46 avenue Félix Viallet, 38031, Grenoble, France

RÉSUMÉ

Le but recherché par cette étude est la diminution du débit observé dans le codeur de parole nommé CELP FS1016 (standard développé par le département de la défense des Etats Unis d'Amérique "DoD"). Plus précisément, on s'est intéressé à la quantification des coefficients LSF (Line Spectrum Frequencies). Dans le standard CELP FS1016, ces coefficients sont calculés pour chaque trame de 30 ms et sont codés sur 34 bits après avoir subi une quantification scalaire. Nous avons jugé ce débit élevé et nous proposons d'appliquer une méthode d'énumération en treillis visant la diminution de celui-ci. Pour cela, nous avons d'abord effectué une analyse fine des différentes combinaisons données par les tables de la quantification initialement employée. L'analyse a révélé que cette quantification n'était pas optimale. En effet, nous avons trouvé qu'il y avait un nombre important de vecteurs LSF qui ne seront jamais utilisés car rendant le filtre de synthèse instable. Nous avons donc exploité l'ordonnement naturel des coefficients LSF pour appliquer un système d'énumération en treillis pour ce quantificateur scalaire sans changer ses tables. On est parvenu ainsi à réduire le nombre de bits alloués aux coefficients LSF de 34 bits à 30 bits. Ainsi, une distorsion spectrale équivalente a été obtenue en utilisant soit le quantificateur classique ou en appliquant le système d'énumération en treillis que nous proposons. Cependant, ce dernier a l'avantage de permettre un gain en débit de 4 bits/trame.

ABSTRACT

In the present study, we were interested in the reduction of the bit-rate observed in the speech coder named CELP FS1016 (federal standard developed by the US department of the defense "DoD"). More precisely, the quantization of the Line Spectrum Frequencies (LSF) parameters was concerned. In the standard CELP FS1016, these coefficients are derived from the input speech signal through linear prediction analysis of each 30-ms frame. A direct scalar quantization at 34 bits per frame is used. We considered the bit rate too high and we propose the use of an enumeration technique in conjunction with a treillis search coding schemes for speech LSF parameters, to reduce this rate to 30 bits per frame. To this end, we first did a thorough analysis of the different possible combinations given by the tables of the scalar quantization initially used in the FS1016. The analysis revealed that this quantization was not optimum. Indeed, we found that many combinations of quantization levels can not be used because they can lead to an unstable synthesis filter. Then, we exploit the natural ordering of the LSF to design an enumeration algorithm with a treillis schemes that reduces the bit rate of the LSF coefficients from 34 to 30 bits without decreasing the performance of the coder (an equivalent spectral distortion is obtained with the two quantizers).

1. INTRODUCTION

La prédiction linéaire est aujourd'hui largement utilisée dans beaucoup d'applications en traitement du signal de la parole, notamment pour représenter l'enveloppe spectrale court-terme de ce signal. C'est le cas du codage bas débit en bande étroite où on modélise souvent cette enveloppe par un filtre tout-pôle d'ordre 10. Les dix coefficients de ce filtre sont estimés à partir d'une trame du signal ayant typiquement une longueur de 10 à 30 ms [1]-[5]. Des algorithmes à base de méthodes d'autocorrélation et de covariance [2]-[4] sont souvent utilisées pour calculer ces coefficients, appelés

aussi coefficients de prédiction ou coefficients LPC (Linear predictive coding) où la complexité algorithmique et la stabilité du filtre sont le souci majeur [6], [8]. Notons que les causes de l'instabilité sont multiples. Ils proviennent souvent du caractère fortement non stationnaire du signal. En effet, la présence de longues périodes de silence ainsi que l'existence de certains sons voisés très pauvres en harmoniques et fortement prédictibles, conditionnent mal les algorithmes d'estimation (les variances des erreurs de prédiction devenant faibles devant l'énergie moyenne du signal). De plus, les régions de transition entre les sons sont aussi d'autres causes d'instabilité du fait de l'accumulation des erreurs numériques et de leur propagation lors du traitement de ces régions [8].

Par ailleurs, en codage bas débit, il est important de quantifier les coefficients LPC en utilisant le minimum de bits possibles sans introduire une distorsion spectrale excessive et tout en gardant une complexité raisonnable. Cependant, dans leur forme directe, ces coefficients s'adaptent mal à une quantification à cause de leur large dynamique. Il est alors nécessaire de les transformer vers des représentations plus adéquates n'exigeant que 18 à 34 bits/trame. On préfère utiliser les coefficients de réflexion (RC) qui sont moins sensibles à la quantification. Ceux-ci sont bornés par l'unité et la stabilité du filtre est mieux assurée en gardant cette borne lors de l'opération de quantification. Des algorithmes de passage des coefficients LPC aux coefficients RC sont alors utilisés. Cependant, au voisinage de l'unité, la sensibilité spectrale des coefficients RC reste gênante et on tend à étendre cette région au moyen de transformations non linéaires. Le rapport d'aires ou log-area ratio (LAR), le sinus inverse (IS) et les paires de raies spectrales, appelées communément coefficients LSF (Line Spectrum Frequencies), sont souvent préférées [4], [9]. Des quantifications non uniformes sont ensuite appliquées et 34 bits/trames sont retenus pour coder l'enveloppe spectrale [7], [9], [10].

Actuellement, la représentation des coefficients LPC par les paires des raies spectrales est très utilisée [11]-[14], car elle possède des propriétés naturelles désirables pour une quantification, comme nous allons le préciser dans la section suivante. Plusieurs schémas de quantification des LSF, aussi bien scalaires (QS) que vectorielles (QV), sont rencontrés dans la littérature [11-15]. Dans ce travail, nous avons opté pour la QS à cause de la complexité moindre qu'elle présente comparativement à la QV, ce qui était une exigence de départ¹. D'où notre intérêt pour le codeur CELP (Code Excited Linear Prediction) de la norme FS1016 [16], [17], standard développé par le département de la défense des Etats Unis d'Amérique "DoD", qui répondait a priori à notre exigence. Ce dernier utilise 10 coefficients LSF pour représenter l'enveloppe spectrale. Ceux-ci sont quantifiés de façon scalaire et emploient 34 bits/trame, répartis respectivement comme suit : 3, 4, 4, 4, 4, 3, 3, 3, 3, 3 (Tableau 1).

Par la suite, une fine analyse de ce tableau révélera que ce codeur utilise une numérotation dont certaines valeurs ne peuvent être retenues car introduisant des combinaisons de coefficients LSF qui rendraient le filtre de synthèse instable.

En effet, les 34 bits précédents peuvent donner $N = \prod_{i=1}^{10} 2^{b_i}$, soit 17179869184 combinaisons possibles où b_i serait donc le nombre de bits accordé pour pouvoir indexer tous les niveaux de quantification de la $i^{ème}$ table. Cependant, l'ordonnement naturel des LSF va nous permettre de

¹ Ce travail entre dans le cadre du projet CNEPRU J1602/02/09/04 incluant la mise au point d'un codeur bas débit fonctionnant sur DSP (TMS 320C25 de Texas Instrument) au voisinage de 4 kbps pour une transmission vocale sécurisée.

réduire ce nombre à seulement un maximum de 554958388 (10000100010011111111000110100 en binaire) combinaisons acceptables (soit environ 3.25% du nombre total). Ce chiffre n'exige que 30 bits pour sa représentation.

En d'autres termes, 96.75% combinaisons conduiront à la réalisation d'un filtre de synthèse instable. Nous avons pu éviter ces combinaisons et réduire ainsi le débit par l'application d'un système d'énumération en treillis.

La quantification par treillis a eu cette appellation suite à son schéma utilisant des chemins sous forme de treillis. Ces chemins passent d'un niveau de reproduction à un autre se trouvant dans des tables différentes.

Plusieurs travaux récents [18]-[21] se sont intéressés à l'application de ce type de quantification aux coefficients LSF. Généralement, la quantification par treillis utilise l'algorithme de Viterbi [22] pour optimiser le parcours à travers les niveaux de reproduction ou mot-codes. Une version de cet algorithme a été proposée dans les travaux de Lahouti [20].

Dans notre cas, on se propose d'appliquer un système d'énumération en treillis sur des tables formées de coefficients LSF et utilisées dans le codeur CELP; ceci en vue de réduire le nombre bits alloués à l'ensemble de ces coefficients tout en gardant une complexité raisonnable. L'algorithme utilisé est une adaptation de l'algorithme de Malone et Fisher [21]. 30 bits seulement (au lieu de 34 bits) ont alors été jugés suffisants pour coder les paramètres LSF de ce codeur tout en gardant la même qualité décodée. 4 bits/trame ont ainsi été récupérés ce qui équivaut à un gain de 133 bits/sec.

Dans la seconde section de cet article on définira les coefficients LSF et on donnera leurs propriétés et leurs avantages par rapport aux coefficients LPC. Dans la section 3 on présentera la quantification scalaire utilisée dans le codeur FS1016. La section 4 décrit le système d'énumération par treillis utilisé et présente les algorithmes de codage et de décodage mis au point. Les résultats obtenus sont discutés dans la section 4.

2. DEFINITION ET PROPRIETES DES PARAMETRES LSF

2.1 Définition

Les coefficients LSF sont une autre représentation des coefficients de prédiction a_k . Ils dérivent du filtre d'analyse par prédiction linéaire $A(z)$ d'ordre p donné par l'équation suivante [1]-[5]:

$$A(z) = 1 + \sum_{k=1}^p a_k z^{-k} \quad (1)$$

La représentation LSF a été introduite pour la première fois par Itakura [11]. Elle est devenue par la suite largement utilisée, notamment dans le codage de la parole où elle est préférée pour effectuer la quantification de

l'information spectrale, comme on peut le constater dans [12]-[17]. En effet, la large dynamique des valeurs des coefficients a_k cause une grande difficulté dans la détermination des régions de leur quantification. Cela provoque des erreurs pouvant conduire à une instabilité du filtre de synthèse $1/A(z)$. Par contre, les coefficients LSF possèdent une gamme de valeurs bornée entre 0 et 0.5, ce qui rend leur quantification plus aisée.

Deux polynômes $P(z)$ et $Q(z)$ sont définis respectivement par l'addition et la soustraction du terme $z^{-(p+1)}A(z)$ [23], [24]:

$$\begin{aligned} P(z) &= A(z) + z^{-(p+1)}A(z) \\ Q(z) &= A(z) - z^{-(p+1)}A(z) \end{aligned} \quad (2)$$

Les deux polynômes $P(z)$ et $Q(z)$ sont, respectivement, symétrique et antisymétrique et ont toutes leurs racines sur le cercle unité. Deux autres polynômes $P'(z)$ et $Q'(z)$ sont formés tels que :

$$\begin{aligned} P'(z) &= (1+z^{-1})P(z) \\ Q'(z) &= (1-z^{-1})Q(z) \end{aligned} \quad (3)$$

Les polynômes $P'(z)$ et $Q'(z)$ sont aussi symétriques et ont les propriétés suivantes [25], [26]:

- Si les racines du filtre $A(z)$ sont à l'intérieur du cercle unité, alors les racines de $P'(z)$ et $Q'(z)$ sont sur ce même cercle. Elles sont entrelacées et commencent par une racine de $P'(z)$.
- Inversement si les racines de $P'(z)$ et $Q'(z)$ sont sur le cercle unité et sont entrelacées et commencent par la racine de $P'(z)$, alors les racines de $A(z)$ sont à l'intérieur du cercle unité.
- Du fait que les racines de $P'(z)$ et $Q'(z)$ sont sur le cercle unité, ces derniers peuvent être complètement définis par la position angulaire de leurs racines. En outre, les polynômes $P'(z)$ et $Q'(z)$ ayant des coefficients réels, leurs racines se présentent sous forme de paires de nombres complexes conjugués. Par conséquent, uniquement les positions angulaires situées sur le demi cercle supérieur du plan Z sont suffisantes pour déterminer complètement $P'(z)$ et $Q'(z)$.

Ainsi, les coefficients LSF sont définis comme les positions angulaires ω_i des racines de $P'(z)$ et $Q'(z)$ situées dans le demi cercle supérieur du plan Z. Ils sont déterminés généralement par l'algorithme de Kabal et Ramachandran [23].

2.2. Propriétés des coefficients LSF.

Les autres propriétés des coefficients LSF [9], [11], [23]-[27] peuvent être résumées ainsi :

- Une condition nécessaire et suffisante pour la stabilité du filtre de synthèse $1/A(z)$ est que les coefficients LSF doivent respecter la condition suivante :

$$0 < LSF_1 < LSF_2 < \dots < LSF_p < 0.5 \quad (4)$$

Où la fréquence normalisée 0.5 correspond à la fréquence de Nyquist.

Cette propriété exprime l'ordonnement de ces coefficients.

- Les coefficients LSF sont toujours concentrés autour des formants. La largeur de bande d'un formant dépend de la proximité de ses coefficients LSF correspondants.
- En général, la sensibilité spectrale de chaque coefficient LSF est localisée. Cela veut dire qu'un changement dans un coefficient causera un changement dans le spectre de puissance près de son voisinage.
- Un espacement court entre deux coefficients LSF correspond à une résonance (formant), tandis qu'un espacement large entre deux coefficients LSF correspond à une vallée (antirésonance).
- La représentation LSF est une représentation fréquentielle dont la quantification peut aisément incorporer des traits très importants pour la perception et la sensibilité de l'oreille humaine en particulier. Un exemple est donné dans [28], où les fréquences basses des LSF ont été quantifiées avec plus de bits que les fréquences plus hautes.

2.3 Utilisation de la représentation LSF

La représentation LSF d'ordre 10 est très utilisée dans les codeurs de parole actuels et en particulier les standards dont le débit est inférieur à 16 kbps. On cite par exemple :

- Le standard ITU-T G.729 CS-ACELP à 8 kbps [29]
- Le standard ITU-T G.723.1, codeur à double débit pour la communication multimédia à 5.3/6.3 kbps [30]
- Le GSM 6.60 EFR du standard européen de télécommunications à 12.2 kbps [31]
- Le standard japonais JDC-2 PSI-CELP à 5.6 kbps [32]
- Le standard américain de la défense pour la téléphonie sécurisée FS1016 CELP à 4.8 kbps [16], [17].
- Le nouveau standard américain pour la téléphonie sécurisée, FS 1017 MELP (Mixed Exited Linear Prediction) à 2.4 kbps [33]-[34].

Dans ces standards, la représentation LSF procure une bonne qualité de quantification pour l'information spectrale, avec moins de bits et une meilleure qualité de parole. Notons que des standards plus anciens tels que le GSM 6.10 FR [38] et le IS-54 [39] utilisent les coefficients LAR. Ces standards ont été remplacés respectivement par le GSM 6.60 [31] et le IS-641 [40], où on préfère les LSF pour une meilleure quantification de l'information spectrale.

Tableau 1 : Tables des niveaux de quantification des coefficients LSF d'après le standard US CELP FS1016. C'est une quantification scalaire codée sur 34 bits. (C1 est la table de LSF1, C2 celle de LSF2, ...).

C ₁		C ₂		C ₃		C ₄	
indice	valeur	indice	valeur	indice	valeur	indice	valeur
0	0.0125	0	0.0262	0	0.0525	0	0.0775
1	0.0213	1	0.0294	1	0.0575	1	0.0825
2	0.0281	2	0.0331	2	0.0625	2	0.0900
3	0.0312	3	0.0369	3	0.0675	3	0.0994
4	0.0350	4	0.0406	4	0.0731	4	0.1100
5	0.0425	5	0.0450	5	0.0800	5	0.1212
6	0.0525	6	0.0500	6	0.0881	6	0.1350
7	0.0625	7	0.0550	7	0.0969	7	0.1462
		8	0.0600	8	0.1062	8	0.1588
		9	0.0650	9	0.1188	9	0.1713
		10	0.0700	10	0.1312	10	0.1838
		11	0.0762	11	0.1438	11	0.1962
		12	0.0838	12	0.1562	12	0.2088
		13	0.0925	13	0.1688	13	0.2212
		14	0.1013	14	0.1812	14	0.2338
		15	0.1100	15	0.1938	15	0.2462
C ₅		C ₆		C ₇		C ₈	
indice	valeur	indice	valeur	indice	valeur	indice	valeur
0	0.1250	0	0.1838	0	0.2250	0	0.2781
1	0.1312	1	0.1962	1	0.2350	1	0.3000
2	0.1412	2	0.2112	2	0.2450	2	0.3156
3	0.1512	3	0.2288	3	0.2625	3	0.3312
4	0.1606	4	0.2500	4	0.2875	4	0.3500
5	0.1688	5	0.2750	5	0.3100	5	0.3688
6	0.1788	6	0.3000	6	0.3375	6	0.3938
7	0.1888	7	0.3250	7	0.3625	7	0.4188
8	0.1988						
9	0.2088						
10	0.2188						
11	0.2312						
12	0.2438						
13	0.2562						
14	0.2688						
15	0.2812						
C ₉		C ₁₀					
indice	valeur	indice	valeur				
0	0.3450	0	0.3988				
1	0.3600	1	0.4088				
2	0.3750	2	0.4188				
3	0.3875	3	0.4275				
4	0.4000	4	0.4362				
5	0.4138	5	0.4488				
6	0.4288	6	0.4638				
7	0.4438	7	0.4788				

3. QUANTIFICATION DES COEFFICIENTS LSF DANS LE CODEUR CELP FS1016

Avant de transmettre les coefficients LSF, il convient d'abord de les quantifier. En effet, ces coefficients contribuent directement à la représentation du spectre vocal

et ne présentent pas tous la même loi de répartition (distribution non uniforme) [16], [17], [28].

Dans le standard américain FS1016 les 10 coefficients LSF sont calculés pour chaque trame de parole de 30 ms et sont codés sur 34 bits (3, 4, 4, 4, 4, 3, 3, 3, 3, 3), après avoir subi une quantification scalaire.

Les tables de quantification pour chaque coefficient LSF_i , $i=1$ à 10 , ont été données au tableau 1. Notons que les valeurs de ces tables représentent les fréquences normalisées entre 0 et 0.5.

Dans ce qui suit, nous exposons l'algorithme de quantification scalaire utilisé par la norme FS1016. Dans cet algorithme, on a utilisé les notations suivantes :

- N_L est le vecteur contenant le nombre de niveaux de quantification des coefficients LSF.
- $\hat{LSF}_{i,j}$ est l'élément de la table de quantification équivalent à une matrice d'ordre $p \times N_L$ (max). N_L (max) représente la valeur maximale du vecteur N_L et p le nombre de coefficients LSF.
- $LSFQ_i$ est le vecteur contenant les coefficients quantifiés.

Algorithme :

$$\begin{aligned} \text{Pour } i = 1, \dots, p \\ J_{min} = \arg \min_{j=1, \dots, N_L} |LSF_i - \hat{LSF}_{i,j}| \\ LSFQ_i = \hat{LSF}_{i, J_{min}} \end{aligned}$$

A noter aussi que cette quantification scalaire (codée sur 34 bits), utilisée par la norme FS1016, ne présente pratiquement pas de distorsions audibles malgré les mesures objectives relativement grandes : 1.73 dB pour la distorsion spectrale moyenne, 13.88% pour les outliers 2-4 dB et 0.22 % pour ceux supérieurs à 4 dB. Cette qualité est due à la bonne allocation des bits entre les tables de la quantification scalaire [16], [17].

4. SYSTEME D'ENUMERATION EN TREILLIS POUR LA QUANTIFICATION DES LSF

La quantification par énumération en treillis est une méthode qui utilise une quantification scalaire (ou QS) à l'origine. La dénomination "treillis" vient du fait que la quantification donne un indice relatif au chemin optimal parcouru à travers les tables de la quantification scalaire en question. Le système en treillis dans la quantification scalaire des coefficients LSF profite de l'ordre ascendant de ces coefficients.

4.1 Quantification par énumération en treillis.

Une fine analyse des combinaisons des différents éléments de la table de la quantification scalaire utilisée dans la norme FS1016 a montré qu'il y avait un nombre très important de vecteurs LSF qui, s'ils étaient utilisés, rendraient le filtre de synthèse instable. Aussi, ce système est conçu pour éviter les combinaisons donnant des vecteurs

non acceptables en sortie du quantificateur (soit par exemple $LSF_2 < LSF_1$).

Dans ce qui va suivre nous proposons d'appliquer une méthode d'énumération utilisant un système en treillis réduisant le nombre de bits nécessaires aux coefficients LSF à 30 bits au lieu de 34 bits, sans changer les tables de la quantification scalaire suscitée. Autrement dit, il s'agit de modifier la façon de les indexer, de sorte à permettre un gain en débit.

4.2 Algorithme de codage

Pour comprendre l'algorithme de quantification qui va suivre, on donne les notations suivantes :

- LSF_i est le coefficient d'ordre i obtenu après analyse.
- Le mot-code du coefficient LSF_i est donné par $\hat{LSF}_{i,j}$.

L'opération de quantification sera notée :

$$QS_{34}(LSF_i) = \hat{LSF}_{i,j}$$

- Pour un ordre de prédiction p , le nombre maximal de combinaisons que les tables de cette QS pourraient donner est : $N_{max} = \prod_{i=1}^p 2^{b_i}$ où b_i est le nombre de bits accordé à la i^{eme} table.

- $N_{i+1}(j)$ est le nombre de chemins possibles allant du mot-code du coefficient LSF_i dans la table C_i à l'indice j (soit $\hat{LSF}_{i,j}$), pour atteindre le mot-code du dernier coefficient de la table C_p , tout en vérifiant la propriété d'ordonnement donnée par l'équation (4).

Ainsi, pour une position j dans la table C_{p-1} , le nombre de combinaisons possibles pour atteindre le dernier coefficient dans la table C_p est donné par :

$$N_p(j) = \sum_{\substack{k=1 \\ \hat{LSF}_{p,k} > \hat{LSF}_{p-1,j}}}^{2^p} 1, \quad 1 \leq j \leq 2^{b_{p-1}} \quad (5)$$

Autrement dit, de la position j se trouvant dans la table C_{p-1} , on peut former $N_p(j)$ transitions possibles vers la table C_p .

La figure 1 donne un exemple avec les deux tables 9 et 10 du FS1016. On dénombre un total de 52 chemins possibles entre de ces deux tables.

Dans cet ordre d'idée, nous donnons un autre exemple montrant le calcul du nombre de chemins possibles lorsqu'on se place dans la dernière position de la table C_8 (figure 2). D'après le tableau 1, on dénombre sept chemins possibles permettant d'atteindre des mot-codes au niveau de la table C_{10} .

On peut refaire le même dénombrement pour les autres positions dans C_8 . Finalement, on obtient respectivement : 52, 52, 52, 52, 44, 36, 20. Avec les 7 chemins précédents, cela donne un total de 315 chemins possibles.

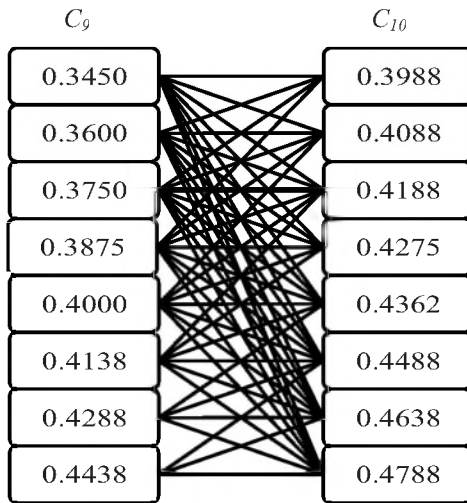


Figure 1 : Exemple de transitions de la table 9 vers la table 10 du QS du CELP FS 1016.

Dans le tableau 2, on donne le nombre de tous les chemins possibles menant de n'importe quelle position dans une table C_i quelconque jusqu'à la table C_{10} en respectant la propriété d'ordonnement.

Pour chaque $\hat{L}SF_{i,j}$, $1 \leq i \leq p-1$, $1 \leq j \leq 2^{b_i}$ on a :

$$N_{i+1}(j) = \sum_{\hat{L}SF_{i+1,k} > \hat{L}SF_{i,j}} N_{i+2}(k) \quad (6)$$

Autrement dit, d'un niveau j dans la table C_i , on aura N_{i+1} chemins possibles correspondant à une bonne succession des coefficients LSF . Ces chemins peuvent se

déduire de la connaissance du nombre de chemin à l'ordre $i+2$.

Notons que la progression est du type arrière (ou *backward*).

• Le total des combinaisons possibles est :

$$N(p, \bar{C}) = \sum_{j=1}^l N_2(j) \quad (7)$$

$$\text{Où } \bar{C} = \{C_1, C_2, \dots, C_p\} \text{ et } l = 2^{b_1}$$

$N_2(j)$ est le nombre de combinaisons possibles de coefficients vérifiant :

$$\hat{L}SF_{1,j} < \hat{L}SF_{2,k}, \quad 1 \leq j \leq 2^{b_1}, 1 \leq k \leq 2^{b_2}.$$

Dans le cas du codeur CELP FS1016, $N(p, \bar{C})$ vaut 554958388.

Formellement, le nombre de bits nécessaires est :

$$K = \lceil \log_2(N(p, \bar{C})) \rceil \quad (8)$$

où $\lceil x \rceil$ désigne le plus petit entier supérieur à x .

K vaut 30 bits dans le cas précédent. On peut remarquer

$$\text{que } K \leq \sum_{i=1}^p b_i.$$

• Dans l'algorithme suivant, m désigne le plus petit ordre dans la table C_i tel que le mot-code d'ordre m soit directement supérieur au coefficient LSF_{i-1} dont l'ordre est j dans C_{i-1} , c-à-d : $\hat{L}SF_{i-1,j} < \hat{L}SF_{i,m}$

Par ailleurs, les valeurs N_i sont celles déterminées précédemment (Tableau 2). On notera par c le mot-code de l'énumération en treillis qui sera initialisé à zéro.

Algorithme de codage :

Etape 1 : Trouver j tel que $QS_{34}(LSF_1) = \hat{L}SF_{1,j}$ avec $1 \leq j \leq 2^{b_1}$ et mettre $j^* = j-1$

Etape 2 : Initialiser $c = 0$ si $j^* = 1$ sinon $c = \sum_{k=1}^{j^*} N_2(k)$

Etape 3 : Pour $i = 2$ à p

$$m = \arg \min_{k=1..2^{b_i}} (\hat{L}SF_{i-1,j^*} < \hat{L}SF_{i,k})$$

Trouver j tel que $QS_{34}(LSF_i) = \hat{L}SF_{i,j}$ avec $1 \leq j \leq 2^{b_i}$ et mettre $j^* = j-1$

$$\text{si } j^* \neq m \text{ calculer } c = c + \sum_{k=m}^{j^*} N_{i+1}(k-1)$$

Etape 4 : Le mot-code d'énumération en treillis est donné par c .

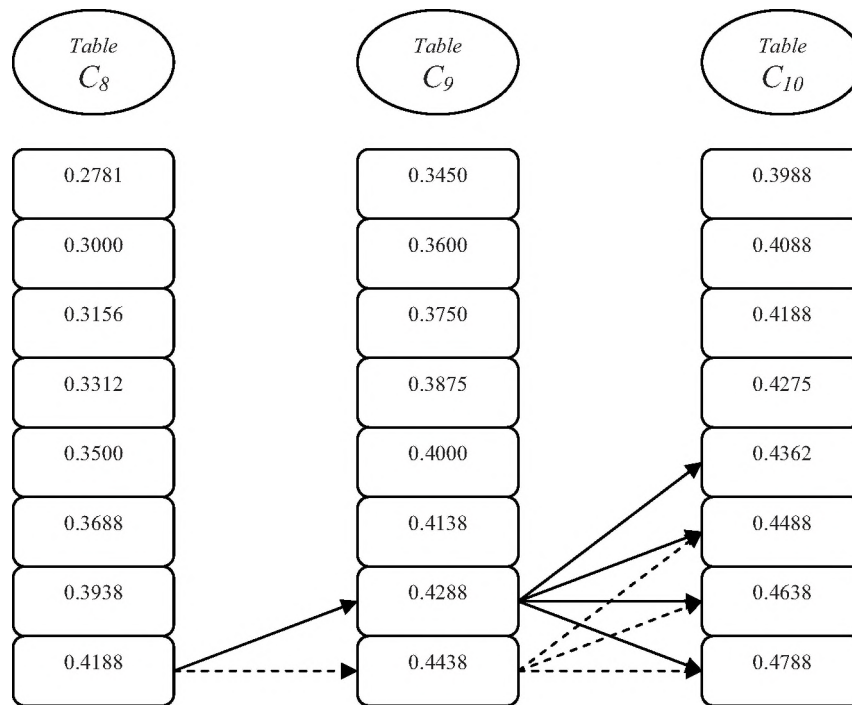


Figure. 2 : Exemple donnant les chemins possibles partant de la 8^{ème} position dans la table C_8 pour atteindre le dernier mot-code au niveau de la table C_{10} .

Tableau 2 : Le nombre des chemins possibles N_i partant de la table C_1 jusqu'à la table C_{10}

Indice dans la table C_{i-1}	N_i : Nombres de chemin possibles partant de la table C_{i-1}									
	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}
1	97762793	8512383	896795	96022	9455	1904	315	52	8	1
2	97762793	8512383	896795	96022	9455	1904	315	52	8	1
3	89250410	8512383	896795	96022	9455	1904	315	52	8	1
4	80738027	8512383	896795	96022	9455	1589	315	52	8	1
5	72225644	8512383	896795	96022	9455	959	263	44	7	1
6	55200878	8512383	800773	96022	9455	644	211	36	6	1
7	38176112	8512383	704751	77112	9455	381	107	20	4	1
8	23841731	7615588	608729	67657	7551	170	63	7	3	1
9	554958388	6718793	512707	58202	5647	9455	1904	315	52	8
10		5821998	416685	39292	5647					
11		4925203	320663	29837	3743					
12		4028408	243551	22286	2154					
13		3227635	175894	10992	2154					
14		2522884	117692	7249	1195					
15		1914155	78400	5095	1195					
16		1401448	48563	2941	551					
		97762793	8512383	896795	96022					

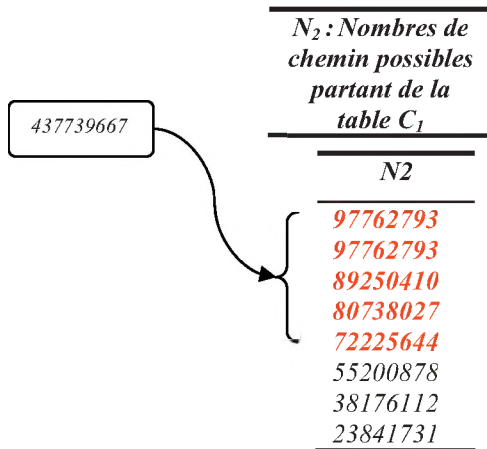


Figure 3 : Exemple de calcul du mot-code *c* pour le coefficient LSF₁

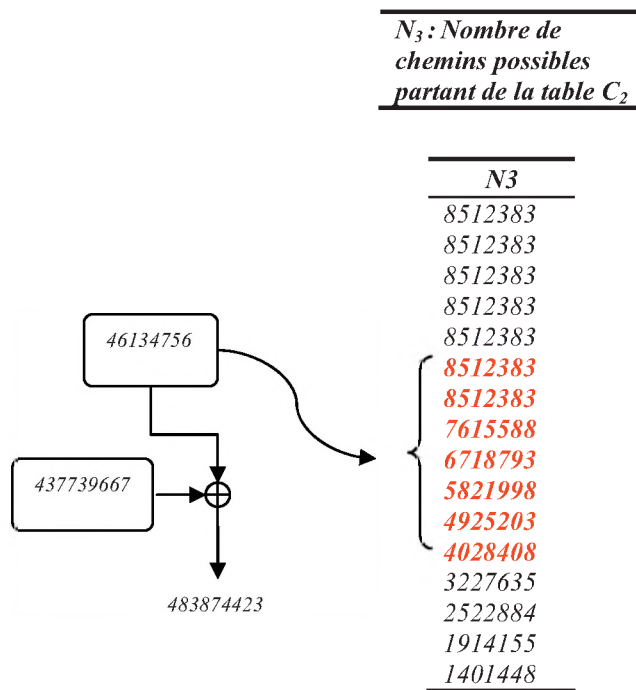


Figure 4 : Calcul du mot-code *c* pour le coefficient LSF₂

Pour une meilleure compréhension de l'algorithme, nous prenons l'exemple de la quantification du vecteur LSF suivant :

$$LSF = [0.0409 \ 0.0869 \ 0.1312 \ 0.1621 \ 0.2340 \ 0.2649 \ 0.3052 \ 0.3321 \ 0.3830 \ 0.4073],$$

obtenu après analyse d'une trame de 30 ms du signal parole.

Le déroulement de l'opération de quantification suit les étapes suivantes : dans les deux premières étapes, on cherche le mot-code du premier coefficient LSF_1 , en appliquant l'algorithme de la quantification scalaire du FS1016. On s'intéresse ici à son indice.

Dans notre cas, il s'agit de l'indice $j=6$ et par conséquent on aura : $j^*=6-1=5$. Comme $j^* \neq 1$, on calculera donc c en utilisant le tableau 2 et selon la figure (3), c'est-à-dire comme suit :

$$c = \sum_{k=1}^5 N_2(k) = 437739667 \quad (9)$$

Dans une troisième étape, on déroule une boucle qui commence au coefficient LSF_2 pour atteindre le coefficient LSF_{10} . Pendant ce déroulement, on détermine d'abord l'ordre m comme indiqué plus haut. Dans le cas de LSF_2 , m vaut 6. Par la suite, on cherche, comme dans la première étape, l'indice j du représentant du coefficient LSF_2 . Dans cet exemple j vaut 13 et on aura alors $j^* = 12$.

Comme $j^* \neq m$, on continue on calcule c comme suit :

$$c = c + \sum_{k=m}^{j^*} N_{i+1}(k-1) \quad (10)$$

Ainsi, pour le cas du LSF_2 et comme indiqué à la figure 4,

$$\text{on aura } c = 437739667 + \sum_{k=6}^{12} N_3(k-1), \text{ soit}$$

$$c = 483874423.$$

Les mêmes opérations seront répétées pour les coefficients qui restent et on obtient les résultats suivants :

LSF_3	$j^* = 10$	$m = 7$	$c = 486117295$
LSF_4	$j^* = 8$	$m = 7$	$c = 486262064$
LSF_5	$j^* = 11$	$m = 5$	$c = 486313017$
LSF_6	$j^* = 5$	$m = 5$	$c = 486313976$
LSF_7	$j^* = 5$	$m = 5$	$c = 486314239$
LSF_8	$j^* = 3$	$m = 3$	$c = 486314291$
LSF_9	$j^* = 3$	$m = 1$	$c = 486314315$
LSF_{10}	$j^* = 1$	$m = 1$	$c = 486314316$

A ce niveau de la quatrième étape, on transmettra seulement l'indice c obtenu à la fin du traitement de la table de LSF_{10} c'est-à-dire $c=483874423$ et qui sera codé sur 30 bits, car c vaut en binaire:

01110011111001001000101001100.

4.3 Algorithme de décodage

Au niveau de la réception, le nombre c sera réceptionné pour être utilisé dans l'opération inverse, telle que donnée par l'algorithme de décodage suivant.

Dans le cas de l'exemple précédent, c a été trouvé égal à 486 314 316. Dans la première étape de l'algorithme, on cherchera l'indice j tel que c vérifie :

$$\sum_{k=1}^{j-1} N_2(k-1) \leq c < \sum_{k=1}^j N_2(k-1) \quad (11)$$

D'après le tableau 2 et la figure 5, on remarque que la valeur $j=5$ vérifie cette relation. En effet, on a :

$$\sum_{k=1}^5 N_2(k-1) = 437\ 739\ 667$$

$$\sum_{k=1}^6 N_2(k-1) = 492\ 940\ 545 \quad (12)$$

Ainsi, le représentant du premier coefficient a un indice égal à 6. D'où $\hat{LSF}_1 = 0.0425$ et, selon l'algorithme, l'indice j^* sera égal à $6 - 1 = 5$.

A la deuxième étape on retranchera de c la valeur $\sum_{k=1}^5 N_2(k-1) = 437739667$ pour avoir une nouvelle valeur de $c = 48574649$.

A l'étape trois, une boucle qui commencera à partir du coefficient LSF_2 jusqu'au dernier coefficient LSF_{10} itérera les opérations indiquées dans l'algorithme. Pour le coefficient LSF_2 , par exemple, on effectuera les opérations suivantes.

- Trouver m tel que $m = M_2(\hat{LSF}_{1,6})$.

D'après la table C_2 de la quantification scalaire du FS1016, m vaut 6 (Tableau 1).

Comme $c \neq 0$, on cherchera j vérifiant

$$\sum_{k=m}^{j-1} N_3(k-1) \leq c < \sum_{k=m}^j N_3(k-1) \quad (13)$$

j vaut donc 13 selon le tableau 2, car les deux bornes inférieure et supérieure sont égales respectivement à 46134756 et à 49362391 (figure 6).

- On met $j^* = j - 1 = 12$ et on attribue à LSF_2 la valeur 0.0838.

Algorithme de décodage :

Etape 1. Trouver j tel que $\sum_{k=1}^{j-1} N_2(\hat{L}SF_{1,k-1}) \leq c < \sum_{k=1}^j N_2(\hat{L}SF_{1,k-1})$ où $1 \leq j \leq 2^{b_2}$ et mettre $j^* = j-1$.

Mettre $\hat{L}SF_1 = \hat{L}SF_{1,j^*}$

Etape 2. Calculer $c = c - \sum_{k=1}^{j^*} N_2(\hat{L}SF_{i,j^*})$

Etape 3. pour $i=2$ à p

$m = \arg \min_{k=1..2^{b_i}} (\hat{L}SF_{i-1,j^*} < \hat{L}SF_{i-1,k})$

Mettre $j = m$ si $c = 0$ sinon trouver j tel que

$\sum_{k=m}^{j-1} N_{i+1}(\hat{L}SF_{1,k-1}) \leq c < \sum_{k=m}^j N_{i+1}(\hat{L}SF_{1,k-1})$ où $1 \leq j \leq 2^{b_i}$ et mettre $j^* = j-1$

Mettre $\hat{L}SF_i = \hat{L}SF_{i,j^*}$

si $j^* \neq m$ calculer $c = c - \sum_{k=1}^{j^*} N_{i+1}(\hat{L}SF_{i,j^*})$

Etape 4. Le vecteur LSF reconstitué est donné par $\{\hat{L}SF_1, \hat{L}SF_2, \hat{L}SF_3, \dots, \hat{L}SF_p\}$

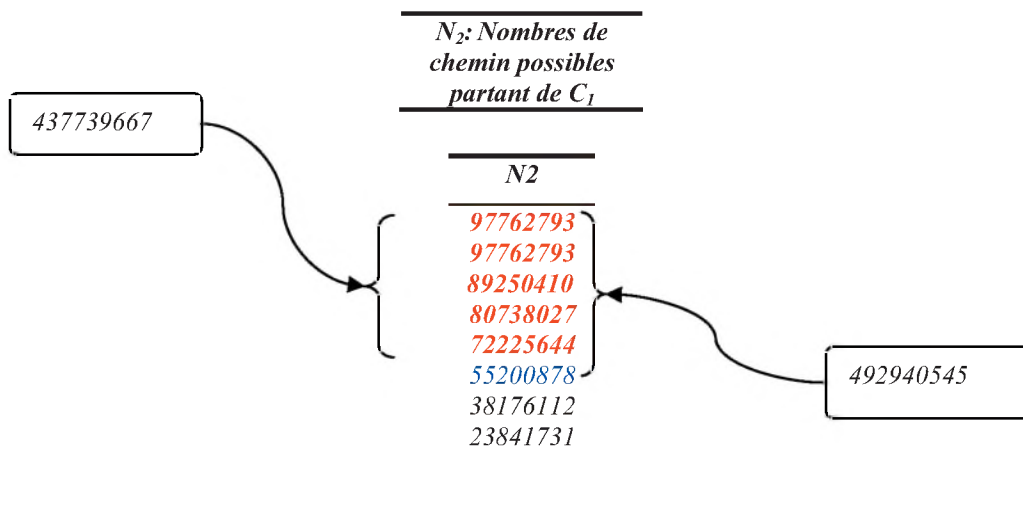


Figure 5 : Détermination de l'indice de LSF_1 (pour $c = 483\ 874\ 423$). Les valeurs sont extraites du tableau 2

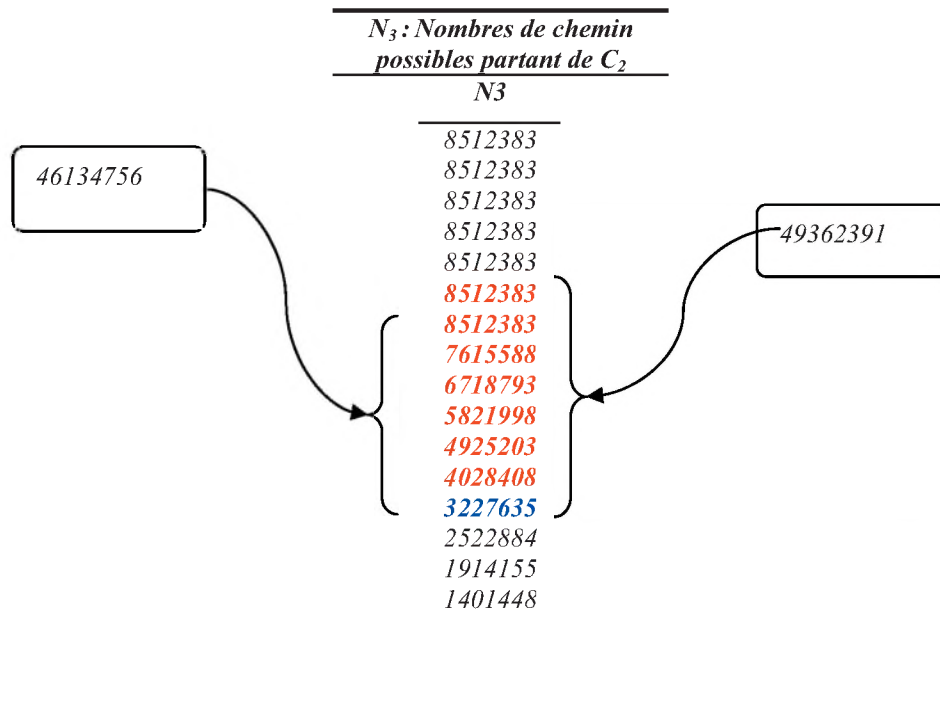


Figure 6 : Détermination de l'indice de LSF_2

Puisque de nouveau, $j^* \neq m$, c sera alors comparé à 46134756 et ce pour avoir un nouveau c égal à 2439893 ($c-46134756 = 2439893$).

Les mêmes étapes que précédemment seront réitérées pour le reste des coefficients LSF .

5. RESULTATS ET INTERPRETATIONS

- *Complexité*

La complexité du codeur FS1016 est de 8.8 Mflops [41]. Dans notre cas, nous ajoutons $(6 \times 8 + 4 \times 16) = 112$ additions et autant de comparaisons pour la partie codage. Ceci correspond 0.007 Mflops, qui reste très faible devant la complexité initiale. Pour le décodeur, 278 additions et 674 comparaisons sont ajoutées, comparativement au décodeur initial, soit 0.032 Mflops.

En ce qui concerne l'occupation mémoire, celle des données augmente légèrement. 110 mots de 32 bits sont nécessaires pour stocker les valeurs entières du tableau 2. Ceci est valable aussi bien pour le codeur que pour le décodeur. Pour la mémoire programme, celle-ci augmente de 10% pour le codeur et de 23% pour le décodeur.

- *Distorsion spectrale*

Des tests sur le standard FS1016, aussi bien avec la

quantification scalaire classique qu'avec l'application du système d'énumération en treillis, ont été effectués sur deux bases de données de parole, l'une extraite de la base Américaine (TIMIT) [42] et l'autre Arabe (PAPE) [43]. Pour les deux bases de données, les trames silencieuses des débuts et des fins de phrases ont été limitées à un maximum de 2 trames.

TIMIT est une base de données de parole échantillonnée à 16 KHz comprenant huit dialectes des Etats-Unis (USA). En pratique, et dans le codage à très bas débit nous utilisons des fichiers de parole échantillonnés à 8 KHz, ce qui nous a conduit à sous-échantillonner les 6300 fichiers utilisés de TIMIT à 8KHz. Un total de 112 000 trames de 30 ms a été utilisé.

Une autre base de données de parole arabe PAPE (Phrases Arabes Phonétiquement Equilibrées) a été aussi utilisée dans les tests objectifs. Son enregistrement a été effectué en chambre sourde à l'Institut ICP de Grenoble par six locuteurs algériens (trois masculins et trois féminins) de niveaux universitaires parlant correctement l'arabe. Les fichiers de parole étaient échantillonnés à 10 KHz. Ils forment un total de 106 334 trames de 30 ms. De même que pour la base de données précédente, nous avons dû opérer un sous-échantillonnage à 8 KHz.

Les performances du quantificateur sont évaluées par la distorsion spectrale moyenne SD (Spectral Distorsion) qui est souvent utilisée comme mesure objective de la performance d'encodage des paramètres LSF. Cette

mesure de distorsion fournit une bonne corrélation avec la perception auditive humaine. Développée et calculée sur une largeur de bande limitée, son expression discrète, pour une trame i , est donnée en *décibels* par [7], [28], [44]

$$SD_i = \sqrt{\frac{1}{n_1 - n_0} \sum_{n=n_0}^{n_1-1} \left[10 \log \frac{S(e^{j2\pi n/N})}{\hat{S}(e^{j2\pi n/N})} \right]^2} \quad (14)$$

Pour un signal de parole échantillonné à 8 kHz avec une largeur de bande de 3 kHz, une FFT sur 256 points est utilisée pour calculer les spectres de puissance originaux $S(e^{j2\pi n/N})$ et quantifiés $\hat{S}(e^{j2\pi n/N})$ du filtre de synthèse LPC de la i ème trame du signal de parole. Pour le calcul de la SD , Ramachandran et al. [45] utilisent la bande 125-3000 Hz, avec une résolution de 31.25 Hz. Paliwal et Atal. [28] considèrent la bande 0-3000 Hz et utilisent une résolution de 15 Hz. Pour leur part, LeBlanc et al. [44] considèrent la bande 125 Hz-3.1 KHz et une résolution de 31.25 Hz. Dans notre cas, nous avons retenu la bande et la résolution de LeBlanc et al. La SD est donc calculée avec une résolution de 31.25 Hz par échantillon. Dans l'équation (14), n_0 et n_1 correspondent donc respectivement à 125 Hz et 3125 Hz et valent 4 et 100. Ainsi, 96 raies spectrales uniformément espacées seront utilisées dans l'équation (14).

La SD moyenne est évaluée pour tout le nombre de trames N_f de la base de donnée parole, soit :

$$\overline{SD} = \frac{1}{N_f} \sum_{i=1}^{N_f} SD_i \quad (15)$$

En général, une \overline{SD} d'environ 1 dB indique que la distorsion perçue pendant la quantification est négligeable. Dans le passé, cette valeur a été suggérée comme seuil pour la transparence spectrale de la parole (une qualité de quantification est dite transparente si aucune distorsion audible due à cette quantification n'est introduite dans la parole codée). Paliwal et Atal [28] ont établi que la SD moyenne n'est pas suffisante pour mesurer seule la qualité perçue. Ils ont introduit la notion des trames spectrales externes (outliers frames). Par conséquent, nous pouvons obtenir une qualité de codage transparent si les trois conditions suivantes sont maintenues 1) la SD moyenne est d'environ 1 dB, 2) le pourcentage des trames «outliers» ayant une SD entre 2 et 4 dB est moins de 2% et 3) aucune trame ne doit avoir une SD qui dépasse 4 dB.

Le traitement de 218 334 trames extraites des deux bases de données précédentes a donné une distorsion spectrale moyenne de 1.73 dB, 13.88 pour les % 2-4 dB et 0.22 pour les % supérieurs à 4 dB, confirmant les

résultats donnés dans la littérature concernant ce standard [9].

- *Sensibilité aux erreurs binaires*

Ce travail est axé essentiellement sur un codage source. Il convient parfaitement pour un canal sans pertes. Cependant, la sensibilité aux erreurs binaires mérite d'être approfondie ultérieurement. En effet, nous avons constaté une légère dégradation des performances de cette modification du codeur (sans protection) dès que le taux d'erreurs binaires ou BER (bit error rate) dépasse 10^{-2} (tableau 3, cas d'un BSC). Les mêmes distorsions ne sont observées qu'aux environs de 5×10^{-3} dans le cas du FS 1016 classique. Ceci était prévisible, vu que l'indice de 30 bits, obtenu par notre approche, est plus sensible que le vecteur obtenu par l'allocation initiale de 34 bits (partagés de façon non uniforme). Cependant, une conduite de notre méthode par une approche telle que celle abordée par Secker et Perkis [46] et portant sur un codage conjoint canal-source permettrait de mieux exploiter notre approche.

Tableau 3. Performances du codeur mis au point (Entre parenthèses, le FS1016 classique)

BER	SD	SD >2 dB
Idéal	1.73 (1.73)	13.88 (13.88)
0.001	1.85 (1.84)	14.91 (14.85)
0.01	4.53 (2.04)	27.28 (18.23)
0.05	7.89 (4.18)	44.12 (28.17)

6. CONCLUSION

Dans cet article, nous avons présenté le système d'énumération en treillis que nous proposons d'appliquer à la quantification des coefficients LSF en vue de diminuer le débit alloué à ces coefficients dans le standard CELP FS1016, sans changer ses tables de quantification. En fait, nous avons noté que seulement 3.25 % des combinaisons totales que l'on peut former avec les éléments de ces tables de quantification, peuvent vérifier la propriété d'ordonnement des coefficients LSF et conduisent ainsi à un filtre stable. L'application d'un système d'énumération en treillis nous a permis d'établir de nouveaux algorithmes respectivement de codage et de décodage. Ces algorithmes ont été explicités par des exemples détaillés. Nous avons montré, que ce système permet de diminuer le nombre de bits nécessaires au codage des LSF de 34 bits à 30 bits. Les performances de ce système d'énumération ont été évaluées sur 218 334 vecteurs LSF et ont donné une distorsion spectrale de 1.73 dB, avec des outliers de 13.88 % pour 2-4 dB et 0.22

% pour ceux >4dB, résultat analogue à celui du standard FS1016, mais avec un gain en débit de 133 bits/s.

REFERENCES

- [1] B.S. Atal, S.L. Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave", J. Acoust. Soc. Amer., vol. 50, pp. 637-655, 1971.
- [2] J. Makhoul, "Linear prediction : A tutorial review", Proc. IEEE, vol. 63, pp. 561-580, Apr. 1975.
- [3] N.S. Jayant, P. Noll, "Digital coding of waveforms", Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [4] J.D. Markel, A. H. Gray, Jr., "Linear prediction of speech", Berlin, Germany, Springer Verlag, 1976.
- [5] J.D. Markel, A. H. Gray, Jr., "Quantization and bit allocation in speech processing", IEEE trans. acoustics, speech and signal processing, Vol. ASSP-24, pp. 459-473, 1976.
- [6] F. Itakura, S. Saito, "Digital filtering techniques for speech analysis and synthesis", 7th International congress on acoustics, 25-C1, Budapest, 1971
- [7] K. Paliwal, W. Kleijn, "Quantization of LPC parameters", (Chapter 12) in speech Coding and Synthesis, ed. by W. Kleijn and K. Paliwal, Elsevier, Amsterdam, pp. 433-466, 1995.
- [8] A. Benallal, A. Gilloire, "Instabilité et stabilité des algorithmes des moindres carrés transversaux rapides excités par la parole", 12^{ème} colloque GRETSI, Juan-Les-Pins, France, pp. 509-512, 1989.
- [9] S. Grassi "Optimised implementation of speech processing algorithm", thèse de doctorat, université de Neuchâtel, faculté de sciences, Février 1998.
- [10] U.Viswanathan, J. Makhoul, "Quantization properties of transmission parameters in linear predictive systems", IEEE Transaction on speech, signal processing, Vol. ASSP-23, pp. 309-321, 1975.
- [11] F. Itakura, "Line spectrum representation of linear predictive coefficients of speech signals," JASA, vol. 57, suppl 1, pp. S35 (A), 1975.
- [12] ICASSP'95, session: "Spectral quantization", 10 articles différents, Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 716-755, 1995
- [13] EUROSPEECH'95, session : "Quantization of spectral parameters", 9 articles différents, Proc. European Conf. on speech communication and technology, Vol. 2, pp. 1029-1064, 1995.
- [14] ICASSP'96, session: "Spectral quantization", 10 articles différents, Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vol. 1, pp. 737-776, 1996
- [15] ICASSP'97, session: "Speech coding at low bit rates", 14 articles différents, Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vol. 2, pp. 1555-1610, 1997
- [16] "Federal standard 1016, Telecommunications analogue to digital conversation radio voice by 4800 bits/s code excited linear prediction (CELP)", National communication systems office of technology and standards. Washington DC. 20305 -20310, 1991.
- [17] J.P. Campbell Jr., V.C. Welch, T.E. Tremain, "An expandable error protected 4800 bps CELP coder (U.S. Federal standard 4800 bps voice coder)," in proc. IEEE Int. Conf. Acoust., Speech., Signal process., Glasgow, U.K, pp. 735-738, 1989.
- [18] M.W. Marcellin & T.R. Fischer, "Trellis coded quantization of memoryless and Gauss-Markov sources", IEEE Transactions on Communications, Vol. 38, pp. 83-93 Jan 1990.
- [19] S. Kang, Y. Shin, T. R. Fisher, "Low-complexity predictive trellis-coded quantization of speech line spectral frequencies", IEEE Transaction on signal processing, Vol. 52, N07, pp. 2070-2079, July 2004.
- [20] F. Lahouti, A.K. Khandani, "Quantization of LSF Parameters using a trellis modeling", IEEE trans. Speech, audio process., vol. 11, pp. 400-412, September 2003.
- [21] T. Malone, T. R. Fisher "Enumeration and trellis searched coding schemes for speech LSP parameters", IEEE Transaction on speech and audio processing, vol.1, N°3, pp. 304-314, July 1993.
- [22] G.D. Forney, Jr., "The Viterbi algorithm ", Proc. IEEE (Invited Paper), vol. 61, pp. 268-278, Mar. 1973.
- [23] P. Kabal and R.P. Ramachandran, "The computation of line spectrum frequencies using Chebyshev polynomials," IEEE trans. acoustics, speech and signal processing, vol. ASSP-34, pp. 1419-1426, 1986.
- [24] F.K. Soong, B. Juang, "Line spectrum pairs (LSP) and speech data compression," Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1.10.1-1.10.4, 1984
- [25] T. Bäckström, P. Alku, T. Paatero, B. Kleijn "A time domain interpolation for the LSP decomposition", IEEE Transaction on Speech and Audio Processing, Vol. 12, No6, pp. 554-560, November 2004.
- [26] F.K Soong et B. Juang "Line spectrum pair LSP and speech data compression", Proc ICASSP, pp.1.10.1 - 1.10.4, 1984.

[27] F.K. Soong, B. Juang, "Optimal quantization of LSP parameters", IEEE Transaction on Speech and Audio Processing, Vol. 1, pp. 15-24, 1993.

[28] K. Paliwal, B. S. Atal, "Efficient vector quantization of LPC parameters at 24bits/frame", IEEE Transaction on Speech and Audio Processing, Vol. 1, pp. 3-14, January 1993.

[29] ITU-T "Coding speech at 8 kbps using conjugate structure Algebraic code excited linear prediction", ITU-T Recommendation G.729, Telecommunication standardization sector of ITU (03/96).

[30] ITU "Dual rate speech coder for multimedia communication transmitting at 5.6 and 6.3kbps", ITU Recommendation G.723.1, Telecommunication standardization sector of ITU (03/96).

[31] K. Järvinen et al., "GSM Enhanced full rate speech coder", Proc. IEEE ICASSP'97, vol.2, pp. 771-774, 1997.

[32] T. Ohya, H. Suda, T. Miki, "5.6 kbps PSI-CELP of the half rate PDC speech coding standard", Proc. IEEE Vehicular technology conference, Vol. 1, pp. 1680 – 1684, 1994.

[33] J. Stachurski, A. McCree, V. R. Viswanathan "High quality MELP coding at bit rate around 4 kbps", IEEE, pp. 485 – 488, March 1999

[34] W. Lin, S. N. Koh, X. Lin, "Mixed excited linear prediction coding of wide band speech at 8 kbps", IEEE, pp. 1137 – 1140, April 2000.

[35] T. W. K. Koishida, V. Cuperman, A. Gersho, J.S. Collura, "A 1200 BPS speech coder based on MELP", IEEE, pp. 1375 – 1378, April 2000.

[36] E. J. Daniel, K. A. Teague, "Federal standard 2.1 KBPS MELP over IP", 43rd IEEE Midwest Symposium on circuits and systems, Lansing MI, pp. 568–571, Aug. 2000

[37] L.I. Wenyuan, M. A. Xin, Z. Yuzhong, Z. Yifang "Study and development of MELP vocoder", IEEE, ICSLP'02 Proceeding, pp. 1684-1688, Juin 2002

[38] European telecommunication standard institute (ETSI), "Full-rate speech transcoding", Recommendation du GSM 06.10.1992.

[39] EIA/TIA Interim Standard, Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard, IS-54-B, Telecommunication Industry Association, April, 1992

[40] T. Honkanen, J. Vainio, K. Jarvinen, P. Haavisto, R. Salami, C. Laflamme, J.P. Adoul, "EFR speech codec for IS-136 digital system", IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vol. 2, pp. 731-734, 1997

[41] M. Mauc, G. Baudoin, M. Jelinek, "Complexity Reduction for FS1016 with multistage Search", proc. IEEE Int. Conf. Acoust., Speech., Signal process., Adelaide, pp. 261-264, Avril, 1994 .


[42] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, N. Dahlgren, and V. Zue, "Timit acoustic-phonetic continuous speech corpus", Linguistic Data Consortium, 1993.

[43] M. Boudraa, B. Boudraa, B. Guerin, "Twenty List of Ten Arabic Sentences for Assessment", Acustica associated with Acta-Acustica, vol.86, pp. 870-882, 2000.


[44] W.P. LeBlanc, B. Bhattacharya, S.A. Mahmoud, V. Cuperman, "Efficient search and design procedures for robust multi-stage VQ of LPC parameters for 4 kb/s speech coding", IEEE Transaction on Speech and Audio Processing, Vol. 1, NO.4, pp. 554-560, October, 1993.

[45] R. P. Ramachandran, M. M. Sondhi, and N. Seshadri, "A Two Codebook Format for Robust Quantization of Lline Spectral Frequencies", IEEE Transactions on Speech and Audio Processing, vol. 3, no. 3, pp. 157--168, May 1995.

[46] P. Secker, A. Perkis, "Joint source and channel treillis coding of line spectrum pair parameters", Speech communication, Vol.11, pp. 149-158, 1992.




**WESTERN
ELECTRO-ACOUSTIC
LABORATORY**

A division of Veneklasen Associates, Inc. 

ACOUSTICAL TESTING & MEASUREMENTS

Laboratory Testing	
Sound Transmission Loss, STC	ASTM E-90* (ISO 140*)
Sound Absorption, NRC	ASTM C-423* (ISO 354*)
Calibration of Microphones	ANSI SI-10*
Acoustic Power	ANSI S12-32
<i>Full Anechoic Chamber Measurements also available</i>	
Field Testing	
Noise Reduction, NIC, FSTC	ASTM E-336*
Impact Sound Transmission, FIIC	ASTM E-1007*
Building Facades	ASTM E-996*
*NVLAP Accredited	



TEL: 661.775.3741 FAX: 661.775.3742
25132 Rye Canyon Loop Santa Clarita, CA 91355
www.weal.com