

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Ecole Nationale Polytechnique



Département d'Électronique
Laboratoire Signal et Communication

Mémoire

pour l'obtention du diplôme de Master en Électronique

Codage à Large Bande de l'Enveloppe Spectrale

Zakaria KHANFRI

Présenté et soutenu publiquement le 20/09/2016

Composition du Jury:

Président	Pr / M .MEHENNI	ENP
Promoteur	Pr / D .BERKANI	ENP
Examineur	Mme / MMA A .MOUSSAOUI	ENP

Dédicace

I Dedicate this work to:
My Parents & My Brothers,
MY Family,
MY Friends,

Zaki

Remerciements

Nous remercions le bon Dieu de nous avoir donné la volonté et la patience qui nous ont permis de mener à bien ce travail.

Nous remercions vivement notre promoteur, P. D.BERKANI, pour sa gentillesse, ses précieux conseils et pour la confiance qu'il nous a témoignée en dirigeant ce travail tout en nous laissant travailler très librement.

On tiens à adresser nous sincères remerciements à tous les membres du Jury chargé d'examiner la soutenance de notre projet de fin d'étude..

Nous remercions tous ceux qui ont de près ou de loin contribué à l'aboutissement et l'amélioration de ce travail.

Résumé

ملخص:

في هذا العمل نهتم بالتشفير ذو الشريط الواسع للعلاف الطيفي، لهذا الهدف قمنا بمقارنة رتبة التعرف: ١٠، ١٦، ١٨، ٢٠ جذر. لذلك قمنا بتكميم خطوات التواترات الطيفية، أخيراً قارنا الفعاليات بواسطة التعوجات الطيفية.

كلمات مفتاحية: تشفير الكلام، خطوط التواترات الطيفية، التكميم الشعاعي، التكميم الخطي.

Abstract:

In this work we are interested in wide band coding of the spectral envelope, for that we compared several orders of prediction: 10, 16, 18 and 20 poles. We proceeded to the quantification of the LSF for the various orders, there after we compared their performances by means of the spectral distortions.

Key Words: Speech coding, lines of spectral frequencies LSF, vectorial quantification, scalar quantification.

Résumé:

Dans cet travail on s'intéresse au codage large bande de l'enveloppe spectrale, pour cela nous avons comparé plusieurs ordres de prédiction : 10, 16,18 et 20 pôles. Nous avons procédé à la quantification des LSF pour les différents ordres, par la suite nous avons comparé leurs performances au moyen des distorsions spectrales.

Mots Clés: Codage de la parole, lignes de fréquences spectrales LSF, quantification vectorielle, quantification scalaire.

Table des Matières

Liste des figures

Liste des tableaux

Introduction Générale	8
I Codage à Large Bande	9
1 Codage à large bande	10
1.1 Intérêts de l'évolution vers le codage large bande	10
1.1.1 La production vocale	10
1.2 La perception de la parole	11
1.3 Codage large bande	13
1.3.1 Codage de la parole large bande	13
1.3.2 Codage de la musique en large bande	13
1.3.3 Les codeurs large bande	14
II Résultats et Simulations	15
2 Résultats et simulations	16
2.1 Conditions d'analyse	16
2.2 L'analyse LPC	17
2.2.1 Représentation des LSP	17
2.3 Codage de l'enveloppe spectrale	18
2.3.1 Principes	18
2.3.2 Création du dictionnaire de quantification	18
2.3.3 Quantification des coefficients LSP	21
2.3.4 Interprétations et commentaires	24
2.4 Calcul de la distorsion spectrale	24
2.4.1 Interprétation et commentaires	25
Conclusion Générale	26
Bibliographie	27

Liste des figures

1.1	Perception auditive	13
2.1	signal parole utilisé	17
2.2	L'enveloppe spectrale des LSF pour M=10, M=16, M=18 et M=20 pôles	18
2.3	Quantification Vectorielle des coefficients LSP pour m=10 pôles	21
2.4	Quantification Vectorielle des coefficients LSP pour m=16 pôles.	22
2.5	Quantification Vectorielle des coefficients LSP pour m=18 pôles.	22
2.6	Quantification Vectorielle des coefficients LSP pour m=20 pôles.	23
2.7	Enveloppe spectrales des LSF et LSF quantifié pour M=10,16,18 et 20 pôles	24
2.8	Distorsions spectrale entre les LSP et les LSP quantifié pour M=10,16,18 et 20 pôles.	25

Liste des tableaux

1.1	codeurs à large bande	14
2.1	Valeurs des (a_i) pour les premières trames pour $m=10,16,18$ & 20 pôles.	17
2.2	(b) Valeur des (LSP_i) pour les premières trames pour $M=10,16,18$ & 20 pôles.	18
2.3	Exemples des deux premières valeurs du dictionnaire obtenu par LBG	19
2.4	Exemples des deux premières valeurs du dictionnaire obtenue par LBG.	20
2.5	Exemples des deux premières valeurs du dictionnaire obtenue par LBG	20
2.6	Exemples des deux premières valeurs du dictionnaires obtenue par LBG	21
2.7	Valeur des LSP quantifiers pour les premiers trames pour $M=10$ et $M=16$	23
2.8	Valeur ISP quantifieres pour les deux premières trames pour $M=18$, et $M=20$	23

Introduction Générale

Ce mémoire étudie le codage des signaux à large bande. La transmission en large bande correspond à l'élargissement de la bande passante utilisée pour la transmission du signal de parole.

En effet, la bande passante utilisée habituellement en téléphonie est 200-3400 Hz, cependant les nouvelles technologies liées aux réseaux permettent une utilisation plus flexible de la transmission de la parole, grâce à un choix assez large de codeur et de bande passante, qui a facilité l'apparition d'une nouvelle bande passante 50-7000 Hz, améliorant la qualité du signal de parole transmis. L'étude de la qualité de la parole est un domaine important de la psychoacoustique pour ses applications dans la synthèse sonore, la médecine ou comme ici, pour la téléphonie.

Le développement des applications multimédia sur l'Internet ainsi que les systèmes de conférence téléphonique feraient bon usage d'un système adaptatif permettant de régler le niveau de qualité du codage selon le débit disponible. Cette étude propose une solution destinée à répondre à ce besoin.

Objectif du travail: L'objectif de notre étude est tout d'abord d'élargir la bande passante tout en optimisant le quantificateur, et avoir une distorsion spectrale minimale afin d'avoir une bonne transmission de signaux audio dans la gamme de fréquence voulue.

Nous avons organisé notre travail en deux chapitres :

Le premier chapitre Donne des notions de bases et principes du codage à large bande.

Le deuxième chapitre Nous avons comparé les résultats du codage de l'enveloppe spectrale pour différents ordres de prédiction $m=10,16,18$ et 20 pôles, à l'aide de la distorsion spectrale.

Partie I

Codage à Large Bande

Codage à large bande

Dans ce chapitre on va étudier le codage de signaux large bande. La transmission à large bande correspond à l'élargissement de la bande passante utilisée pour la transmission du signal de parole. En effet, la bande passante utilisée habituellement en téléphonie est 300-3400 Hz, elle définit le débit de base d'une ligne téléphonique qui est de 64 kbits/s. Cependant les nouvelles technologies liées aux réseaux permettent une utilisation plus flexible de la transmission de la parole, grâce à un choix assez large de codeur et de sa bande passante, qui a facilité l'apparition d'une nouvelle bande passante 50-7000 Hz, améliorant la qualité du signal de parole transmis. Le développement des applications multimédia sur l'Internet ainsi que les systèmes de conférence téléphonique feraient bon usage d'un système adaptatif permettant de régler le niveau de qualité du codage selon le débit disponible. Cette étude propose une solution destinée à répondre à ce besoin

Intérêts de l'évolution vers le codage large bande

La téléphonie peut être décrite comme un système de communication permettant la transmission de la parole. Elle peut, en premier lieu être vue comme un système créant une liaison entre deux personnes et permettant la transmission d'un message. Son but (comme celui de tout système comportant de la parole) devrait être de maximiser la compréhension de ce message. Le message étant défini comme un signal de parole transmis de la bouche d'un locuteur jusqu'à l'oreille d'un auditeur, le rôle des deux personnes changeant au cours de la conversation. Ainsi, la compréhension du message délivré par un locuteur dépend de : [2],[3]

- La compréhensibilité du message, liée directement au locuteur ou au système et à sa capacité à donner une information, de transmettre les phonèmes (souvent en fonction de son articulation, tout dépendant du contexte de locution).
- L'intelligibilité, qui correspond à la possibilité d'établir un sens au message transmis avec l'ensemble des phonèmes du message.
- La communicabilité, qui est la compréhension de l'ensemble des messages, dans les deux sens de la liaison.

Il est important de remarquer que, la compréhension du message dépend du locuteur, du contexte de locution, ainsi que des connaissances du sujet sur le message.

L'étude de la large bande va principalement porter sur le premier facteur. Or, pour connaître en quoi l'augmentation de la bande passante permet d'améliorer la qualité globale de la téléphonie, il est nécessaire d'étudier le message vocal transmis. Le prochain paragraphe porte donc sur l'étude de la parole : sa production par un locuteur et sa perception par un auditeur.

La production vocale

Tout d'abord, un son a une forme physique qui se propage dans un milieu par le biais d'ondes. Ces ondes sont liées au canal de transmission (air, câble) mais surtout au producteur de ce son. Elles peuvent alors être quantifiées sur une échelle de fréquences. Le signal de la parole a donc des caractéristiques temporelles mais aussi fréquentielles. Dans le cas de la parole humaine, les fréquences dépendent de la forme et de la position

de certains organes du corps humain. La parole peut alors être vue comme un signal source (corde vocales, glotte) qui est filtré par des tuyaux formés par les conduits vocaux (comme le conduit nasal). La manière dont le son d'origine est filtré dépend de la signification que veut lui donner le locuteur. Ainsi on peut voir que la production vocale est composée de sons ayant des composantes fréquentielles très spécifiques, dont on donne souvent comme valeur la fréquence fondamentale (f_0 , correspondant au signal porteur), et les premiers formants (f_i , piques dans l'amplitude spectrale dus aux résonances du conduit vocal). Certains phonèmes comme les voyelles se caractérisent très facilement par ses formants. Les consonnes sont produites de manières différentes [?], elles peuvent être sonores (« l », « r »), nasales (« m », « n ») ou fricatives (« h », « f »). Ces dernières, comme le « s » ou le « f », produisent de l'énergie essentiellement dans les hautes fréquences ainsi que dans un formant très bas, autour de 150 Hz. De même la résonance nasale se situe au alentour de 250 Hz.

Les fréquences utilisées par la parole humaine, peuvent donc être comprises entre 110 et 7 kHz (speech communications). La bande étroite 300-3400 Hz utilisée par la téléphonie permet de faire passer les 3 premiers formants, et ainsi de garantir une intelligibilité de la parole du locuteur, mais ne permet pas de transmettre l'intégralité des fréquences présentes dans un signal de parole. Par exemple, il est très difficile de différencier un « s » d'un « f » prononcé seul, lors d'une conversation téléphonique.

La perception de la parole

Il est nécessaire de rappeler que les fréquences audibles par une oreille humaine sont habituellement situées entre 20 Hz et 20 kHz : ce qui à première vu semble très loin de la bande étroite de la téléphonie. De plus, suite à de nombreuses utilisations, le cerveau humain, a créé une référence de la qualité « sonore » de la voix humaine, transmise à travers un système de téléphonie [?]. L'évaluation de la qualité est alors biaisée par la référence de la téléphonie fixe, fortement liée aux fréquences de la bande étroite. Afin de mieux appréhender le choix de la bande élargie pour la téléphonie, le paragraphe suivant rappelle la notion de « bande critique » ; Les bandes critiques correspondent à une répartition dans le spectre des fréquences d'un ensemble de bandes de fréquences. Ces bandes sont des regroupements des excitations sonores ayant des fréquences voisines et perceptivement proches au sein de certaines bandes fréquentielles. Il est possible de passer de l'échelle des fréquences à celle des bandes critiques grâce à la fonction suivante [6] :

$$z_{(barks)} = 13 \cdot \tan^{-1} \cdot (0.76 \cdot f_{(kHz)}) + 35 \cdot \tan^{-1} + \left(\frac{f_{(kHz)}}{7.5} \right)^2 \quad (1.1)$$

Cette échelle en bande critique est une échelle perceptive (correspondant à la Tonie), dont l'unité, le « Bark » regroupe un ensemble variable de fréquences, 1 Bark correspondant à une bande passante de 100 Hz à 3500 Hz. Une bande passante peut alors être obtenue en Barks :

$$z_{bp} = z(f_h) - z(f_b) \quad (1.2)$$

Les fréquences audibles vont de 0 (20 Hz) jusqu'à 24 Barks (16 kHz). La téléphonie bande étroite représente 14 Barks (de 3 à 16 Barks), soit plus de la moitié de l'échelle audible. L'utilisation des bandes critiques permet de connaître la valeur perceptive d'une bande passante. Par exemple, dans [5] une bande passante de 180-2800 Hz est considérée de même qualité qu'une seconde bande passante de 280-3550 Hz. Un simple calcul suffit pour voir que la différence en Barks de ces deux bandes passantes est 0, car les deux fréquences de coupures subissent une simple translation d'environ 1 Bark.

Cette échelle permet alors d'analyser le choix de la bande étroite en téléphonie classique et l'apport du large bande. La fréquence de coupure basse, baisse de 300 Hz jusqu'à 50 Hz, soit une augmentation de la bande passante de 2,5 Barks, tandis que la fréquence de coupure haute, permet une augmentation de 4 Barks.

La fréquence centrale, permet également de connaître le poids de fréquences basses et hautes dans les deux bandes passantes.

$$f_c = \sqrt{f_b \cdot f_h} \quad (1.3)$$

Pour la bande étroite : $f_c = 1010$ Hz, et pour la bande élargie : $f_c = 590$ Hz.

Ces deux valeurs montrent que la bande élargie comporte perceptivement plus de basse fréquence que la bande étroite. Cette description des bandes passantes par le biais de la fréquence centrale et de la largeur spectrale en Barks est utilisée dans [1], afin de lier la perception fréquentielle à la qualité d'un signal de parole. Par ailleurs, il a été vu précédemment que l'intelligibilité était fortement liée aux premiers formants. Dans [5] il est montré que la sensation naturelle de la voix dépend fortement du premier formant et l'intelligibilité du second formant. En effet, les différents phonèmes sont perçus en fonction :

- Du rapport entre les formants.
- De leurs variations dans le temps.

De plus, le premier formant étant approximativement entre 270 et 730 Hz pour les hommes, et entre 310 et 850 Hz pour les femmes, le choix de la fréquence de coupure basse à 300 Hz a été choisi judicieusement. La bande étroite permet un bon compromis entre intelligibilité et qualité du son.

Pour autant, celle-ci ne permet pas de transmettre la fréquence fondamentale f_0 , qui semble être liée à la sensation naturelle de la voix. Celle-ci comprise entre 110 et 200 Hz pour un adulte, et montant jusque 300 Hz pour un enfant, permet de transmettre la prosodie, comme les intonations ou les émotions.

Théoriquement, la perception humaine permet de reconstruire cette fréquence fondamentale en son absence, et de percevoir tout de même les différentes intonations exprimées par le locuteur. Malgré cela, dans [5] on voit également que la perception de la parole à travers un système de téléphonie dépend énormément de la présence des fréquences basses ; Une petite différence de 45 Hz, de 225 à 180 Hz, sur la fréquence de coupure basse améliore nettement la sensation naturelle de la voix. La perception de la fréquence fondamentale semble donc avoir une importance dans l'évaluation de la qualité de la parole.

Enfin, lors de l'augmentation de la fréquence de coupure basse de 123 à 208 Hz, une dégradation est ressentie sur la perception de la voix humaine, celle-ci semble moins naturelle. Une dégradation est obtenue également lors d'une diminution de la fréquence de coupure de 5500 à 3500 Hz. Mais il montre également qu'il est nécessaire d'améliorer la bande étroite aux deux extrémités. En effet, il montre que pour une fréquence de coupure basse proche de 300 Hz, le changement de la fréquence de coupure haute (de 7000 à 3500 Hz) n'a que peu d'effets. De même, pour une fréquence de coupure haute à 3500 Hz, un changement de la fréquence basse de 55 à 300 Hz a peu d'effets également. Il est donc impossible de compenser une coupure trop forte, en haut ou en bas spectre, en agrandissant l'autre côté.

La bande étroite introduit donc une dégradation de la sensation naturelle de la voix par :

- L'atténuation du premier formant.
- L'absence de transmission de f_0
- L'absence de transmission des hautes fréquences.

La parole humaine produit des fréquences qui en partie ne sont pas comprises dans la bande étroite, et qui sont nécessaires pour obtenir une voix humaine naturelle. La téléphonie large bande, qui permet la transmission de la majorité des fréquences produites par la voix, nous espérons que celle-ci permettra de rendre la voix d'un interlocuteur plus naturelle.

L'oreille ne peut percevoir que certains sons. Fig 1.1 donne une représentation du domaine audible pour un être humain. On remarque tout d'abord que le niveau de perception dépend grandement de la plage de fréquences considérée ainsi que du niveau sonore. On définit alors deux courbes dans le plan fréquence/intensité : un seuil d'audibilité et un seuil de confort. La zone ainsi définie est le domaine dans lequel les sons peuvent être perçus. Tout signal en dehors de cette plage est inaudible, gênant ou même dangereux [2].

La bande d'audition est composée des fréquences de 20 Hz à 20 000 Hz. En pratique une telle largeur de bande n'est conservée que pour un codage de très haute fidélité (qualité CD). Selon la nature du signal à coder (parole ou musique) on filtre le signal en sélectionnant soit la bande téléphonique, suffisante pour la parole, soit une bande plus large pour traiter des sons plus complexes.

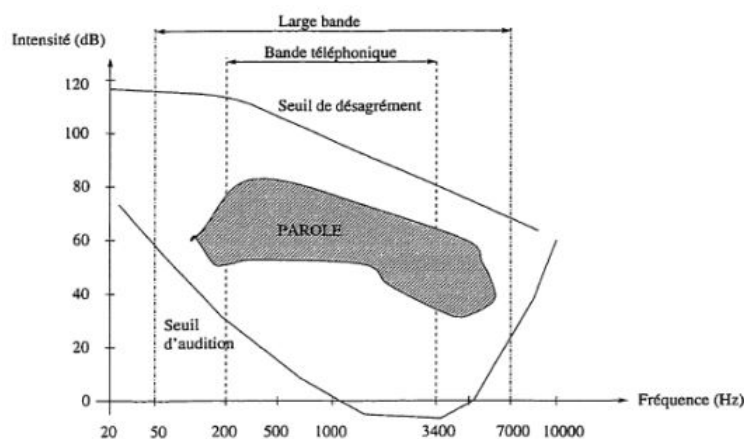


Fig 1.1: Perception auditive

Codage large bande

Le codage large bande considère théoriquement aussi bien les signaux de parole que les signaux de musique. Les techniques de codage peuvent alors différer selon le type d'application du codeur. On sait que le premier système large bande est apparu en 1988, ce type de transmission est aujourd'hui très peu répandu. En effet, il a fallu une suite d'évolution dans les télécommunications pour permettre l'utilisation de CoDec large bande dans l'industrie.

Codage de la parole large bande

L'agrandissement de la bande passante pour la parole n'apporte pas grand chose du point de vu de l'information. Contrairement à la musique, où la bande élargie peut comporter des phénomènes supplémentaires (notes aigües), l'information de parole (l'intelligibilité) est intégralement contenue dans la bande téléphonique. On peut néanmoins espérer deux améliorations :

- Pour les phonèmes voisés, l'addition de la bande de fréquences 50 Hz - 140 Hz donne une meilleure représentation des premières harmoniques. On remarque surtout cela pour un locuteur masculin pour lequel la fréquence de pitch est assez faible. D'une manière plus générale les basses fréquences procurent une sensation de confort et un sentiment de parler « face à face ».
- L'apport des hautes fréquences, supérieures à 3400 Hz, n'a de l'importance que pour les phonèmes plus complexes tels que les fricatives non voisées (ex: « S », « CH », « F »), les fricatives voisées (ex: « Z ») ou encore Les plosives (ex: « T », « D »).

Un codeur large bande optimisé pour la parole pourrait prendre soin de bien représenter les formants ainsi que la structure harmonique en limitant le codage des hautes fréquences lorsque le son est voisé. En revanche, un débit plus conséquent pourrait être attribué à la partie supérieure du spectre lorsque le phonème est non voisé ou composé.

Codage de la musique en large bande

La musique n'est intéressante à coder que lorsque l'on dispose d'une largeur de bande suffisante. Le codage large bande permet d'offrir une telle qualité. Contrairement à la parole, il n'existe pas réellement de modèle permettant de représenter le signal. En revanche les sons en musique sont beaucoup plus stationnaires que les phonèmes en parole. Pour cette raison, on est porté à utiliser des trames d'analyse plus grandes qu'en parole. En pratique on peut travailler avec des blocs d'au moins 20 ms.

Comme il vient d'être précisé, il est difficile de prévoir à l'avance l'allure de l'enveloppe spectrale. On peut cependant admettre que la musique est une combinaison de bruit et de "tons". Un ton pur est une concentration d'énergie sur une raie spectrale donnée, avec un plancher de bruit faible. A titre d'exemple, une note de musique isolée conduit à un spectre comprenant seulement un ton pur, localisé à la fréquence de la note.

Les codeurs large bande

Idéalement un codeur large bande doit pouvoir traiter sans préférences aussi bien les sources de parole que celles de musique. Deux approches sont alors possibles: soit on améliore un codeur de parole (type ACELP par exemple) pour qu'il traite au mieux la musique, soit on part d'un codeur mieux conçu pour la musique (codage par transformée type TCX par exemple) que l'on adapte afin de mieux coder la parole. La seconde stratégie semble plus prometteuse dans la mesure où un codeur ACELP est conçu presque exclusivement pour la parole (pour la musique, le débit consacré au pitch est parfois du gaspillage) tandis qu'un codeur par transformée assure toujours une contribution minimale quelle que soit la source considérée.

Il n'existe pas encore beaucoup de normes en large bande pour le moment. La référence à considérer est encore la norme UIT G.722. Ce codeur est un codeur de forme d'onde temporelle de type ADPCM. Il utilise un débit de 45 kbits/s à 64 kbits/s. Une seconde norme devrait bientôt pouvoir remplacer ce dernier. Les nouveaux débits à considérer seront probablement 16 kbits/s à 32 kbits/s. La technique de codage utilisée est cette fois un codage par transformée MLT avec un codage entropique sur les indices de quantification en bout de ligne (un peu comme MPEG1 layer Le tableau 1.1 donne une brève description des deux codeurs large bande qui viennent d'être évoqués.

codeur	G.722	Nouveau large bande G.7XX
Année	1988	1998
Débit	3modes: 48 Kb/s,56Kb/s et 64 Kb/s	3 modes: 16 Kb/s,24 Kbs et 32 Kb/s
Délai	0,125 ms(+1,5 ms lookahead)	20 ms (+20 ms lookahead).
Modèle	<ul style="list-style-type: none"> • codage en deux sous-bandes(QMF) • <i>bullet</i> Codage ADPCM d'ordre 4 dans chacune des bandes • Hautes fréquences quantifiés à 2 bits par échantillon • Basses fréquences quantifiées à 3 bits, 4 bits ou 5 bits par échantillon selon le débit choisi, les trois modes étant encapsulés 	<ul style="list-style-type: none"> • Codage par transformé MTL • Quantification scalaire des raies de la transformée • Attribution du budget par catégorisation selon les bandes de fréquence • Codage entropique (huffman) sur les indices de quantification

Tableau 1.1: codeurs à large bande

Conclusion

Le codage large bande n'est pas un domaine d'étude très récent, le premier CoDec utilisant une bande passante élargie, le ITU G.722 , fut développé dans les années 1980 pour être utilisé sur le réseau RNIS. Mais le codage de la parole en large bande entraîne des techniques différentes de la bande étroite. En effet, il y a une plus forte dynamique spectrale pour la parole en bande large. De plus la voix est plus inharmonique dans les hautes fréquences, comme pour les fricatives, en raison des caractéristiques morphologiques. Mais, plusieurs études, efforts de développement et standardisation ont permis de créer quelques CoDecs de meilleure qualité et moins coûteux en débit, dans la suite de notre travail, on va s'intéresser aux codeurs utilisant le principe du codage de l'enveloppe spectrale, c'est ce qui va être abordé dans le prochain paragraphe.

Partie II

Résultats et Simulations

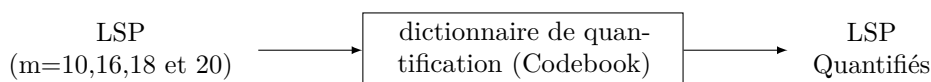
Résultats et simulations

L'objectif de nos tests et simulation est d'étudier le codage large bande. Ces codeurs, du type d'analyse-par-synthèse, doivent transmettre des signaux possédant une gamme de fréquences limitée à 8 kHz et ayant été échantillonnés à 16 kHz. Pour cela on va procéder de la manière suivante :

- En premier lieu, on code l'enveloppe spectrale d'un signal possédant une large bande. L'étude se fait à l'aide des Lignes de Prédiction Spectrales (LSP), qui sont obtenus par l'analyse LPC :
 - Extraction des coefficients a_i .
 - conversion des a_i en coefficients LSP (Lineare Spectral Prediction)

L'extraction des LSP se fait pour un ordre de prédiction (m) variable entre 10 et 20 pôles (m=10 pôles, m=16 pôles, m=18 pôles, m=20 pôles).

- Application de la méthode LBG sur les Lignes de Prédiction Spectrales (LSP), afin d'obtenir le dictionnaire de quantification (Code- Book) pour m=10, 16, 18, et 20 pôles.
- Faire passer les coefficients LSP par le dictionnaire de quantification, pour extraire les LSP quantifiés pour m=10, 16, 18, et 20.



- Calculer la distorsions spectrale entre les LSP et les LSP quantifiés en fonction du nombre de bite qui vari entre (40bits-80 bits).
- Pour la partie programmation on a utilisé Matlab pour les représentations des graphes.

Conditions d'analyse

Le signal utilisé dans notre simulation est signal parole qu'on échantillonnera avec une fréquence $f_e=16$ KHz, afin de satisfaire la condition de Schanon ($f_e \geq 2BP$).

Le signal échantillonné ainsi obtenu sera découpé en trames de 480 échantillons, dans notre exemple on aura 27 trames.

Le signal parole utilisé dans notre simulation est schématisé à l'aide de Matlab comme suit :

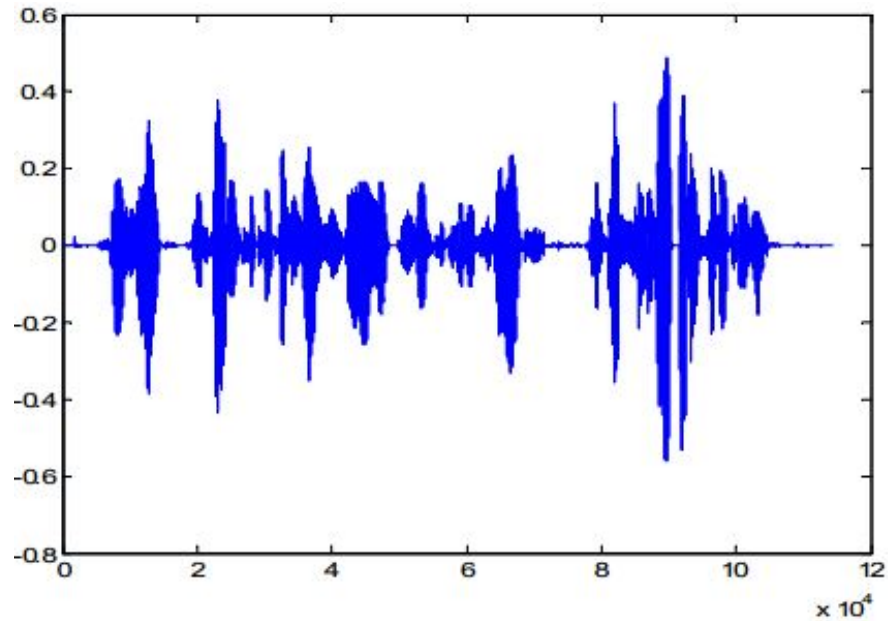


Fig 2.1: signal parole utilisé

L'analyse LPC

Après avoir échantillonné le signal parole, ce dernier va être sujet de l'analyse LPC afin d'extraire les coefficients a_i et cela pour un ordre de prédiction de 10, 16, 18 et 20 pôles.

Le tableau suivant donne les valeurs des a_i pour les deux premières trames de notre signal parole déjà échantillonné dans la première partie de notre simulation.

Coef	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	
10	1.000000	-0.287399	-0.240764	0.233074	0.123018	-0.024625	0.018431	-0.17185	0.109583	-0.031213											
16	1.000000	0.024929	-0.0235596	-0.265476	-0.011110	-0.205724	0.155493	0.092839	-0.127162	0.200685	0.149240	0.014280	-0.181261	-0.007304	-0.006202	0.043103					
18	1.000000	0.024799	-0.243996	-0.260993	-0.011367	-0.142888	-0.204687	0.169360	-0.139902	0.209183	0.164311	-0.00510	-0.193058	-0.007475	-0.007464	0.018111	-0.111816	-0.000542			
20	1.000000	-0.38247	-0.259045	-0.257825	-0.066038	-0.013869	-0.172994	0.071197	0.134452	-0.079627	0.110242	0.043871	0.011563	-0.130037	-0.17304	0.054208	-0.029663	-0.022421	0.032027	0.017209	

Tableau 2.1: Valeurs des (a_i) pour les premières trames pour $m=10,16,18$ & 20 pôles.

Représentation des LSP

Afin d'obtenir une précision optimale, on utilise les coefficients LSP (les paires de raies spectrales) qui est la représentation la plus répandue des coefficients a_i , afin de stabiliser le filtre de synthèse LP.

Dans les tableaux suivants nous avons donné les valeurs des LSP des deux premières trames de notre signal pour $m=10, 16, 18$ et 20 pôles:

2.3. Codage de l'enveloppe spectral

Coef	lsp_1	lsp_2	lsp_3	lsp_4	lsp_5	lsp_6	lsp_7	lsp_8	lsp_9	lsp_{10}	lsp_{11}	lsp_{12}	lsp_{13}	lsp_{14}	lsp_{15}	lsp_{16}	lsp_{17}	lsp_{18}	lsp_{19}	lsp_{20}	
10	0.949780	0.865249	0.701371	0.525716	0.247921	-0.088602	-0.418647	-0.663515	-0.875379	-0.56497											
16	0.991562	0.959491	0.868061	0.776024	0.702251	0.42179	0.258875	0.152963	-0.012120	-0.204207	-0.391750	0.489356	-0.756081	-0.886430	-0.933906	-0.988969					
18	0.991148	0.963702	0.921106	0.814724	0.665668	0.566103	0.40089	0.228094	0.093596	-0.042830	-0.348309	-0.460598	-0.567045	-0.660782	-0.7680334	-0.890212	-0.948920	-0.982302			
20	0.991672	0.97090	0.926047	0.861874	0.746488	0.633356	0.522155	0.385774	0.221405	0.085516	0.0788037	-0.290370	-0.415835	-0.518835	-0.61832	-0.709264	-0.823632	-0.908868	-0.954786	-0.985187	

Tableau 2.2: (b) Valeur des (LSP_i) pour les premières trames pour M=10,16,18& 20 pôles.

Codage de l'enveloppe spectral

Principes

Le principe du codage de l'enveloppe spectrale est basé sur la quantification des lignes de fréquences spectrales (LSF), afin de transmettre l'information de cette dernière à large bande. Pour avoir une idée sur l'enveloppe spectrale, on a va schématiser cette dernière on utilisant les LSF obtenus auparavant :

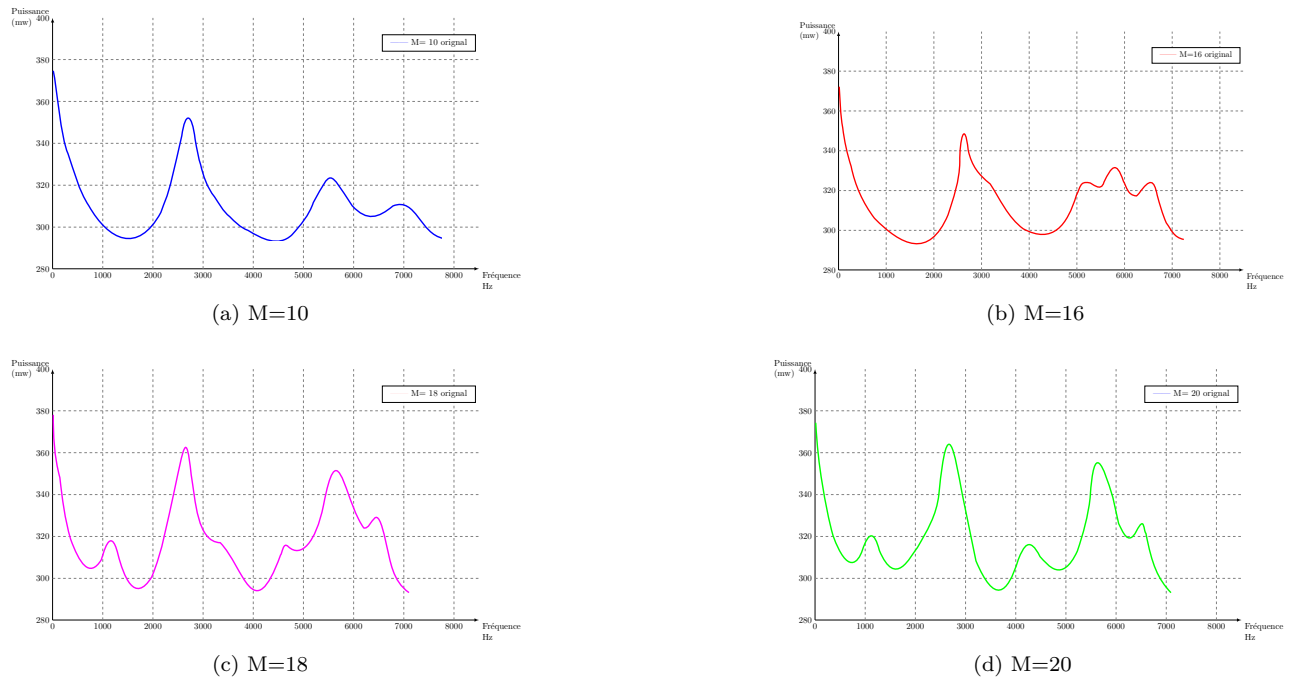


Fig 2.2: L'enveloppe spectrale des LSF pour M=10, M=16, M=18 et M=20 pôles

Création du dictionnaire de quantification

Afin de pouvoir quantifier les coefficients LSP on crée un dictionnaire optimal par l'application de la méthode LBG (Algorithme de LINDE BUSO GRAY).

Dans le but de faire une comparaison précise, on fait varier le nombre de bits par trame pour chaque m (10, 16,18 et 20 pôles).

Dans notre cas on a pris pour :

- **m=10**

On trouver le dictionnaire de quantification pour :

2.3. Codage de l'enveloppe spectral

- 40bits par trame. (Dictionnaire de 16entrées).
- 60bits par trame. (Dictionnaire de 64entrées).
- 80bits par trame. (Dictionnaire de 256entrées).

On cite comme exemple les deux premières valeurs du code book (dictionnaire) obtenu pour les $m=10, 16, 18$ et 20 pôles dans les tableaux suivants :

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9521	0.9480
Dico2	0.8770	0.8768
Dico3	0.203	0.7384
Dico4	0.5420	0.5243
Dico5	0.2659	0.2604
Dico6	-0.0147	0.528
Dico7	-0.3878	-0.3891
Dico8	-0.7045	-0.6935
Dico9	-0.8652	-0.8671
Dico10	-0.9625	-0.9653

(a) 40 bits/trame

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9498	0.9455
Dico2	0.8652	0.8697
Dico3	0.7014	0.7512
Dico4	0.5257	0.5354
Dico5	0.2479	0.2510
Dico6	-0.0886	-0.0467
Dico7	-0.4186	-0.3853
Dico8	-0.6635	-0.6901
Dico9	-0.8754	-0.8644
Dico10	-0.9565	-0.9643

(b) 60 bits/trame

Dico	1 ^{ème} valeur	2 ^{ème} valeur
Dico1	0.498	0.9465
Dico2	0.852	0.8726
Dico3	0.7014	0.7465
Dico4	0.5257	0.5274
Dico5	0.2479	0.2503
Dico6	-0.0886	-0.0456
Dico7	-0.4186	-0.3867
Dico8	-0.6635	-0.6914
Dico9	-0.8754	0.843
Dico10	-0.9565	-0.9643

(c) 80 bits/trame

Tableau 2.3: Exemples des deux premières valeurs du dictionnaire obtenu par LBG

- $m=16$

On trouver le dictionnaire de quantification pour :

- 48bits par trame. (Dictionnaire de 8entrées).
- 64bits par trame. (Dictionnaire de 16entrées).
- 80bits par trame. (Dictionnaire de 32entrées).

2.3. Codage de l'enveloppe spectral

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9921	0.9906
Dico2	0.9651	0.9614
Dico3	0.9341	0.9139
Dico4	0.7153	0.7664
Dico5	0.5752	0.6112
Dico6	0.4278	0.4421
Dico7	0.2168	0.2329
Dico8	0.1186	0.1221
Dico9	-0.0607	-0.0408
Dico10	-0.4211	-0.4110
Dico11	-0.5026	-0.4989
Dico12	-0.6231	-0.6182
Dico13	-0.7081	-0.7062
Dico14	-0.8466	-0.8438
Dico15	-0.9428	-0.9419
Dico16	-0.9819	-0.9814

(a) 48 bits/trame

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9927	0.9905
Dico2	0.9648	0.9604
Dico3	0.9400	0.9153
Dico4	0.7073	0.7714
Dico5	0.5577	0.5979
Dico6	0.4322	0.446
Dico7	0.2101	0.2372
Dico8	0.1172	0.1214
Dico9	0.0703	-0.0378
Dico10	-0.4202	-0.4097
Dico11	-0.5048	-0.4996
Dico12	-0.223	-0.6175
Dico13	-0.7084	-0.7055
Dico14	-0.8452	-0.8425
Dico15	-0.9438	-0.9425
Dico16	-0.9438	-0.9425

(b) 64 bits/trame

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9928	0.9908
Dico2	0.9648	0.615
Dico3	0.9416	0.9178
Dico4	0.7041	0.7601
Dico5	0.5536	0.5970
Dico6	0.4333	0.4471
Dico7	0.2082	0.2340
Dico8	0.1169	0.1213
Dico9	-0.0733	-0.0400
Dico10	-0.4200	-0.4119
Dico11	-0.5053	-0.5000
Dico12	-0.6221	-0.6183
Dico13	-0.7086	-0.7059
Dico14	-0.8449	-0.8436
Dico15	-0.9440	-0.9424
Dico16	-0.920	-0.9814

(c) 80 bits/trame

Tableau 2.4: Exemples des deux premières valeurs du dictionnaire obtenue par LBG.

m=18 :

On trouver le dictionnaire de quantification pour :

- 54bits par trame. (Dictionnaire de 8 entrées).
- 72bits par trame. (Dictionnaire de 16 entrées).

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9932	0.9921
Dico2	0.99690	0.9675
Dico3	0.9415	0.9308
Dico4	0.8449	0.8597
Dico5	0.6348	0.6551
Dico6	0.5584	0.5498
Dico7	0.3467	0.3723
Dico8	0.2254	0.2281
Dico9	0.0973	0.1017
Dico10	-0.0034	-0.0069
Dico11	-0.4027	0.3808
Dico12	-0.4741	-0.4743
Dico13	-0.5860	-0.5747
Dico14	-0.6673	-0.6670
Dico15	-0.7678	-0.7585
Dico16	-0.8861	-0.8862
Dico17	-0.9532	-0.9528
Dico18	-0.9849	-0.9850

(a) 54 bits/trame

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9932	0.9923
Dico2	0.9691	0.9677
Dico3	0.9410	0.9319
Dico4	0.402	0.8483
Dico5	0.6416	0.6630
Dico6	0.5609	0.5600
Dico7	0.3482	0.3695
Dico8	0.2250	0.2277
Dico9	0.0989	0.1036
Dico10	-0.0035	-0.004
Dico11	-0.4017	-0.3855
Dico12	-0.4737	-0.4734
Dico13	-0.5854	-0.5766
Dico14	-0.6673	-0.66677
Dico15	-0.7678	-0.7611
Dico16	-0.8859	-0.8861
Dico17	-0.9530	-0.9524
Dico18	-0.9849	-0.9848

(b) 72 bits/trame

Tableau 2.5: Exemples des deux premières valeurs du dictionnaire obtenue par LBG

m=20 :

On trouver le dictionnaire de quantification pour :

- 40bits par trame. (Dictionnaire de 4 entrées).
- 60bits par trame. (Dictionnaire de 8 entrées).
- 80bits par trame. (Dictionnaire de 16 entrées).

2.3. Codage de l'enveloppe spectral

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.927	0.9917
Dico2	0.9712	0.9676
Dico3	0.9338	0.9270
Dico4	0.8809	0.8674
Dico5	0.7524	0.7783
Dico6	0.6220	0.6559
Dico7	0.5092	0.5241
Dico8	0.3639	0.3812
Dico9	0.02189	0.2293
Dico10	0.0953	0.1018
Dico11	-0.0614	-0.534
Dico12	-0.3128	-0.2986
Dico13	-0.4232	-0.4192
Dico14	-0.5267	-0.5214
Dico15	-0.6209	-0.6186
Dico16	-0.7108	-0.7078
Dico17	-0.9081	-0.9070
Dico18	-0.9081	-0.9070
Dico19	-0.9570	-0.9564
Dico20	-0.9875	-0.9870

(a) 40 bits/trame

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	à.9925	0.9920
Dico2	0.9713	0.9683
Dico3	0.9713	0.9688
Dico4	0.8866	0.849
Dico5	0.7564	0.7707
Dico6	0.6104	0.6671
Dico7	0.5096	0.5277
Dico8	0.3645	0.3774
Dico9	0.2211	0.2271
Dico10	0.0895	0.1051
Dico11	-0.0599	-0.0519
Dico12	-0.3094	-0.3028
Dico13	-0.4235	-0.4187
Dico14	-0.5264	0.5219
Dico15	-0.6207	-0.6185
Dico16	-0.7097	-0.7090
Dico17	0.8094	-0.8109
Dico18	-0.9088	-0.9064
Dico19	-0.9573	-0.9563
Dico20	-0.9875	-0.9871

(b) 60 bits/trame

Dico	1 ^{ère} valeur	2 ^{ème} valeur
Dico1	0.9926	0.9922
Dico2	0.9715	0.9685
Dico3	0.9340	0.267
Dico4	0.8871	0.8607
Dico5	0.7532	0.7639
Dico6	0.6102	0.6576
Dico7	0.5093	0.5375
Dico8	0.3638	0.3841
Dico9	0.2207	0.2278
Dico10	0.0901	0.1058
Dico11	-0.0597	-0.0488
Dico12	-0.3108	-0.3009
Dico13	-0.4238	-0.4184
Dico14	-0.5268	-0.5207
Dico15	-0.6208	-0.6181
Dico16	-0.7100	-0.7084
Dico17	-0.8093	-0.8115
Dico18	-0.9088	-0.9061
Dico19	0.9573	-0.9559
Dico20	-0.9876	-0.9868

(c) 80 bits/trame

Tableau 2.6: Exemples des deux premières valeurs du dictionnaires obtenue par LBG

Quantification des coefficients LSP

Dans cette étape on introduit les LSP pour (m=10, 16, 18,20 pôles) dans un quantificateur qui a comme code book les dictionnaires obtenus.

Les coefficients LSP vont subir une quantification vectorielle afin de réduire le nombre de bits à envoyer. A partir d'un dictionnaire déjà trouvé, on cherche des indices qui minimisent l'erreur quadratique de la manière suivante : **Pour m=10 :**

- Dans un dictionnaire de 16, 64, et 256 entrées trouvé. dans le paragraphe 2.3.2 et de dimension 10, chercher une entrée qui se rapproche des coefficients LSP « $\text{Min}(\text{ErrQuadVlsp}, T[\text{indice}])$ ».
- Coder l'indice i trouvé sur 4,6, et 8 bits pour l'envoyer au décodeur.

Et voici le schéma qui explique le procédé :

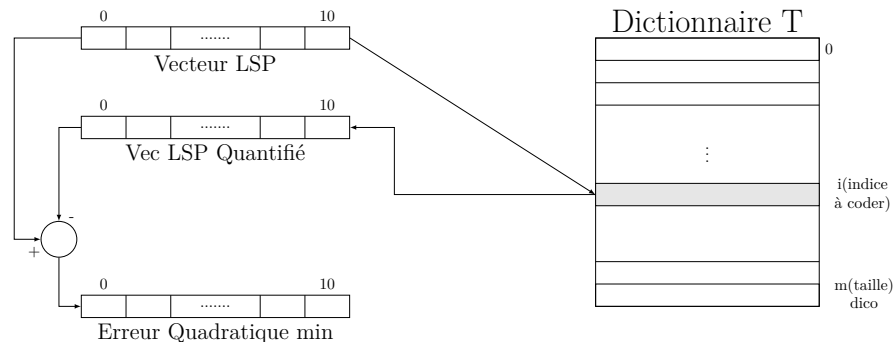


Fig 2.3: Quantification Vectorielle des coefficients LSP pour m=10 pôles

On refait la même étape pour $m=16, 18$ et 20 pôles, on change les dictionnaires utilisés (déjà trouvé) et on introduisant les changements suivants :

Pour $m=16$:

- Dans un dictionnaire de 8, 16, et 32 entrées trouvé dans le paragraphe 2.3.2 et de dimension 16, chercher une entrée qui se rapproche des coefficients LSP « $\text{Min}(\text{ErrQuadVlsp}, T[\text{indice}])$ ».
- Coder l'indice i trouvé sur 3, 4 et 5 bits pour l'envoyer au décodeur.

Et voici le schéma qui explique le procédé :

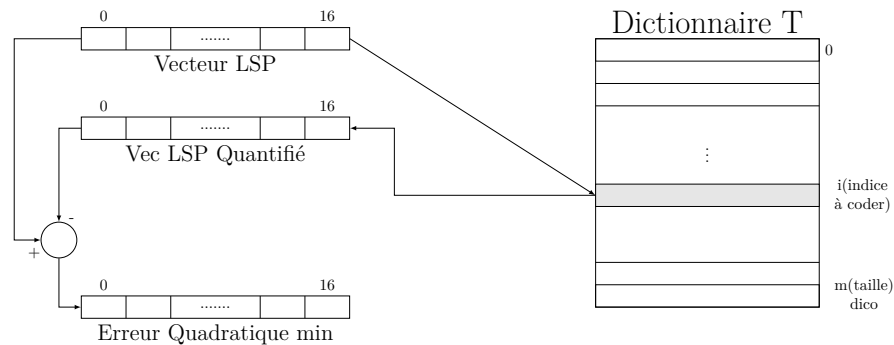


Fig 2.4: Quantification Vectorielle des coefficients LSP pour $m=16$ pôles.

Pour $m=18$:

- Dans un dictionnaire de 8 et 16 entrées trouvé dans le paragraphe 2.3.2 et de dimension 18, chercher une entrée qui se rapproche des coefficients LSP « $\text{Min}(\text{ErrQuadVlsp}, T[\text{indice}])$ ».
- Coder l'indice i trouvé sur 3, 4 et 5 bits pour l'envoyer au décodeur.

Et voici le schéma qui explique le procédé :

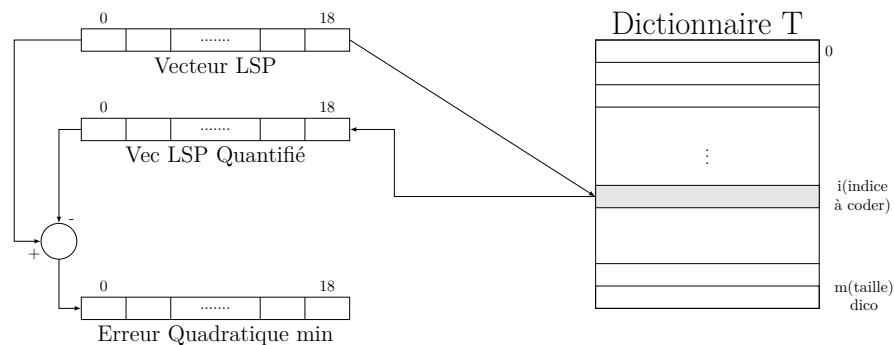


Fig 2.5: Quantification Vectorielle des coefficients LSP pour $m=18$ pôles.

Pour $m=20$:

- Dans un dictionnaire de 4, 8, et 16 entrées trouvé dans le paragraphe 2.3.2 et de dimension 20, chercher une entrée qui se rapproche des coefficients LSP « $\text{Min}(\text{ErrQuadVlsp}, T[\text{indice}])$ ».
- Coder l'indice i trouvé sur 2, 3 et 4 bits pour l'envoyer au décodeur.

2.3. Codage de l'enveloppe spectral

Et voici le schéma qui explique le procédé :

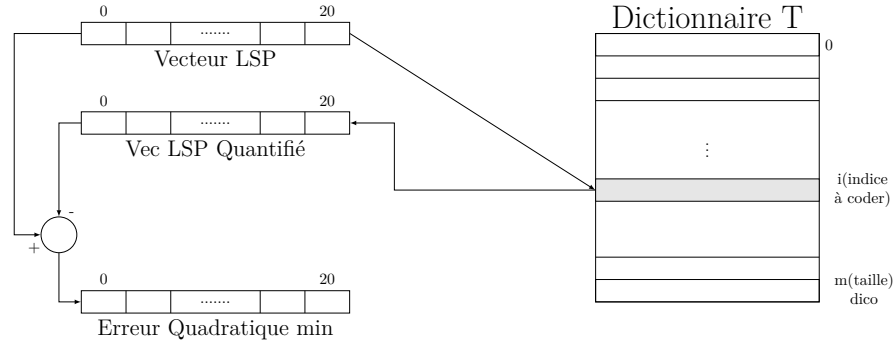


Fig 2.6: Quantification Vectorielle des coefficients LSP pour $m=20$ pôles.

Le tableau ci dessous, donne à titre d'exemple les deux premières trames des LSF quantifiées par la méthode de quantification énoncée ci dessus en utilisant le premier dictionnaire obtenu pour chaque ordre de prédiction.

M=10	Trame1	Trame2
LSP1q	0.994200	0.914089
LSP2q	0.024788	0.814741
LSP3q	-0.2396865	0.813419
LSP4q	-0.267542	0.542672
LSP5q	-0.011459	0.223811
LSP6q	-0.124534	0.073174
LSP7q	-0.205724	-0.352625
LSP8q	0.156676	-0.83512
LSP9q	0.097537	-0.817702
LSP10q	-0.956497	-0.903347

(a) M=10

M=16	Trame1	trame2
LSP1q	0.991562	0.911949
LSP2q	0.967922	0.915410
LSP3q	0.923356	0.912468
LSP4q	0.865425	0.848075
LSP5q	0.745355	0.803351
LSP6q	0.636547	0.623979
LSP7q	0.524999	0.139223
LSP8q	0.382963	0.139245
LSP9q	-0.222120	0.007498
LSP10q	-0.084207	-0.70638
LSP11q	-0.071750	-0.458140
LSP12q	-0.299356	-0.254369
LSP13q	-0.416081	-0.841249
LSP14q	-0.516430	-0.847229
LSP15q	-0.613906	-0.921217
LSP16q	-0.78969	-0.927561

(b) M=16

Tableau 2.7: Valeur des LSP quantifiers pour les premiers trames pour M=10 et M=16.

M=18	Trame1	Trame2
LSP1q	0.961148	0.967548
LSP2q	0.963702	0.976543
LSP3q	0.961106	0.943216
LSP4q	0.854724	0.884359
LSP5q	0.695668	0.687654
LSP6q	0.508103	0.564328
LSP7q	0.522089	0.398765
LSP8q	0.201094	0.276547
LSP9q	0.099596	0.094783
LSP10q	-0.042830	-0.006132
LSP11q	-0.380989	-0.434565
LSP12q	-0.460598	-0.489659
LSP13q	-0.590544	-0.234589
LSP14q	-0.626547	-0.665434
LSP15q	-0.700439	-0.787654
LSP16q	-0.810988	-0.887660
LSP17q	-0.934549	-0.965434
LSP18q	-0.912345	-0.998765

(a) M=18

M=20	Trame1	Trame3
LSP1q	0.987654	0.982739
LSP2q	0.909874	0.998765
LSP3q	0.998765	0.934987
LSP4q	0.898765	0.878765
LSP5q	0.796547	0.758765
LSP6q	0.667654	0.623454
LSP7q	0.423247	0.324450
LSP8q	0.398765	0.398765
LSP9q	0.123459	0.876543
LSP10q	0.012345	à.23345
LSP11q	-0.987654	0.544328
LSP12q	-0.298564	-0.397654
LSP13q	-0.498765	-0.435423
LSP14q	-0.899495	-à.544235
LSP15q	-0.432233	-0.609887
LSP16q	-0.987654	-0.797654
LSP17q	-0.854323	-0.988764
LSP18q	-0.909877	-0.999876
LSP19q	-0.967548	-0.912334
LSP20q	-0.967548	-0.912334

(b) M=20

Tableau 2.8: Valeur LSP quantifiers pour les deux premières trames pour M=18, et M=20

Après avoir quantifié les LSF on utilisant le dictionnaire du paragraphe 2.3.2, pour $m=10,16,18$ et 20 pôles. On va schématiser l'enveloppe spectrale 2.7 des LSF et des LSF quantifiés dans un même graphe afin de faire une analyse comparative de la performance du codage large bande pour différents ordres de prédiction .

2.4. Calcul de la distorsion spectrale

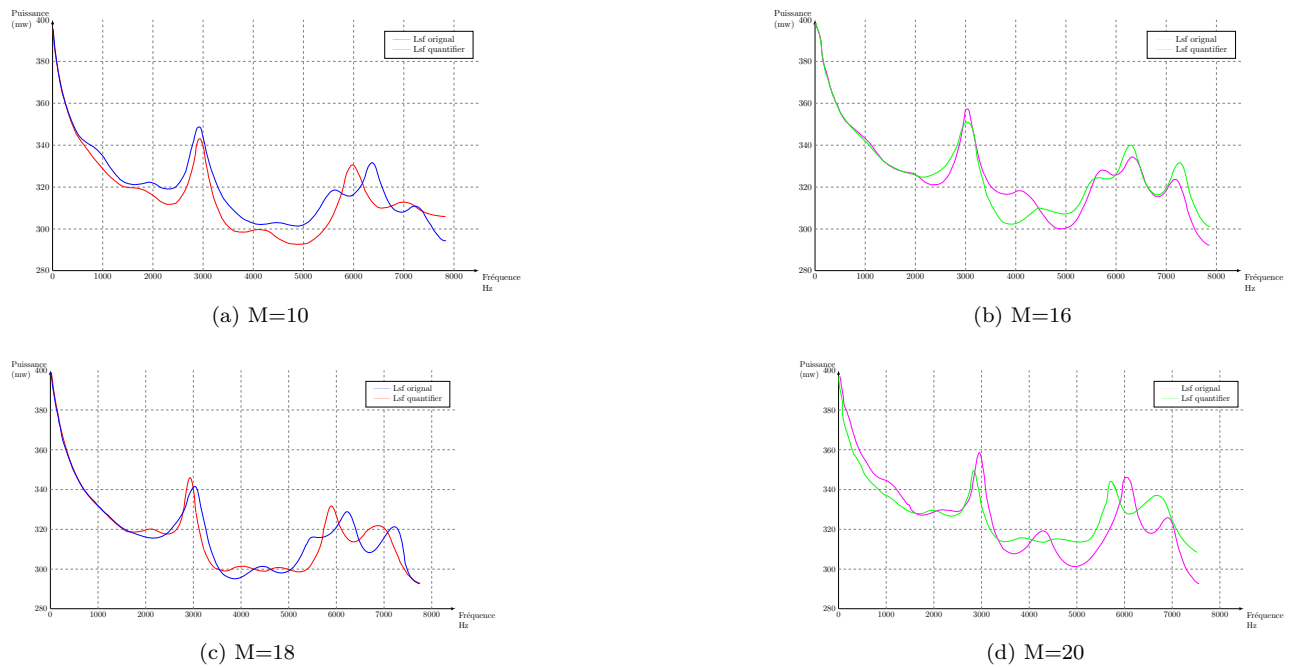


Fig 2.7: Enveloppe spectrales des LSF et LSF quantifié pour M=10,16,18 et 20 pôles

Interprétations et commentaires

En analysant la figure 2.7 on constate que l'enveloppe spectrale des LSF originales et des LSF quantifiées se rapprochent dans la plage de 0-4000 Hz (bande étroite) et cela pour les différents ordres de prédiction, dépassé cette plage là (c.a.d dans la bande élargie :4000-8000 Hz) on remarque que la différence entre l'enveloppe originale et quantifiée augmente pour un ordre de prédiction de 10 et 20 pôles ce qui provoque la perte de l'authenticité du signal original entraînant une diminution de la performance du quantificateur. Quant aux ordres de prédiction de 16 et 18 pôles l'enveloppe spectrale quantifiée garde la même forme que l'enveloppe spectrale originale dans la bande élargie augmentant ainsi la performance du quantificateur.

Calcul de la distorsion spectrale

Après l'extraction des LSP quantifiés, on a calculé la distorsion spectrale en dB, pour m=10,16,18 et 20 en fonction du nombre de bits (variant entre 40 et 80bits).

Afin de faire une comparaison des résultats trouvés entre la variation de la distorsion spectrale pour m=10, 16,18 et 20 pôles, on schématisé les valeurs obtenus sur même graphe, comme le montre la figure suivante 3.10 :

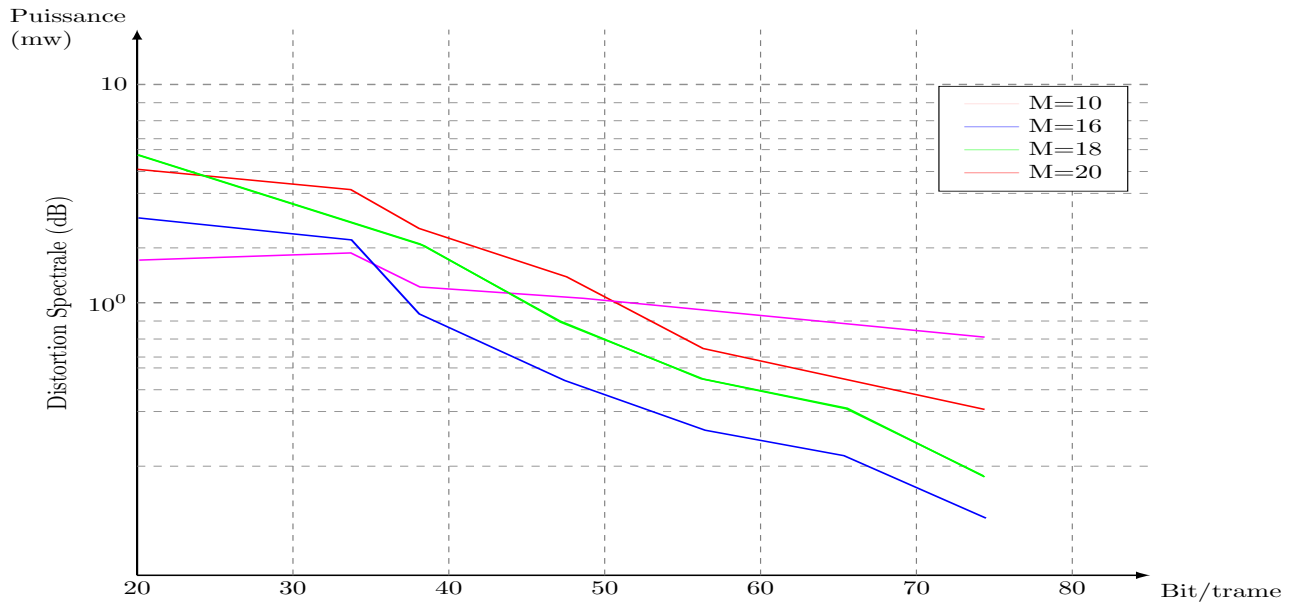


Fig 2.8: Distorsions spectrale entre les LSP et les LSP quantifié pour M=10,16,18 et 20 pôles.

Interprétation et commentaires

La figure 2.8 représente la distorsion spectrale calculée en dB pour différents ordres de prédictions en fonction du nombre de bits par trame, on voit que la distorsion spectrale est inversement proportionnel aux nombre de bits par trame car pour un nombre réduit de bits le dictionnaire de quantification (code book) est moins riche et donc le quantificateur est moins performant .

On constate aussi que pour l'ordre de prédiction m=16 pôles la distorsion est moindre par rapport à celles calculées pour m=18,20 pôles, ce qui valorise la quantification pour un ordre de m=16 pôles dans la bande élargie

Conclusion Générale

La technique standard de codage utilisée dans la téléphonie s'effectue dans une bande de fréquences étroite (200-3400 Hz) utilisant un ordre de prédiction du filtre analyse-synthèse égale à 10, cette dernière ne permet pas la transmission de signaux musicaux.

Nous avons vu le long des chapitres précédents que pour avoir une gamme de fréquences large s'étalant jusqu'à 8000 Hz (la large bande), et donc avoir plus de flexibilité et de performance, il faut augmenter l'ordre de prédiction des filtres. Pour une analyse comparative fiable nous avons pris un ordre de prédiction de 10, 16, 18, 20 pôles.

En premier lieu nous avons extraits les paires de raies spectrales LSP du signal parole échantillonné utilisé dans notre simulation, le codage se fait sur des trames de 80 échantillons chacune.

En deuxième lieu nous avons appliqué l'algorithme de Linde Buser Gray (LBG) afin de constituer un dictionnaire de quantification (code book) pour un nombre de bits par trame allant de 20 à 80 bits/trame.

En dernier lieu nous avons quantifié les LSP en utilisant les code books trouvés, par la suite on a schématisé les enveloppes spectrales des LSP originale et quantifiée et on a calculé les distorsions spectrales entre celles-ci pour des ordres de 10, 16, 18, 20 pôles.

L'étude citée ci-dessus nous informe sur l'état du quantificateur pour des ordres supérieurs et nous permet de constater que le codage à large bande est meilleur pour un ordre de prédiction de 16 pôles.

On conclut que pour le codage à large bande il est conseillé d'utiliser un ordre de prédiction égale à 16 pôles permettant d'avoir une quantification performante et une distorsion acceptable.

Bibliographie

- [1] T.Dutoit, "Introduction au Traitement Automatique de la Parole", Faculté Polytechnique de Mons 1989.
- [2] R.Boite et M.Kunt,"Traitement de la parole", Presses Polytechniques Romandes, première édition.
- [3] M. Xie et D.Berkani. "Amélioration des performances des codeurs de parole"Août 1997
- [4] F.Itakura and S. Saito, "Analysis synthesis telephony based upon the maximum likelihood method" in Rep 6 th Int. Congr. on acoustics, Kohasi, Ed. Tokyo, Japan Aug. 21-28, 1968, C-5-5.
- [5] J. D Markel and A. H. Gray, Jr "A linear prediction vocoder simulation based upon the autocorrelation method", IEEE Trans Acoust. Speech. Signal Processing vol ASSP622, PP.124-134, Apr. 1974.
- [6] P.Kroon and B.S. Atal, "Predictive coding of speech using analysis-by-synthesis techniques", in Advances in Speech Signal Processing S. Furui and M.M. Sondhi, Eds New York: Markel- Dekker, pp 141-164. 1991.
- [7] A. H. Gray, and J. D. Markel ?Quantization and bit allocation in speech processing?, IEEE Trans, on Acoustic, Speech Signal Processing, vol. ASSP-24, pp. 459-473, Oct. 1976. ENP2006 73
- [8] ITU-T Rec. G.722.2. Wideband Coding of Speech at Around 16 kbits/s Using Adaptative Multi-Rate Wideband (AMR-WB). International Telecommunication Union, CHGeneva. 2002.