HAL open science

Les paires de raies spectrales pour le codage de la parole à bas débit

Samir Saoudi, Jean-Marc Boucher, Alain Le Guyader

To cite this version:

Samir Saoudi, Jean-Marc Boucher, Alain Le Guyader. Les paires de raies spectrales pour le codage de la parole à bas débit. GRETSI 91 : 13ème colloque sur le traitement du signal et des images, Sep 1991, Juan-Les-Pins, France. hal-02160921

HAL Id: hal-02160921

https://hal.science/hal-02160921v1

Submitted on 20 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LES PAIRES DE RAIES SPECTRALES POUR LE CODAGE DE LA PAROLE A BAS DEBIT

S. Saoudi *, J.M. Boucher * et A. Le Guyader **

* Ecole Nationale Supérieure des Télécommunications
de Bretagne, BP 832, 29285, Brest - FRANCE
** Centre National d'Etudes des Télécommunications
BP 40, Route de Trégastel, Lannion - FRANCE

RÉSUMÉ

Dans les méthodes d'analyse-synthèse utilisées en prédiction linéaire, les paires de raies spectrales (LSP) réalisent une bonne modélisation du signal de parole. L'algorithme de Levinson éclaté est utilisé pour calculer les LSP. Afin de réaliser un codage optimal de ces paramètres, une méthode d'estimation non paramétrique des densités de probabilité est utilisée. Plusieurs méthodes de quantifications scalaires sont comparées en utilisant l'allocation optimale de bits. Un codeur prédictif excité par codes CELP à 8 kbits/s est utilisé pour comparer la qualité de la parole synthétisée par les LSP et par les rapports d'aires logarithmiques.

Mots clés :
Algorithme de Levinson éclaté, paires de raies spectrales, estimateur à noyau, rapport d'aires logarithmiques, Quantificateur de Lloyd-Max, CELP.

ABSTRACT

The Line Spectrum Pairs (LSP) provide an efficient representation of the synthesis filter used in Linear Predictive Coding of speech. The split Levinson algorithm is used to derive the LSP parameters. The probability density functions are evaluated by a nonparametric estimation method. Several algorithms for the scalar quantization are compared using the optimal bit allocation. A code Excited Linear Predictive (CELP) coder at 8 kbits/s is used for comparison of the LSP and the Log Area Ratios encoding schemes.

Keywords :
Split Levinson algorithm, Line Spectrum Pairs, Kernel estimator, Log area ratio, Lloyd-Max quantizer, CELP.

1. INTRODUCTION

Plusieurs méthodes de codage de la parole utilisant la prédiction linéaire ont été étudiées en vue de la transmission de la parole à bas débit. L'analyse de la parole par les paires de raies spectrales (LSP) est connue comme étant un mode de représentation très intéressant car ces paramètres sont équivalents, mathématiquement, aux coefficients de prédiction linéaire LPC ou aux coefficients de corrélation partielle parcor et possèdent quelques propriétés supplémentaires qui les rendent beaucoup plus attractifs que les coefficients LPC ou parcor pour leur quantification ou leur interpolation .

Cet article est organisé comme suit : le paragraphe 2 contient une description de l'analyse de la parole par les LSP. Dans le paragraphe 3 on rappelle l'utilisation de l'algorithme de Levinson éclaté pour l'extraction des LSP. Dans le paragraphe 4, une méthode d'estimation est utilisée pour évaluer les densités de probabilité des LSP. Différentes méthodes de quantification scalaire des LSP seront discutées au paragraphe 5. L'allocation optimale de bits sera décrite au paragraphe 6. Un codeur prédictif excité par codes à 8 kbits/s sera utilisé au paragraphe 7 suivi de l'analyse des performances.

2. ANALYSE DE LA PAROLE PAR LES LSP

Pour un ordre donné p, l'analyse de la parole en prédiction linéaire donne un filtre tout pôle $1/A_p(z)$, avec :

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

Les paramètres $\{a_i\}_{i=1,2,\ldots,p}$, sont les coefficients LPC [7]. On définit alors les deux polynômes suivants :

$$P_{p+1}(z) = A_p(z) + z^{-(p+1)} A_p(z^{-1})$$

$$P^*_{p+1}(z) = A_p(z) - z^{-(p+1)} A_p(z^{-1})$$

Les deux polynômes prédictifs ainsi construits sont respectivement symétrique et anti-symétrique et possèdent des propriétés très intéressantes qui peuvent être résumées comme suit [13,14,15] :
- pour un filtre stable Ap(z), toutes les racines sont sur le cercle unité
- les racines des deux polynômes s'alternent deux à deux sur le cercle unité.
On montre facilement que ces polynômes peuvent être exprimés comme :

pour p pair :

$$P_{p+1}(z) = (1 + z^{-1}) \prod_{i=1,3,\ldots,p-1} (1 - 2\cos\omega_i z^{-1} + z^{-2})$$

$$P^*_{p+1}(z) = (1 - z^{-1}) \prod_{i=2,4,\ldots,p} (1 - 2\cos\omega_i z^{-1} + z^{-2})$$

où :

$$0 < \omega_1 < \omega_2 < \ldots < \omega_p < \pi .$$

Les paramètres $\{\omega_i\}_{i=1,2,\ldots,p}$ sont définis comme les paires de raies spectrales.

3. EXTRACTION DES LSP

Pour obtenir les racines des deux polynômes prédictifs (symétrique et anti-symétrique), le nombre d'opérations peut être réduit par des algorithmes exploitant le fait que les deux polynômes ont des racines complexes conjuguées de module égal à 1.
Delsarte et Genin [1] ont proposé récemment l'algorithme de Levinson éclaté, version symétrique et anti-symétrique de façon à réduire la complexité de l'algorithme de Levinson classique, afin de calculer le polynôme prédicteur $A_p(z)$. Ils ont étendu cette méthode aux calculs des fréquences de Pisarenko.
Ici, on utilise une formulation analogue pour déterminer et extraire les coefficients LSP $(\omega_i)_{i=1,\ldots,p}$. Dans une première partie, les polynômes prédictifs singuliers symétrique et anti-symétrique, calculés tous les deux par la même version de l'algorithme de Levinson éclaté, peuvent être résolus directement par la méthode FFT ou celle du cosinus inverse. Ensuite à l'aide d'un changement de variable qui s'inspire de celui proposé par Delsarte et Genin pour le calcul des fréquences de Pisarenko, les coefficients LSP peuvent être déterminés à partir des valeurs propres de certaines matrices symétriques tridiagonales dont les éléments sont calculés uniquement à l'aide d'une seule version de l'algorithme de Levinson éclaté [11]. Les valeurs propres sont calculées par la méthode de la bissection pour les matrices tridiagonales [9].

4. ESTIMATION DES DENSITES DE PROBABILITE

Afin de réaliser un codage optimal des coefficients LSP, on a besoin d'estimer leurs densités de probabilité. Pour cela, une nouvelle méthode a été développée pour l'estimation non paramétrique des densités de probabilité basée sur la méthode du noyau et adaptée aux variables aléatoires à support borné [4], ce qui est le cas des coefficients LSP.

5. QUANTIFICATION SCALAIRE DES LSP

La comparaison de plusieurs méthodes de quantification scalaire basées sur la minimisation de la distorsion de quantification moyenne tout en conservant l'ordre entre les paramètres LSP, relation nécessaire et suffisante pour la stabilité du filtre de synthèse de parole, nous a permis de sélectionner la quantification adaptative non uniforme comme étant la plus performante. Afin de résoudre le problème du nombre trop important de densités de probabilité conditionnelles du paramètre courant sachant la donnée du quantifié du paramètre précédent, nous avons approché celle-ci par la densité des différences des LSP consécutifs translatée par la valeur du quantifié du paramètre précédent [12], ce qui est plus facile à obtenir et donne alors un quantificateur adaptatif non uniforme différentiel.

Pour étudier les performances des paramètres LSP, les différentes méthodes de quantification ont été testées sur un signal de parole de 40 doubles phrases d'une durée totale de trois minutes et trente secondes prononcées par dix locuteurs différents (quatre phrases par locuteur).

La distorsion spectrale moyenne a été utilisée comme mesure pour comparer les différentes méthodes de codage des paramètres LSP pour un nombre total de bits variant de 10 à 40 en utilisant la procédure de l'allocation optimale de bits (paragraphe 6). Dans ce sens, la quantification adaptative non uniforme différentielle (utilisant l'approximation des densités conditionnelles) est la plus performante. Un gain de 2 à 4 bits a été noté, pour une même distorsion spectrale moyenne, entre cette méthode et la plus proche méthode en performance, qui est la quantification non uniforme de Lloyd-Max.

6. ALLOCATION OPTIMALE DE BITS

Etant donné un nombre total de bits B à allouer pour l'ensemble des variables du vecteur en codage scalaire, il est préférable (dans le sens de la minimisation de l'erreur quadratique moyenne) d'allouer différents nombres de bits b_i pour chaque variable du vecteur.

Beaucoup de procédures ont été développées pour optimiser l'allocation des bits. Fox [3] a présenté un algorithme pour

l'allocation de bits optimale dans le cas général où les composantes peuvent avoir des densités de probabilité quelconques et distinctes.

La méthode de l'allocation de bits optimal consiste à calculer tout d'abord la variation de la distorsion quand on ajoute un bit, ceci pour chaque composante du vecteur et pour chaque bit b : $\Delta_i(b) = D_i(b - 1) - D_i(b)$, pour $i = 1,...,p$ et $b = 1,...$ Accorder, ensuite, les B bits un par un suivant un ordre décroissant obtenu par les valeurs des variations $\Delta_i(b)$.

Cet algorithme donne une solution optimale dans le cas où les fonctions de distorsions de quantifications $D_i(b)$ sont monotones, décroissantes et convexes [3]. Ceci a été vérifié dans le cas des coefficients LSP, des différences des LSP consécutifs, ainsi que dans le cas des LAR.

Cette procédure est aussi appliquable dans le cas où on veut attribuer un nombre de niveaux pour chaque composante au lieu d'un nombre entier de bits. La variable b de la procédure représente alors la variable nombre de niveaux.

Pour un nombre de bits total égal à 32, l'allocation de bits optimal est (2,3,4,4,4,3,3,3,3,3) pour les coefficients LSP, (3,3,3,4,4,3,3,3,3,3) pour les différences des LSP consécutifs et (4,4,3,3,3,3,3,3,3,3) pour les rapports d'aires logarithmiques. La dimension du vecteur des composantes étant égale à l'ordre de prédiction p qu'on a fixé à 10.

7. PERFORMANCES ET RESULTATS

Actuellement, la plupart des systèmes de codage prédictif utilise la méthode d'analyse par synthèse pour la détermination du signal d'excitation. La figure 1 montre le schéma du codeur prédictif à excitation par codes [5] utilisé pour introduire les algorithmes de codage décrits auparavant.

L'algorithme de quantification adaptative Forward non uniforme optimale a été introduit dans un codeur prédictif excité par codes CELP à 8 kbits/s, ainsi que l'algorithme de quantification optimale (Lloyd-Max) des rapports d'aires logarithmiques (LAR). Le gain de prédiction a été utilisé comme critère objectif pour comparer la qualité de la parole synthétisée par les LSP ou par les LAR. Pour un débit des coefficients de prédiction en dessus de 32 bits par trame, la qualité de la parole obtenue par des algorithmes utilisant les LSP est de même ordre que celle fournie par les algorithmes utilisant les LAR. En dessous de 32 bits par trame, le gain de prédiction, le rapport signal à bruit global et segmental montrent la supériorité des algorithmes utilisant les LSP.

8. CONCLUSIONS

Après un rappel sur l'analyse de la parole par les paires de raies spectrales, l'algorithme de Levinson éclaté est utilisé pour l'extraction des LSP. Après avoir utilisé une méthode d'estimation non paramétrique des densités de probabilité des LSP, nous avons comparé, dans le cadre de la quantification scalaire, différentes méthodes de codages. La distorsion spectrale moyenne a été utilisée comme critère de comparaison. Un codeur prédictif à excitation par codes a été utilisé pour comparer la qualité de la parole synthétisée par les LSP et par les LAR.

REFERENCES

[1] P. Delsarte and Y. Genin,"The split Levinson algorithm," IEEE Trans. Acoust., speech, and signal processing, Vol. ASSP-34, N°.3, pp. 470-478, June 1986.

[2] P. Delsarte and Y. Genin,"On the splitting of classical algorithms in linear prediction theory," IEEE Trans. Acoust., speech, and signal processing, Vol. ASSP-35, N°.5, pp. 645-653, May 1987.

[3] B. Fox,"Discrete optimization via marginal analysis," Mang. Sci., Vol.13, pp. 210-216, November 1966.

[4] F. Ghorbel, S. Saoudi et A. Hillion, "Estimation non paramétrique des densités de probabilité à support borné," XXIIIèmes Journées de statistique de l'ASU, Strasbourg, France, Mai 27-30, 1991, pp. 175-177.

[5] A. Le Guyader, D. Massaloux and J.P. Petit, "Robust and Fast Code-Excited Linear Predictive Coding of Speech," Proc. of ICASSP, pp. 120 - 123, 1989.

[6] S.P. Lloyd,"Least square quantization in PCM," IEEE Trans. Infor. Theory, pp. 129-137, March 1982.

[7] J.D. Markel and A.H. Gray, *Linear prediction of speech*,Springer Verlag, 1976.

[8] J. Max,"Quantizing for minimum distortion," IEEE Trans. Info., Theory, pp. 7-12, March 1960.

[9] N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, Series in computational mathematics, 1980.

[10] S. Saito and K. Nakata, *Fundamentals of speech signal processing*, Academic press, 1985.

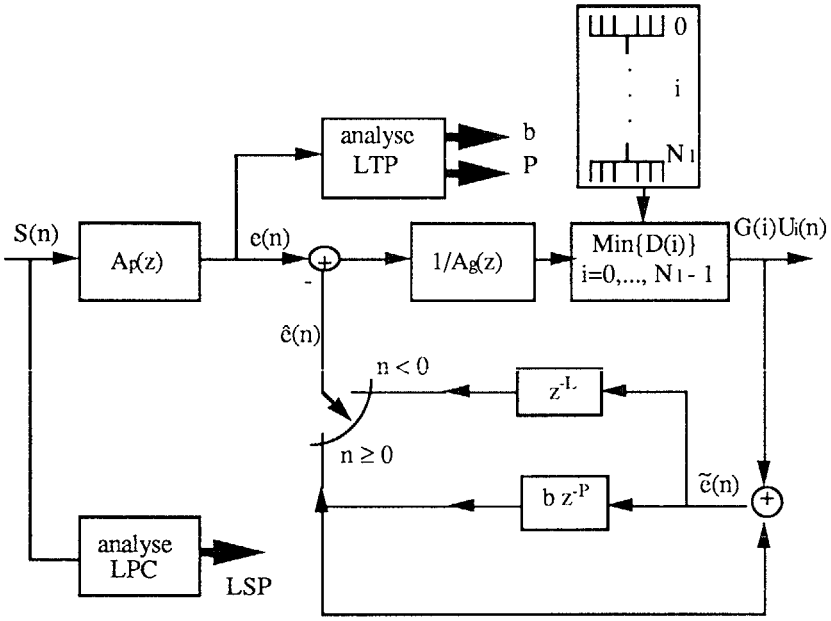
[11] S. Saoudi, J.M. Boucher and A. Le Guyader, "The Split Levinson Algorithm for extracting the Line Spectrum Pairs," Proc. of EUSIPCO - 90, Barcelona, Spain, September 18 - 21, 1990.

[12] S. Saoudi, J.M. Boucher and A. Le Guyader,"Optimal scalar quantization of the LSP and the LAR for speech coding," Proc. of Intern. Conf. on Spoken Language processing-90, Kobe, Japan, November 18-22, 1990, pp. 101-104.

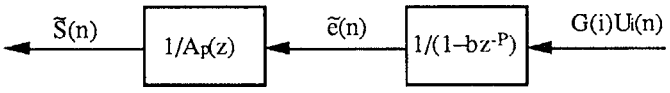
[13] F. K. Soong and B.H. Juang,"Line spectrum pair (LSP) and speech data compression," in Proc. 1984 IEEE, Intern. Conf. Acoust., speech, and signal processing, San Diego, CA, pp.1.10.1-1.10.4, 1984.

[14] N. Sugamura and F. Itakura,"Speech analysis and synthesis methods developed at ECL in NTT - form LPC to LSP-," Speech communication, North-Holland, Vol.5,N°2, pp.199-215, June 1986.

[15] N. Sugamura and N. Favardin, "Quantizer design in LSP speech analysis-synthesis," IEEE Journal on Selected Areas in communications, Vol. 6, N° 2, February 1988.



Codeur



Décodeur

Fig. 1. Codeur prédictif à excitation par codes