

RÉDUCTION DE LA COMPLEXITÉ DE LA QUANTIFICATION VECTORIELLE DES COEFFICIENTS LSP POUR LE CODAGE DE LA PAROLE

B. KÖVESI* #, S. SAOUDI*, JM. BOUCHER* and Z. REGULY#

* ENSTBr, Dépt. SC., Technopôle de Brest Iroise, BP 832, 29285 Brest Cedex, France

Université Technique de Budapest, Dépt. MMT., Mûegyetem rkp. 9, 1521 Budapest, Hongrie

RÉSUMÉ

Dans cet article nous étudions la quantification vectorielle des coefficients LSP (Line Spectrum Pairs) du codeur CELP (Code Excited Linear Predictive). Des méthodes de réduction du temps de recherche du plus proche voisin (ppv) sont proposées. La base de ces méthodes est le réordonnement du dictionnaire suivant l'ordre croissant des normes des mots de code. De cette manière la norme du vecteur à coder permet de localiser son voisinage dans le dictionnaire où la recherche du ppv sera effectuée.

Dans la première partie nous introduisons une norme qui donne une zone de recherche plus petite qu'avec les autres normes déjà publiées pour la même distorsion. Ensuite, nous montrons que la combinaison de deux normes différentes peut encore réduire considérablement la zone de recherche.

1. INTRODUCTION

Beaucoup de travaux ont été consacrés ces dernières années à l'utilisation de la quantification vectorielle (QV) comme un moyen de compression de données. En codage de la parole, l'utilisation de la quantification vectorielle (QV) permet d'atteindre les mêmes performances que la quantification scalaire (QS) en réduisant le débit nécessaire pour le codage des paramètres à quantifier. Par contre la QV nécessite plus de temps de calcul que la QS. Il est donc important de trouver des algorithmes de codage rapide afin de pouvoir implémenter la QV en temps réel.

Dans le codeur CELP la quantification vectorielle est utilisée pour le codage des coefficients LSP. Dans cet article notre objectif est de réduire la complexité de la QV des LSP pour cette application.

2. RÉORDONNEMENT DU DICTIONNAIRE

Dans [1,2], Naja a étudié les différentes méthodes de réduction de la complexité de la QV. Il a introduit une

ABSTRACT

In this paper, vector quantization of the LSP (Line Spectrum Pairs) coefficients in the CELP (Code Excited Linear Predictive) coder is addressed. Methods to reduce the searching time of the nearest neighbour is proposed. The basis of these methods is the reordering of the codebook following increasing norms of its codewords. In this way the norms of the vector to be coded permits to localize its neighbourhood in the codebook where the search of the nearest neighbour will be made.

In the first part of this paper we introduce a norm that results in a smaller searching zone for the same distortion than other norms already published. Then we show that by logical combination of two different norms the search zone can be significantly further reduced.

méthode de localisation du voisinage de recherche du plus proche voisin (ppv) qui consiste à associer à chaque vecteur une norme définie suivant une mesure de distorsion. Il réordonne le dictionnaire suivant l'ordre croissant des normes des mots de code: celles-ci sont stockées pour éviter de les recalculer par la suite. Dans la phase de codage, on cherche d'abord le mot de code ayant la norme la plus proche au vecteur à coder. Ce mot de code et son voisinage dans le dictionnaire ordonné forment la zone de recherche de longueur souhaitée. Le ppv du vecteur à coder est ensuite déterminé par une recherche exhaustive dans cette zone. Pour un dictionnaire de M mots de code et pour une longueur de la zone de recherche L fixée, il existe donc M-L+1 régions de recherche différentes. Parmi plusieurs normes examinées dans [1], la norme euclidienne (1) est celle qui a donné le meilleur résultat.

$$n_e = \sum_{i=1}^p (x_i)^2 \quad (1)$$



3. LA NOUVELLE NORME

Étant donné que les paramètres LSP constituent une séquence ordonnée [3], les dernières composantes ont une influence beaucoup plus importante sur la valeur de la norme euclidienne que les premières composantes. Le tableau 1 montre la moyenne de chaque coefficient LSP, dans le cas d'une analyse court terme à l'ordre 10. Ces valeurs ont été estimées par une séquence d'apprentissage (B1) de 66000 vecteurs LSP. Cette base a été obtenue par analyse LSP de 200 phrases, qui représente environ 22 minute de parole. Les phrases ont été prononcées par plusieurs locuteurs: 5 femmes, 5 hommes.

m ₁	0.3091
m ₂	0.4576
m ₃	0.7551
m ₄	1.0480
m ₅	1.3516
m ₆	1.5863
m ₇	1.8815
m ₈	2.0988
m ₉	2.3963
m ₁₀	2.5661

Tableau 1. Moyenne des coefficients LSP

Comme l'ordre du dictionnaire n'est donc pratiquement déterminé que par les dernières composantes des mots de codes. On utilise très peu l'information portée par les premiers coefficients. C'est pourquoi nous avons plutôt choisi d'utiliser comme norme la somme des valeurs absolues des composantes (la norme L1). Comme les coefficients LSP sont positifs, nous avons pris la somme algébrique (2).

$$n_L = \sum_{i=1}^P x_i \quad (2)$$

Avec cette norme, l'influence de chaque coefficient est plus équilibré. Son utilisation a déjà montré l'obtention d'une distorsion plus faible que la norme euclidienne.

On définit une seconde norme pondérant la norme L1 par la variance de chaque coefficient LSP. Supposons qu'un coefficient soit constant (sa variance est nulle), ce n'est pas la peine d'en tenir compte dans le calcul de la norme car il aurait toujours la même influence: il ne créerait qu'une translation. Plus grande est la variance d'une

composante, mieux on peut organiser le dictionnaire avec cette composante.

Le tableau 2 montre l'estimation de la variance des coefficients LSP sur la base d'apprentissage B1 (première colonne) et par le dictionnaire D1 de taille 256 (deuxième colonne). Ce dernier dictionnaire D1 était obtenu par l'algorithme RGSKAε, que nous avons proposé récemment [4]. Pour la construction du D1, les 25600 premiers vecteurs de la base B1 étaient utilisés comme base d'apprentissage.

Variance LSP	de la base	du dictionnaire
v ₁	0.0063	0.0053
v ₂	0.0172	0.0159
v ₃	0.0329	0.0357
v ₄	0.0328	0.0302
v ₅	0.0331	0.0365
v ₆	0.0266	0.0310
v ₇	0.0152	0.0126
v ₈	0.0179	0.0111
v ₉	0.0140	0.0084
v ₁₀	0.0141	0.0125

Tableau 2. Variance des coefficients LSP

Il est alors préférable de donner un poids plus fort aux coefficients ayant une variance plus importante. Un tel facteur de pondération de la norme L1 peut être tout simplement la valeur de la variance respective de chaque coefficient. (3):

$$n_p = \sum_{i=1}^P v_i \cdot x_i \quad (3)$$

La longueur de la zone de recherche est indépendante du choix de la base d'apprentissage ou du dictionnaire pour l'estimation de la variance. Pour nos simulations nous avons choisi le dictionnaire, parce que le calcul est plus rapide et le dictionnaire est toujours disponible pour faire ces estimations.

Si le dictionnaire est déjà ordonné selon un autre critère (par exemple un dictionnaire robuste vis-à-vis des erreurs de transmission, voir [4]) on introduit une mémoire tampon de longueur M qui fournit le numéro des mots de code selon la norme.

4. COMPARAISON DES NORMES

La complexité de la norme pondérée est identique à celle de la norme euclidienne. Nous avons effectué plusieurs comparaisons pour savoir comment on peut ordonner un dictionnaire selon la norme euclidienne et la norme pondérée. La séquence de test T1 correspond à 160 s de signal, soit 8000 vecteurs LSP, que nous n'avons pas utilisés dans la phase d'apprentissage. Les coordonnées d'un point p sur les figures 1 et 2 sont:

$$p = \left(\frac{\|\bar{v}\| - \text{MinN}}{\text{MaxN} - \text{MinN}}; \frac{\|Q(\bar{v})\| - \text{MinN}}{\text{MaxN} - \text{MinN}} \right) \quad (4)$$

où $\|\bar{v}\|$ est la norme d'un vecteur à quantifier,
 $\|Q(\bar{v})\|$ est la norme du vecteur quantifié,

MinN et MaxN sont les norme minimale et maximale des mots de codes du dictionnaire.

Figure 1. Norme Euclidienne

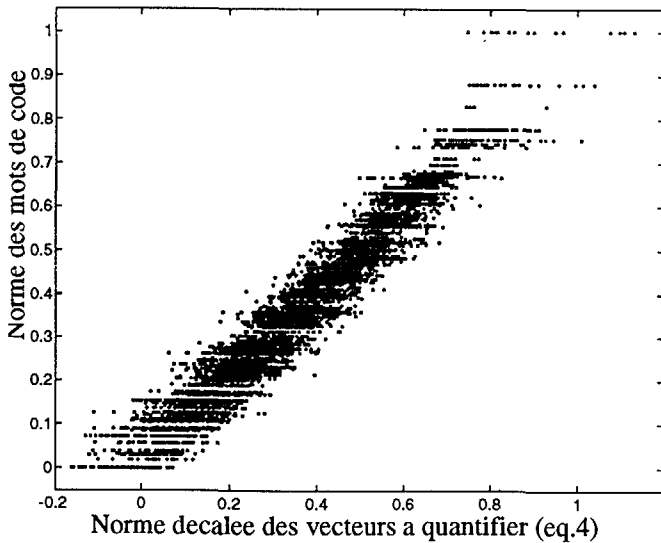
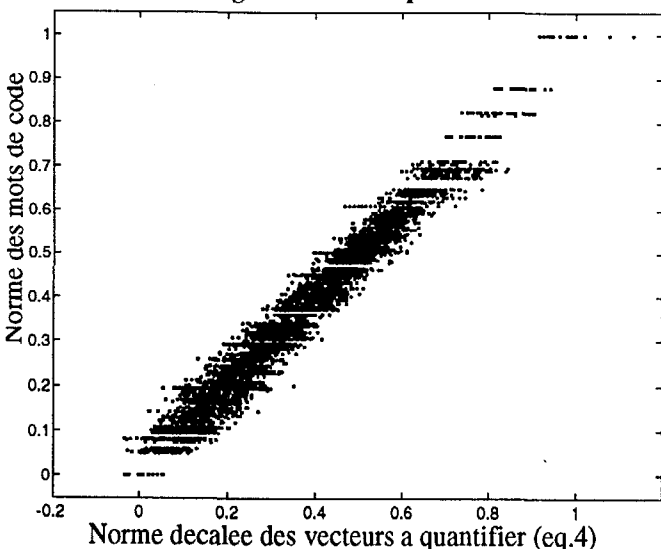
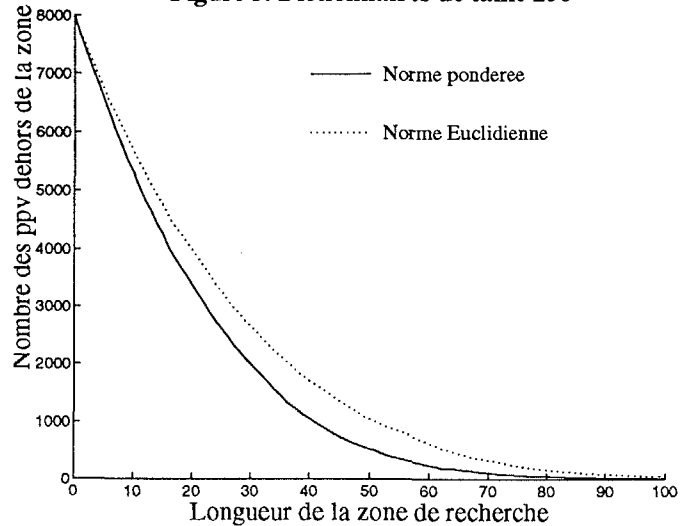


Figure 2. Norme ponderee



Plus le nuage de points est étroit, meilleure est la norme. La figure 3 montre combien de fois le ppv se trouve en dehors de la zone de recherche en fonction de la longueur de celle-ci. La figure 4 donne la distorsion de la QV en fonction de la longueur moyenne de la zone considérée. Tous ces résultats obtenus par simulations ont permis de valider la nouvelle norme.

Figure 3. Dictionnaires de taille 256



5. COMBINAISON DE DEUX NORMES

Nous avons également étudié la combinaison de deux normes différentes. Pour un vecteur à coder, chaque norme définit une région de recherche de longueur fixée préalablement. La zone de recherche définitive du ppv est alors déterminée par l'intersection de ces deux régions. Ainsi la longueur de la zone de recherche diminue mais en même temps la distorsion augmente.

Les zones ainsi obtenues ont une longueur variable, leur longueur moyenne sera considérée. Étant donné la longueur souhaitée de la zone de recherche définitive, il serait intéressant de fixer la longueur des zones de recherche initiales en fonction de la longueur moyenne de la zone d'intersection.

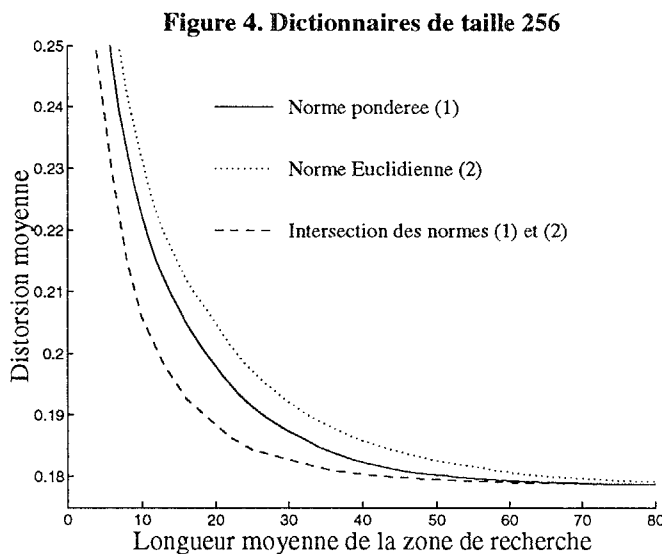
Dans la phase de codage, on prend la région déterminée par l'une des deux normes, et on n'effectue de recherche de ppv que parmi les mots de code qui se trouvent aussi dans la région déterminée par l'autre norme. Pour pouvoir rapidement faire cette décision on remplit préalablement une matrice de taille $M \times M$ qui indique si un mot de code donné se trouve dans la région de recherche de longueur fixée autour d'un autre mot de code selon la



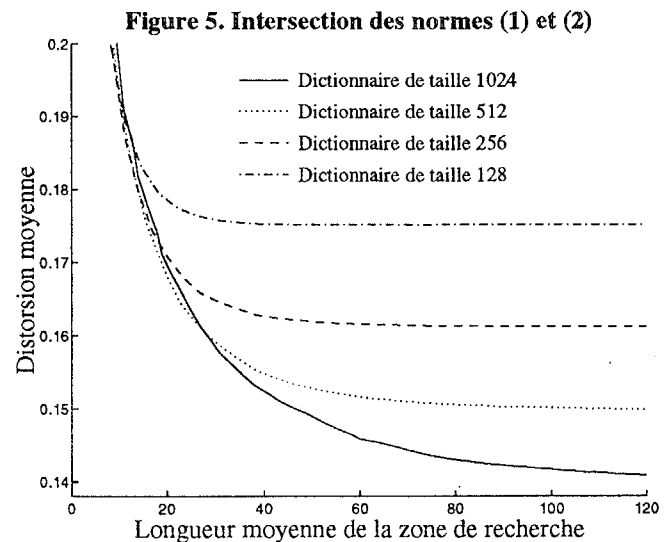
deuxième norme. Une telle matrice augmente l'exigence en capacité de mémoire du codeur. Si l'intersection de deux régions est vide, on prend le mot de code ayant la plus proche norme au vecteur à coder selon la première mesure de distorsion.

Pour avoir un bon résultat avec cette méthode, il faut prendre deux bonnes normes qui définissent des régions de recherche suffisamment différentes. Nous avons testé des combinaisons de plusieurs normes. Les simulations montrent qu'en utilisant l'intersection de la norme euclidienne et la nouvelle norme pondérée on peut considérablement diminuer la distorsion de la QV pour une région de recherche de longueur fixée.

La figure 4 donne les courbes des distorsions moyennes, obtenues dans les différents cas: norme euclidienne, norme pondérée et l'intersection des deux normes, en fonction de la longueur moyenne de la zone de recherche, dans le cas de la quantification vectorielle de la séquence de test T1 utilisant le dictionnaire D1.



La figure 5 a pour but de montrer que utilisant la méthode de recherche rapide, pour la même longueur de la zone de recherche avec l'augmentation de la taille du dictionnaire la distorsion diminue. Nous avons utilisé la même base d'apprentissage pour construire ces dictionnaires. Donc dans une implémentation en temps réel, quand le temps de calcul de recherche de ppv est limité et qu'il y a suffisamment de mémoire, il peut être intéressant d'utiliser un dictionnaire plus grand avec la méthode de recherche rapide.



6. CONCLUSION

Nous avons présenté une nouvelle norme pondérée pour la localisation du voisinage de recherche du ppv pour la QV des coefficients LSP dans le codeur CELP. Avec une complexité identique, la distorsion obtenue utilisant cette norme est plus faible que celle qui utilise la norme euclidienne. Nous avons obtenu la distorsion la plus faible en cherchant le ppv dans l'intersection des zones de recherche de ces deux normes.

7. RÉFÉRENCES:

- [1] N. Naja, "Construction de dictionnaires et quantification vectorielle pour les codeurs de parole LSP-CELP", Thèse de Doctorat, Université de Rennes I, Juin 1994.
- [2] N. Naja, J.M. Boucher and S. Saoudi, "Fast LSP Vector Quantization Algorithms Comparison," Proc. MELECON-94, Antalya, Turkey, pp. 1127-1130, April 12-14, 1994.
- [3] S. Saoudi, "Codage de la parole par les paires de raies spectrales", Thèse de Doctorat, Université de Rennes I, Décembre 1990.
- [4] B. Kövesi, S. Saoudi, J.M. Boucher and Z. Reguly, "A Fast Robust Stochastic Algorithm for Vector Quantizer Design for Nonstationary Channels," Proc. ICASSP-95, Detroit, Michigan, USA, Vol 1, pp. 269-273, Mai 8-12, 1995