

OPTIMISATION D'UN QUANTIFICATEUR VECTORIEL
À CODES-PRODUITS À ANALYSE LINÉAIRE

Amel BENZAÏA, Georges TZIRITAS

Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Electricité, CNRS/UPS
Plateau de Moulon, 91192 Gif-sur-Yvette, France.

RÉSUMÉ

Cet article étudie la quantification vectorielle à codes-produits (QVCP) et à analyse linéaire. Une transformation linéaire partitionne le vecteur source en plusieurs sous-vecteurs qui sont ensuite quantifiés. Nous montrons qu'une condition nécessaire et suffisante sur la transformation doit être vérifiée pour avoir des quantificateurs indépendants. L'optimisation des quantificateurs sous la contrainte de débit fixe est réalisée. Enfin, deux QVCP particuliers opérant sur des images naturelles, sont étudiés. Le premier est généré par extraction de la moyenne des blocs. Le second QVCP fonctionne dans le domaine transformé.

1. INTRODUCTION

La quantification vectorielle (QV) est une technique performante pour la compression des images [1]. Toutefois, la complexité opératoire du codage augmente exponentiellement avec le débit et la dimension des vecteurs et rend rhédibitoire la QV. Pour diminuer cette complexité, une quantification vectorielle à codes-produits (QVCP) à analyse linéaire est envisagée [2]. Du vecteur source \mathbf{x} , sont extraites p caractéristiques pertinentes z_i de dimension k_i :

$$\mathbf{z} \triangleq \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_p \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_p \end{pmatrix} \mathbf{x} = \mathbf{A}\mathbf{x}. \quad (1)$$

Pour que l'opération de synthèse soit possible, il faut que $k = \sum_{i=1}^p k_i$. La seconde étape est la QV, *a priori* non-séparée, des z_i en vecteurs \hat{z}_i . A cette fin, p dictionnaires de débit R_i sont alors utilisés. Le dictionnaire final est leur produit cartésien. L'avantage de la QVCP est que le débit global R peut être élevé sans que les débits partiels R_i soient prohibitifs puisque :

$$R = \sum_{i=1}^p R_i. \quad (2)$$

ABSTRACT

This article deals with product-codes vector quantizers with linear analysis. The source vector is partitioned into several subvectors, by means of a linear transformation. These subvectors are then quantized. In order to obtain independent quantizers, we show that a necessary and sufficient condition on the transformation must be satisfied. In this case, the optimization of the vector quantizers of the product-codes under a fixed rate constraint is solved. Finally, two particular product-codes are used to encode natural images. First is the Mean Removed Vector Quantizer. The second works in the transform-domain.

On montre aisément que la complexité opératoire du QVCP est plus faible que celle du QV standard. Par exemple, le nombre de multiplications requises est égal à $\sum_{i=1}^p k_i 2^{R_i}$ pour le QVCP et à $k 2^R$ pour le QV standard. En effet, puisque $R > R_i$, il vient $k_i 2^R > k_i 2^{R_i}$, soit encore $(\sum_{i=1}^p k_i) 2^R > \sum_{i=1}^p k_i 2^{R_i}$, ce qui établit le résultat annoncé.

Notre étude se propose de déterminer les débits R_i^* qui minimisent l'Erreur Quadratique Moyenne (EQM) quand le débit global R est fixé. La partie 2 fournit la solution de ce problème d'optimisation sous contrainte. La partie 3 présente deux cas particuliers de QVCP, utilisés en compression d'images. Enfin, la partie 4 fournit les résultats de simulations sur des images naturelles.

2. OPTIMISATION DU QVCP

L'EQM D est donnée par l'équation :

$$D \triangleq E[(\mathbf{z} - \hat{\mathbf{z}})^T (\mathbf{A}\mathbf{A}^T)^{-1} (\mathbf{z} - \hat{\mathbf{z}})]. \quad (3)$$

avec $\hat{\mathbf{z}} = (\hat{z}_1, \dots, \hat{z}_p)^T$. Montrons que D est la somme des EQM partielles résultant de la QV séparée des z_i , si et seulement si :

$$\mathbf{A}_i \mathbf{A}_i^T = \mathbf{0} \quad \forall i \neq j. \quad (4)$$



En effet, si l'équation (4) est satisfaite, $(\mathbf{A}\mathbf{A}^T)^{-1}$ est une matrice diagonale par blocs. L'EQM devient alors :

$$D = \sum_{i=1}^p E[(\mathbf{z}_i - \hat{\mathbf{z}}_i)^T (\mathbf{A}_i \mathbf{A}_i^T)^{-1} (\mathbf{z}_i - \hat{\mathbf{z}}_i)]. \quad (5)$$

Or, $(\mathbf{A}_i \mathbf{A}_i^T)^{-1}$ est une matrice symétrique définie positive, diagonalisable en Δ_i^2 dans une base orthonormale de vecteurs propres \mathbf{P}_i : $(\mathbf{A}_i \mathbf{A}_i^T)^{-1} = \mathbf{P}_i \Delta_i^2 \mathbf{P}_i^T$. L'EQM apparaît comme la somme des contributions séparées de p QV indépendants, opérant sur les transformés \mathbf{u}_i :

$$D = \sum_{i=1}^p E[(\mathbf{u}_i - \hat{\mathbf{u}}_i)^T (\mathbf{u}_i - \hat{\mathbf{u}}_i)], \quad (6)$$

$$\text{où } \mathbf{u}_i \triangleq \mathbf{B}_i \mathbf{x} \text{ et } \mathbf{B}_i \triangleq \Delta_i \mathbf{P}_i^T \mathbf{A}_i. \quad (7)$$

Réciproquement, si D est la somme des distorsions partielles dans p espaces transformés linéaires de l'espace source :

$$D = \sum_{i=1}^p E[(\mathbf{u}_i - \hat{\mathbf{u}}_i)^T (\mathbf{u}_i - \hat{\mathbf{u}}_i)], \quad \mathbf{u}_i \triangleq \mathbf{B}_i \mathbf{x}. \quad (8)$$

La matrice inverse généralisée $\mathbf{A}_i^\# \triangleq (\mathbf{A}_i^T \mathbf{A}_i)^{-1} \mathbf{A}_i^T$ permet d'écrire :

$$D = \sum_{i=1}^p E[(\mathbf{z}_i - \hat{\mathbf{z}}_i)^T (\mathbf{B}_i \mathbf{A}_i^\#)^T \mathbf{B}_i \mathbf{A}_i^\# (\mathbf{z}_i - \hat{\mathbf{z}}_i)]. \quad (9)$$

Nous en déduisons que $(\mathbf{A}\mathbf{A}^T)^{-1}$ est une matrice diagonale par blocs. Il en va de même pour $\mathbf{A}\mathbf{A}^T$ (formule d'inversion des matrices-blocs). Ceci est la traduction de l'équation (4).

Dans toute la suite, nous considérons vraie l'équation (4). Dans ce cas, D est la somme des distorsions partielles de p QV. Or, dans les conditions asymptotiques (R_i élevés), il existe une expression explicite de la valeur minimale de l'EQM [3] :

$$D^* = \sum_{i=1}^p k_i d(k_i) \|p(\mathbf{u}_i)\|_{\frac{k_i}{k_i+2}} 2^{-\frac{2R_i}{k_i}}. \quad (10)$$

$$\text{où } \|p(\mathbf{u}_i)\|_{\frac{k_i}{k_i+2}} \triangleq \left(\int_{R^{k_i}} [p(\mathbf{u}_i)]^{(\frac{k_i}{k_i+2})} d\mathbf{u} \right)^{\frac{k_i+2}{k_i}}, \quad (11)$$

et $d(k_i)$ est un coefficient indépendant de la densité de probabilité $p(\mathbf{u}_i)$ et tabulé [3].

Ainsi, l'optimisation du QVCP consiste à chercher les R_i^* qui minimisent D^* sous la contrainte (2). Le problème posé ressemble à un problème d'allocation de bits à chacun des p QV du code-produit [4]. Etant

une somme de fonctions convexes, la fonction D^* est une fonction convexe du p -uplet (R_1, \dots, R_p) .

Elle admet un unique minimum sous les contraintes considérées, donné par les conditions de Kuhn-Tucker [5]. A cette fin, le multiplicateur de Lagrange λ est introduit définissant la fonction $f(\cdot)$ comme suit :

$$f(R_1, \dots, R_p) \triangleq D^* - \lambda (R - \sum_{i=1}^p R_i). \quad (12)$$

Les conditions nécessaires et suffisantes d'optimalité s'écrivent alors :

$$\begin{cases} \frac{\partial f(R_1, \dots, R_p)}{\partial R_i} = 0 & \forall i \text{ tel que } R_i > 0 \\ \frac{\partial f(R_1, \dots, R_p)}{\partial R_i} \geq 0 & \forall i \text{ tel que } R_i = 0 \end{cases} \quad (13)$$

Un simple calcul montre que :

$$\frac{\partial f(R_1, \dots, R_p)}{\partial R_i} = 2 \ln(2) (\mu - a_i 2^{-\frac{2R_i}{k_i}}), \quad (14)$$

$$\text{où on note } \mu \triangleq \frac{\lambda}{2 \ln(2)} \text{ et } a_i \triangleq d(k_i) \|p(\mathbf{u}_i)\|_{\frac{k_i}{k_i+2}}. \quad (15)$$

On en déduit donc que :

$$\begin{cases} \text{si } a_i \leq \mu \text{ alors } R_i = 0 \\ \text{si } a_i > \mu \text{ alors } R_i = \frac{k_i}{2} \log_2 \left(\frac{a_i}{\mu} \right) > 0 \end{cases} \quad (16)$$

Le multiplicateur μ doit satisfaire l'équation :

$$R = \log_2 \left[\prod_{a_i > \mu} \left(\frac{a_i}{\mu} \right)^{\frac{k_i}{2}} \right]. \quad (17)$$

Ayant résolu l'équation (17), l'équation (16) donne par la suite, tous les débits R_i^* . La solution obtenue ne garantit pas des débits entiers. Tous les R_i^* doivent être arrondis par l'entier $[R_i^*]$ immédiatement inférieur. Si après cet ajustement,

$$\delta R \triangleq R - \sum_{i=1}^p [R_i^*] > 0, \quad (18)$$

il faut assigner séquentiellement aux p QV les δR bits restants. Un bit est alloué au QV qui maximise la baisse de la distorsion partielle $\delta_i = k_i a_i 2^{-\frac{2}{[R_i^*]}} (1 - 2^{-\frac{2}{k_i}})$ jusqu'à ce que la somme des débits partiels soit égale à R .



3. EXEMPLES DE QVCP

La moyenne des intensités des pixels des blocs \mathbf{x} représente une information qu'il faut coder avec soin. Par ailleurs, centrer les vecteurs permet de diminuer leur dynamique.

De tels avantages justifient la quantification séparée de la moyenne scalaire z_1 de $\mathbf{x} = (x_1, \dots, x_k)^T$ et du vecteur résiduel $\mathbf{z}_2 = (x_1 - z_1, \dots, x_{k-1} - z_1)^T$ de dimension $k-1$ [6]. Nous avons montré [7] que les matrices $\mathbf{B}_1, \mathbf{B}_2$ optimisent ce QVCP spatial (QVCP1):

$$\mathbf{B}_1 = \left[\begin{array}{c|c} \sqrt{\frac{1}{k}} \mathbf{e}^T & \sqrt{\frac{1}{k}} \end{array} \right], \quad (19)$$

$$\mathbf{B}_2 = \left[\begin{array}{c|c} \mathbf{v}_1^T & -\sqrt{\frac{1}{k}} \\ \mathbf{v}_2^T & 0 \\ \vdots & \vdots \\ \mathbf{v}_{k-1}^T & 0 \end{array} \right]. \quad (20)$$

$$\mathbf{v}_j \triangleq \left[\sqrt{\frac{2 - \delta_{1,j}}{k-1}} \cos \left(\frac{(2i-1)(j-1)\pi}{2(k-1)} \right) \right] \forall i \in [1, k-1]. \quad (21)$$

Le second système étudié (QVCP2) fonctionne dans le domaine de la Transformation en Cosinus Discrète (TCD). Le coefficient DC z_1 représente la moyenne du bloc. Pour les mêmes raisons que précédemment, il est préconisé de coder z_1 séparément du vecteur \mathbf{z}_2 constitué des $k-1$ autres coefficients. La matrice de la TCD étant unitaire, l'équation (4) est vérifiée.

4. RÉSULTATS DE SIMULATIONS

Nos simulations ont porté sur les images des séquences "Voiture" et "White Trevor". Chaque image est divisée en blocs carrés $k=16$ pour le QVCP1 et $k=64$ pour le QVCP2. A l'issue de l'analyse, les 2 séquences \mathbf{u}_1 et \mathbf{u}_2 sont quantifiées séparément, par l'algorithme de LBG [8]. Plusieurs couples (R_1, R_2) correspondent à un débit global donné. De nos simulations, seul le couple assurant la plus faible distorsion globale est retenu puis comparé au couple (R_1^*, R_2^*) fourni par l'étude théorique. Notons que pour le calcul de (R_1^*, R_2^*) , pour le QVCP1, nous avons supposé que \mathbf{u}_1 et \mathbf{u}_2 suivent des lois normales de variance σ_1^2 et Γ_2 . Dans le cas gaussien, il est montré dans [9] que :

$$\begin{cases} \|p(\mathbf{u}_1)\|_{\frac{1}{3}} &= 2\pi 3^{\frac{3}{2}} \sigma_1^2 \\ \|p(\mathbf{u}_2)\|_{\frac{k-1}{k+1}} &= 2\pi \left(\frac{k+1}{k-1} \right)^{\frac{k+1}{2}} |(\Gamma_2)^{\frac{1}{k-1}}| \end{cases} \quad (22)$$

Dans le domaine transformé, on suppose que \mathbf{u}_1 suit une loi gaussienne et que les coefficients restants sont indépendants ayant chacun une distribution laplacienne

de paramètre λ_i [10]. Dans ce cas, un simple calcul montre que :

$$\|p(\mathbf{u}_2)\|_{\frac{k-1}{k+1}} = 4 \left(\frac{k+1}{k-1} \right)^{k+1} \prod_{j=2}^k \lambda_j^{-\frac{2}{(k-1)}}. \quad (23)$$

Les débits du QVCP1 sont proches de ceux obtenus par le calcul théorique (Tab. 1). Les résultats sont surtout probants pour un débit global supérieur à 10 bits/vecteur, correspondant mieux aux conditions asymptotiques. La figure 1 illustre les performances du QVCP fournissant la plus faible distorsion pour un débit global donné et les performances théoriques au débit considéré.

Les tableaux 2 et 3 consignent les résultats obtenus, dans le domaine transformé. Pour mémoire, nous rappelons également les performances du QV standard équivalent.

La différence enregistrée entre le débit théorique et scalaire peut s'expliquer par l'écart par rapport aux conditions asymptotiques. En effet, le débit maximal mis en oeuvre s'élève à 12 bits/vecteur soit 0.1875 bpb. De plus, rappelons que l'algorithme de LBG ne fournit pas systématiquement le minimum absolu de l'EQM.

Une autre approche consiste à calculer les coefficients a_1 et a_2 directement d'après les histogrammes de \mathbf{u}_1 et \mathbf{u}_2 . L'histogramme de la séquence \mathbf{u}_1 présente un pic prépondérant autour de la valeur moyenne de l'arrière-plan. Cela indiquant que la séquence scalaire contient des valeurs fortement corrélées. Un codage prédictif en mode différentiel (MICD) est préconisé afin d'exploiter avantageusement cette caractéristique. Pour ce faire, on utilise le prédicteur linéaire d'ordre 3, de fonction de transfert $P(z_1, z_2) = 0.95z_1^{-1} + 0.95z_2^{-1} - 0.9025z_1^{-1}z_2^{-1}$. L'erreur de prédiction (calculée en boucle ouverte) présente une dynamique plus faible et se prête à une meilleure quantification que la moyenne. Le tableau 4 fournit le débit scalaire expérimental et les débits scalaires théoriques (réel et entier).

5. RÉFÉRENCES

- [1] A. Gersho, R.M.Gray. *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1992.
- [2] S.E. Budge, R.L. Baker. "Compression of color digital Images using Vector Quantization in Product Codes", *Proc. IEEE Int. Conf. on Acoustics, Speech, Signal Processing*, pp. 129-132, 1985.
- [3] P.L. Zador, "Asymptotic Quantization of Continuous Signals and the Quantization Dimension", *IEEE Trans. on Inform. Theory*, Vol. 28, No. 2, pp. 139-148, 1982.
- [4] A. Segall, "Bit Allocation and Encoding for Vector Sources", *IEEE Trans. Inform. Theory*, Vol. 22, No. 2, pp. 162-169, 1976.



[5] W.W. Kuhn, A.W. Tucker, "Nonlinear Programming", *Proc. 2nd Berkeley Symp. on Mathematical Statistics and Probability, University of California, Berkeley*, pp. 481-492, 1951.

[6] T. Murakami, K. Asai, E. Yamazaki, "Vector Quantizer on Video Signals", *Electronic Letters*, Vol. 18, No. 23, pp. 1005-1006, 1982.

[7] A. Benazza, G. Tziritas, "Optimisation d'un Quantificateur à Codes-Produit sous une Contrainte de Débit Fixe", *Actes Colloque GRETSI*, pp. 849-852, 1991.

[8] Y. Linde, A. Buzo, R.M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Trans. on Commun.* (1980), Vol. 28, No. 1, pp. 84-95.

[9] T.D. Lookabaugh, R.M. Gray, "High-Resolution Theory and the Vector Quantizer Advantage", *IEEE Trans. on Inform. Theory*, Vol. 35, No. 5, pp. 1020-1033, 1989.

[10] R.C. Randall, J.D. Gibson, "Distribution of the Two-Dimensional DCT Coefficients for Images", *IEEE Trans. on Commun.*, Vol. 31, No. 6, pp. 835-839, 1983.

Débit Global (bits/bocs)	Débit Scalaire Théorique	Débit Scalaire Expérimental
6.0	4.07	3.0
7.0	4.14	3.0
8.0	4.20	3.0
9.0	4.26	3.0
10.0	4.32	4.0
11.0	4.39	4.0
12.0	4.45	4.0

TAB. 1 - Débits scalaires expérimentaux et théoriques pour le QVCP1, "Voiture"

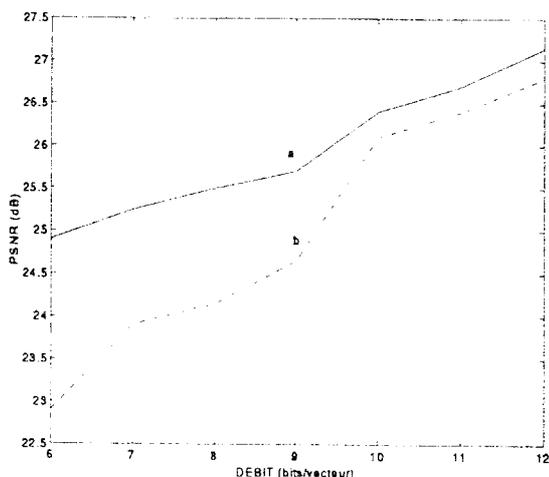


FIG. 1 - (a): PSNR théorique, (b): PSNR expérimental du QVCP1, "Voiture".

Débit Global (bits/bocs)	Débit Scalaire Théorique	Débit Scalaire Expérimental
6.0	5.11	3.0
7.0	5.13	3.0
8.0	5.14	3.0
9.0	5.16	3.0
10.0	5.17	3.0
11.0	5.19	3.0
12.0	5.20	3.0

TAB. 2 - Débits scalaires expérimentaux et théoriques pour le QVCP2, "Voiture"

Débit Global (bits/bocs)	PSNR Théorique (dB)	PSNR Expérimental (dB)	PSNR Standard (dB)
6.0	25.92	20.72	22.02
7.0	26.01	21.35	22.72
8.0	26.10	21.96	23.49
9.0	26.20	22.49	24.42
10.0	26.29	23.07	25.88
11.0	26.39	23.67	28.05
12.0	26.48	24.47	33.28

TAB. 3 - PSNR expérimentaux et théoriques pour le QVCP2, "Voiture".

Débit Global (bits/bocs)	Débit Scalaire Théorique	Débit Scalaire Théorique Entier	Débit Scalaire Expérimental
2.0	0.8716	1	1
3.0	0.9341	1	2
4.0	0.9966	1	3
5.0	1.0591	1	3
6.0	1.1216	2	3
7.0	1.1841	2	4
8.0	1.2466	2	4
9.0	1.3091	3	4
10.0	1.3716	3	4

TAB. 4 - Débits scalaires expérimentaux et théoriques pour le QVCP1 avec codage MICD de u_1 , "White Trevo".