

# Optimisation du Dictionnaire pour la Quantification Vectorielle d'Images Couleur

Christophe CHARRIER   Hocine CHERIFI

Université Jean Monnet de Saint-Étienne  
Équipe Ingénierie de la Vision – UMR CNRS 5516

3, rue Javelin Pagnon BP 505

42007 Saint-Étienne Cedex 01 - FRANCE

Phone : 33 477 923030 – Fax : 33 477 923039

{charrier|cherifi}@vision.univ-st-etienne.fr

<http://www.univ-st-etienne.fr/~iupvis/tsivision.html>

## RÉSUMÉ

Lors de la compression d'images couleur par la technique de la Quantification Vectorielle (QV), il est primordial d'obtenir un dictionnaire représentant au mieux les images. Une des techniques la plus couramment utilisée pour cette étape est la méthode LBG. Or cette solution ne permet pas de prendre en compte la totalité de la répartition spatiale du nuage couleur associé à l'image. Pour remédier à cet handicap, nous avons introduit une étape dans l'algorithme LBG basée sur l'algorithme du Max-Min. Les résultats obtenus à partir de mesures objectives et psychophysiques ont permis de conclure à la supériorité de notre méthode même si le temps de calcul est légèrement plus important.

## ABSTRACT

Using the Vector Quantization (VQ) technique to compress color images, we need to obtain the best dictionary representing the training set. The most widely used method for the codebook generation is the LBG algorithm. Yet, this technique does not allow to take into account the color spatial repartition associated to the color image. To solve this problem, we introduced in the LBG algorithm a step based on the Max-Min algorithm. Both objective and psychophysical measures let one demonstrate the improvement of the VQ reconstructed quality images, even if the increase of the computational time is not very important.

## 1 Introduction

Les physiologues et les psychophysiciens ont montré que la perception des couleurs peut être décrite par un vecteur couleur à trois dimensions [1]. En effet les êtres humains sont seulement dotés de trois types de récepteurs couleurs, appelés cônes, disposés à l'intérieur de la rétine.

Or la plupart des techniques de compression d'images couleur repose sur le fait qu'une image couleur peut être décomposée selon ses trois plans couleur R, G et B. Chacun de ces plans est alors traité indépendamment. Cependant une telle approche ne peut ni tenir compte des corrélations existantes entre les 3 plans couleurs ni de facteurs liés à la perception des couleurs.

Une des techniques les plus employées lors de la construction du dictionnaire  $\mathbf{D}$  lors de la compression par QV est l'algorithme LBG [2]. Cependant les images couleur compressées par QV à l'aide du dictionnaire  $\mathbf{D}$  souffrent d'un problème important : « la perte des couleurs » [3]. La raison principale de cette dégradation est due au fait que les vecteurs code  $\mathbf{y}$  issus de l'algorithme LBG sont produits en prenant le barycentre des blocs appartenant à une même partition. Dans ce cas, on ne calcule que des couleurs moyennes et les couleurs situées aux extrémités du nuage couleur ne sont pas prises en compte.

Ceci se traduit donc par l'apparition de fausses couleurs dans l'image reconstruite.

Pour remédier à ce problème, nous proposons une nouvelle méthode pour la génération du dictionnaire  $\mathbf{D}$ . Cette technique repose sur l'algorithme LBG auquel nous avons introduit une étape d'exploitation des dissymétries couleurs.

## 2 La QV

La QV consiste à représenter tout vecteur  $\mathbf{x}$  de dimension  $k$  par un vecteur  $\mathbf{y}$  de dimension analogue appartenant à un ensemble fini appelé *dictionnaire*  $\mathcal{D}$ . Lors du codage par QV, l'étape primordiale réside dans l'élaboration du dictionnaire  $\mathcal{D}$ . Sa création s'effectue à partir d'une séquence d'apprentissage.

### 2.1 L'algorithme LBG

Cet algorithme itératif proposé par Linde, Buzo et Gray [2] correspond à une extension de l'algorithme initialement introduit par Lloyd et connu sous le nom de Lloyd-Max [4]. Ce dernier est utilisé pour la création de dictionnaire dans le cas de la quantification scalaire. Pour un dictionnaire initial

donné, cet algorithme optimise le codeur et le décodeur de façon à obtenir la meilleure partition possible du signal.

Le codeur est défini selon la règle du *plus proche voisin* :

$$\forall j, \text{ si } d(\mathbf{x}, \mathbf{z}_i) < d(\mathbf{x}, \mathbf{z}_j) \Rightarrow C(\mathbf{x}) = \mathbf{z}_i. \quad (1)$$

où  $C$  représente une région de Voronoï associée à l'espace  $\mathbb{R}^k$ . Cette règle détermine donc la partition de  $\mathbb{R}^k$  en cellules de Voronoï.

Lors de la sélection d'un vecteur  $\mathbf{y}$  destiné à représenter un vecteur  $\mathbf{x}$ , la distance séparant les vecteurs est calculée au moyen de la distance Euclidienne  $d$  :

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^k |x_i - y_i|^2. \quad (2)$$

La partie la plus sensible de cette technique réside donc dans le choix et la construction du dit dictionnaire. Cet algorithme nécessite un dictionnaire initial. Tout utilisateur averti de la méthode LBG a remarqué qu'une initialisation judicieuse contribuait fortement aux performances de cette technique. En effet chaque itération ne provoquant qu'un changement local du dictionnaire, l'algorithme peut converger vers un minimum local.

Il existe diverses techniques pour choisir le dictionnaire initial. La méthode proposée dans la version initiale de l'algorithme LBG est une méthode par dichotomie vectorielle plus connue sous le nom de « splitting » [2]. Cette technique consiste à diviser chaque vecteur représentant  $\mathbf{y}$  en deux nouveaux vecteurs  $\mathbf{y} + \varepsilon$  et  $\mathbf{y} - \varepsilon$  où  $\varepsilon$  est un vecteur aléatoire de perturbation de faible énergie. On applique ensuite les itérations de LBG sur ce nouveau dictionnaire. Le dictionnaire initial est alors le barycentre de la séquence d'apprentissage. L'algorithme génère ensuite une succession de dictionnaires dans lesquels le nombre de vecteur est multiplié par 2 à chacune des itérations.

L'inconvénient majeur de cette technique est qu'elle a tendance à recentrer le nuage des vecteurs du dictionnaire  $\mathcal{C}$  autour de l'axe principal du nuage couleur de la séquence d'apprentissage. La figure 1 illustre ce handicap sur une image originale de textile (a). Nous voyons nettement que plus le taux de compression augmente (Fig. 1(c) et Fig. 1(d)) et que plus le nuage a tendance à se resserrer autour de l'axe principal du nuage couleur initial (Fig. 1(b)).

Pour pallier cet inconvénient, nous introduisons, lors de la construction du dictionnaire, une étape supplémentaire basée sur la dissymétrie couleur et permettant une meilleure prise en compte de la répartition spatiale des couleurs.

## 2.2 La Dissymétrie Couleur dans l'algorithme LBG modifié

La méthode proposée utilise la technique du Max-Min [5]. Cette méthode fournit un partitionnement optimal du nuage couleur puisqu'il maximise les distances entre les couleurs représentatives du dit nuage.

Nous considérons les modifications suivantes apportées à l'algorithme LBG :

- Les  $n$  premiers vecteurs code  $\mathbf{y}$  sont générés en utilisant la technique usuelle de l'algorithme LBG.

- Au lieu d'utiliser une perturbation aléatoire sur les  $n$  premiers vecteurs code, les  $n$  vecteurs suivants sont générés en utilisant la méthode du Max-Min. On recherche les vecteurs  $\mathbf{x}$ , appartenant à l'image, les plus éloignés des  $n$  premiers vecteurs code  $\mathbf{y}$ . La sélection des couleurs les plus représentatives s'effectue selon le processus suivant :

soit  $V_n = \{q_1, \dots, q_n\}, n < N, V_n \subseteq \mathbf{D}$ , les  $n$  premières couleurs représentatives obtenues avec l'algorithme LBG.

Chaque couleur  $q_{n+i}, \forall i \in [1, 2, \dots, n]$ , est donc déterminée par :

$$\min_{j=1, \dots, n} \|q_{n+1} - q_j\|^2 \geq \min_{j=1, \dots, n} \|x - q_j\|^2 \quad (3)$$

$\forall x \in \{I - V_n\}$ ;  $I$  représente le training set.

Cette approche permet donc de prendre en compte les couleurs les plus excentrées du nuage couleur.

## 3 Résultats

Pour valider notre méthode, nous présentons des résultats sous forme de mesures objectives et mesures issues d'expériences psychophysiques. La figure 1 est une illustration typique de la différence d'évolution du nuage couleur par les deux méthodes. Les mesures objectives usuelles (SNR, PSNR, etc.) ne permettent pas de mettre en évidence les artefacts observés lors de la compression. Les mesures objectives utilisées reposent alors sur l'évolution du nuage couleur, les temps de calcul et de la distorsion moyenne entre l'image compressée et l'image originale.

### 3.1 Mesures objectives

La figure Fig.1 illustre les performances de la méthode proposée sur l'image couleur de textile Fig.1(a). La figure Fig.1(b) représente le nuage des couleurs associé à l'image Fig.1 non compressée. La figure Fig.1(c) montre la prise en compte des couleurs avec l'algorithme LBG, tandis que la figure Fig.1(e) montre la disposition du nuage de couleur avec la nouvelle approche, et ce pour un même taux de compression égal à 10 :1. On voit nettement que les couleurs situées aux extrémités du nuage ont été sélectionnées avec la nouvelle méthode. De même pour un taux de compression plus élevé (20 :1) nous arrivons à la même conclusion : nous obtenons une meilleure prise en compte de la répartition spatiale du nuage couleur lors de la construction du dictionnaire par la méthode du LBG modifié (Fig. 1.(f)) que par la méthode LBG classique (Fig. 1.(d)).

Le tableau 1 vient confirmer l'amélioration globale induite par notre technique sur le schéma classique de l'algorithme LBG. Ce tableau présente la distorsion moyenne entre l'image originale et l'image compressée selon les deux méthodes. L'algorithme LBG modifié engendre une distorsion moindre par

TAB. 1 — Distorsion moyenne entre l'image originale et l'image compressée de la figure 1 selon les deux méthodes.

Méthode	Taux	
	10 : 1	20 : 1
LBG classique	0.207 dB	0.258 dB
LBG modifié	0.169 dB	0.231 dB

rapport à la solution LBG pour les deux taux de compression 10 : 1 et 20 : 1.

L'utilisation de l'algorithme Max-Min permet une amélioration de la répartition des vecteurs code  $\mathbf{y}$  dans l'espace. Cette amélioration est obtenue au détriment du temps de calcul qui est accru dans des proportions de l'ordre de 30 % (Tab. 2). Dans ce tableau, le temps est exprimé en temps CPU, sur une station SUN SPARCstation 20.

TAB. 2 — Temps d'exécution des algorithmes pour l'image de la figure 1.

Méthode	Taux	
	10 : 1	20 : 1
LBG classique	20.09 u	12.64 u
LBG modifié	26.07 u	17.43 u

## 3.2 Mesures Psychophysiques des résultats

### 3.2.1 Méthodologie

Lors de cette expérience, les images traitées sont codées sur 24 bits, soit 8 bits par canal. Ces images sont affichés sur un écran CRT SUN 24 bits ayant une résolution spatiale de 1152x900 et géré par une carte vidéo de type GC14/SX. Lors de l'affichage sur cet écran, les couleurs des images subissent des déviations colorimétriques du fait de la non-linéarité de la réponse du moniteur. Cette déviation est corrigée à l'aide de la correction gamma calculée au préalable. Cette correction est appliquée systématiquement pour l'affichage des images.

Nous avons utilisés des mesures subjectives pour montrer l'amélioration de la qualité des images parmi un large panel d'images couleur, telles que des images de textiles, de scènes naturelles, des tableaux, ayant différents degrés de complexité. Les 10 observateurs sélectionnés ne connaissaient pas le but de ces expériences, et possédaient une vision normale des couleurs.

Nous avons demandé à ces observateurs de classer deux images suivant leur qualité visuelle : ces deux images sont en fait l'image reconstruite avec l'algorithme LBG et celle issue de la QV utilisant la nouvelle approche. Ces deux images sont présentées 160 fois, en affichant aléatoirement les images à droite ou à gauche l'une de l'autre. L'ordre d'apparition est aléatoire mais sous la contrainte suivante : l'image originale est affichée 80 fois à droite et 80 fois à gauche.

En utilisant la méthode de décision binaire, nous avons

calculé une mesure de discriminabilité  $d'$  [6] :

$$d' = \frac{1}{\sqrt{2}} [z(D) - z(F)] \quad (4)$$

où  $z(D)$  and  $z(F)$  sont les fonctions inverses de la loi de distribution normale [6] et  $D$  et  $F$  représentent les probabilités associées aux méthodes à comparer.

Cette valeur nous permet de mesurer la différence perceptuelle entre l'image traitée par l'algorithme LBG et celle traitée suivant l'approche proposée. Les images qui sont parfaitement différenciables ont un taux de détection de 1 et un taux de fausse alarme de 0. Si les images sont de moins en moins différenciables, les deux taux deviennent similaires.

### 3.2.2 Résultats

Les résultats issus de cette expérimentation sont présentés dans le Tableau 3 pour tous les observateurs. Ce tableau résulte de la comparaison entre l'image originale et de l'image reconstruite.

TAB. 3 — Réponses de tous les observateurs.

Stimuli	LBG	LBG Modifié
	à gauche	à droite
<LBG,LBG modifié>	72	8
<LBG Modifié,LBG>	10	70

Le couple <LBG,LBG modifié> représente la présentation de l'image compressée par l'algorithme LBG classique à gauche et de l'image compressée par notre approche à gauche.

Le calcul de la sensibilité de l'observateur est effectué selon l'équation 4 et nous obtenons alors  $d' = 1.71$ .

La valeur de sensibilité obtenue signifie, statistiquement, que l'observateur peut pratiquement toujours distinguer l'image compressée par la méthode LBG classique de l'image compressée par notre approche.

Les résultats de cette analyse permettent de mettre en évidence la supériorité de la méthode proposée.

## 4 Conclusion

Nous avons montré que cette nouvelle technique permet d'optimiser le dictionnaire par une meilleure prise en compte de la répartition des couleurs. L'effet moyenné des couleurs par l'algorithme LBG est ainsi évité, ce qui se traduit par une meilleure définition des couleurs représentatives du dictionnaire.

La contrepartie de cette technique est une augmentation du temps de calcul du dictionnaire. Si on considère le compromis temps de calcul/qualité des images, l'augmentation du coût calculatoire n'est de toute façon pas un handicap.

## Références

- [1] A. B. Watson, K. R. Nielsen, A. Poirson, A. Fitzhugh, A. Bilson, K. Nguyen, J. Albert and J. Ahumada, "Use of raster framebuffer in vision research," *Behavior Research Methods, Instruments and Computers*, vol. 18, no. 6, pp. 587–594, 1986.
- [2] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1991.
- [3] C. Bei and R. Gray, "An improvement of the minimum distortion encoding algorithm for vector quantization," *IEEE Trans.*, vol. 10, pp. 1132–1133, 1985.
- [4] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. IT-28, pp. 129–137, March 1982.
- [5] G. Houle and E. Dubois, "Quantization of color images for display on graphic terminals," in *IEEE, Global Telecomm Conference, GLOBE-COM'86*, pp. 1138–1142, 1986.
- [6] N. A. Macmillan and C. D. Creelman, *Detection theory : a user's guide*. Cambridge University Press, 1990.

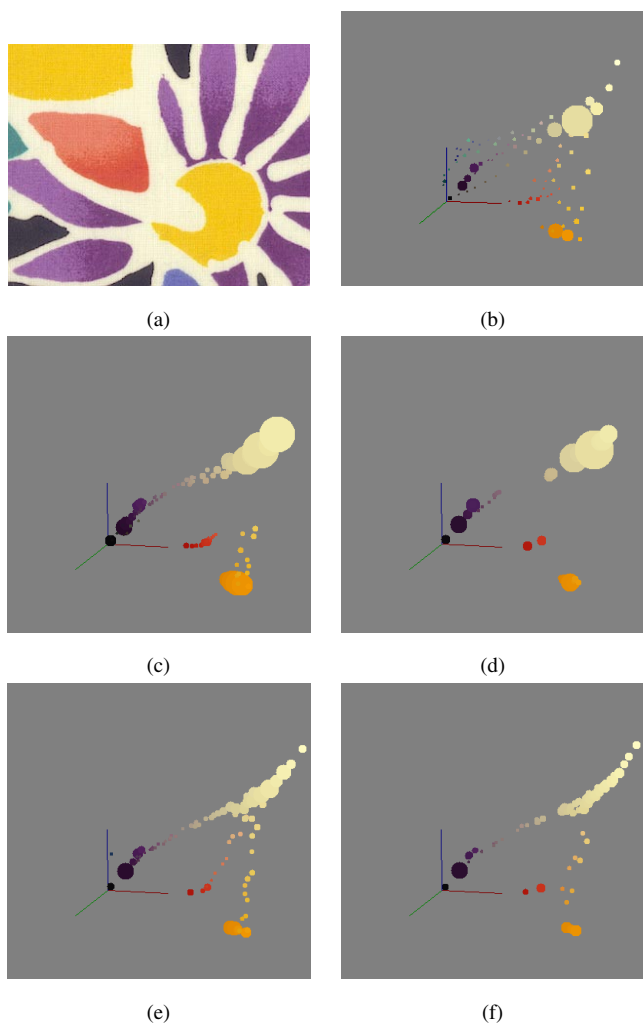


FIG. 1 — Évolution du nuage couleur initial (b) de l'image couleur de textile (a) lors de la compression par l'algorithme LBG à un taux de compression de 10 :1 (c) et 20 :1 (d) et par l'algorithme LBG modifié à même taux de compression de 10 :1 (e) et 20 :1 (f)