

Construction du Dictionnaire par Coopération de Modèles Statistiques et Perceptuels pour la Quantification Vectorielle d'Images Couleur

Christophe CHARRIER¹ Hocine CHERIFI²

¹Université Laval Département de Génie Électrique et Génie Informatique
Laboratoire de Radiocommunications et de Traitement du Signal
Sainte-Foy, Québec, Canada – G1K 7P4

²Université Jean Monnet – EIV
3 rue Javelin Pagnon, 42007 saint-Étienne cedex 01, France
charrier@gel.ulaval.ca, cherifi@vision.univ-st-etienne.fr

Résumé – La qualité des schémas de compression des images numériques couleur est fortement limitée par l'absence de coopération entre une approche statistique et une approche perceptuelle; ces deux domaines étant très souvent « ignorés » l'un par l'autre. Cependant une coopération entre ces deux approches permettrait un net gain de qualité des images reconstruites. Dans le contexte de la compression des images couleur par quantification vectorielle, nous définissons un algorithme de construction perceptuellement amélioré du dictionnaire basé l'exploitation conjointe des redondances statistiques et perceptuelles des images. Les résultats obtenus à l'aide de mesures de fidélité et psychophysiques permettent de conclure à un gain substantiel vis-à-vis du rendu des couleurs et des contours, pour une complexité et un taux de compression comparables à ceux de l'algorithme LBG.

Abstract – The quality of color digital image compression scheme is highly limited by a lack of cooperation between both statistical approach and perceptual exploitation; these two approaches being often “ignore” one from the other. Nevertheless, a cooperation between these two axis allow a gain of the quality of images. In the context of color images vector quantization, we propose to construct a perceptually tuned codebook. This is based on a jointly exploitation of statistical and perceptual redundancies. Results obtained from fidelity and perceptual measures let us obtain both the best color rendering and edge preservation, without an increase in complexity.

1 Introduction

D'une manière globale lors de la compression des images numériques, les approches basées sur l'exploitation des redondances statistiques et perceptuelles des images n'interagissent pas suffisamment. Ce manque d'interaction est d'autant plus marqué lorsque l'on traite des images couleur. En effet, la perception visuelle prend toute sa dimension (et sa complexité) dès lors que l'on se situe dans le domaine spatio-colorimétrique.

Ainsi l'utilisation de la technique de compression par la quantification vectorielle des image couleur peut être singulièrement améliorée en prenant en compte les exigences dictées par le système visuel humain.

Nous pouvons alors, soit tenter d'intégrer les caractéristiques fonctionnelles de la vision humaine dans le schéma de compression, soit optimiser l'algorithme de compression à l'aide d'expérimentations psychophysiques. Cette dernière approche présente l'avantage de n'utiliser qu'un nombre très restreint de connaissances *a priori* sur le fonctionnement du système visuel.

Nous utilisons alors une telle approche de façon à améliorer la qualité de reconstruction des images dans notre schéma de compression. Chacun des paramètres nécessaires au fonctionnement de l'algorithme LBG de construc-

tion du dictionnaire est ajusté, dans un premier temps, à partir de critères statistiques. Ce réglage est ensuite repris et affiné en fonction des impératifs induits par le système visuel.

2 Coopération des modèles statistiques et perceptuels

La construction du dictionnaire à partir d'une base d'images couleur en utilisant l'algorithme LBG [1] présente trois inconvénients majeurs, à savoir un temps de construction prohibitif, un phénomène de perte des couleurs et une dégradation rapide de la qualité de reproduction des contours [2].

2.1 Les étapes existantes

Dans des travaux précédents [2], nous avons adressé le problème de la complexité, de la perte des couleurs et de la dégradation de la qualité des contours.

Afin de réduire la complexité du schéma de construction du dictionnaire, nous avons réduit la redondance de la base d'apprentissage en recherchant un sous-ensemble de vecteurs représentatifs de cette base d'apprentissage

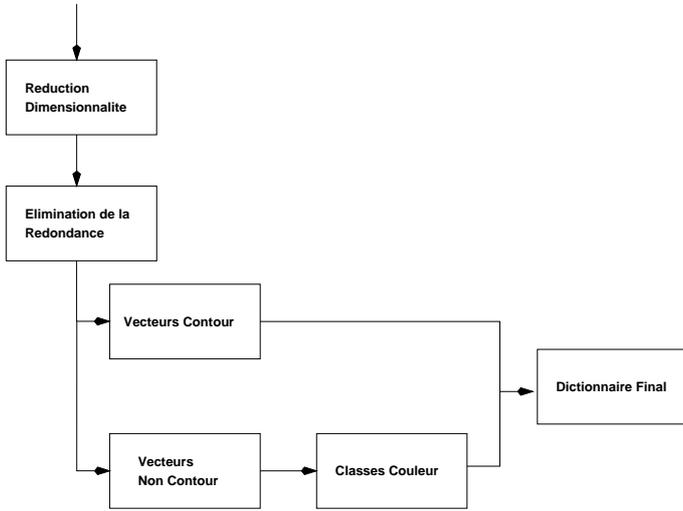


FIG. 1: Principe de l'algorithme perceptuellement amélioré

via une mesure de pertinence scalaire η perceptuellement optimisée [2]. Ceci nous a également permis de réduire la dimensionnalité du problème.

Pour améliorer le rendu des couleurs, nous avons effectué une classification de la mesure de pertinence η basée sur une modélisation d'un mélange fini de loi gaussienne, à l'aide de l'algorithme SEM. Ceci permet de générer autant de sous-dictionnaire que de classe dans la partition.

Enfin, nous avons freiné la dégradation de la qualité des contours en nous basant sur une carte de contour déterminée via le gradient de CUMANI [3], nous avons utilisé une classification de la base d'apprentissage en deux ensembles distincts composés respectivement de vecteurs contour et de vecteurs non contour.

Dans chacune des solutions algorithmiques apportées, nous avons déterminé les valeurs des différents paramètres à partir d'expérimentations psychophysiques [2, 4]. Ces optimisations ont eu lieu dans le domaine spatio-colorimétrique de façon à utiliser la sensibilité du système visuel humain en l'incluant dans le schéma de construction du dictionnaire.

Toutes ces solutions ont été abordées indépendamment les uns des autres.

2.2 Principe d'élaboration du dictionnaire

Nous proposons un algorithme *perceptuellement amélioré* de construction du dictionnaire intégrant chacune de ces étapes. La figure 1 est une représentation schématique de notre algorithme.

2.2.1 La dimensionnalité

L'étape de réduction de la dimensionnalité s'effectue à l'aide du calcul de la mesure de pertinence scalaire η . Cette mesure de pertinence a été définie comme une combinaison linéaire de la moyenne \hat{m} d'un vecteur \vec{x} et de l'écart-type $\hat{\sigma}$ du même vecteur: $\hat{\eta} = \alpha \hat{m} + (1 - \alpha)\hat{\sigma}$, $0 \leq \alpha \leq 1$. La valeur de α a été déterminée à l'aide de tests psychophysiques [2].

Dans une deuxième étape, nous construisons l'échan-

tillon représentatif Λ_0 de la base d'apprentissage scalaire Λ .

2.2.2 Les contours

À partir de l'échantillon Λ_0 , nous pouvons déterminer l'échantillon vectoriel associé. Il est alors possible de prendre en compte l'information contour en déterminant deux ensembles Θ_1 et Θ_2 qui correspondent respectivement à l'ensemble des *vecteurs contours* et des *vecteurs régions*. À partir de l'ensemble Θ_1 , nous déterminons la quantité de vecteurs code μ_1 du dictionnaire final, en prenant en compte que l'observateur humain requiert une quantité supplémentaire d'information contour. Ainsi la proportion finale de vecteurs code dans le dictionnaire \mathcal{D} est donnée par :

$$\mu_1 = \left(\frac{\text{card}(\Theta_1)}{\text{card}(\Theta_0)} + 0.1 \right) \cdot \text{card}(\mathcal{D}). \quad (1)$$

Par conséquent la quantité de vecteurs code μ_2 déterminés à partir de l'ensemble Θ_2 est égale à $(\text{card}(\mathcal{D}) - \mu_1)$.

2.2.3 La couleur

Nous allons ensuite appliquer l'étape de préclassification sur l'ensemble Θ_2 . Nous n'appliquons pas cette étape sur l'ensemble Θ_1 car il n'est constitué que de contours. Dans ce cas, une classification des contours n'apporterait pas d'information supplémentaire quant à cet ensemble. Les vecteurs qui constituent l'ensemble Θ_2 peuvent représenter des régions de complexité différente (*i.e.* textures, régions homogènes, etc.). Dans ce cas, nous considérons que la distribution de l'ensemble Θ_2 peut être modélisée par une loi mélange de gaussiennes. Nous appliquons alors sur cet ensemble la stratégie de préclassification.

Nous déduisons une partition des vecteurs de l'ensemble Θ_2 . Cependant le calcul des vecteurs code ne va pas s'effectuer proportionnellement à la taille finale M du dictionnaire, mais proportionnellement à la taille de l'ensemble μ_2 .

Le dictionnaire final \mathcal{D} s'obtient alors par réunion de tous les vecteurs code ainsi déterminés.

3 Performances de l'algorithme

De façon à évaluer les performances de l'algorithme LBG perceptuellement amélioré, nous utilisons conjointement des mesures de fidélité basés sur une combinaison de descripteurs de fidélité [2], ainsi que des tests psychophysiques de la qualité. Ceci permet de quantifier les différences de performances entre l'algorithme LBG et l'algorithme perceptuellement amélioré.

3.1 Environnement expérimental

Toutes ces mesures ont été calculées sur une même base d'images. Nous avons utilisé trois taux de compression différent : 1:10, 1:20 et 1:30. Ces taux correspondent respectivement à une très bonne, une passable et une mauvaise qualité d'image. La base d'apprentissage est constituée de 20 images couleur de complexité variable (tableaux, scènes naturelles, images de synthèse, etc.). La taille des vecteurs utilisée est de 4×4 .

Durant les tests psychophysiques, une correction gamma a été effectuée sur chacune des images présentées à l’observateur. De plus, les 10 observateurs qui ont participé à ces tests avaient tous vision normale des couleurs ainsi qu’une acuité visuelle correcte avec ou sans correction. L’écran utilisé lors de ces tests psychophysiques est un écran SUN CRT piloté par une carte graphique de type GC14/SX. La résolution spatiale de cet écran est de 1152 par 900 pixels avec un codage des couleurs de 8 bits par canal.

3.2 Mesure de fidélité

La figure 2 représente l’évolution des valeurs obtenues avec les descripteurs locaux de fidélité [5], en fonction de l’évolution du taux de compression. Plus le taux augmente, et plus la différence de qualité des images est évidente, au profit de l’algorithme LBG perceptuellement amélioré. Plusieurs constatations s’imposent :

- pour le taux de compression égal à 1:10, nous constatons que les différences de performances, en terme de qualité de reconstruction des images, entre les deux algorithmes ne sont pas encore significatives.

En effet, l’axe de courbure du nuage de points de la figure 2(a) est pratiquement plat. À ce niveau de compression, nous nous situons dans un cadre d’étude où la qualité des images est très bonne, pour les deux algorithmes.

- Plus le taux de compression augmente et plus la différence de qualité s’accroît.

En observant les figures 2(b) et 2(c), nous notons que l’axe de courbure du nuage de points est de plus en plus prononcé et qu’il se rapproche de plus en plus de l’axe des ordonnées, *i.e.*, pour des notes de qualité élevées pour les images compressées avec l’algorithme LBG perceptuellement amélioré, nous obtenons des notes de basse qualité pour les mêmes images compressées avec l’algorithme LBG.

Ceci signifie que nous obtenons une image de meilleure qualité lorsque nous utilisons l’algorithme LBG perceptuellement amélioré pour compresser les images.

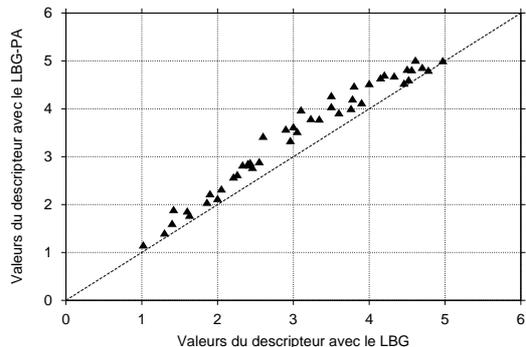
Si nous ne considérons que la mesure de fidélité pour juger les performances de l’algorithme perceptuellement amélioré, le gain de qualité ainsi obtenu est très significatif au-fur-et-à-mesure que le taux de compression augmente.

3.3 Mesure psychophysique

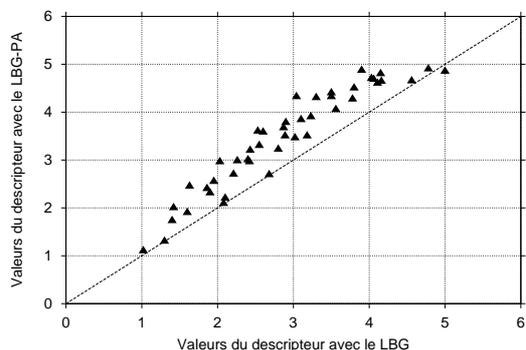
Les tests psychophysiques mis en place ont permis de calculer la sensibilité de discrimination des observateurs par rapport aux images présentées. Le test consiste à montrer à l’observateur deux images reconstruites : l’une compressée avec l’algorithme perceptuellement amélioré et la seconde compressée avec l’algorithme LBG. L’observateur devait répondre à la question suivante : « parmi les deux images qui vous sont présentées, où se situe l’image de meilleure qualité? »

La valeur de la mesure de la sensibilité est définie par [6] :

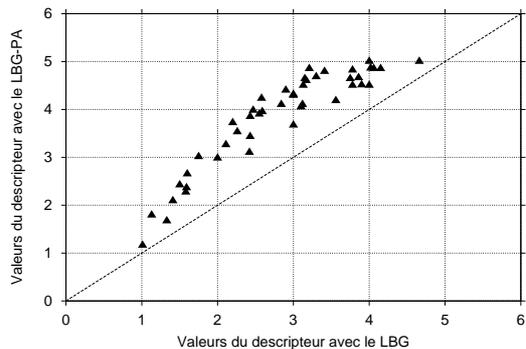
$$d' = \frac{1}{\sqrt{2}} [z(P(D)) - z(P(FA))] \quad (2)$$



(a) Comparaison des algorithmes pour un taux de compression de 1:10



(b) Comparaison des algorithmes pour un taux de compression de 1:20



(c) Comparaison des algorithmes pour un taux de compression de 1:30

FIG. 2: Résultats de la comparaison de l’algorithme LBG et de l’algorithme perceptuellement amélioré selon le combinaison des descripteurs de fidélité.

où $z(\cdot)$ représente la fonction inverse de la distribution normale, D une détection correcte et FA une fausse alarme. À partir des données obtenues pour tous les observateurs, nous avons mesuré la sensibilité de l'observateur pour les trois taux de compression (La propriété d'homogénéité de la population [2] a été testée et validée). Le tableau 1 donne les résultats des différentes sensibilités calculées.

Taux de compression	Sensibilité d'
1:10	1.34
1:20	1.97
1:30	2.28

TAB. 1: Résultats du calcul de la valeur de sensibilité de l'observateur.

De même, nous notons que les trois valeurs d' sont supérieures à 1.0, ce qui signifie que dans la majorité des cas, l'observateur a désigné comme méthode restituant les images avec la meilleure qualité, la méthode basée sur une préclassification de la base. La mesure d' calculée pour un taux de compression égal à 1:30 est supérieure à 2. Ceci signifie que la différence entre les deux images est importante.

Les résultats ainsi obtenus indiquent tous que la méthode proposée permet d'obtenir un dictionnaire de meilleure qualité que celui de l'algorithme LBG (et donc une image de meilleure qualité) au fur et à mesure que le taux de compression augmente.

3.4 Temps de calcul

Les résultats précédents ont montré que la qualité d'une image codée par l'algorithme LBG perceptuellement amélioré est meilleure que celle obtenue à partir de l'algorithme LBG standard. Puisque la construction du dictionnaire s'effectue sur un échantillon représentatif de la base d'apprentissage, l'algorithme proposé doit être plus rapide, sinon au moins aussi rapide que l'algorithme LBG classique (et ce même si nous utilisons une étape de séparation des contours suivie d'une seconde étape de préclassification dans l'algorithme LBG perceptuellement amélioré).

La figure 3 représente l'évolution du temps de calcul en fonction de la taille du dictionnaire pour les deux algorithmes considérés, à savoir :

1. *LBG 1* : l'algorithme LBG classique,
2. *LBG 2* : l'algorithme LBG perceptuellement amélioré dans sa globalité.

Le fait d'introduire différentes étapes permettant une meilleure prise en compte de la sensibilité de l'être humain, certes ralentit le processus (*LBG 2*), mais permet une amélioration de la qualité des images par rapport à l'algorithme *LBG 1*. Qui plus est, ce ralentissement du temps de calcul n'est pas dramatique dans ce sens où l'algorithme perceptuellement amélioré est pratiquement aussi « long » que l'algorithme LBG standard.

L'algorithme ainsi développé est de complexité équivalente à l'algorithme LBG.

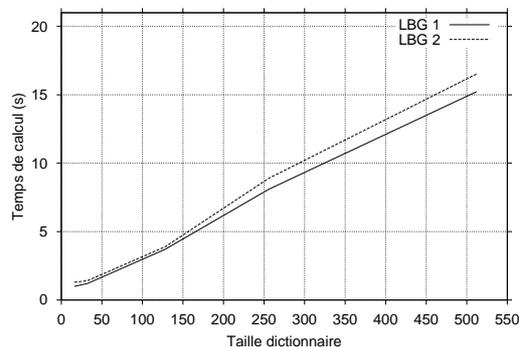


FIG. 3: Évolution du temps d'exécution de l'algorithme LBG perceptuellement amélioré, de l'algorithme LBG classique, et de l'algorithme perceptuellement amélioré privé de l'étape de prise en compte des contours.

4 Conclusion

Dans le schéma de construction du dictionnaire, nous avons utilisé conjointement une approche statistique et perceptuelle. Ceci nous a permis de constamment prendre en compte la sensibilité de l'être humain en vue d'une amélioration perceptuelle du schéma de codage. Les résultats obtenus montrent que cette approche permet un gain significatif de la qualité des images reconstruites.

Une telle approche conduit non seulement à des résultats probants, mais ouvre également un vaste champ d'applications.

Références

- [1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1991.
- [2] C. Charrier, *Vers l'optimisation statistique et perceptuelle de la qualité pour la compression des images couleur par quantification vectorielle*. PhD thesis, Université Jean Monnet, Saint-Etienne, France, 1998.
- [3] A. Cumani, P. Grattoni, and A. Giuducci, "An edge-based description of color images," vol. 53, pp. 313–323, July 1991.
- [4] V. Bedat, *Aspects psychovisuels de la perception des couleurs : Application au codage d'images couleur fixes avec compression de l'information*. PhD thesis, IRESTE - Université de Nantes, 1998.
- [5] C. Charrier and H. Cherifi, "Évaluation de métriques de qualité couleur pour la compression par quantification vectorielle," in *CORESA '99*, pp. 92–99, 14-15 juin 1999.
- [6] N. A. Macmillan and C. D. Creelman, *Detection theory: a user's guide*. Cambridge University Press, 1990.