

Critère de qualité visuelle d'images couleur – application à l'évaluation réaliste de codeurs d'images

Patrick LE CALLET, Dominique BARBA

IRCCyN UMR n°6597 CNRS

EPUN, rue Christian Pauc La Chantreterie

BP 50609 44306 NANTES CEDEX 3

France

plecalle@ireste.fr, dbarba@ireste.fr

Résumé – Nous proposons un critère objectif bi variant de qualité visuelle pour l'évaluation de méthodes de codage. Le critère est basé sur les propriétés du système visuel humain de manière à obtenir la meilleure correspondance avec les jugements humains. Contrairement à d'autres critères de la littérature, nous n'utilisons aucune information sur le type de dégradations introduites par les méthodes de codage. Le critère de qualité comporte deux fonctions principales, la première construit une représentation visuelle de l'image et des erreurs perceptuelles (basée sur les résultats d'expériences de psychophysique conduites au laboratoire), la seconde effectue la combinaison de ces erreurs perceptuelles pour produire une note de qualité. Nous proposons une approche originale pour cette seconde fonction dans la mesure où nous avons montré précédemment son importance. Nous comparons les notes fournies par ce critère avec celles issues de jugements humains sur une base de données de 140 images dégradées avec trois codeurs différents : JPEG, JPEG2000, une méthode basée sur des régions d'intérêt (ROI). Les résultats obtenus montrent une très bonne corrélation entre les jugements humains et les notes fournies par notre critère, beaucoup plus importante que celles liées au rapport Signal à Bruit ou qui s'en dérivent.

Abstract – We propose an (bi-variant) objective visual quality criterion for the evaluation of coding schemes. The criterion is based on human visual system properties in order to fit best with human judgements. Contrary to some others objective criteria, it doesn't use any information on the type of degradations introduced by the coding schemes. The objective criterion uses two main stages : the first one computes the visual representation of the image and the visual distortion too (based on results from psychophysics experiments made in the laboratory) and the second one performs the pooling of the perceptual errors into a single value: the quality rating. Since we have previously showed its importance we propose a new approach for modeling the second stage, extended here to color images. We compare results of this new criterion with human judgments on a database of 140 images distorted by three types of compression schemes (JPEG, JPEG2000, a ROI-based image coding scheme). Results show very good correlation between this quality criterion and human judgements, far above those given by SNR criteria or derived from it.

1. Introduction

L'appréciation de la qualité visuelle d'une image en sortie d'un traitement est une information précieuse en traitement d'images. Par exemple, la régulation de la qualité en sortie d'un codeur permettant d'adapter le débit est une application qui nécessite une évaluation de qualité visuelle de l'image codée puis décodée. Dans ce papier, nous nous intéressons à l'évaluation des schémas de codage avec pertes. En effet, ces schémas introduisent des distorsions visuelles entre l'image originale et l'image décodée qui dépendent du taux de compression, du type d'image et du schéma de codage. Bien qu'il soit connu que cette mesure soit faiblement corrélée avec les jugements humains, le Rapport Signal à Bruit (RSB) ou ceux qui s'en dérivent, est généralement utilisé pour mesurer la qualité d'une image en sortie d'un codeur (image décodée). L'autre technique pour accéder à une évaluation de qualité consiste à mener des tests subjectifs de qualité où des observateurs humains notent des images sur une échelle de qualité. Toutefois, cette méthode est si coûteuse en temps et exigent des conditions de mise en œuvre si particulières que très peu de concepteurs de schémas de codage l'utilisent en pratique. Nous proposons donc de définir un outil algorithmique permettant de fournir une mesure de la qualité

d'une image en sortie d'un codeur quelconque, hautement corrélé avec le jugement humain, afin d'en mesurer les performances.

Compte tenu de l'application visée, nous ne pouvons pas introduire de connaissance a priori sur la façon dont peuvent être codées les images. Dans le cas où la méthode de codage utilisée est connue, il existe déjà des critères spécifiques dans la littérature. Dans [1] et [2] on mesure la dégradation visuelle due aux effets de blocs (typique des méthodes de compression par bloc et transformation DCT) postulant que ce type de dégradations est introduit. L'utilisation de telles métriques, dites spécialisées, est intéressante dans le cas de la régulation de qualité d'un codeur donné mais n'est pas adaptée à notre cas. De plus, ce type de métriques négligent trop souvent les aspects purement psycho-visuels ce qui limite leur capacité à reproduire le jugement humain.

D'autres critères [3][4][5] incluent des modèles de vision humaine relativement élaborés permettant d'obtenir des cartes de différences perceptuelles entre deux images. La note globale qualifiant les différences entre les deux images (l'une est considérée comme la référence) est obtenue en combinant de façon généralement simpliste les informations des cartes d'erreurs perceptuelles. Ces métriques, plus génériques,

conviennent relativement pour l'évaluation des performances des codeurs. Les deux défauts principaux des métriques de ce style proposées dans la littérature sont les suivants:

- les modèles proposés ne reposent pas complètement sur des résultats expérimentaux ;
- les modèles de cumul des erreurs sont trop simplistes.

Le critère décrit dans ce papier propose des solutions originales afin de remédier à ces deux faiblesses.

2. Structure générale du critère

La structure générale du critère, illustrée figure 1, est composée de deux fonctions principales.

La première fonction permet de calculer une représentation visuelle de chaque image. Elle fournit plusieurs cartes qui définissent selon les canaux fréquentiels, l'information telle qu'elle est perçue en sortie des premiers traitements du système visuel humain. Pour le critère proposé dans ce papier, nous nous sommes beaucoup inspirés du modèle de perception de Daly [6] que nous avons étendu à la perception des images couleur. De plus, nous avons entièrement redéfini certains aspects du modèle à partir de résultats d'expériences psychophysiques, ce n'est pas le cas de la version initiale proposée par Daly.

Avec les représentations visuelles calculées pour l'image originale et l'image dégradée, nous obtenons par soustraction des cartes obtenues les cartes d'erreurs perceptuelles.

La deuxième fonction importante dans la construction du critère est celle du cumul des erreurs qui consiste à calculer une note globale de qualité à partir des cartes d'erreurs perceptuelles. Cette fonction est difficile à modéliser dans la mesure où elle met en jeu des aspects du système visuel humain encore mal identifiés. La plupart des auteurs utilise une sommation de Minkowski qui permet de pondérer les erreurs selon leur intensité. Selon nous, ce type d'approche est trop simpliste comparée au raffinement de la fonction de représentation visuelle. Nous avons déjà montré à quel point la fonction de cumul était cruciale en proposant des approches plus raffinées [7]. Ainsi, en prenant des modèles de représentation visuelle très simples associés à des méthodes de cumul évoluées, nous obtenons de meilleurs résultats en terme de correspondance des notes du critère avec les jugements subjectifs qu'en utilisant un modèle de représentation visuelle complexe associé à une fonction de cumul simpliste.

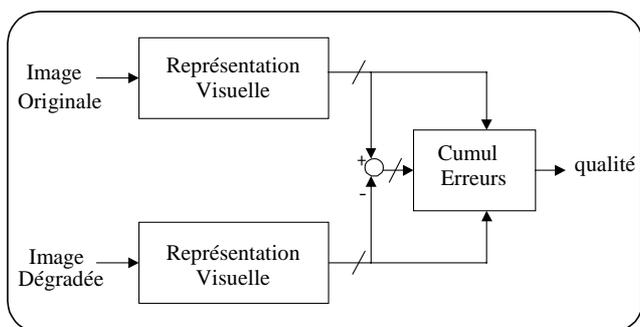


FIG. 1 : Bloc diagramme général du critère

Nous proposons dans ce papier un modèle plus réaliste de cumuls d'erreurs que les approches classiques.

3. Représentation visuelle

La figure 2 illustre la façon d'obtenir la représentation visuelle d'une image. Elle comporte quatre étapes :

- transformation de l'espace couleur RVB en A/CR1/CR2;
- calcul d'une fonction de sensibilité aux contrastes sur chaque composante de l'espace couleur ;
- décomposition en canaux perceptuels (n sous-bandes);
- calcul des effets de masquage sur chaque composante.

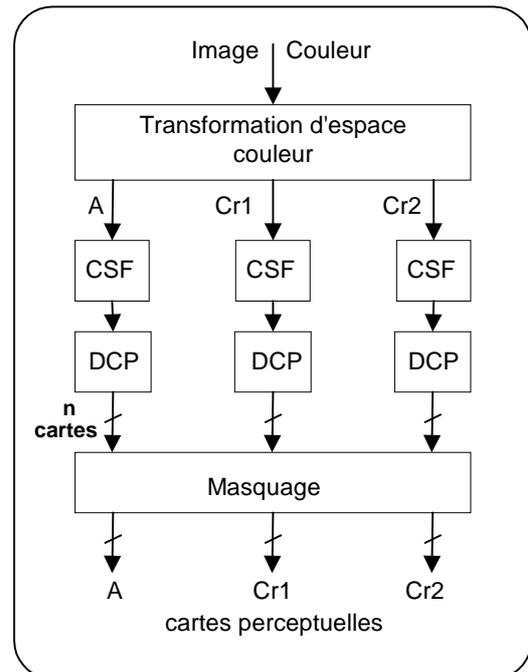


FIG. 2 : Construction des cartes perceptuelles

3.1 Espace couleur

Une théorie bien connue en vision humaine couleur consiste à combiner les réponses des trois types de cônes de la rétine (L, M et S) de manière à créer un espace de représentation des couleurs à 3 composantes antagonistes. Cette théorie repose sur l'existence de champs récepteurs au sein du système visuel humain présentant deux zones, l'une excitatrice et l'autre inhibitrice.

Dans le cadre de cette étude, nous utilisons l'espace défini par les composantes de Krauskopf [8]. Cette espace, validé par plusieurs expériences menées au laboratoire, est caractérisé par une composante achromatique A (antagonisme blanc/noir), et deux composantes Cr1 (antagonisme rouge/vert) et Cr2 (antagonisme jaune/violet).

3.2 Fonction de sensibilité aux contrastes

La fonction de sensibilité aux contrastes décrit les variations de la sensibilité d'un signal en fonction de sa fréquence spatiale. Pour la composante A, nous utilisons la fonction définie par Daly. Pour les composantes couleur Cr1 et Cr2, nous proposons deux modèles de filtres passe-bas à

partir des résultats d'expériences menées au laboratoire. Ces filtres incluent une variation de la sensibilité en fonction de l'orientation.

3.3 Décomposition en canaux perceptuels

Comme dans la plupart des approches proposées dans la littérature, nous utilisons une décomposition en sous-bandes afin de prendre en compte la spécialisation des champs récepteurs du système visuel. Le succès de l'utilisation des décompositions en canaux perceptuels tient dans leur capacité à faciliter la modélisation des effets de masquage.

L'originalité de la décomposition utilisée ici est qu'elle est complètement caractérisée à partir de résultats expérimentaux sur les trois composantes. Le pavage fréquentiel de la décomposition de la composante achromatique est illustré sur la figure 3, le pavage des composantes couleur est identique mais limité au domaine I et II. Nous obtenons ainsi 17 canaux pour la composante achromatique et 5 canaux pour chaque composante chromatique.

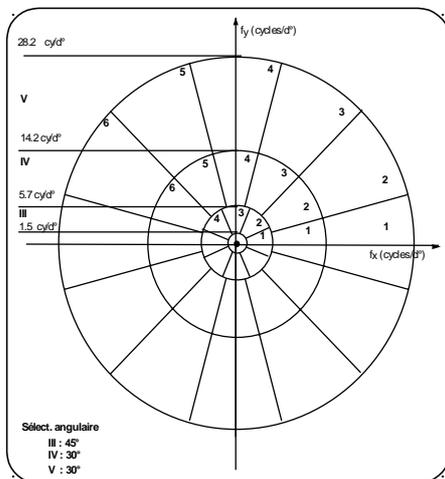


FIG. 3 : Pavage fréquentiel de la décomposition en sous-bandes pour la composante achromatique.

3.4 Masquage spatial

L'effet de masquage est un phénomène bien connu qui se traduit par des changements de visibilité d'un signal en fonction de son contexte. On sait par exemple, que les contours n'ont pas le même effet que les textures sur la visibilité d'un signal. Dans le cadre d'une décomposition en composantes et canaux perceptuels, on peut distinguer plusieurs effets de masquage. Nous nous intéressons ici uniquement aux effets de masquage intra composante et limités aux aspects intra canal. La fonction proposée est implémentée en adaptant le modèle de Daly à notre décomposition en canaux perceptuels. Ce modèle caractérise l'élévation relative du seuil de visibilité en chaque site de chaque canal en fonction de l'information en ce site. Les sorties de la décomposition en sous-bandes, grâce à la fonction de sensibilité aux contrastes, étant normalisées par rapport au seuil de visibilité sans signal masquant, il suffit de diviser ces sorties par l'élévation relative du seuil de visibilité pour prendre en compte les effets de masquage.

Daly ne proposant des paramètres que pour la composante achromatique, nous avons déterminé les paramètres pour les composantes chromatiques à partir d'expériences complémentaires.

4. Cumul des erreurs

La seconde fonction de cumul des erreurs permet de générer une note unique à partir des 27 cartes d'erreurs perceptuelles. Nous décomposons cette opération en trois étapes :

- le cumul fréquentiel qui consiste à regrouper toutes les cartes perceptuelles par composante correspondant chacune à un domaine spatio-fréquentiel particulier,
- le cumul composante qui consiste à regrouper les informations provenant des trois composantes,
- le cumul spatial qui consiste à fusionner les informations réparties spatialement dans une carte.

Concernant l'ordre ces étapes, nous avons montré précédemment qu'il était plus judicieux de réaliser le cumul fréquentiel avant le cumul spatial. La place du cumul composante n'a pas été étudiée, nous avons choisi de l'insérer entre le cumul fréquentiel et le cumul spatial.

4.1 Cumul fréquentiel

Le cumul fréquentiel est lui même décomposé en deux étapes :

- le cumul angulaire qui consiste à regrouper les informations de tous les canaux de chaque bande radiale en une seule carte. Nous utilisons alors une sommation de Minkowski,
- Le cumul radial qui consiste à cumuler les informations issues de chaque carte des différentes bandes de fréquences radiales. Nous avons montré précédemment [7] que la gêne liée à une dégradation, même normalisée par rapport au seuil de visibilité, dépend fortement de la bande radiale considérée. Nous utilisons donc une combinaison linéaire entre les sous-bandes pour pondérer leur importance.

4.2 Cumul spatial

Le cumul composante n'a pas été étudié pour l'instant, il se limite donc à une sommation des cartes provenant des trois composantes.

En revanche, Le cumul spatial fait l'objet d'une attention particulière car il s'agit d'un aspect sensible d'un critère de qualité. Nous avons proposé [7] plusieurs manières d'opérer. Nous proposons pour ce critère une pondération des erreurs selon les deux caractéristiques suivantes :

- la densité des erreurs : des erreurs éparpillées sont beaucoup moins gênantes que des erreurs concentrées dans une même région. Pour cela nous calculons la densité des erreurs sur des fenêtres d'extension spatiale correspondant au champ fovéal ;

la structure : les déformations de structure sont particulièrement gênantes. Pour cela nous avons distingué quatre classes d'erreurs, l'apparition de nouveaux contours, la

perte de contours, les erreurs sur les contour, enfin les autres types d'erreurs.

Chaque site est donc pondéré par un facteur de densité et un facteur lié à sa classe structurelle. Pour obtenir la note finale, nous utilisons une sommation de Minkowski.

5. Résultats et perspectives

5.1 Base de donnée d'évaluations subjectives

Nous avons élaboré une base de données d'évaluations subjectives afin de comparer les performances du critère avec les jugements humains. Nous avons utilisé 10 images de scènes naturelles que nous avons dégradées en utilisant 3 schémas de codage différents, JPEG, JPEG2000 et le codage LAR [9]. Ces trois codeurs ont été choisis d'une part pour la nature très différente des artefacts introduits, d'autre part parce qu'ils représentent trois approches principales en matière de codage d'images. En utilisant plusieurs taux de compression, nous aboutissons à 140 images évaluées.

Les évaluations subjectives ont été menées dans les conditions normalisées définies par l'ITU à 6 fois la hauteur d'écran et en utilisant une procédure DSIS (Double Stimulus Impairment Scale) à 5 catégories. Chaque image a été notée deux fois par 14 observateurs âgés de 20 à 45 ans ayant une vision normale (sans ou avec correction optique)

5.2 Evaluation du critère

L'évaluation du critère est réalisée en mesurant le coefficient de corrélation (cc) et le coefficient de corrélation de rang (ccr) entre les mesures subjectives et les notes du critère. Dans le tableau 1, nous comparons les performances du critère (critère complet) avec celles du PSNR et celles du critère utilisant une simple sommation de Minkowski comme fonction de cumul (critère simple).

TAB. 1 : comparaison des performances

	PSNR	critère simple	Critère complet
cc	0,533	0,758	0,859
ccr	0,657	0,761	0,866

Les performances du critère simple (approche typique de la littérature) qui caractérisent l'apport de la représentation visuelle, montrent une amélioration significative par rapport au PSNR. Toutefois, les performances ne sont pas très élevées. Il faut noter que les conditions d'évaluations sont critiques ici compte tenu de la diversité des dégradations et des images. De plus, les paramètres des fonctions de cumul n'ont pas été adaptés au jeu d'images de la base. Le critère complet affiche des performances correctes et nettement supérieures à celle du critère simple confirmant l'importance de la fonction de cumul.

5.3 Conclusion et perspectives

Les résultats obtenus sont très encourageants comme première approche de conception d'un critère de qualité

d'images couleur visant à remplacer le PSNR classiquement utilisé dans la littérature. Ils confirment notamment la pertinence de concevoir l'approche non seulement à partir d'un modèle psycho-visuel de représentation des images mais aussi en intégrant des aspects structurels.

Plusieurs évolutions sont envisageables afin d'améliorer les performances du critère, les deux principales concernent les aspects suivants :

- le modèle de masquage proposé reste relativement basique puisqu'il n'intègre ni les interactions inter canaux ni les interactions inter composantes. De plus, il repose sur des expériences de visibilité de signaux simples, il peut être intéressant d'utiliser plutôt un modèle basé sur la détection du bruit de quantification qui a donné de meilleurs résultats dans le cas de l'évaluation de qualité des images monochromes [7].
- Le cumul composante nécessite une étude approfondie afin de définir à la fois la méthode et sa place dans la fonction de cumul globale.

Références

- [1] Y. Horita, M. Katayama, T. Murai, and M. Miyahara, "Objective Picture Quality Scale for Video Coding", ICIP, Lausanne, Vol. III of III pp. 319-322, 1996.
- [2] A. B. Watson, "DCTune: A Technique for visual optimization of DCT quantization matrices for individual images " Society for Information Display Digest of Technical Papers XXIV, 946-949, 1993.
- [3] Stefan Winkler: "A perceptual distortion metric for digital color video." in Proc. SPIE Human Vision and Electronic Imaging Conference, vol. 3644, pp. 175-184, San Jose, California, 1999.
- [4] C. J. van den Branden Lambrecht, "Perceptual models and architectures for video coding applications", Phd thesis, Lausanne, 1996.
- [5] S. Comes and, B. Macq, "Human visual quality criterion", ICIP, Lausanne, Vol I of III, pp. 869-872, 1996.
- [6] S. Daly, "The visible different predictor : an algorithm for the assessment of image fidelity", in Proc. of SPIE, Vol. 1666, Human vision, visual processing and digital display III, pp 2-15, 1992.
- [7] P. Le Callet, and D. Barba, "Image Quality Assesment : From Site Errors to A Global Appreciation of Quality", Picture Coding Symposium, pp 105-108, Seoul, 2001.
- [8] D. R. Williams, J. Kraukopf, and D. W. Heeley, "Cardinal directions of color space", Vision Research, Vol. 22, pp. 1123-1131, 1982.
- [9] O. Déforges and, J. Ronsin, "Locally Adaptive Method for Progressive Still Image Coding", IEEE International Symposium on Signal Processing and its Applications (ISSPA), Brisbane, Australia, 22-25 august 1999.