

Une métrique psychovisuelle d'évaluation de la qualité des images couleur intégrant les effets de masquage inter-canal

Christophe CHARRIER¹, Thierry EUDE^{2,3}

¹Université de Caen - Basse Normandie
LUSAC EA-2607, Groupe Vision et Analyse d'Images
120, route de l'exode, 50000 Saint-Lô, France

²Université Laval, LSVN, cité universitaire, Québec, QC, Canada, G1K 7P4

³Université de Bourgogne, Le2i, UMR CNRS 5158, BP 47870, 21078 DIJON Cedex, France
c.charrier@chbg.unicaen.fr, eude@gel.ulaval.ca

Résumé – Tout traitement appliqué sur des images numériques doit, au final, être validé par un processus de qualité. Dans le cadre de la compression des images, ce processus repose sur la mesure de la qualité de l'image reconstruite. Or, le recours à des études subjectives est très coûteux car fastidieux à mettre en œuvre. Dès lors, de nombreuses métriques de qualité sont apparues avec pour principal objectif la corrélation avec le jugement humain. Dans ce papier, une métrique psychovisuelle de la qualité basée sur une décomposition en ondelettes et sur une mesure des effets de masquage intra et inter-bande est développée. Une étude des performances est réalisée en terme de corrélation et de robustesse au type d'images.

Abstract – Any process applied to digital images has to be validated by a performance measure. In the compression area, this measure of performance means a quality measure of the reconstructed images. Then, the use of psychophysical tests to measure the quality is quite time consuming. Therefore, many quality metrics have been defined in order to reach a high correlation degree with the human judgment. In this paper, a perceptually tuned metric based on a wavelet transform and a measure of the intra- and inter-channel visual masking effect is developed. A performance measure is then computed in terms of correlation and robustness to the kind of image.

1 Introduction

Les contraintes de stockage, de distribution et de présentation de l'information nécessitent l'utilisation d'algorithmes de compression réduisant la taille du message original. Dans la phase de compression, il est nécessaire de disposer d'une métrique de qualité pour évaluer la distance perceptuelle entre l'image originale et l'image compressée afin d'optimiser le taux de compression en fonction de la qualité requise. Une métrique de qualité doit évaluer le seuil de perception des dégradations pour les applications où les différences entre l'image originale et l'image reconstruite ne sont pas tolérées. De plus elle doit pouvoir fournir une échelle de dégradation perceptible pour les applications qui se satisfont d'une qualité moindre (qualité à la demande).

De nombreuses métriques de qualité développées spécifiquement pour l'évaluation de la qualité d'images compressées sont présentes dans la littérature. Les plus récentes intègrent une modélisation du système visuel humain. Ainsi, elles se proposent de prendre en compte les phénomènes connus comme la représentation de la couleur, la sensibilité au contraste et les effets de masquage. Le schéma communément admis consiste tout d'abord à représenter les images à évaluer dans un espace de représentation décorrélé des couleurs. Chacune des composantes résultantes de cette transformation est alors décomposée selon plusieurs canaux perceptuels. Une erreur est alors estimée pour chacun de ces canaux. La combinaison de ces erreurs, par sommation souvent pondérée, permet d'obtenir une

évaluation de la qualité. Cependant, ce traitement des erreurs en intra-canal ne prend pas en compte les effets de masquage dus aux interactions entre canaux différents. En effet FOLEY [1] a montré l'existence d'une interaction entre les différents canaux. WATSON et SALOMON ont proposé un modèle intégrant les interactions entre les canaux liés à la luminance et ceux liés à la chrominance [2]. ROSS et SPEED [3], ont proposé quant à eux, un modèle basé sur un terme adaptatif dépendant d'un ou de plusieurs canaux correspondants aux fréquences spatiales et à l'orientation.

Dans ce papier une nouvelle métrique d'évaluation de la qualité des images couleur compressées basée sur des modèles de vision (prenant en compte l'effet de masquage inter-canal) est présentée.

2 La mesure de la qualité

Le modèle perceptuel utilisé dans la mise en place de la métrique de qualité est décrit Figure 1. Il repose tout d'abord sur une transformation du triplet (R,V,B) dans l'espace YC_bC_r [4]. À partir de ce triplet (Y, C_b, C_r) , une décomposition en ondelettes est opérée de manière à obtenir une décomposition multi-canal selon différentes orientations et fréquences, telle que réalisée par le système visuel humain. Cette transformation est effectuée au moyen d'un filtre Daubechies 9/7 [5]. L'erreur est alors déterminée pour chacune des sous-bandes obtenues, et pour chaque coefficient w à la position (i, j) présent dans la

sous-bande b considérée avec :

$$e_b(i, j) = |w_b(i, j) - \hat{w}_b(i, j)| \quad (1)$$

La mesure de la distortion $d_b(i, j)$ est alors calculée en fonction d'un coefficient de masquage $m_b(i, j)$:

$$d_b(i, j) = \log \left(\frac{\alpha \cdot e_b(i, j)}{m_b(i, j)} \right) \quad (2)$$

où α est un paramètre déterminé en fonction de la sensibilité au contraste couleur, et $m_b(i, j)$ prend en compte le masquage intra-bande et le masquage inter-bande tel que :

$$m_b(i, j) = C_{\text{intra},b}(i, j) \cdot C_{\text{inter}}(i, j) \quad (3)$$

où $C_{\text{intra},b}(i, j)$ correspond à un seuil de visibilité appliqué à un coefficient localisé au point (i, j) dans la sous-bande b considérée et $C_{\text{inter}}(i, j)$ est une mesure des effets de masquage inter-bande.

On obtient ainsi pour chaque sous-bande, une mesure de distortion. Plus l'effet de masquage est important, meilleure sera la qualité (equ. 2). Finalement, une note globale est obtenue en combinant l'ensemble des mesures de distortion de chaque sous-bande et pour chaque composante colorimétrique, en utilisant une sommation de Minkowski :

$$E = \left(\sum_{Y, C_b, C_r} \sum_b \sum_{i, j} |d_b(i, j)|^\beta \right)^{1/\beta} \quad \text{avec } \beta = 5 \text{ [2]}. \quad (4)$$

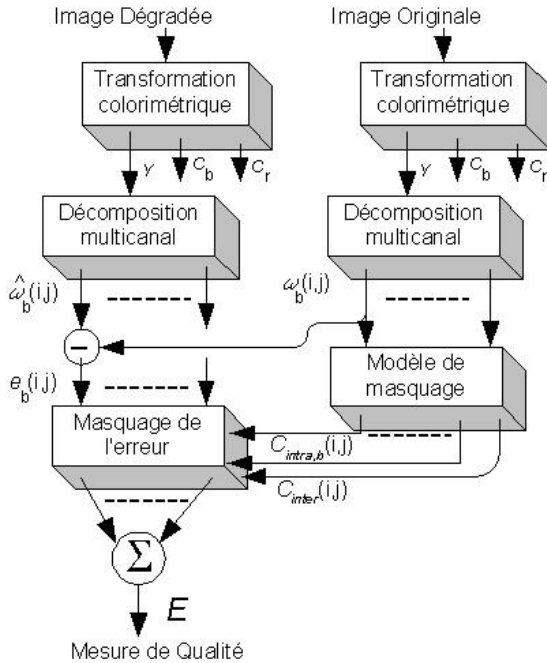


FIG. 1: modèle perceptuel utilisé.

2.1 Le changement de référentiel

La transformation des coordonnées (R,V,B) dans l'espace YC_bC_r est réalisée afin de réduire la corrélation entre les composantes colorimétriques. La composante Y représente l'axe de luminance tandis que le couple (C_b, C_r) est défini en tant que

composantes chromatiques. Cet espace est basé sur une transformation linéaire des coordonnées (R,V,B) gamma corrigées (R', V', B') définies dans l'intervalle [0,1] et codées sur 8 bits. Ainsi, la transformation linéaire suivante permet d'obtenir les nouvelles coordonnées (Y, C_b, C_r) :

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{pmatrix} \begin{pmatrix} R' \\ V' \\ B' \end{pmatrix} \quad (5)$$

2.2 Le modèle de masquage

Le masquage est un phénomène visuel très important, puisqu'il permet de décrire les interactions existantes entre les stimuli. Le terme de masquage est utilisé lorsque la perception du stimulus S1 est masquée par la présence d'un stimulus S2. Les effets de masquage se produisant dans les bandes de fréquence spatiale limitées en fréquence radiale et en orientation, il convient de prendre en compte le masquage intervenant dans une même bande de fréquence (masquage intra-bande) et le masquage induit par les interactions entre les bandes de fréquences différentes (masquage inter-bande), tels que décrits par l'équation 3.

2.2.1 Le masquage intra-bande

Si l'on considère un stimulus visible seul, il peut ne pas être perçu à cause de la présence d'un autre stimulus. En d'autres termes, le masquage augmente le seuil de visibilité en fonction du contraste du masque. Une étude a tenté d'exploiter les informations sur le niveau local de luminance afin de définir une nouvelle mesure du contraste local d'une image [6], mais aucune amélioration n'a pu être observée [7]. Aussi, la non linéarité de la fonction de seuil de visibilité $C_{\text{intra},b}$ peut-elle être approximée par :

$$C_{\text{intra},b}(i, j) = \max(1, w_b(i, j)^\epsilon) \quad (6)$$

où $\epsilon = 0.75$ [7].

2.2.2 Le masquage inter-bande

Le masquage dû aux coefficients présents dans les autres sous-bandes doit également être modélisé. Pour ce faire, il convient de pouvoir mesurer les interactions inter-bande selon les trois orientations et le support spatial obtenus par la décomposition en ondelettes. En d'autres termes, cela revient à introduire un coefficient de prédiction de l'effet de masquage selon les trois orientations. Ce coefficient, noté $h_{i,j}$, permet ainsi de calculer le masquage inter-bande localisé en une position (i, j) dans la bande u . La technique consiste alors à utiliser une fenêtre d'étude de taille $M \times M$ centrée sur le coefficient situé en (i, j) pour chacune des trois orientations. Un masque de pondération Gaussien ω_s est ensuite appliqué sur chacune des trois fenêtres (une par orientation). En effet, l'utilisation de noyaux Gaussien permet de caractériser la fonction de sommation simplement par leur variance. Les valeurs obtenues sont ensuite

sommées. Le résultat est alors pondéré par un facteur ω_f afin d'obtenir le coefficient $h_{i,j}$ selon la formule :

$$h_{i,j} = \sum_{u \in [1,2,3]} \omega_f(u) \sum_{(i,j) \in V_{i,j}} \omega_s(i,j) \hat{w}_{i,j,u}^2 \quad (7)$$

où u représente la bande utilisée selon les trois orientations (0° , 45° et 90°), et $V_{i,j}$ représente le voisinage $M \times M$ autour du coefficient centré en (i,j) . Dans cette étude, la valeur de M est fixée à 7, et ce pour deux raisons :

1. pour des valeurs inférieures à 7, la taille du voisinage est trop petit pour permettre une évaluation correcte et totale des interactions des sous-bandes selon les trois orientations,
2. pour des valeurs supérieures à 7, la taille du voisinage permet effectivement de prendre en compte la totalité des effets de masquage inter-bande selon les trois orientations. Cependant, le gain de précision obtenu par l'utilisation d'un masque de taille supérieure à 7 induit une telle augmentation (de type exponentielle) de la complexité en terme d'accès aux données et d'opérations arithmétiques, que son utilisation devient rédhibitoire.

Dans l'équation (7), les coefficients de pondération sont normalisés par

$$\sum_{u \in [1,2,3]} \omega_f(u) \sum_{(i,j) \in V_{i,j}} \omega_s(i,j) = 1, \quad (8)$$

équation dans laquelle les valeurs $\omega_f(u)$ et $\omega_s(i,j)$ sont décrites par la distribution gaussienne

$$\omega(R) = \frac{1}{(2\pi)^{\frac{2}{3}} \sqrt{\|\Lambda\|}} e^{-\frac{1}{2}(R^T)\Lambda^{-1}(R)} \quad (9)$$

où $\Lambda = E[(R - m)(R^T - m^T)]$ est la matrice de variance-covariance du masque de pondération.

Dès lors, la modélisation du masquage inter-bande peut être formulée comme suit :

$$C_{\text{inter}}(i,j) = \max(1, h_{i,j}^\gamma) \quad (10)$$

où γ est la pente de la courbe.

3 Évaluation des performances

3.1 Méthodologie

Plusieurs critères peuvent caractériser une métrique de qualité visuelle selon ses performances, ceci en rapport avec des évaluations subjectives. Ces critères sont la *justesse*, la *monotonie* et la *robustesse* en fonction du type d'images traitées (homogènes ou fortement texturées).

La *justesse* J_P d'une métrique dans la prédiction d'un jugement subjectif (avec une erreur moyenne minimale) peut être estimée en utilisant le coefficient de corrélation linéaire de PEARSON. Ainsi, pour un ensemble de données composé de N couples (x_i, y_i) , la justesse est donnée par :

$$J_P = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}} \quad (11)$$

où \bar{x} et \bar{y} sont les moyennes respectives des ensembles de données.

La *monotonie* m_S mesure quant à elle si l'accroissement d'une variable est associé à l'accroissement d'une autre variable. Idéalement, les variations de la valeur de la métrique pour différentes images devraient toujours avoir le même signe que les variations du jugement subjectif correspondant. Le degré de monotonie peut être évalué par la corrélation de rang de Spearman donnée par :

$$\begin{aligned} m_S &= \frac{\sum(\chi_i - \bar{\chi})(\psi_i - \bar{\psi})}{\sqrt{\sum(\chi_i - \bar{\chi})^2} \sqrt{\sum(\psi_i - \bar{\psi})^2}} \\ &= 1 - \frac{6(\psi_i - \chi_i)^2}{N(N^2 - 1)} \end{aligned} \quad (12)$$

où χ_i est le rang de x_i , ψ_i le rang de y_i et $\bar{\chi}$ et $\bar{\psi}$ sont respectivement les rangs médians. La corrélation de rang de Spearman est non-paramétrique, *i.e.*, aucune hypothèse sur la forme du rapport entre les x_i et les y_i n'est formulée.

Aussi, afin d'évaluer les performances de la métrique présentée, avons-nous d'une part utilisé les résultats de précédents travaux [8] qui consistaient à fournir une évaluation comparative de métriques proposées dans la littérature et d'autre part utilisé une métrique développée récemment [9]. Ainsi, dans [8], il a été montré que la métrique proposée par KARUNASEKERA *et al.* offre un bon compromis entre justesse et robustesse quant au type d'image (allant d'homogène à fortement texturée). Dans [9], LAI *et al.* ont développé une métrique de qualité basée sur une décomposition du signal en ondelettes de Haar en modélisant certains facteurs de la vision bas niveau (sensibilité au contraste, effet de masquage intra-bande, etc.). Cette métrique est corrélée avec les notes obtenues lors de tests psychophysiques de la qualité.

L'évaluation de la métrique proposée dans ce papier consiste ainsi à comparer cette dernière avec les deux métriques décrites précédemment, selon les trois critères de comparaison cités.

3.2 Résultats

Les images utilisées sont issues de la base d'image couleur « LIVE Quality Assessment Database » [10], dont la figure 2 représente un échantillon. La base est composée de 29 images originales (*i.e.*, non compressées) dont la complexité varie de faiblement texturée à fortement texturée. Ces dernières (toutes codées sur 8 bits par canal) ont été compressées à l'aide du standard JPEG2000 à des taux de compression différents de manière à obtenir une distribution des notes de qualité des images reconstruites relativement uniforme sur l'échelle de notation définie dans la recommandation UIT-R BT.500-10 [11]. Ainsi, la base d'image utilisée compte 198 images au total.

Afin de mesurer la corrélation existante entre les notes obtenues via la métrique proposée et les notes moyennes de qualité (*Mean Opinion Score*) obtenues par un groupe d'observateurs, les résultats de l'évaluation de la qualité fournis avec la base d'images LIVE ont été utilisés. Ceci a permis d'éviter un processus de notation par les observateurs très fastidieux à mettre en œuvre.

Le tableau 1 présente les résultats obtenus pour la métrique proposée, celle de KARUNASEKERA *et al.*, ainsi que celle de LAI *et al.*

Les résultats font apparaître une justesse moyenne supérieure à 97% ainsi qu'une monotonie moyenne de 96%, ce qui dé-



FIG. 2: Échantillon de la base LIVE.

| Métrique | justesse j_P | monotonicité m_S |
|--------------|----------------|--------------------|
| proposée | 0.972 | 0.964 |
| KARUNASEKERA | 0.927 | 0.921 |
| LAI | 0.951 | 0.944 |

TAB. 1: Justesse et monotonicité moyennes mesurées pour chacune des trois métriques.

note une forte corrélation entre les valeurs obtenues subjectivement et celles obtenues *via* la métrique.

La figure 3 présente le tracé des valeurs de qualité obtenues par la métrique proposée en fonction des notes de qualité fournies avec la base de données. On note que la métrique reste corrélée avec les notes moyennes des observateurs. Ceci denote une robustesse de la métrique vis à vis du type d'image à évaluer (fortement texturées ou homogènes).

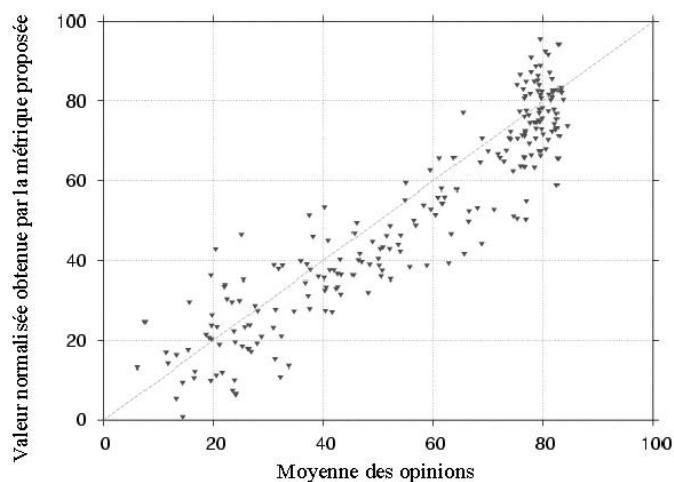


FIG. 3: Valeurs moyennes des opinions des observateurs en fonction de la qualité évaluée par la métrique proposée.

4 Conclusion

Une métrique de qualité des images couleur compressées intégrant une modélisation des effets de masquage (intra- et inter-bande) a été développée. Un changement de référentiel spatio-colorimétrique suivi d'une décomposition en ondelettes de l'image sont opérés au préalable sur l'image originale et l'image compressée. La modélisation intra-bande utilisée permet de quantifier le masquage sur une seule bande, tandis que la modélisation inter-bande développée intervient sur chacune des trois orientations obtenues. Ceci permet alors d'obtenir une mesure globale de l'effet de masquage sur les trois composantes colorimétriques.

L'utilisation d'une telle modélisation dans la mesure des effets de masquage permet ainsi d'obtenir une métrique fortement corrélée avec la perception humaine.

Références

- [1] J. M. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," *Journal of the Optical Society of America*, vol. 11, pp. 1710–1719, June 1994.
- [2] A. B. Watson and J. Solomon, "Model of visual contrast gain control and pattern masking," *Journal of the Optical Society of America*, vol. 14, pp. 2379–2391, Sept. 1997.
- [3] J. Ross and H. D. Speed, "Spatial and temporal contrast sensitivity function of the visual system," *Proceedings of the Royal Society of London*, vol. 246, pp. 61–70, 1991.
- [4] M. Nadenau and J. Reichel, "Opponent color, human vision and wavelets for image compression," in *Proc. of the Seventh Color Imaging Conference, IS&T*, (Scottsdale, Arizona), pp. 237–242, Nov. 1999.
- [5] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, pp. 205–220, 1992.
- [6] E. Peli, "Contrast in complex images," *Journal of the Optical Society of America*, vol. 7, pp. 2032–2040, Oct. 1990.
- [7] M. Nadenau and J. Reichel, "Image compression related contrast masking measurements," in *SPIE, Human Vision and Electronic Imaging*, (San-Jose, CA), pp. 188–199, Jan. 2000.
- [8] A. Mayache and T. Eude, "An evaluation of quality metrics for compressed images based on human visual sensitivity," in *Fourth International Conference on Signal Processing (ICSP'98)*, (Beijing, China), pp. 779–782, Oct. 1998.
- [9] Y.-K. Lai and C.-C. J. Kuo, "A haar wavelet approach to compressed image quality assessment," *Journal of Visual Communication and Image Representation*, vol. 11, pp. 17–40, Mar. 2000.
- [10] Laboratory for Image & Video Engineering, University of Texas (Austin), "LIVE Image Quality Assessment Database," <http://live.ece.utexas.edu/research/Quality>, 2002.
- [11] R. U.-R. BT.500-10, "Méthodologie d'évaluation subjective de la qualité des images de télévision," tech. rep., UIT, 2000.