

Elaboration de critères de mesure de qualité objective pour la qualité subjective

Ahmed TAMTAOUI^{1,3}, Marc RIBAGNAC², Driss ABOUTAJDINE³

¹INPT

Avenue Allal Al Fassi, Rabat Instituts, 10100 Rabat, Maroc

²ENIC

201, rue de la Cimaïse, 59650 Villeneuve d'Ascq, France

³GSC-LEESA

Avenue Ibn Batouta B.P. 1014, Rabat Maroc

tamtaoui@ibnsina.inpt.ac.ma, ribagnac@elv.enic.fr, aboutaj@fsr.ac.ma

Résumé – L'évaluation de la qualité est indispensable dans le codage d'image. Cette évaluation est souvent réalisée par des critères subjectifs tels que le Mean Opinion Score (MOS)[1]. Cette mesure reste lourde et ne peut être en aucun cas utilisée dans un schéma de codage d'images. Les mesures classiques objectives telles que PSNR ou EQM sont de bons indicateurs pour les distorsions de type erreurs aléatoires. Mais elles donnent des résultats moins bons pour les erreurs structurées (ex: effet de bloc, erreurs de contraste ou de structure spatiale). De nouvelles mesures objectives, utilisant une référence, ont été introduites [2][3][4][5] pour remédier à ces défauts structurels. Ces méthodes restent globales et masquent les erreurs locales visibles et gênantes pour un observateur humain. Afin d'y palier, nous proposons dans cet article une nouvelle méthode de mesure tenant compte de ces erreurs locales.

Abstract – The evaluation of picture quality is indispensable in image coding. Subjective assessment tests, such as the Mean Opinion Score (MOS)[1], are widely used to evaluate the picture quality of coded images. These measures are experimentally difficult, lengthy and cannot be integrated in a coding scheme. The simple distortion scales such as the Mean Square Error (MSE) or even the Peak Signal to Noise Ratio (PSNR) are good distortion indicators for random errors, but not for structured errors. However, there are many studies dealing with the construction of objective scales which are introduced [2][3][4][5] to alleviate these structured errors. Generally, these methods evaluate the quality in global fashion and mask the visible and strong local errors. This paper proposes a new methodology for the determination of objective quality metrics with respect to local errors.

1. Introduction

Dans les schémas de codage les plus répandus, l'erreur entre l'image codée-décodée et la référence est mesurée par l'Erreur Quadratique Moyenne(EQM) ou son corollaire, le PSNR. L'avantage de ce genre de méthode est la simplicité et l'efficacité. Les temps de calcul sont réduits ce qui convient bien aux codages ayant pour but de fonctionner en temps réel.

Toutefois, si l'EQM est une bonne mesure des bruits aléatoires, elle est parfois totalement inefficace lorsqu'il s'agit d'erreurs structurées. Nous pouvons citer les effets de blocs généralement produits par JPEG pour des taux de compression élevés ou par MPEG lors de mouvements rapides du sujet.

Si la mesure subjective, le MOS(Mean Opinion Score)[1], est réputée fiable, elle possède toutefois des inconvénients d'ordre pratique car la procédure est longue et donc totalement incompatible avec son intégration dans un schéma de codage. Dans le cadre de la surveillance d'un réseau de télévision[4], il serait possible de travailler sur des

échantillons mais à condition d'embaucher le nombre voulu de personnes ce qui est incompatible avec les contraintes de temps et de budget.

Il a donc fallu introduire de nouvelles mesures objectives, utilisant une référence pour remédier aux défauts de la mesure classique EQM et donnant des résultats aussi proches que possible du MOS[2][3][4][5]. Ces méthodes restent globales et masquent les erreurs locales visibles et gênantes pour un observateur humain.

Afin d'y palier, cet article met en oeuvre une nouvelle méthode de mesure de la qualité des images numériques monochromes fixes. L'amélioration par rapport aux méthodes existantes porte en particulier sur la prise en compte de la localité (concentration) des erreurs et notamment les erreurs ponctuelles qui deviennent visuellement gênantes par suite de leur concentration dans une même zone.

2. Critère global de mesure de qualité

Notre méthode se base sur l'application des critères de mesure locales sur des zones de l'image. Nous avons choisi

comme critères de mesure objective ceux proposés par Pasi Fränti pour leur simplicité et efficacité en codage. En effet, on obtient une meilleure corrélation par rapport à la méthode de Picture Quality Scale[3] (PQS : mesure qui est prise comme référence de comparaison pour les mesures objectives).

Notre méthode se décompose en trois étapes :

- la première étape consiste à calculer les différentes qualités locales et extraire deux catégories de blocs: bloc d'erreur pertinente et bloc de bonne qualité,
- la deuxième étape consiste à introduire une notion de densité d'erreur permettant de faire ressortir la compacité des erreurs,
- la troisième étape consiste à donner à toute l'image une note globale en intégrant les mesures locales calculées.

2.1 Critères de mesure locaux

Ces critères sont ceux utilisés par Fränti [2] dans son article. L'image est analysée par un bloc 3x3 pixels sur laquelle on applique trois critères de mesure :

- Critère de mesure de l'erreur de contraste [6]:

$$d_1 = \frac{(\sigma_A - \sigma_B)^2}{\max(1, \sigma_A)} \quad (1)$$

- Critère de mesure de l'erreur de structure spatiale :

$$d_2 = \frac{(|G_{x_A} - G_{x_B}| + |G_{y_A} - G_{y_B}|)}{2 \max(1, \sigma_A)} \quad (2)$$

- Critère de mesure de l'erreur de quantification :

$$d_3 = (Q_A - Q_B)^2 \quad (3)$$

où A et B désignent des blocs 3x3 de l'image originale et de l'image dégradée respectivement. σ est la déviation standard. G_x et G_y sont les réponses horizontales et verticales aux détecteurs de contours. Q_A et Q_B sont le nombre de niveaux de gris distincts.

Le critère de mesure locale est la pondération de ces trois critères :

$$d = w_1 d_1 + w_2 d_2 + w_3 d_3 \quad (4)$$

Les valeurs des pondérations sont choisies à $w_1=0.45$, $w_2=0.30$, $w_3=0.25$. Ces valeurs ont été optimisées en maximisant la corrélation des notes finales avec les mesures subjectives.

La note finale est la moyenne des notes locales. Cette note est ensuite remise à l'échelle [0, 1] par une fonction d'échelle[2].

2.2 Densité d'erreur

Même si l'erreur est faible, du fait d'une concentration suffisante, cette erreur deviendra gênante pour l'oeil humain. D'où l'intérêt de l'introduction de la notion de densité

d'erreur dans la nouvelle mesure globale.

La densité d'erreur se calcule sur des fenêtres de taille carrée. Si dans cette fenêtre il y a suffisamment de blocs dont la qualité est inférieure à un seuil S_1 (nous qualifierons dans ce cas ces blocs d'erreurs comme pertinents), nous dirons que l'erreur est concentrée dans cette fenêtre et donc détectable par l'oeil humain.

Nous représentons la densité d'erreur par le rapport N_p/N où N_p est le nombre de blocs 3x3 pixels d'erreurs pertinentes dans une fenêtre 7x7 blocs. $N=7 \times 7 = 49$ est le nombre de blocs dans cette fenêtre.

2.3 Méthode finale de mesure de qualité

L'analyse de qualité se déroule de la manière suivante : L'image est parcourue avec une fenêtre glissante de 49 (7x7) blocs. La détermination de ce nombre a été choisie de manière à obtenir une meilleure corrélation de la mesure finale avec les tests subjectifs, le MOS. Pour attribuer une note à la fenêtre i nous utilisons la formule suivante :

$$N_L^i = Q_M^i (1 - \alpha \frac{N_p}{N}) \quad (5)$$

où $\frac{N_p}{N}$ est la densité de l'erreur dans la fenêtre i et α est un paramètre fixe d'adaptation à la qualité subjective globale de l'image.

Q_M^i est la qualité moyenne sur la fenêtre i calculée suivant la méthode de Fränti (équation (5)).

A chaque calcul de l'équation (5) nous regardons si le résultat permet de minimiser la note obtenue. La note globale qui sera attribuée à l'image est la note de la zone de plus mauvaise qualité (minimum) :

$$Note_globale = \min_i (N_L^i) \quad (6)$$

3. Simulations et résultats

Nous avons basé nos simulations sur les images *Lena* et *Barbara* (figure 1) dégradées par :

- l'algorithme de compression de JPEG avec différents taux de compression exprimés en pourcentage de qualité d'images initiales (80%, 60%, 40%, 20%, 15%, 10%, 5%) (figures 2),
- des bruits ponctuels simulés plus ou moins perceptibles (figures 3).
- du bruit de type « Poivre et sel » (figure 4)

Pour les mesures subjectives nous avons mobilisé 17 observateurs et nous nous sommes référés aux conditions qui avaient été utilisées par Pasi Fränti. L'échelle de note retenue est le MOS étendu (0 à 10). Par la suite, la note

finale sera normalisée entre 0 et 1. Nous avons aussi accordé la possibilité aux observateurs de donner des notes par demi valeurs.

3.1 Optimisation des paramètres de la nouvelle méthode

Les paramètres qui entrent en jeu dans notre méthode sont la taille de la fenêtre d'observation locale, le seuil S_1 de pertinence des erreurs des blocs de 3x3 pixels et le paramètre d'adaptation à la mesure MOS (α).

La taille de la fenêtre d'observation locale (facteur de regroupement des blocs 3x3 pixels) a été optimisée à 7x7 (49 blocs) suivant une meilleure corrélation de la qualité objective au MOS. La qualité mesurée obtient des notes très basses (parfois zéro) lorsque le facteur de regroupement est minimal car nous détectons par ce moyen des erreurs très locales mais fortes. Ensuite il augmente jusqu'à un certain point qui correspond à la qualité mesurée sur toute l'image.

Le seuil S_1 permet de déterminer la concentration des blocs d'erreurs visibles dans une fenêtre d'observation locale. Comme la note finale (équation (6)) de l'image est la qualité de la fenêtre d'observation dont l'erreur est la plus visible, S_1 fixe le degré de concentration des blocs erronés dans cette fenêtre. En maximisant la corrélation de cette note finale à la mesure subjective, notre choix est porté sur la valeur 0.3

La valeur de α dans la formule (5) est un paramètre d'adaptation des mesures calculées par notre méthode aux mesures subjectives obtenues. Cette valeur était arbitraire dans un premier temps pour être précisée suite aux nombreux tests. Nous avons choisi $\alpha=0.3$ car elle donne un maximum de corrélation des notes objectives aux mesures MOS.

3.2 Résultats de la nouvelle méthode

Les résultats obtenus sont nettement améliorés en comparaison avec la méthode de Fränti[2] (TAB. 1).

Dans le cas de la figure 3 (Barbara dégradée avec un bruit ponctuel) la méthode de Fränti donne une qualité supérieure à 0.99 alors que le MOS est de 0.6625. La note obtenue par la nouvelle méthode est de 0.7134, l'écart est d'environ 0.05 au lieu de 0.32 dans le cas de la méthode de Fränti.

Le graphique suivant (figure 5) montrent la corrélation entre les notes objectives obtenues par la méthode de Fränti et la nouvelle méthode proposée avec le MOS.

La corrélation obtenue par Fränti est de 0.835 et celle obtenue par notre méthode est de 0.9101.

Les résultats de la figure 5 montre que la méthode de Fränti (aussi pour les autres méthodes existantes) donnent de bons résultats dans le cadre des dégradations pas très visibles (ex : JPEG à taux de compression faibles). Nous avons pu

constater que pour certains groupes d'images une corrélation de 0.985 avec le MOS dans le cas de la méthode de Fränti. Les inconvénients des méthodes basées sur un calcul final par la moyenne sur toute l'image, se remarquent dans le cadre des dégradations ponctuelles très visibles.

TAB. 1 Comparaison entre les mesures subjectives(MOS) et les deux méthodes objectives (nouvelle méthode et méthode de Fränti)

	MOS	Nouvelle Méthode	Fränti
fig.2	0.529	0.4589	0.7618
fig.3	0.6625	0.7134	0.998

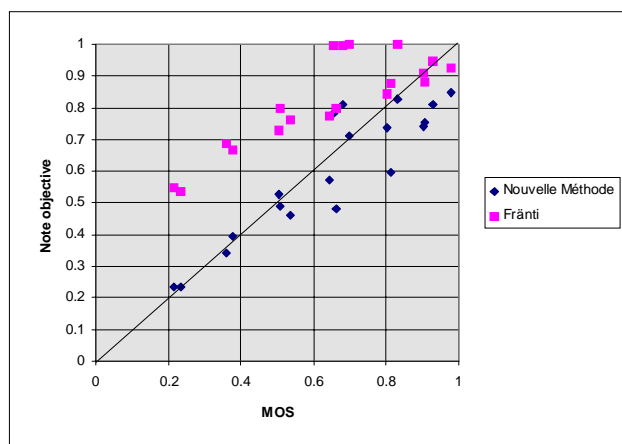


Figure 5 : Corrélation avec le MOS

La nouvelle méthode apporte plus d'améliorations sur la mesure des images dégradées par un bruit ponctuel (MOS faible) sans altérer la qualité des mesures pour celles qui étaient déjà bonnes avec la méthode de Fränti.



Figure 1 : Image originale Barbara



Figure 2 : Barbara, compression JPEG à 15%



Figure 3 : Barbara dégradée par un bruit ponctuel



Figure 4 : Barbara dégradée par un bruit « Poivre et Sel »

4. Conclusion et perspectives

Dans cet article nous avons présenté une amélioration des mesures de qualité objective par rapport à la mesure de Fränti[2] dans le cadre de la compression d'images fixes monochromes. Notre méthode s'inspire de la localité (densité) des erreurs qui tient compte de la vision humaine. Elle conserve les avantages et les bons résultats des méthodes existantes tout en améliorant la prise en compte de la localité des erreurs.

Toutefois il semble apparaître des obstacles aux développements futurs des méthodes objectives d'estimation de la qualité subjective. Il serait difficile de vouloir concevoir une méthode universelle, prenant en compte tous les cas possibles de dégradations d'images. Nous pouvons toutefois envisager des développements futurs dans ce domaine.

Notre méthode est encore perfectible dans la prise en compte des défauts locaux dans les images de qualité globalement homogènes. Il reste deux voies à explorer.

Premièrement, la modulation de la taille de la fenêtre d'observation des erreurs locales suivant la taille de l'image.

Deuxièmement, envisager une modification de la formule (5) qui tiendrait compte de plus d'un seuil. Des tests ont été faits mais il ne sont pas encore probants.

Références

- [1] CCIR, « *Method for the subjective assessment of the quality of television pictures* », Recommendations and Report of the CCIR, Rec. 500-2, Genève 1982
- [2] Pasi Fränti, « *Blockwise distortion measure for statistical and structural errors in digital images* », Signal processing: Image communication 13, pp 89-98, 1998.
- [3] M. Miyahara, K. Kotani, V. Ralph Algazi, « *Objective picture quality scale (PQS) for image coding* », IEEE transactions on communication, 1998.
- [4] S. Olsson, M. Stroppiana, J. Baña, « *Objective methods for assessment of video quality: State of the art* », IEEE transactions on broadcasting, vol. 43, n°4, Décembre 1997.
- [5] Wilfried Osberger, Anthony J. Maeder, « *Automatic identification of perceptual important regions in an image* », ICPR IEEE, 1998
- [6] Eli Peli, « *Contrast in complex images* », J Optical Society of America/ Vol. 7, N° 10/ Octobre 1990..