

Contribution d'un modèle d'attention visuelle à l'évaluation sans référence de la qualité des images compressées JPEG

Rémi Barland

Abdelhakim Saadane

IRCCyN-IVC, UMR n°6597 CNRS

École Polytechnique de l'Université de Nantes, rue Christian Pauc, La Chantrerie, BP 50609, 44 306, Nantes, France.

{remi.barland ; abdelhakim.saadane}@univ-nantes.fr

Résumé

A bas débit, les standards de compression actuels (JPEG) génèrent des dégradations visuelles pouvant gêner un observateur humain. De tels artéfacts sont généralement exploités pour évaluer la qualité d'une image sans référence. Cependant, même si les dégradations sont réparties dans l'image entière, certaines régions d'intérêt attirent le Système Visuel Humain et donc contribuent fortement à la détermination de la qualité perçue de l'image. Dans ce papier, nous proposons d'utiliser un algorithme simple d'attention visuelle pour pondérer des mesures de distorsions liées respectivement aux effets de bloc et de flou, principaux artéfacts créés par la compression JPEG. Ces mesures sont ensuite combinées afin de prédire une note de qualité. Une étude comparative des résultats prédits avec des notes de qualité issues de tests subjectifs, démontrent l'efficacité de l'approche.

Mots clefs

Évaluation de la qualité image sans référence, attention visuelle.

1 Introduction

L'apparition de nouvelles technologies et la recherche d'algorithmes de compression de plus en plus efficaces ont largement contribué au développement de nouveaux services d'acquisition et de diffusion d'image numérique. Pour tout nouveau service proposé, les standards actuels de compression image tels que JPEG, sont intégrés afin de minimiser la quantité d'information contenue dans une image, tout en offrant à l'utilisateur final, un maximum de qualité. La mesure en ligne de cette qualité perçue, qui dépend aussi bien des méthodes de codage utilisées que des erreurs de transmission, constitue aujourd'hui une des clés de l'explosion du tout numérique.

Les tests subjectifs représentent l'approche naturelle dans l'évaluation de la qualité : ils déterminent la note de qualité moyenne (MOS : « Mean Opinion Score ») à partir des jugements humains de qualité. Certaines

recommandations [1-3] de l'Union Internationale des Télécommunications (« International Telecommunication Union », ITU) spécifient pour chaque type d'applications les conditions d'observations, le choix des observateurs, le matériel à utiliser, les procédures à utiliser et les méthodes d'analyse des données. Bien que demeurant la référence dans le domaine, les tests subjectifs sont toutefois coûteux, longs et fastidieux à mettre en œuvre.

Les métriques de fidélité simples comme le PSNR, ou élaborées comme les métriques perceptuelles [4-8], qui connaissent également un essor considérable, ne sont pas du tout appropriées car elles nécessitent l'utilisation d'une image de référence qui n'est pas toujours disponible. Les métriques de qualité sans référence (NR : « No Reference ») représentent une alternative intéressante pour toutes les applications où il est nécessaire de contrôler en ligne, l'impact des dégradations induites par le processus de quantification. En général, les métriques NR combinent des mesures distinctes de dégradations en une seule [9], afin de prédire la qualité perçue. Dans le cas de la compression JPEG, deux défauts majeurs sont générés : les effets de blocs et de flou.

Dans la littérature, les mesures quantifiant les effets de blocs sont les plus nombreuses [10-16]. Cependant, peu intègrent des propriétés du Système Visuel Humain (SVH). Dans [13], la mesure d'effets de blocs prend en compte l'effet de masquage, sous la forme de pondérations calculées à partir des moyennes et écart-types des pixels situés aux frontières des blocs. Wang et al [11, 12] proposent de mesurer l'effet de blocs en analysant fréquemment le spectre de puissance d'un signal 1D constitué de la différence absolue de pixels consécutifs. Des effets de masquage liés à l'adaptation à la lumière et à la complexité des textures sont incorporés. En ce qui concerne l'effet de flou, la plupart des métriques proposées sont basées sur une approche signal [17-19]: Marziliano et al [18] proposent de mesurer le flou contenu dans une image de luminance, en calculant l'augmentation de la taille des contours à partir des points d'inflexion définissant le début et la fin de ces contours. Une autre approche [17] consiste à mesurer la précision des contours, en calculant le kurtosis local de chaque contour.

Dans [19], Marichal et al mesure l'effet de flou en exploitant l'information issue de l'histogramme des coefficients DCT.

Dans nos travaux antérieurs [20, 21], nous avons défini des mesures des effets de flou et de blocs, basées sur approche purement signal. Nous avons intégré dans le calcul de ces mesures de distorsion, des pondérations générées par un algorithme classant les régions contenues dans une image par importance perceptuelle [22]. Dans ce papier, nous proposons d'étendre ces travaux en intégrant une nouvelle carte d'importance perceptuelle. Un observateur humain n'effectue pas une analyse complète de l'image pour déterminer la qualité perçue mais plutôt, sélectionne certaines zones, apparaissant comme perceptuellement plus importantes. Ces zones ne correspondent pas forcément à des régions : des pixels appartenant à une même région de l'image n'ont pas forcément la même importance perceptuelle. A partir de ces deux observations, nous proposons d'identifier les zones perceptuellement importantes à l'aide d'un algorithme simple simulant l'attention visuelle pré attentive. La carte d'importance résultante va donc permettre d'identifier ces zones et d'intégrer dans le calcul des mesures de distorsion, leur contribution dans l'établissement de la qualité perçue des images compressées JPEG. Le papier est organisé de la manière suivante : la section 2 présente la structure de la métrique de qualité sans référence proposée, en détaillant la génération de la carte d'importance. La section 3 présente les expériences et les résultats. Enfin, des conclusions seront données dans la section 4.

2 Evaluation de la qualité sans référence des images JPEG

2.1 Structure de la métrique proposée

La figure 1 présente la structure de la métrique proposée : après une conversion couleur dans l'espace de Krauskopf [23], une carte d'importance est générée. Celle-ci est utilisée dans le calcul des mesures distinctes d'effets de blocs (BIM) et de flou (BM). La dernière étape de l'algorithme consiste à cumuler ces mesures de distorsion afin de déterminer une note de qualité prédite (pMOS).

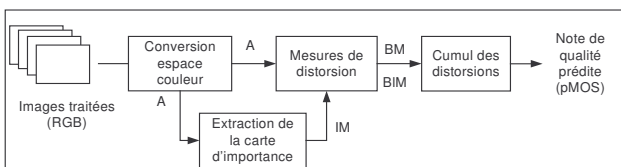


Figure 1 - Structure de la métrique sans référence proposée

2.2 Génération de la carte d'importance perceptuelle

Les observateurs humains focalisent leur attention sur certaines zones de l'image afin de déterminer la qualité perçue. Deux mécanismes sont mis en jeu : le premier appelé « bottom up » dépend du contenu de l'image, tandis que le second, « top down », se réfère à la tâche à accomplir. Dans cette section, la carte d'importance perceptuelle est générée par une approche « bottom up ». Dans la littérature, ce mécanisme de la vision humaine est simulé par des approches multi résolution décomposant l'image en différents canaux perceptuels liés aux sélectivités angulaire et radiale du SVH [24-27].

La figure 2 présente le fonctionnement de la génération de la carte d'importance. L'espace de couleur de Krauskopf [23] décompose l'image en trois canaux couleur : un achromatique A et deux chromatiques Cr_1 et Cr_2 . Ici, seule la composante achromatique sera utilisée. La composante achromatique est ensuite décomposée en cinq résolutions grâce à une pyramide laplacienne [28] : cette opération permet d'obtenir des images de résolutions plus petites, tout en éliminant la corrélation spatiale existante dans l'image de départ.

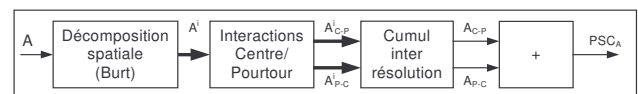


Figure 2 - Algorithme de génération de la carte d'importance perceptuelle

Ensuite, les interactions centre/pourtour sont incorporées : ces mécanismes de la vision humaine, réalisés par les champs récepteurs de l'aire corticale, détectent les discontinuités spatiales de l'image. Ainsi, pour chaque résolution i , l'émulation des champs récepteurs s'effectue de la manière suivante :

$$A^i_{C-P} = R(FPB(A^i, \sigma_1) - FPB(A^i, \sigma_2))$$

$$A^i_{P-C} = R(FPB(A^i, \sigma_2) - FPB(A^i, \sigma_1))$$

Où R est la fonction définie par $R(x)=0$ si $x \leq 0$, sinon $R(x)=x$. $FPB(., \sigma)$ est un filtre passe-bas, défini par la convolution de la composante achromatique A^i à la résolution i , par un filtre gaussien ($\sigma_1=0.4$ et $\sigma_2=2.4$).

Pour chaque résolution i , nous obtenons deux images distinctes modélisant les interactions Centre/Pourtour (A^i_{C-P}) et Pourtour/Centre (A^i_{P-C}). Afin d'obtenir une carte de saillance de même résolution que l'image passée en entrée mais aussi pour cumuler les informations

obtenues par chaque résolution i , un cumul inter résolution est réalisé :

$$A_{X-Y}^i(m,n) = \frac{A_{X-Y}^i(m,n)^2}{\text{Max}(A_{X-Y}^i(m,n), A_{X-Y}^{i-1}(m,n))^2 + \sigma^2}$$

Où σ est une constante assurant la non-nullité du dénominateur. Ce cumul inter résolution entre deux images de même interaction (Centre/Pourtour ou Pourtour/Centre) mais de résolutions différentes, permet de simuler les mécanismes excitateur et inhibiteur du SVH. A la sortie de ce contrôle de gain, nous obtenons deux images distinctes A_{C-P} et A_{P-C} , qui sont sommées linéairement, afin d'obtenir une carte 2D de points saillants (PSC_A).

La carte de saillance finale est générée à partir de cette carte de points saillants. Les zones où il existe une forte concentration de points saillants, sont les régions ayant une forte probabilité d'attirer le regard de l'observateur humain. Ainsi, un filtrage itératif est réalisé afin de mettre en évidence ces zones. Il est effectué de la manière suivante :

$$IM_{k+1} = \frac{IM_k + FPB(IM_k, \sigma_{PS})}{\text{Max}(IM_k + FPB(IM_k, \sigma_{PS}))}$$

Où $FPB(., \sigma_{PS})$ définit un filtre passe-bas réalisant une convolution avec un filtre gaussien ($\sigma_{PS} = 5$). Au départ du schéma itératif, IM_0 est initialisée avec les valeurs de PSC_A ; au bout d'une vingtaine d'itérations, nous obtenons la carte finale d'importance (IM). La figure 3 présente un exemple de carte d'importance perceptuelle, générée par l'algorithme proposé.

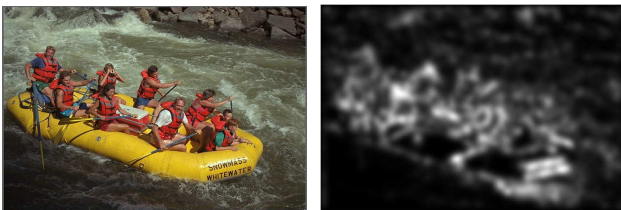


Figure 3 – A gauche, image originale « rapids » ; à droite, la carte d'importance perceptuelle associée

2.3 Mesures de distorsions

2.3.1 Mesure de l'effet de blocs

L'effet de blocs peut être défini comme une discontinuité artificielle entre deux blocs voisins appartenant à une image compressée par le standard JPEG. Cet artéfact résulte d'une quantification sévère et indépendante des coefficients DCT de chaque bloc. Il est d'autant plus visible si le contraste avec les blocs voisins est important. A partir de ces deux constatations, nous avons défini dans [21], une mesure locale ($LBLM(k,l)$) de l'effet de blocs, pour le bloc (k,l) :

$$LBIM(k,l) = \frac{R_H(k,l) + R_V(k,l)}{2} \cdot S(k,l)$$

Où $S(k,l)$ définit la sévérité de quantification et $R_H(k,l)$ (respectivement $R_V(k,l)$) est une valeur de renforcement liée au contraste horizontal (respectivement vertical) des blocs voisins.

Pour le bloc (k,l) , la sévérité de quantification est exprimée de la manière suivante :

$$S(k,l) = \frac{1}{1 + a \times STD(k,l)}$$

Où $STD(k,l)$ est la valeur de l'écart type du bloc considéré et a , une constante. Le renforcement de contraste horizontal est calculé par la formule suivante :

$$R_H(k,l) = 1 + C_H(k,l)$$

$$\text{Avec } C_H(k,l) = \frac{|A(k,l) - A(k,l-1)| + |A(k,l) - A(k,l+1)|}{2 \times \text{Max}\{|A(k,l) - A(k,l-1)|, |A(k,l) - A(k,l+1)|\}}$$

$A(k,l)$ (respectivement $A(k,l-1)$ et $A(k,l+1)$) est la valeur moyenne achromatique du bloc (k,l) (respectivement des blocs $(k,l-1)$ et $(k,l+1)$). Le renforcement de contraste vertical utilise la même formulation mais basée sur les valeurs des blocs verticaux voisins.

La mesure finale de l'effet de blocs est obtenue en sommant les mesures locales d'effets de blocs pondérées par la carte d'importance IM :

$$BIM = \left(\frac{1}{NB_V \times NB_H} \sum_{k=1}^{NB_V} \sum_{l=1}^{NB_H} (IM(k,l) \times LBIM(k,l))^p \right)^{1/p}$$

Où NB_V (respectivement NB_H) représente le nombre de blocs verticaux (respectivement horizontaux) contenus dans l'image traitée. Dans notre implémentation, la valeur du paramètre p est égale à 2.

2.3.2 Mesure de l'effet de flou

L'effet de flou est engendré par l'atténuation des coefficients hautes fréquences, durant l'étape de quantification JPEG. Cet artéfact est visuellement défini, comme une distorsion globale sur l'image entière, caractérisée par une augmentation de l'étalement des contours et des détails spatiaux. Ainsi, en tenant des pondérations perceptuelles fournies par la carte d'importance IM, nous avons formulé dans [20], la mesure d'effet de flou pour l'image de taille $M \times N$, de la manière suivante :

$$BM = \frac{\sum_{i=1}^M \sum_{j=1}^N IM(i,j) \cdot A'_{Edge}(i,j) \cdot I_A^2(i,j) \cdot \frac{N(A'_{Edge})}{M \times N}}{\sum_{i=1}^M \sum_{j=1}^N IM(i,j) \cdot A_{Edge}(i,j) \cdot I_A^2(i,j) \cdot \frac{N(A_{Edge})}{M \times N}}$$

Où A_{Edge} est l'image binaire issue d'une détection de contours (par filtrage Sobel), A'_{Edge} est l'image complémentaire de A_{Edge} . $I_A(i,j)$ définit l'intensité du pixel (i,j) de la composante achromatique. Enfin, $N(A_{Edge})$

(respectivement $N(A'_{Edge})$) est le nombre de pixels non nuls de A_{Edge} (respectivement A'_{Edge}).

2.4 Prédiction de la qualité

Afin de déterminer la qualité perçue de l'image traitée, nous effectuons un cumul des mesures de dégradations BIM et BM. Celui-ci est défini comme une combinaison linéaire des mesures d'effets de blocs et de flou, auquel on ajoute le terme croisé associé. Ainsi, la note de qualité prédite (pMOS) est obtenue par :

$$pMOS = a_0 + a_1 \cdot BIM + a_2 \cdot BM + a_3 \cdot BIM \cdot BM$$

Où les a_i ($i=0..3$) sont des pondérations dont les valeurs ont été ajustées pour optimiser la performance de prédiction.

3 Expériences et résultats

Pour déterminer la performance d'une métrique de qualité, la corrélation entre notes prédites et notes subjectives (MOS) est analysée. Actuellement, VQEG (« Video Quality Expert Group ») préconise certains indicateurs [29] quantifiant des propriétés adéquates à l'analyse de la performance. Le coefficient de Spearman et l'erreur quadratique moyenne (RMSE : « Root Mean Square Error ») définissent l'exactitude de prédiction ; le coefficient de Spearman, la monotonie ; le pourcentage de points aberrants (« Outlier ratio »), l'uniformité et le coefficient Kappa, l'agrément entre notes prédites et subjectives. Cette section présente les résultats de corrélation basés sur le traitement d'une base d'images compressées JPEG [30]. Elle est constituée de 29 images couleurs originales (24 bits/pixel, RGB, 768x512 pixels) et des versions compressées respectives.

Pour les besoins de notre étude, cette base d'images est divisée en deux. La première partie sert à optimiser les pondérations a_i du cumul inter distorsion (section 2.4) : ces poids sont estimés en minimisant l'erreur quadratique moyenne entre les notes prédites à partir des mesures de distorsion (BM et BIM) et les MOS correspondants. La deuxième partie (elle comprend 75 images) sert à analyser la performance de la métrique proposée en calculant les différents indicateurs statistiques de performance.

Le tableau 1 présente les résultats de corrélation de chaque mesure de distorsion liée respectivement aux effets de blocs et de flou. On distingue deux cas :

- le calcul des mesures de distorsion n'intègre pas les pondérations de la carte d'importance perceptuelle (BIM et BM) ;
- ces pondérations sont prises en compte (BIM_{IM} et BM_{IM}).

En comparant la performance de chaque mesure de distorsion, on constate que la plupart des indicateurs proposés par VQEG sont meilleurs lors de l'intégration

des pondérations perceptuelles. La corrélation de Pearson et l'erreur quadratique moyenne (RMSE) permettent de quantifier l'habilité à prédire la qualité perçue. Plus la corrélation de Pearson est proche de 1 et plus la RMSE est petite, meilleure est l'habilité de prédiction. En comparant les valeurs de la corrélation de Pearson et de la RMSE, on constate de meilleurs résultats pour les mesures de distorsion intégrant les pondérations perceptuelles. Ceci nous permet d'affirmer que l'utilisation de ces poids permet d'améliorer l'habilité de prédiction. La monotonie de prédiction est analysée à l'aide du coefficient de Spearman. Plus le coefficient de Spearman est proche de 1, meilleure est la monotonie. Les mesures de distorsion intégrant les pondérations perceptuelles obtiennent de meilleures valeurs, la monotonie de prédiction est donc meilleure. Le pourcentage de points aberrants est relativement constant pour toutes les mesures de dégradation. Enfin, le coefficient Kappa représente une mesure d'accord entre notes prédites et notes subjectives. S'il est supérieur à 0.4, on peut affirmer qu'il existe un bon agrément, résultat constaté pour chaque mesure de distorsion. De plus, nous remarquons que la valeur d'agrément est meilleure si les mesures de dégradation intègrent les pondérations perceptuelles.

	Pearson	RMSE	Spearman	Outlier	Kappa
BIM	0.910	0.865	0.921	12%	0.448
BIM _{IM}	0.932	0.626	0.943	13.3%	0.678
BM	0.912	0.711	0.90	13.3%	0.518
BM _{IM}	0.935	0.625	0.949	14.6%	0.732

Tableau 1 - Performances respectives des mesures de distorsion liées aux effets de blocs (BIM) et de flou (BM), tenant compte ou non, de la carte d'importance perceptuelle IM

Le tableau 2 présente les résultats des indicateurs de performance appliquée à la mesure conjointe des effets de blocs et de flou. Cette mesure est obtenue en réalisant le cumul inter distorsion défini dans la section 2.4. La métrique CM est basée sur les calculs des mesures de distorsion BM et BIM, n'intégrant pas les pondérations de la carte d'importance perceptuelle IM. Au contraire, la métrique CM_{IM} les intègre, puisqu'elle prend en entrée les mesures de distorsion BIM_{IM} et BM_{IM}.

En comparant les résultats de CM_{IM} par rapport à ceux de CM, on constate une nette augmentation de tous les indicateurs de performance. CM_{IM} présente une meilleure exactitude dans la prédiction de la qualité perçue (meilleurs coefficients de Pearson et RMSE) que CM. Pour les deux métriques, la relation de monotonie est respectée. De plus, CM_{IM} obtient une meilleure uniformité dans la prédiction de la qualité : le pourcentage de points aberrants (Outlier) est très faible. Enfin, la métrique CM_{IM}

obtient un meilleur accord entre notes prédites et subjectives.

	Pearson	RMSE	Spearman	Outlier	Kappa
CM	0.930	0.70	0.922	6.67%	0.518
CM _{IM}	0.965	0.485	0.954	1.3%	0.803

Tableau 2 – Performances de la métrique proposée : CM utilise le cumul proposé de BIM et BM (mesures de distorsion n'intégrant pas les pondérations perceptuelles IM) ; CM_{IM} utilise le cumul proposé de BIM_{IM} et BM_{IM} (mesures de distorsion intégrant les pondérations perceptuelles IM)

L'étude des tableaux 1 et 2 permet de confirmer l'apport de la carte d'importance perceptuelle dans un schéma de prédiction de la qualité. Cette carte permet d'identifier et de quantifier l'importance de certaines zones dans une image, susceptibles d'attirer le regard d'un observateur humain. La prise en compte de ces pondérations perceptuelles permet de mieux intégrer les corrélations spatiales existantes entre les pixels de l'image, dans le calcul des mesures de distorsion.

4 Conclusion

Dans ce papier, nous avons présenté la contribution d'une carte d'importance perceptuelle, générée par un algorithme simple simulant l'attention visuelle pré attentive, dans un schéma d'évaluation sans référence de la qualité des images compressées JPEG. La métrique proposée est basée sur la mesure respective des deux artefacts les plus gênants, engendrés par la compression JPEG : les effets de blocs et de flou. L'apport de cette carte a été démontré en comparant les résultats de prédiction de métriques intégrant ou non, ces pondérations perceptuelles dans le calcul des mesures de distorsion. L'intégration de ces pondérations perceptuelles permet d'identifier des zones susceptibles d'attirer le regard humain mais aussi de quantifier leur importance perceptuelle. Afin d'affiner cette carte et peut être améliorer le schéma d'évaluation de la qualité présenté, nos travaux futurs ont pour but d'intégrer d'autres caractéristiques influençant la vision humaine (couleur, orientation, etc.) et d'analyser leurs apports respectifs.

Références

[1] ITU-R Recommendation BT.500-10, Methodology for the Subjective Assessment of the Quality of Television Pictures. 2000, ITU: Geneva.
 [2] ITU-T Recommendation P.910, Subjective Video Quality Assessment Methods for Multimedia Application. 1999, ITU: Geneva, Switzerland.

[3] ITU-T Recommendation P.930, Principles of a Reference Impairment System for Video. 1996, ITU: Geneva, Switzerland.
 [4] Teo, P.C. and D.J. Heeger. Perceptual Image Distortion. in IEEE International Conference on Image Processing. 1994. Austin, Texas USA.
 [5] Lubin, J., A Visual Discrimination Model for Imaging System Design and Evaluation. Vision Models for target detection and recognition, 1995: p. 245-283.
 [6] van den Branden Lambrecht, C. and J. Farrell. Perceptual Quality Metric for Digitally Coded Colour Images. in European Signal Processing Conference. 1996. Trieste, Italy.
 [7] Winkler, S. A perceptual distortion metric for digital color images. in International Conference of Image Processing. 1998. Chicago, Illinois USA.
 [8] Fontaine, B., A. Saadane, et al. Perceptual Quality Metrics: Evaluation of Individual Components. in International Conference on Image Processing. 2004. Singapore.
 [9] Farias, M., S.K. Mitra, et al. Perceptual contributions of blocky, blurry and noisy artifacts to overall annoyance. in International Conference on Multimedia and Expo. 2003. Balitmore, Maryland USA.
 [10] Meesters, L. and J.-B. Martens, A Single-ended Blockiness Measure for JPEG-Coded Images. Signal Processing, 2002. 82(3): p. 369-387.
 [11] Wang, Z., H. Sheikh, et al. No Reference Perceptual Quality Assessment of JPEG Compressed Images. in IEEE International Conference on Image Processing. 2002.
 [12] Wang, Z., A. Bovik, et al. Blind Measurement of Blocking Artifacts in Images. in IEEE International Conference on Image Processing. 2000.
 [13] Wu, H.R., Z. Yu, et al. Impairment metrics for MC/DPCM/DCT encoded digital video. in Picture Coding Symposium. 2001. Seoul, Korea.
 [14] Caviedes, J. and J. Jung. No-Reference Metric for a Video Quality Control Loop. in World Multi-Conference on Systems Cybernetics and Informatics Broadcasting Convention. 2001.
 [15] Vlachos, T., Detection of Blocking Artifacts in Compressed Video. Electronics Letters, 2000. 36(13): p. 1106-1108.
 [16] Gao, W., C. Mermer, et al., A De-Blocking Algorithm and Blockiness Metric for Highly Compressed Images. IEEE Transactions on Circuits and Systems for Video Technology, 2002. 12(12): p. 1150-1159.
 [17] Caviedes, J. and S. Gurbuz. No-Reference Sharpness Metric Based on Local Edge Kurtosis. in IEEE International Conference on Image Processing. 2002. Rochestern, New York, USA.

- [18] Marziliano, P., F. Dufaux, et al., A No-Reference Perceptual Blur Metric. *Image Communication*, 2002. 19: p. 163-172.
- [19] Marichal, X., W.-Y. ma, et al. Blur Determination in the Compressed Domain Using DCT Information. in *IEEE International Conference on Image Processing*. 1999. Kobe, Japan.
- [20] Barland, R. and A. Saadane. Reference-Free Quality Metric Using a Region-Based Attention Model for JPEG-2000 Compressed Images. in *IS&T/SPIE Symposium on Electronic Imaging*. 2006. San Jose, California, USA.
- [21] Barland, R. and A. Saadane. A Reference Free Quality Metric for Compressed Images. in *Second International Workshop on Video Processing and Quality Metrics for Consumer Electronics*. 2006. Scottsdale, Arizona, USA.
- [22] Osberger, W. and A.M. Rohaly. Automatic Detection of Regions of Interest in Complex Video Sequences. in *SPIE Human Vision and Electronic Imaging*. 2001. San Jose, California, USA.
- [23] Williams, D.R., J. Krauskopf, et al., Cardinal Directions of Color Space. *Vision Research*, 1982. 22: p. 1123-1131.
- [24] Itti, L., C. Koch, et al., A Model of Saliency -Based Visual Attention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998. 20(11): p. 1254-1259.
- [25] Le Meur, O., P. Le Callet, et al. From Low Level Perception to High Level Perception, a Coherent Approach for Visual Attention Modeling. in *SPIE Human Vision and Electronic Imaging IX*. 2004. San Jose, California, USA.
- [26] Koch, C. and S. Ullman, Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry. *Human Neurobiology*, 1985. 4: p. 219-227.
- [27] Parkhurst, D. and E. Niebur, Texture Contrast Attracts Overt Visual Attention in Natural Scenes. *European Journal of Neuroscience*, 2004. 19: p. 783-789.
- [28] Burt, P.J. and E.H. Adelson, The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, 1983. 31(4): p. 532-540.
- [29] Rohaly, A.M., P. Corriveau, et al. Video Quality Experts Group: Current Results and Future Directions. in *Proceedings of Visual Communications and Images Processing*. 2000.
- [30] Sheikh, H., Z. Wang, et al., LIVE Image Quality Assessment Database.
<http://live.ece.utexas.edu/research/quality>.