

Amélioration de la qualité visuelle d'images couleur dans le cadre du standard de compression JPEG2000

Visual quality enhancement for color images in the framework of the JPEG2000 compression standard

Adrian Stoica, Mohamed-Chaker Larabi, Christine Fernandez-Maloigne

Laboratoire SIC – CNRS FRE 2731 – Université de Poitiers, Bat. SP2MI, Téléport 2, BP 30179 – 86962 FUTUROSCOPE Cedex – FRANCE

Manuscrit reçu le 15 juin 2004

Résumé et mots clés

Durant les dernières années, la prise en compte de modèles du Système Visuel Humain (SVH) dans l'évaluation de la qualité visuelle des images couleur compressées, est devenu un sujet de recherche majeur. Il semble naturel d'intégrer davantage les connaissances récentes sur la perception et la modélisation de la vision humaine, dans les systèmes de compression d'images. Ainsi, dans cet article, nous proposons une méthode d'intégration de ces connaissances pour l'augmentation de la qualité visuelle d'images compressées JPEG2000. Cette méthode consiste en deux parties : une évaluation de laboratoire pour la modélisation du SVH par la Fonction de Sensibilité au Contraste (CSF) et une technique de calcul de facteurs de pondération visuelle pour la compression JPEG2000, utilisant le modèle SVH évalué, dans le domaine de Fourier de l'image couleur.

Images couleur, qualité d'images couleur, système visuel humain, fonction de sensibilité au contraste, compression JPEG2000, facteurs de pondération visuelle.

Abstract and key words

In the last years, the consideration of different models of the Human Visual System (HVS) in the final perceived quality of the compressed images becomes a major research subject. Therefore, it is natural to take advantage of the recent knowledge on both perception and models of the human vision in an image compression system. Thus, in this paper we propose an integration method of that knowledge for the improvement of perceptual JPEG2000 image compression quality. This method consists on two parts : a laboratory evaluation of the HVS model by the Contrast Sensitivity Function (CSF), and an implementation technique of visual weightings for the JPEG2000 scheme, using the evaluated HVS model in the Fourier domain of the color image.

Color image, color image quality, human visual system, contrast sensitivity function, JPEG2000 image compression, visual weightings.

Remerciement

Ces travaux font partie d'un projet RNRT intitulé EIRE (Études d'optimisation algorithmiques de JPEG2000), financé par le Ministère de la Recherche et de l'Industrie, impliquant des partenaires universitaires ainsi que des industriels, dirigé par le groupe Thalès Communications. Les partenaires impliqués sont des membres actifs de la délégation française de normalisation AFNOR pour JPEG2000.

1. Introduction

L'augmentation continue du nombre d'applications concernant les images couleur et les vidéos, conduit à une demande très importante de stockage, d'affichage ou de transmission des données images générées. Evidemment, la compression d'images est une nécessité incontestable et différentes techniques de compression avec ou sans perte ont été implémentées, spécialement dans les standards de la famille JPEG. L'amélioration de la qualité visuelle finale d'images compressées est alors une demande naturelle, pour obtenir de meilleures performances en terme de taux de compression, en gardant dans le même temps une bonne qualité visuelle. Ainsi, des recherches récentes ont été réalisées sur le nouveau standard de compression d'images couleur JPEG2000, d'un point de vue psychovisuel [ISO00, ZDL00], pour obtenir ce qu'on pourrait appeler une compression sans perte visuelle.

L'une des premières pistes pour modéliser le Système Visuel Humain (SVH) consiste en la modélisation de la Fonction de Sensibilité au Contraste (CSF) [GH73, MS74, Mul85, Bar99]. Cette dernière a fait l'objet de plusieurs études et de plusieurs types d'utilisation. L'introduction d'un modèle du SVH caractérisé par la CSF, dans le schéma de compression JPEG2000 est une de ces voies qui permettent d'obtenir des résultats visuels encourageants [NRK03], notamment dans le cas des images couleur. Les différents tests et techniques proposés dans la littérature ont montré un grand potentiel pour l'amélioration de la qualité visuelle du standard JPEG2000.

Cet article est organisé comme suit : la deuxième section présente une brève description du nouveau standard de compression d'images, JPEG2000 ; la troisième section décrit la modélisation du Système visuel Humain (SVH) par la Fonction de Sensibilité au Contraste (CSF) ; la quatrième section présente l'interaction du modèle SVH avec le codeur JPEG2000 et finalement dans la cinquième section sont présentés des résultats expérimentaux sur la compression d'images couleur utilisant des codeurs JPEG2000 avec et sans facteurs de pondération. Enfin, cet article se termine par des conclusions et quelques perspectives.

2. Aperçu de JPEG2000

JPEG2000 est le nouveau standard de compression d'images, qui fournit un ensemble d'outils primordiaux pour les applications émergentes dans le domaine de l'image [SCE01].

JPEG2000 réalise de fortes compressions avec une qualité d'image supérieure à toutes les techniques standards d'encodage. Suite aux réunions du comité JPEG2000 de l'Organisation Internationale de Standardisation à la Nouvelle Orléans en Décembre 2000, JPEG2000 Partie 1, a été officiellement déclarée standard international.

Comme le montre la figure 1, JPEG2000 est basé sur une transformée en ondelettes discrètes (DWT), une quantification scalaire, une modélisation contextuelle, un codeur arithmétique et enfin une allocation de débit post-compression. La DWT est dyadique et peut être effectuée soit avec le filtre réversible LeGall (5,3), qui fournit un codage sans perte, soit avec le filtre bi-orthogonal non-réversible type Daubechies (9,7), qui fournit de meilleurs résultats pour les fortes compressions, mais sans pouvoir faire des compressions sans perte. Le quantificateur suit une approche scalaire basée sur une zone-morte, tout en étant indépendant pour chaque sous-bande. Chaque sous-bande est divisée en blocs rectangulaires (appelés blocs de code dans JPEG2000), généralement de taille 64×64 , codés entropiquement en utilisant la modélisation contextuelle et le codeur arithmétique.

Les données codées sont organisées dans des couches représentant des niveaux de qualité, utilisant l'allocation de débit post-compression et dirigées ensuite vers le code-stream en paquets. Le code-stream généré est divisible en différentes entités à savoir : une résolution, une couche (*i.e.* SNR – Signal-to-Noise Ratio), une position ou un composant progressif, ou bien n'importe quelle combinaison possible.

Le standard JPEG2000 supporte aussi un certain nombre de fonctionnalités, parmi lesquelles beaucoup sont inhérentes à l'algorithme lui-même. Un exemple réside en l'accès aléatoire, qui est possible grâce au codage indépendant des différents blocs de code et à la structure de paquets du code-stream. D'autres fonctionnalités supportées sont la possibilité de définir une Région d'Intérêt (ROI), l'erreur de résilience pour être robuste aux erreurs aléatoires, le traitement d'images multi-composantes, la palettisation des couleurs, la simple rotation et la compression sans perte, pour en mentionner quelques-unes. JPEG2000 [ISO00] est un codeur basé sur la décomposition en ondelettes où les coefficients de chaque sous-bande d'ondelette sont divisés dans des blocs de mêmes dimensions (appelés blocs de code) et chaque bloc de code est ensuite codé indépendamment. La structure du standard, introduit le concept de couches de qualité qui permettent un processus d'optimisation post-compression, où différents sous-bitstreams de chaque bloc de code sont assemblés dans un certain ordre (*e.g.*, optimisation en fonction du rapport débit-distorsion), pour former le bitstream final. Ce processus de création des couches de qualité est une



Figure 1. Schéma bloc de la compression JPEG2000.

composante essentielle dans JPEG2000. Ainsi, après la génération d'un sous-bitstream pour chaque bloc de code, c'est à l'encodeur de déterminer la façon de les assembler pour créer les couches de qualité. Autrement dit, c'est l'encodeur qui doit déterminer comment tracer les lignes des couches de qualité ou choisir les points de troncature. Cette flexibilité permet une allocation de bits adaptative en fonction des blocs de code et rend exploitable un certain nombre d'outils visuels dans JPEG2000.

3. Modélisation du système visuel humain

Les différentes recherches sur le Système Visuel Humain proposent des modèles concernant la façon dont les humains perçoivent le monde environnant [SB82, Dau80, KZ85]. Par exemple, des modèles ont été créés pour caractériser la sensibilité humaine à la luminosité et aux couleurs, par rapport aux fréquences spatiales [ZG81, DAT82, Gla83]. Un de ces modèles est la fonction de sensibilité au contraste (CSF) [Wan95, MM96, WYSV97]. Pour décrire cette fonction, il faut d'abord introduire la notion de seuil de contraste et de sensibilité au contraste.

Le contraste est un concept de base pour la science de la vision. Cela s'explique par le fait que l'information représentée dans le système visuel ne correspond pas au niveau de luminosité absolu, mais au contraste, qui est le rapport entre l'intensité locale et l'intensité moyenne de l'image entière. Les définitions les plus représentatives pour le contraste sont celles proposées par *Michelson* [Mic27] et *Weber* [OT86] à travers sa loi. Ces définitions ont été énoncées à partir d'expériences simples sur la luminance, représentant, par exemple, une grille sinusoïdale en niveaux de gris.

$$C = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}, \quad (1)$$

avec C le contraste et L_{\max} , L_{\min} respectivement les valeurs maximales et minimales de la luminance. Les neurones répondent seulement aux stimuli au-dessus d'un certain contraste. Le contraste nécessaire pour provoquer une réponse des neurones est défini comme le seuil de contraste. L'inverse de ce seuil est la sensibilité au contraste. La sensibilité au contraste varie avec les fréquences spatiales et temporelles, ce qui conduit au concept de CSF.

Il est généralement admis que l'œil humain est plus sensible aux basses fréquences spatiales qu'aux hautes fréquences. *Campbell* et *Robson* [CR68] ont été les premiers à essayer de démontrer l'effet de la CSF avec une représentation comme celle présentée par la figure 2. La luminance des pixels varie d'une façon sinusoïdale dans la direction horizontale. La fréquence de cette

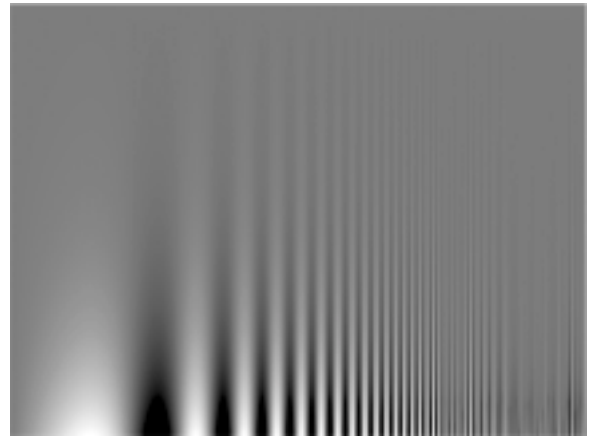


Figure 2. L'expérience de Campbell et Robson pour la construction de la CSF: modulation d'une grille sinusoïdale pour la luminance.

modulation augmente exponentiellement de gauche à droite, et l'amplitude décroît de façon exponentielle vers le haut de l'image. Les valeurs *min* et *max* selon l'axe horizontal restent constantes. Mais cela n'est pas visible quand nous observons cette image. L'observateur a plutôt l'impression que, sur les côtés gauche et droit, la modulation de l'amplitude a été plus petite. Ce que l'observateur voit réellement est sa propre fonction de sensibilité au contraste. S'il pouvait désigner une ligne parcourant les points où le stimulus se confond avec le fond de l'image, il aurait finalement sa propre courbe CSF de la luminance, comme celle présentée dans la figure 3.

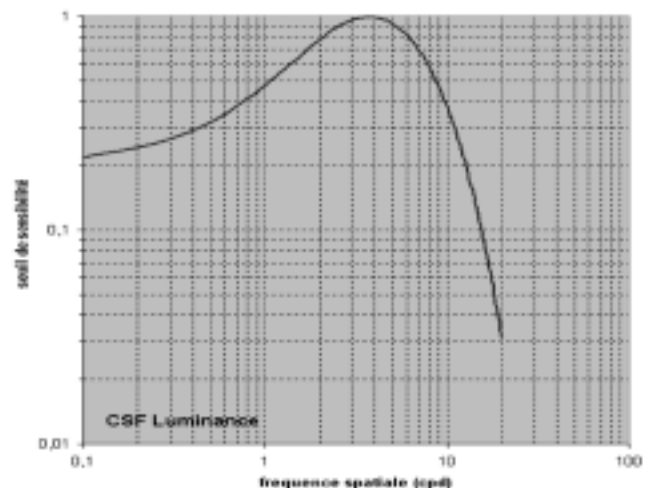


Figure 3. CSF de la luminance calculée à partir de l'équation 2.

Les études physiologiques et psychophysiques ont montré un comportement passe-bas de l'œil et un comportement passe-haut de la rétine 663. Le comportement général est donc, un passe-bande avec un pic de la sensibilité entre 2 et 8 *cpd* (cycles par degré) et une fréquence de coupure au-dessous de 64 *cpd*. Le modèle le plus utilisé pour cette dépendance, a été proposé

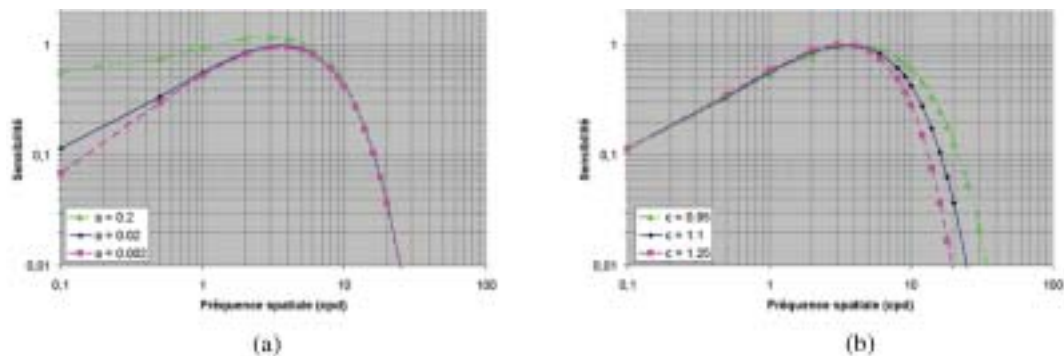


Figure 4. Illustration de l'influence des paramètres a (a) et c (b) pour une CSF de luminance.

par *Mannos et Sakrison* en 1974 pour le cas de la luminance [MS74], *ie* relie la sensibilité spatiale S à la fréquence spatiale f par l'équation suivante :

$$S(f) = b \left(a + \frac{f}{f_p} \right) e^{-\left(\frac{f}{f_p} \right)^c}, \tag{2}$$

où les paramètres a, b, c et f_p doivent être estimés. Le paramètre f_p représente la fréquence de pic et b est une simple constante de gain. L'influence des paramètres a et c est moins intuitive, comme le montre la figure 4. Les courbes de la figure 4-a montrent que le paramètre a influence l'allure de la courbe CSF pour les basses fréquences. Tandis que celle de la figure 4-b montrent que le paramètre c intervient pour les hautes fréquences. Les paramètres finaux définis par les auteurs, décrivent une courbe avec une valeur de pic de 1.0 correspondant à une fréquence $f_p = 8 \text{ cpd}$ et une extrapolation de la fréquence zéro correspondant à une sensibilité de 0.5.

Même si ce modèle est très populaire en traitement d'images, le choix des paramètres reste évidemment très important. Par exemple, dans le cas de la compression d'images avec le nouveau standard JPEG2000, l'allure de la courbe CSF, et plus spécifiquement la position de la valeur de pic, a une influence directe sur le résultat final. Cela s'explique par le fait que la fonction CSF est utilisée comme un filtre en fréquence appliqué sur les différentes sous-bandes de la décomposition en ondelettes (voir section 4). Il est donc évident, que si le pic est situé en dessous de 4 cpd , nous filtrons plus d'informations, que s'il était situé au dessus.

En effet, dans une décomposition en ondelettes, nous retrouvons une bonne partie de l'information dans les sous-bandes des basses fréquences, et donc, plus la fréquence de pic est faible et plus les informations de basses fréquences sont filtrées. Cela est une des raisons pour lesquelles dans la compression d'images, la construction de la CSF reste une étape déterminante. Ainsi, pour définir la courbe CSF la plus adaptée, nous utilisons des expériences psychovisuelles. La CSF finale est ainsi estimée à partir de toutes les réponses des observateurs humains impliqués dans le protocole de test, pour une certaine distance d'observation.

Cette nécessité de construction d'une nouvelle CSF se justifie par le fait que les modèles existants sont soit incomplets soit inadaptés à la compression JPEG2000. D'une part, le modèle de *Mannos et Sakrison* [MS74] ou celui de *Barten* [Bar99] n'adressent pas le problème des CSF chromatiques sachant que la démarche que nous adopterons pour la suite de cette contribution consiste en l'exploitation à la fois de la CSF de luminance et des CSF de chrominance. D'autre part, le pic du modèle de *Mannos et Sakrison* est situé à 8 cpd ce qui est incompatible avec la construction du codeur JPEG2000. L'intégration de ce modèle dans le processus de pondération visuelle n'apporterait aucune amélioration qualitative car l'intégralité de l'information serait préservée lors de la phase de filtrage. Ceci aboutirait à des facteurs de pondération identiques pour toutes les sous-bandes.

3.1. Construction des stimuli

La CSF représente le seuil de contraste transmis au cortex visuel comme une somme de fonctions harmoniques. C'est en effet une fonction multi-variables, dépendant de la fréquence spatiale, de l'orientation, de la dimension du stimulus et de la distance d'observation. Ces variables sont parfois liées entre elles. C'est le cas de la fréquence spatiale qui est liée à la distance d'observation. Cette liaison peut être très facilement vérifiée en observant l'image présentée dans la figure 2, à différentes distances. On la retrouve également par des calculs mathématiques, dans le calcul de la fréquence maximale utilisant l'équation 3 :

$$f_{\max} = R \times d \times \tan\left(\frac{\alpha}{2}\right), \tag{3}$$

avec R la résolution de l'écran d'affichage (calculée en fonction des dimensions physique de l'écran), d la distance d'observation et α l'angle visuel.

La taille critique d'un stimulus est déterminée par la dimension minimale au-delà de laquelle il y a invariance du seuil de sensibilité. Cette taille critique dépend de la fréquence spatiale examinée. En général, elle est située entre 10 et 20 fois la longueur de la période de la fréquence spatiale du signal sinusoidal. Cela

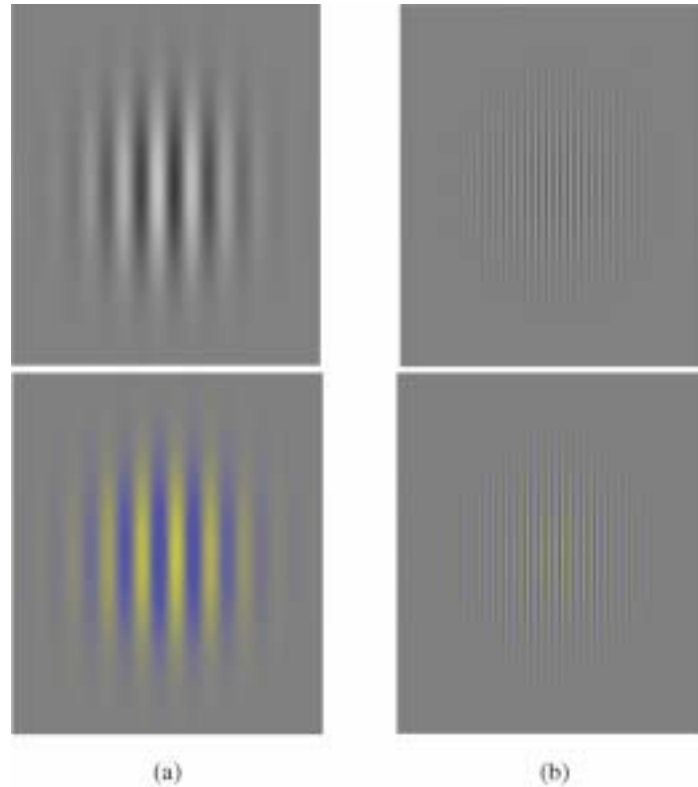


Figure 5. Grilles sinusoidales Gabori pour la construction de la CSF: basses fréquences à gauche; hautes fréquences à droite (les deux images ont le même niveau maximum de gris: 200 pour le blanc et 56 pour le noir).

est vérifié pour des fréquences entre 0.5 et 20 *cpd*. Pour des fréquences plus basses, comme 0.1 *cpd*, la taille correspond seulement à 4 fois la longueur de la période.

Ainsi, une taille de stimulus, plus grande ou égale à la taille critique, conduit à des mesures stables de la CSF. L'allure de la CSF dépend aussi de l'orientation du stimulus. La sensibilité pour les orientations horizontales et verticales des stimuli est similaire.

Les stimuli utilisés dans nos expériences sont des grilles sinusoidales de type Gabor [Dau80, PR83, KZ85]. Leur construction a été strictement liée aux caractéristiques techniques de l'écran d'affichage, comme les dimensions physiques et la résolution d'affichage, mais aussi à la distance d'observation. Les dimensions du stimulus et les fréquences spatiales dépendent de la distance d'observation. Dans la figure 5 sont présentées quelques exemples de grilles Gabori utilisées dans notre protocole d'évaluation.

Tous les stimuli présentés dans la figure 5 ont été construits dans l'espace des couleurs opposées (Opponent Color Space = OCS) [PW93, PW96]. Nous avons donc 3 types de stimuli représentant la luminance et les deux oppositions des couleurs bleu-jaune et rouge-verte, que nous appelons chrominances. La colonne (a) présente des stimuli basses fréquences et (b) des stimuli hautes fréquences. Nous pouvons déjà estimer, à partir de ces simples exemples, que la sensibilité diminue pour les hautes fréquences, car le contraste dans la colonne (b) semble

plus faible que dans la colonne (a), bien qu'ils soient égaux.

La construction mathématique de ces stimuli est effectuée à partir de l'équation 4 :

$$L(x, y) = L_{\text{moyen}} \{1 + C_{\text{max}} \sin [2\pi f(x \cos \theta + y \sin \theta) + \Phi] * \exp [-(x^2 + y^2)/2\sigma^2]\}, \quad (4)$$

avec L_{moyen} le niveau de gris moyen, C_{max} le contraste maximal, f la fréquence spatiale de la grille sinusoidale, θ l'orientation de la grille (0 pour la direction verticale), Φ la phase de la grille et enfin σ l'écart type de l'enveloppe Gaussienne.

3.2 Expériences psychovisuelles

Tous les stimuli ont été présentés sur un fond gris neutre. Ce gris moyen a été choisi pour éviter toutes les erreurs de perception liées à la corrélation des couleurs. Les mesures ont été effectuées pour 8 fréquences spatiales différentes: 1, 2, 3, 4, 5, 6, 8 et 9 *cpd*. La distance d'observation a été fixée à 85 cm. Cela appartient au domaine normal des distances de travail devant un écran d'ordinateur CRT, situé généralement entre 50 et 100 cm. Les expériences subjectives ont été effectuées par un panel de 30 observateurs, avec différents niveaux d'expertise dans le traitement d'images. L'acuité visuelle et la vision normale des cou-

leurs de tous les observateurs ont été vérifiées respectivement avec le test de *Snellen* et le test d'*Ishihara*. Les stimuli ont été présentés sur un écran *Sony FW900* de 24 pouces, avec un tube cathodique *Trinitron*, à une résolution de 1600×1200 pixels. La calibration couleur du tube *CRT* est réalisée avec le *EYE-ONE* monitor Mach 1.1 color calibrator du *GretagMacbeth*. Les stimuli ont été présentés avec le logiciel *Presentation* du *Neurobehavioral Systems*. La configuration utilisée permet une représentation des images sur l'écran, de type pixel à pixel, sans redimensionner l'image ou effectuer d'autres interpolations. La salle d'évaluations psychovisuelles a été construite en conformité avec les standards internationaux ITU-R 500-10 et ISO 3664 [ITU00]. Les murs de la salle d'évaluation sont d'une couleur neutre permettant de minimiser l'effet des lumières parasites. Ils sont aussi d'une chromaticité approximativement égale au « point blanc » de l'écran.

L'éclairage de la salle a été totalement contrôlé en utilisant un illuminant *D65* d'intensité variable, mesuré avec un luminancemètre *T-10M Minolta*. Toutes les sources de lumière, autres que celles utilisées pour l'éclairage de la salle, doivent être évitées car elles dégradent significativement la qualité d'image. L'éclairage de la salle et la luminosité de l'écran sont particulièrement importants car la sensibilité de la rétine aux différentes longueurs d'ondes varie en fonction de la luminosité ambiante. Cette variation se traduit par une plus grande sensibilité aux courtes longueurs d'ondes au fur et à mesure que l'éclairement décroît. Cela s'explique par le fait que la rétine contient deux types de cellules photosensibles : les cônes et les bâtonnets. Les bâtonnets permettent la vision nocturne (vision scotopique) tandis que les cônes permettent la vision diurne (vision photopique). Ces deux types de cellules n'ont pas la même sensibilité spectrale. La figure 6 présente les fonctions d'efficacité spectrale en vision photopique et scotopique, définies par la CIE (Commission Internationale de l'Éclairage) en 1924 respectivement 1951.

La courbe de luminosité spectrale montre comment la sensibilité de l'œil aux différentes longueurs d'ondes varie lorsqu'il est

adapté à la lumière. La courbe en pointillés montre que le maximum de sensibilité change le long du spectre lorsque l'œil est adapté à l'obscurité (conditions de vision scotopique). La courbe en ligne continue montre la sensibilité de l'œil en adaptation à la lumière du jour (photopique). Cette variation de la sensibilité de la rétine est probablement due à la modification des contributions relatives des cônes et des bâtonnets. Entre les domaines photopique et scotopique, se situe le mésopique, domaine intermédiaire où les cônes cessent progressivement de fonctionner tandis que les bâtonnets prennent le relais : espace mal défini, correspondant à la vision dite *crépusculaire*, où le maximum de sensibilité se déplace vers les petites longueurs d'ondes donnant une vision bleutée. D'un point de vue intensité lumineuse, le domaine scotopique est situé en dessous de 0.034 cd/m^2 , le photopique correspondant aux intensités supérieures à 3.4 cd/m^2 , tandis que le mésopique est situé entre les deux limites, supérieure et inférieure.

Dans ces conditions, nous avons opté pour une ambiance scotopique et une stimulation (surface d'évaluation) en domaine photopique. L'ambiance scotopique nous permet d'éviter l'influence de la vision périphérique stimulée par les différents objets qui peuvent être présents dans l'environnement d'observation (bordure d'écran, support de l'écran, murs, etc.). Ces objets peuvent être visibles dans une ambiance photopique, voire même mésopique. Ainsi, toute l'attention de l'observateur sera concentrée sur la zone de test. Comme notre objectif est de mesurer la sensibilité au contraste du système visuel humain pour la luminance ainsi que pour les chrominances, il est évident que la stimulation a été effectuée dans le domaine photopique.

La procédure d'évaluation doit être bien définie aussi. Plusieurs modalités de présentation des stimuli et de traitement des réponses des observateurs peuvent être employées. Nous avons opté pour une augmentation constante du contraste pour chaque fréquence spatiale, en partant d'un contraste zéro. Ainsi le stimulus a la même couleur que le fond, ce qui répond ainsi au phénomène dit de « *crispensing* » du SVH. Chaque image correspondant à un contraste et une fréquence donnée est présentée pendant 500 ms. Ce temps a été choisi pour tenir compte de la vitesse de réaction de l'observateur, qui est située en général entre 300 et 400 ms, le temps que l'information image arrive au cerveau et que la commande d'exécution soit lancée.

Entre deux fréquences différentes, il est présenté un fond uniforme pendant 5 secondes afin de donner à l'œil le temps nécessaire d'effacer la rémanence rétinienne des images antérieures. Les stimuli à différentes fréquences sont présentés d'une façon aléatoire, à savoir qu'un stimulus à une fréquence de 0.5 cpd peut être suivi par une fréquence de 8 cpd . L'observateur doit appuyer sur une touche lorsqu'il perçoit le stimulus, c'est-à-dire lorsqu'il perçoit distinctement les bandes blanches et noires (pour le cas de la luminance), rouges et vertes, ou bleues et jaunes (pour les chrominances). Globalement, 8 différentes fréquences sont présentées à l'observateur. Ce sont les fréquences mentionnées au début de cette section. De plus, chaque expérience à une fréquence spatiale donnée est renouvelée 5 fois

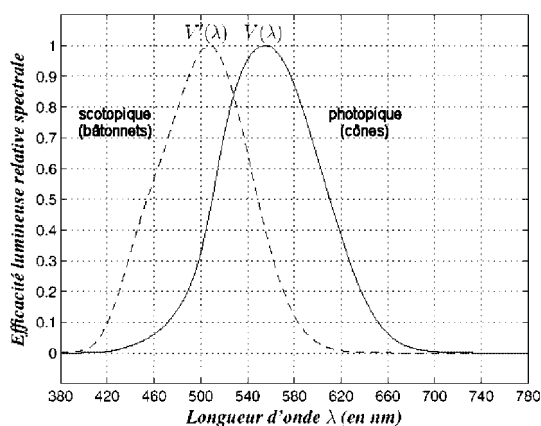


Figure 6. Les fonctions d'efficacité spectrale en vision photopique et scotopique.

dans la même séquence d'évaluation, afin de calculer le seuil de sensibilité final. Le fait de répéter 5 fois la même expérience nous donne la possibilité de vérifier la cohérence des résultats pour chaque observateur. Ceci nous a donc permis d'obtenir des résultats fiables pouvant être intégrés dans la chaîne d'extraction des pondérations visuelles décrite plus loin. Enfin, le facteur 5 entre dans la définition du temps maximal d'observation qui ne doit en aucun cas être dépassé au risque de fausser les observations.

3.3. Interpolation mathématique des résultats

Les différentes valeurs du seuil de sensibilité, données par les observateurs, ont été interpolées à partir des équations 5 et 6, respectivement pour la luminance et la chrominance :

$$S_L(f) = a_1 f^2 \exp(b_1 f^{c_1}) + a_2 \exp(b_2 f^{c_2}) \quad (5)$$

$$S_C(f) = a_3 \exp(b_3 f^{c_3}) \quad (6)$$

avec S_L le seuil de sensibilité (l'inverse du contraste) pour la luminance, S_C celui pour la chrominance et f la fréquence spatiale en *cpd*. Les paramètres a_1, \dots, a_3 , b_1, \dots, b_3 et c_1, \dots, c_3 ont été réglés pour interpoler la courbe de la CSF. Un exemple des interpolations finales des courbes CSF pour la luminance et les deux chrominances, utilisées pour filtrer les coefficients de la décomposition en ondelette de JPEG2000, est présenté dans la figure 7. Les seuils de sensibilité pour les très basses fréquences spatiales (inférieures à 1 *cpd*) sont obtenus par extrapolation des résultats. La mise en oeuvre de ces expériences pour des fréquences comme 0.1 *cpd* est quasiment impossible.

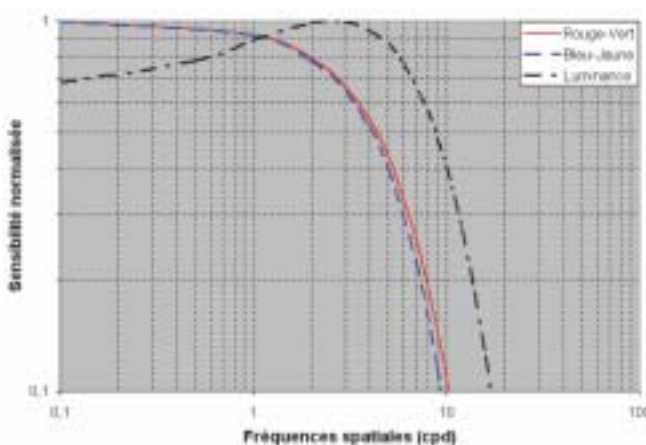


Figure 7. CSF pour la luminance, pour la chrominance rouge-vert et pour la chrominance bleu-jaune.

4. Intégration du SVH dans la compression JPEG2000

Après la modélisation du comportement du Système Visuel Humain, la prochaine étape est d'imaginer comment nous pouvons exploiter ces propriétés d'une manière efficace dans un système de compression d'images. Certains systèmes de compression permettent une exploitation plus directe des propriétés du SVH que d'autres. Par exemple, JPEG2000 supporte des caractéristiques comme les pondérations visuelles progressives (en fréquence et en couleur), le masquage de voisinage, l'auto masquage, l'adaptation locale de la luminosité, les fréquences temporelles et excentriques. Le JPEG classique supporte seulement deux de ces caractéristiques : les pondérations en fréquence et en couleurs.

Une des stratégies d'optimisation visuelle les plus communes dans la compression d'images est l'utilisation de la fonction de sensibilité au contraste (CSF) décrite précédemment, qui caractérise la variation de la sensibilité du système visuel par rapport aux fréquences spatiales 2D [WYSV97, LK98]. En général, l'œil humain est moins sensible aux erreurs de hautes fréquences qu'aux erreurs de basses fréquences. La CSF peut être utilisée pour déterminer la précision relative nécessaire pour les différentes fréquences spatiales, où le terme de pondération est utilisé pour décrire la précision proportionnelle demandée. Pour utiliser la CSF, qui est généralement décrite en fréquences visuelles comme le cycle par degré (*cpd*), elle doit être mappée sur le domaine des fréquences digitales de la compression, comme les cycles par pixel (*cpp*). Une pondération correcte en fréquence ou en couleur peut conduire généralement à une préservation importante des détails et textures, sans l'introduction de distorsions couleurs. En général, les pondérations visuelles sont plus efficaces pour de grandes distances d'observation ou pour de larges résolutions d'affichage ou d'impression. En effet, les pondérations peuvent être utilisées aussi pour réduire les artefacts de flickering dans le Motion JPEG2000 qui se traduisent par un bruit lié au mouvement de la scène.

Il existe deux types de pondérations pouvant être utilisés dans l'algorithme de JPEG2000 : les pondérations fréquentielles fixes et les pondérations visuelles progressives. La méthode présentée dans cet article est basée sur la première technique. Notre choix est motivé par le fait que toutes les modifications réalisées n'affectent pas le schéma du codeur, car celui-ci peut utiliser un fichier externe contenant les pondérations visuelles calculées par notre méthode.

En général, la courbe CSF est une fonction continue des fréquences spatiales, comme nous venons de le voir. Cependant, pour une décomposition en ondelettes discrète, il est habituel qu'une seule pondération CSF soit choisie pour chaque sous-

bande afin de simplifier l'implémentation. Cette technique d'application des pondérations fréquentielles est connue comme étant une pondération fréquentielle fixe. L'ensemble des pondérations CSF peut être incorporés de deux façons dans le JPEG2000: en modifiant le pas de quantification, ou en modifiant l'ordre d'importance des coefficients au codage. Dans les deux cas, les pondérations CSF ne doivent pas être transmises explicitement au décodeur.

Notre proposition est basée sur la deuxième technique d'implémentation, concernant la modification de l'ordre de codage. Dans cette implémentation, les dimensions des pas de quantification ne sont pas modifiées, mais les facteurs de distorsion introduits dans l'optimisation *R-D* (Rate/Distorsion) sont pondérés suivant une proportionnalité linéaire aux facteurs de pondérations CSF correspondant à chaque sous-bande. Cela contrôle d'une manière efficace la création des couches de qualité du bitstream. Cette implémentation est recommandée car elle produit des résultats similaires à la première implémentation, tout en étant compatible avec la compression sans perte et aussi avec la pondération visuelle progressive.

Il est aussi possible d'intégrer une pondération CSF à l'aide de cellules adaptatives [NR99], qui permettent une meilleure adaptation des facteurs de pondération au spectre du signal dans une sous-région (*e.g.*, bloc de code) ou une sous-bande. D'une manière basique, un facteur de pondération adaptatif par rapport aux données, peut être déterminé pour une sous-région en faisant un filtrage des coefficients d'ondelette appartenant à cette sous-région, à l'aide d'un filtre CSF de la sous partie correspondante de la CSF. La pondération peut être réalisée seulement dans la partie encodeur. Théoriquement, en considérant le contenu actuel d'une sous-région, cette approche peut générer des facteurs de pondération plus cohérents qu'en choisissant seulement la pondération CSF correspondant aux fréquences moyennes de la sous-bande. Cependant, les filtres CSF utilisés dans [NR99], sont basés sur un filtre FIR de dimension fixe, appliqué directement sur les coefficients d'ondelette de la sous-bande, comme s'ils composaient une image 2D classique, ce qui n'est pas forcément le cas.

En conséquence, notre méthode a été conçue pour éviter tous ces problèmes et pour réaliser en même temps une meilleure corrélation entre le domaine fréquentiel de la fonction CSF et les domaines correspondant aux différentes sous-bandes de la décomposition en ondelettes. Pour cela, nous avons opté pour une approche différente. L'idée est de filtrer la transformée de Fourier de l'image originale avec la CSF 2D, et ensuite de calculer les facteurs de pondération visuelle à partir de la transformée filtrée par rapport à celle non-filtrée. La correspondance avec les différentes sous-bandes de la décomposition en ondelettes sera effectuée dans l'étape de calcul des facteurs de pondération.

4.1. Filtrage CSF dans le domaine de Fourier

De cette manière, les différents domaines fréquentiels contenus dans l'image originale sont clairement structurés. Nous pouvons donc imaginer une mise en correspondance plus évidente de ces domaines fréquentiels avec ceux de la fonction CSF 2D. La figure 8 présente un exemple du spectre de Fourier pour la composante luminance d'une image couleur, ainsi qu'une représentation de la CSF 2D échantillonnée relativement aux dimensions de l'image originale.

Dans la figure 8, le spectre de Fourier présenté est évidemment recalé. Ainsi, les basses fréquences se retrouvent au centre de

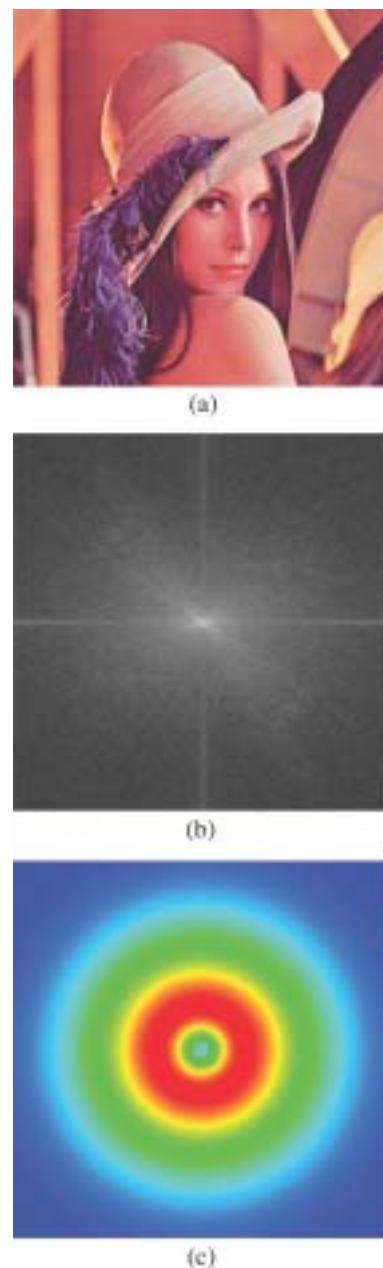


Figure 8. a) Image originale couleur; b) spectre de Fourier de la composante luminance dans l'espace YCbCr; c) CSF 2D pour la luminance échantillonnée relativement aux dimensions de l'image originale.

l'image. Pour l'image de la CSF, le même principe est respecté, donc la fréquence dite zéro (la CSF étant représentée sur des axes logarithmiques, une fréquence égale à zéro n'existe donc pas) est au centre. De cette manière, la signification des différents domaines fréquentiels sera strictement conservée. Par exemple, toutes les basses fréquences de l'image originale seront filtrées seulement avec la partie basses fréquences correspondante de la fonction CSF. Ainsi, nous éviterons le filtrage de certaines basses fréquences par des valeurs de la CSF correspondant aux domaines des hautes ou moyennes fréquences, comme c'est le cas en filtrant directement les coefficients des différentes sous-bandes d'ondelettes.

Dans la compression JPEG2000, les sous-bandes de la décomposition en ondelettes correspondantes aux basses fréquences, contiennent le plus d'information utile pour le codage final. Ainsi, pour obtenir les meilleurs résultats en termes de qualité visuelle, l'allure de la courbe CSF utilisée pour le filtrage, sera modifiée de telle façon que les basses fréquences de l'image originale situées en dessous de la fréquence du pic de la CSF ne soient pas filtrées. La CSF finale aura donc une forme comme celle présentée dans la figure 9, où les valeurs correspondant aux fréquences inférieures à la fréquence du pic, seront mises à 1.

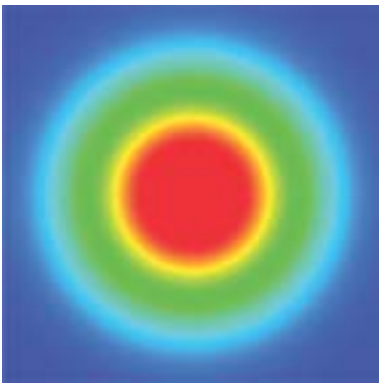


Figure 9. La CSF 2D pour la luminance normalisée, avec les valeurs inférieures à la fréquence du pic, mises à 1.

Par la suite, pour le calcul des facteurs de pondération visuelle, nous avons utilisé, pour des raisons de simplicité, seulement la partie réelle de la transformée de Fourier. Ce choix a été motivé par le fait que les facteurs de pondérations finaux calculés pour la partie réelle ou bien pour le spectre de Fourier, soient quasiment identiques. Cela nous permet aussi de reconstruire l'image originale filtrée par la CSF, en mettant en œuvre la transformée de Fourier inverse. La figure 10 présente un exemple d'image filtrée dans le domaine de Fourier, avec une CSF 2D construite pour une distance d'observation de 114 cm et une résolution d'écran de 1024×768 pixels.



(a)



(b)

Figure 10. a) Image originale ; b) Image filtrée avec la CSF 2D pour une distance d'observation de 114 cm et une résolution de 1024×768 pixels.



(a)



(b)

Figure 11. Zoom sur les images de la figure 10: a) Image originale ; b) Image filtrée.

Les différences entre les deux images présentées dans la figure 10, ne sont pas vraiment perceptibles, même en les regardant à partir d'une distance plus faible que celle pour laquelle le filtrage a été effectué. Le filtrage CSF a affecté principalement les hautes fréquences de l'image, donc des détails pour lesquels l'œil humain n'est pas très sensible. Un zoom sur les deux images, (figure 11), permet d'observer l'effet réel du filtrage CSF sur une image couleur.

4.2. Mise en correspondance entre le spectre de Fourier et les sous-bandes d'ondelettes

La deuxième étape consiste en la mise en correspondance entre les différentes sous-bandes de la décomposition en ondelettes appartenant aux différents niveaux de décomposition et différentes régions du spectre de Fourier. Nous allons décrire notre technique de calcul pour une décomposition en ondelettes en 5 niveaux et une image de taille 512x512 pour la facilité de l'exemple. Pour pouvoir calculer des facteurs de pondération correspondant aux sous-bandes d'ondelettes HL (détails verticaux), LH (détails horizontaux) et HH (détails diagonaux) à différents niveaux de décomposition, il faut mettre en correspondance ces niveaux et les sous-bandes dans la transformée de Fourier de l'image. La décomposition en ondelettes utilisée pour la compression JPEG2000 peut être décrite comme un banc de filtres séparables. Ainsi, l'application successive de ces filtres aboutit à un domaine passe-bande des fréquences spatiales, horizontales et verticales, spécifique à chaque sous-bande. Théoriquement, pour une décomposition en ondelettes à 1 seul niveau, les domaines passes-bandes par rapport à un spectre de Fourier standard sont représentés comme dans la figure 12

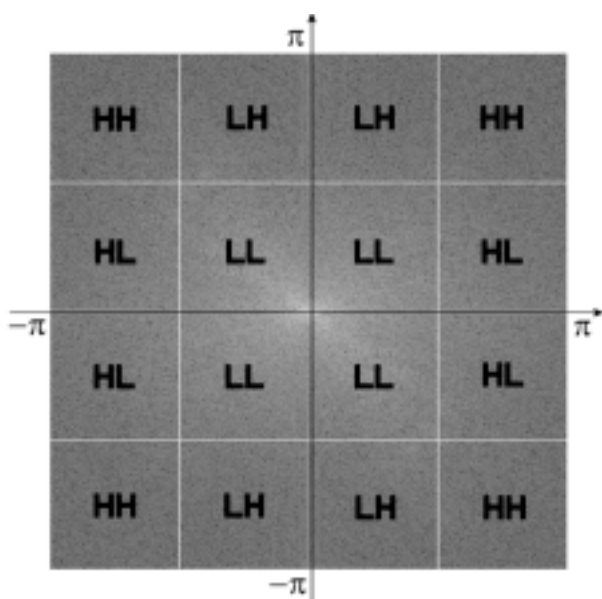


Figure 12. Domaines passes-bandes pour un niveau de décomposition en ondelettes, par rapport au domaine du spectre de Fourier de l'image.

[TM02]. Chaque sous-bande (HL, LH, HH et LL) appartenant au niveau 1 de décomposition a donc une bande de fréquences de $\pi/2$. Pour une décomposition en 2 niveaux, la sous-bande LL présentée dans la figure 12 est à son tour répartie en 4 nouvelles sous-bandes: HL, LH, HH et LL, chacune ayant une bande des fréquences de $\pi/4$. Ainsi, dans le cas d'une décomposition en ondelettes en n niveaux, les bandes fréquentielles ont une largeur de $\pi/2^n$.

Cette représentation théorique ne correspond pas vraiment aux conditions réelles, car par exemple dans un spectre de Fourier, il est très difficile de situer exactement l'information correspondant aux détails verticaux ou bien diagonaux. Ce qui met en cause la structure rectangulaire des trois sous-bandes HL, LH et HH. Un autre aspect qui pose problème réside dans le fait que le spectre de Fourier d'une image a plutôt une structure radiale, qu'une structure rectangulaire. Ainsi par exemple, un filtre passe-haut idéal en 2D aurait une forme ellipsoïdale et non pas rectangulaire. Donc, la délimitation d'un niveau de décomposition en ondelettes par rapport au spectre de Fourier, réalisée par un rectangle, n'est pas conforme à la réalité. En tenant compte de ces constatations, nous avons donc choisi de délimiter les zones correspondant aux différents niveaux de décomposition, par des ellipses (cercles si les dimensions de l'image sont égales).

La figure 13 présente un exemple de délimitation des zones du spectre de Fourier, par des ellipses concentriques, correspondant aux 5 niveaux de décomposition en ondelettes. Les largeurs des bandes fréquentielles de chaque niveau selon les deux axes centraux, respectent le cas théorique de $\pi/2^n$, où n correspond au nombre de niveaux de décomposition, en l'occurrence 5 pour l'exemple présenté dans la figure 13.

Afin de délimiter ensuite les zones correspondantes aux 3 sous-bandes de chaque niveau, HL, LH et HH, nous avons choisi une

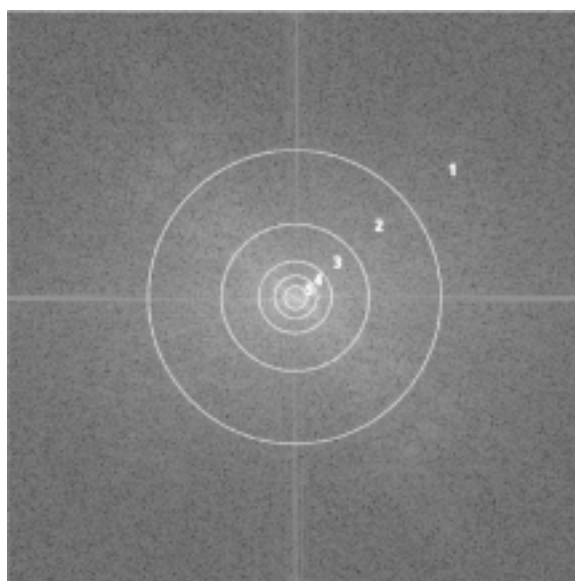


Figure 13. Domaines passe-bande pour 5 niveaux de décomposition en ondelettes, par rapport au domaine du spectre de Fourier de l'image.

technique basée sur des régions angulaires. Ainsi, nous avons établi des régions délimitées par des droites ayant des pentes de 30° par rapport aux axes centraux, pour les détails horizontaux et verticaux, et de 15° pour les diagonaux. Nous avons donc considéré que l'information correspondant aux détails horizontaux se retrouve dans une région de 30° d'ouverture d'un côté et de l'autre de l'axe vertical (Idem pour les détails verticaux) et *vice versa*. Car la frontière entre les détails diagonaux et verticaux, ou diagonaux et horizontaux, ne représente pas une droite mais au contraire, une portion de recouvrement. Ce recouvrement a une ouverture de 15° et de sa valeur dépendent les facteurs de pondération. La figure 14 présente un exemple du recouvrement choisi pour notre application, ainsi que les régions correspondantes à la sous-bande LH du niveau 2 de décomposition en ondelettes.

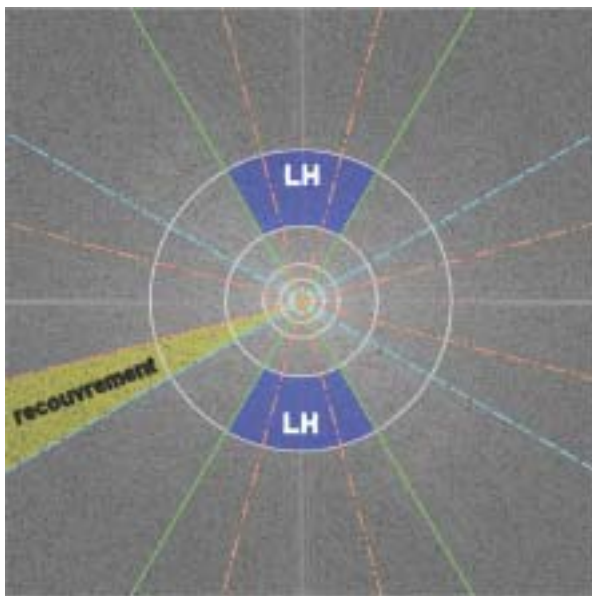


Figure 14. Exemple de zone de recouvrement de 15° entre une région des détails verticaux et une région des détails diagonaux (zone en jaune); et les régions correspondant à la sous-bande LH du niveau 2 de décomposition (zones en bleu).

Dans la figure 14, les lignes en vert (continues) délimitent les régions des détails horizontaux (verticaux dans le spectre de Fourier), les lignes en bleu (pointillées) séparent celles des détails verticaux (horizontaux dans le spectre de Fourier) et les lignes en couleur orange (ligne-point) délimitent les régions des détails diagonaux. Les cercles blancs concentriques bornent les différents niveaux par rapport à la décomposition en ondelettes. Les deux zones hachurées en bleu, représentent les régions correspondant à la sous-bande LH (détails horizontaux) du niveau 2 de la décomposition en ondelettes.

La zone de recouvrement présentée dans la figure 14 représente la frontière imaginaire entre une région correspondant aux détails verticaux de l'image et une région correspondant aux détails diagonaux. Cette zone permet d'éviter les problèmes d'indécision de l'appartenance de certains détails à une des deux

régions. Les valeurs des angles d'ouverture des différentes régions ont été choisies de telle sorte que chacune de trois régions, verticale, horizontale et diagonale, ait une ouverture de 60° .

4.3. Calcul des facteurs de pondération visuelle

Ainsi, pour chaque niveau de décomposition, nous localisons les régions correspondant aux trois sous-bandes d'ondelettes, HL, LH et HH. Une fois cette étape terminée, les facteurs de pondération visuelle finaux, sont calculés en utilisant l'équation 7 :

$$w_x = \sqrt{\frac{\sum [c_{i,j}^{csf}]^2}{\sum c_{i,j}^2}} \quad \text{pour } i, j \in M \quad (7)$$

avec $c_{i,j}$ les coefficients de la transformée de Fourier originale, $c_{i,j}^{csf}$ les coefficients de la transformée de Fourier filtrés par la fonction CSF 2D et M l'ensemble des coefficients. Ces facteurs de pondérations sont utilisés dans le codeur JPEG2000 pour pondérer l'énergie de l'erreur de quantification. Dans les approches classiques, cette pondération est réalisée par un seul facteur indépendant du signal. Ainsi, la fonction de distorsion D (la contribution à la distorsion due à l'énergie de l'erreur de quantification) est donnée par l'équation (8) :

$$\Delta D_{HSV} = w_{csf}^2 \cdot \Delta E_q \quad (8)$$

Dans notre approche le facteur de pondération w_{csf} est remplacé par un facteur différent pour chaque sous-unité. Une sous-unité peut être un bloc de 8×8 coefficients d'ondelette par exemple. Ainsi, l'équation 8 peut être écrite comme suit :

$$\Delta D_{HSV} = w_1^2 \Delta E_{q_1} + w_2^2 \cdot \Delta E_{q_2} + \dots + w_n^2 \cdot \Delta E_{q_n} \quad (9)$$

avec ΔE_{q_n} les contributions à l'énergie de distorsion de chaque sous unité. Une telle technique de pondération, avec plusieurs facteurs au lieu d'un seul, offre évidemment de meilleurs résultats d'un point de vue qualitatif, surtout pour la compression d'images couleurs.

Un exemple de facteurs de pondération visuelle pour la compression JPEG2000 d'images couleurs, obtenus avec la méthode proposée dans cet article, est présenté en annexe.

5. Résultats et discussions

Afin de justifier la méthode proposée dans cet article, nous avons utilisé deux implémentations différentes du standard

JPEG2000, le Kakadu (version 3.4) [Tau03] et une implémentation appelée codeur-EIRE, issue du projet RNRT qui a motivé les travaux présentés dans cet article. D'autres implémentations existent comme Jasper [AW04], JJ2000 [pro00], etc. Cependant, notre choix ne s'est pas tourné vers ces derniers car ils n'intègrent pas la fonctionnalité de pondération visuelle.

La qualité des images compressées avec les deux codeurs, avec et sans facteur de pondération visuelle, a été évaluée à l'aide d'expériences subjectives mais aussi par des mesures mathématiques fréquemment utilisées en analyse d'images. Toutes les évaluations ont été réalisées sur 17 images appartenant à la base d'images de Kodak (images couleur RGB, de dimensions 768×512 pixels, sur 24 bpp). Chaque image a été compressée à plusieurs taux de compression, variant entre 4 et 0.062 bpp , et décompressée en conformité avec le standard JPEG2000. Les évaluations subjectives ont été effectuées par un panel de 20 observateurs, avec plus ou moins d'expérience dans le domaine de l'imagerie couleur.

Les premières évaluations ont été effectuées pour juger la qualité d'images compressées avec pondération visuelle par rapport

à celles compressées sans pondération. La procédure d'évaluation subjective a été basée sur une technique d'ordonnement qualitatif. L'observateur est appelé à sélectionner les différentes images, de la plus mauvaise à la meilleure, par rapport à l'image originale, non compressée. Cette technique permet le calcul d'une valeur mathématique appelée Score d'Option Moyen (MOS – Mean Opinion Score), en utilisant une échelle qualitative de 1 à 5, avec 5 correspondant à la meilleure image.

Ensuite, il est possible de calculer la répartition des différentes réponses constituant le MOS, à l'aide de l'Écart-Type. La figure 15 présente le MOS et l'Écart-Type associé, pour trois configurations des codeurs cités précédemment et pour seulement huit images de la base Kodak, par souci de clarté du graphique. Les configurations des codeurs utilisées sont : le codeur EIRE sans pondération visuelle, le codeur EIRE avec pondération de notre méthode et le codeur Kakadu qui utilisent les pondérations par défaut de JPEG2000.

Les valeurs données dans la figure 15, représentent les moyennes des réponses de tous les observateurs pour chacune des images et chacun des codeurs. Comme nous pouvons l'observer, le MOS offre une classification des préférences qualitatives des observateurs pour les trois codeurs. Nous pouvons voir facilement qu'en général, les observateurs ont préféré les images compressées avec des facteurs de pondération visuelle, notamment celles résultant du codeur EIRE avec pondération. Néanmoins, ces préférences sont strictement liées à la nature de l'image analysée. Cela nous permet aussi de faire une classification des images, en fonction de leur contenu. L'écart-type correspondant à ces réponses offre une information supplémentaire concernant les observateurs. Cela nous permet d'observer et de juger la cohérence des réponses données par les observateurs, mais aussi leurs degrés de concentration. Pour que le MOS final soit validé, l'écart-type doit être le plus faible possible, en général une valeur proche de 1 est satisfaisante. Il est donc possible d'éliminer de l'étape finale du calcul les cas atypiques, afin d'avoir un MOS correct.

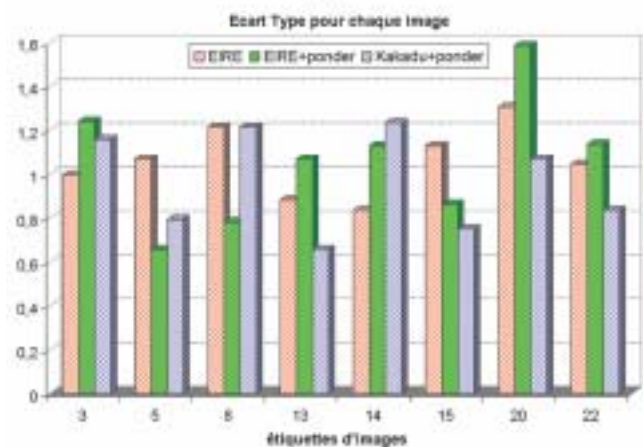
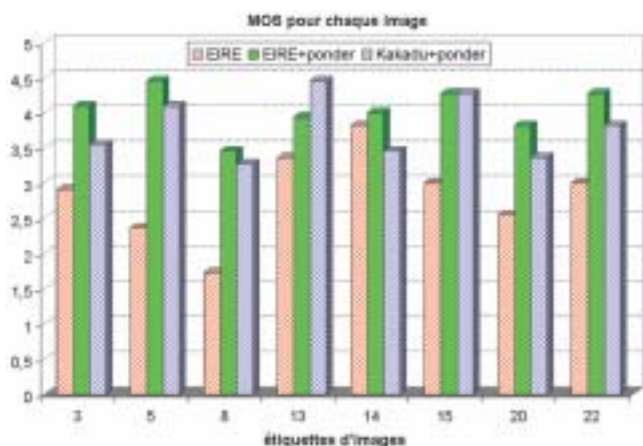


Figure 15. a- Le MOS et b- l'écart-type, pour 8 images de test, à plusieurs taux de compression et 3 implémentations JPEG2000.

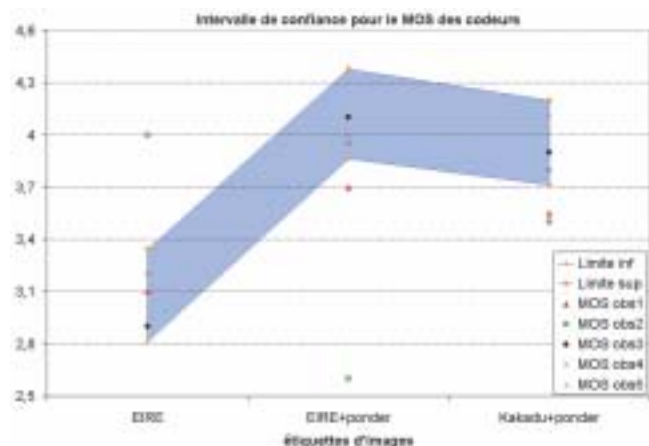


Figure 16. Intervalle de confiance du MOS final, pour toute la base d'images et les 3 codeurs, à plusieurs taux de compression.

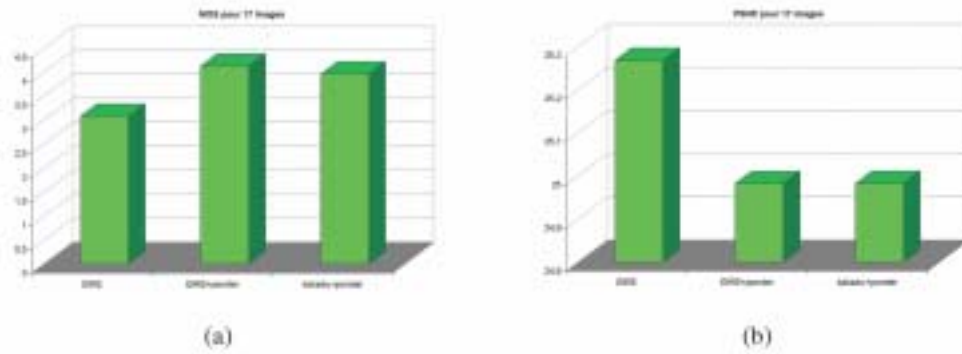


Figure 17. a) MOS moyen et b) PSNR, pour les 17 images de test, à plusieurs taux de compression et 3 implémentations JPEG2000.

Les deux valeurs mathématiques, à savoir le MOS et l'écart-type, nous permettent aussi de calculer l'intervalle de confiance des réponses. Ainsi, nous pouvons juger les réponses d'un observateur par rapport à un domaine des réponses correctes. La figure 16 présente un exemple de cet intervalle de confiance. Ce dernier est calculé en fonction des réponses données par l'ensemble des observateurs pour toutes les expériences menées. Il nous offre le domaine des réponses considérées comme correctes, c'est-à-dire ayant un MOS cohérent. Pour la clarté de l'exemple, nous avons représenté sur la figure 16, seulement les MOS de 5 observateurs. Comme nous pouvons l'observer, cet intervalle nous permet de raffiner le MOS final, en analysant les MOS des différents observateurs. Ainsi, par exemple dans le cas présenté, nous pouvons rejeter du calcul final, le jugement des observateurs 1 et 2. Ceci car leur MOS n'appartient pas à l'intervalle souhaité. Ces trois valeurs mathématiques (MOS, écart-type et intervalle de confiance) représentent des outils très importants dans l'analyse des résultats des évaluations psycho-visuelles.

Le MOS permet aussi de faire une comparaison plus claire avec

les mesures mathématiques classiques, comme le PSNR (Rapport Signal - Bruit). La figure 17 présente le MOS moyen et le PSNR moyen pour les 17 images de test, à plusieurs taux de compression. Les configurations des codeurs utilisés sont les suivantes : le codeur EIRE sans pondération visuelle, le codeur EIRE avec pondération et le codeur Kakadu avec pondération par défaut.

Les résultats de cette évaluation subjective, notamment le MOS présenté dans la figure 17-a, montrent que les observateurs ont préféré en général les images compressées avec des facteurs de pondération visuelle. Nous pouvons constater également que les résultats obtenus avec les facteurs de pondération calculés d'après la méthode proposée dans cet article, utilisés avec le codeur EIRE, sont meilleurs que ceux fournis par le codeur Kakadu. Il est important de mentionner le fait que les évaluations subjectives sont en totale contradiction avec les mesures mathématiques classiques, notamment le PSNR, comme nous pouvons le voir dans la figure 17-b. En effet, d'un point de vue mathématique, les compressions sans facteur de pondération offrent de meilleurs résultats en terme de quantité d'information

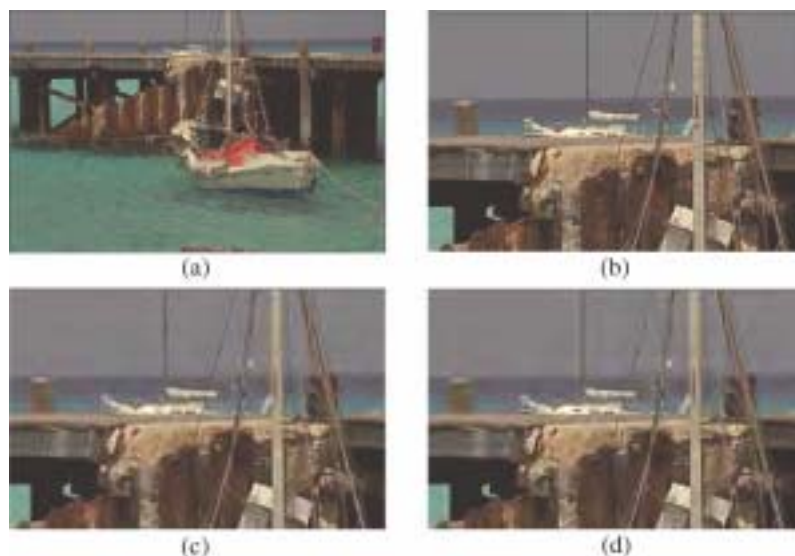


Figure 18. Exemple de compression JPEG2000 avec pondérations visuelles (image c) et sans pondérations (image d); l'image b) est un zoom sur la région concernée de l'image originale a).

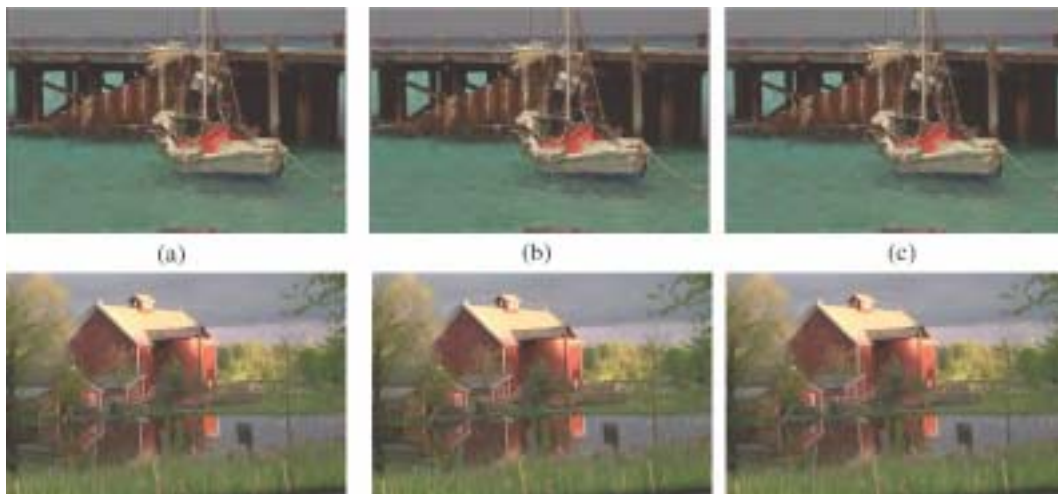


Figure 19. Résultats de compression avec le codeur EIRE avec les pondérations : a) de Nadeneau, b) du standard et c) proposées.

originale présente dans l'image finale compressée. Cependant, comme nous pouvons le voir par exemple dans la figure 18, l'observateur humain ne juge pas la qualité d'une image produite par rapport à la quantité du signal d'origine, mais évidemment par rapport à sa qualité visuelle.

La figure 18 présente une situation d'évaluation subjective, d'où nous pouvons mieux comprendre les choix qualitatifs des observateurs humains. D'un point de vue visuel, l'image compressée avec pondération visuelle (figure 18-c) a été choisie par les observateurs comme étant meilleure que celle compressée sans pondération (figure 18-d). Cela peut être expliqué par le fait que l'image sans pondération présente des défauts de structure pour certains objets de l'image. Rappelons que ce choix est en contradiction avec le PSNR. Ainsi, si nous voulons produire des images compressées qui seront visualisées par des observateurs, l'utilisation des mesures mathématiques classiques pour juger la qualité d'image n'est pas une très bonne option, notamment dans le cas des images couleur. Cependant, il apparaît clairement qu'il est préférable d'utiliser des facteurs de pondération visuelle, pour la compression afin d'améliorer la qualité des images couleur.

Dans les expériences précédentes, deux voies ont été explorées. D'une part, l'étude de l'amélioration qualitative par l'introduction de pondération visuelle et d'autre part, l'étude comparative entre le codeur Kakadu utilisant les facteurs par défaut et le codeur EIRE avec les pondérations issues de la modélisation proposée. Cependant, un aspect reste à explorer, à savoir quantifier l'apport du codeur par rapport à celui des pondérations proposées. Pour ce faire, une étude a été menée impliquant uniquement le codeur EIRE mais cette fois-ci avec différents facteurs de pondération.

Les facteurs retenus sont ceux par défaut de JPEG2000 [TM02], ceux de Nadeneau [NR99] et les facteurs proposés. Des images issues de la compression avec ces différents facteurs sont données

dans la figure 19. Nous pouvons remarquer à partir de cette figure qu'il est difficile de juger visuellement la différence entre les images. Afin de pallier cette difficulté, nous avons fait appel à la mesure PQS (Picture Scale Quality) [MKA98]. Cette dernière a été calibrée par rapport à la base d'images utilisée et aux expériences psychophysiques menées. Les résultats de cette mesure sont présentés par la figure 20. À partir de cette figure, il est possible de noter que les facteurs de pondération proposés permettent un meilleur apport qualitatif par rapport aux autres facteurs et ce pour tous les taux de compression. Les facteurs du standard JPEG2000 donnent les résultats les plus mauvais.

À partir de cette dernière étude, il est possible de tirer les premières conclusions par rapport à l'amélioration de la qualité. En effet, en plus du codeur, les pondérations proposées permettent d'améliorer les résultats finaux en termes de qualité.

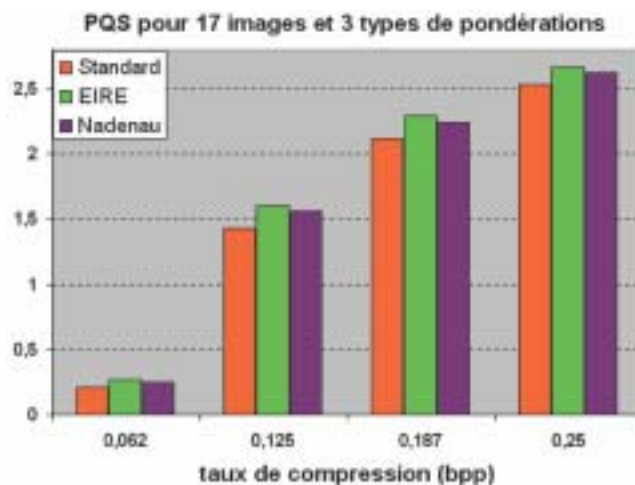


Figure 20. PQS pour codeur EIRE avec les différentes pondérations.

6. Conclusions

Dans cet article, nous avons proposé une méthode de calcul des facteurs de pondération visuelle pour la compression d'images couleur JPEG2000. Cette méthode est basée sur un filtrage des coefficients de la transformation de Fourier, par une fonction modélisant le Système Visuel Humain, appelée Fonction de Sensibilité au Contraste (CSF). La technique proposée réalise une meilleure prise en compte de l'interaction et de la correspondance entre la CSF 2D et les coefficients de décomposition en ondelettes, pour le calcul des facteurs de pondération visuelle.

Les évaluations psychovisuelles de la qualité des images couleur traitées, montrent que notre méthode fournit des facteurs de pondération qui, utilisés avec le codeur EIRE, produisent de meilleurs résultats visuels par rapport à l'implémentation Kakadu utilisant les pondérations JPEG2000 standard. Les résultats obtenus pour la compression avec pondérations sont également meilleurs que ceux pour la compression sans pondérations. On montre ainsi qu'il est préférable d'utiliser des facteurs de pondérations visuelles, pour obtenir une meilleure qualité visuelle d'images couleur compressées. Cet article offre aussi une description détaillée de la modélisation du Système Visuel Humain par la Fonction de Sensibilité au Contraste, dans des conditions de laboratoire qui sont totalement reproductibles. Les travaux décrits dans ce papier vont faire l'objet d'une extension au Motion-JPEG2000, notamment pour le cinéma numérique, ce qui apportera un peu plus de complexité que l'implémentation pour JPEG2000 standard. Cette étude entrera dans le cadre du nouveau groupe de travail défini lors du dernier meeting ISO/JPEG2000 qui a eu lieu à Madrid en mars 2004.

Références

- [AW04] M.D. ADAMS, R.K. WARD, JASPER: a portable flexible open-source software tool kit for image coding/processing, 2004.
- [Bar99] P.G.J. BARTEN, *Contrast sensitivity of the human eye and its effects on image quality*, Spie optical engineering press, Washington, 1999.
- [CR68] F. CAMPBELL, J. ROBSON, Application of fourier analysis to the visibility of gratings, *Journal of Physiology*, 197: 551–566, 1968.
- [DAT82] R.L. DEVALOIS, D.G. ALBRECHT, L.G. THORELL, Spatial frequency selectivity of cells in the macaque visual cortex, *Vision Research*, 22:545–559, 1982.
- [Dau80] J.G. DAUGMAN, Two-dimensional spectral analysis of the of cortical receptive field profiles, *Vision Research*, 20(10): 847–856, 1980.
- [GH73] E. M. GRANGER, J.C. HEURTLEY, Visual chromaticity-modulation transfer funcion. *J. Opt. Soc. Am.*, 63(9): 1173–1174, 1973.
- [Gla83] D.L. GLANZMAN, Spatial properties of cells in the rabbit's striate cortex, *Journal of Physiology*, 340: 535–553, 1983.
- [ISO00] ISO, Information technology – JPEG 2000 image coding system. Technical report, ISO/IEC FDIS15444-1: 2000, 2000.
- [ITU00] ITU, ITU-R recommendation BT.500-10: Methodology for the subjective assessment of the quality of television pictures. Technical report, ITU, Geneva, Switzerland, 2000.
- [KZ85] R.E. KRONAUER, Y.Y. ZEEVI, Reorganization and diversification of signals in vision. *IEEE Transactions on Systems, Man and Cybernetics*, 15(1): 91–101, 1985.
- [LK98] Y.-K. LAI, C.-C. J. KUO, Wavelet image compression with optimized perceptual quality, In *Applications of Digital Image Processing XXI*, San Diego, CA, 1998. SPIE.
- [Mic27] A. MICHELSON, *Studies in Optics*, Phd thesis, University of Chicago, 1927.
- [MKA98] M. MIYAHARA, K. KOTANI, V.R. ALGAZI, Objective picture quality scale (pqs) for image coding, *IEEE Transactions on Communications*, 46(9): 1215-1226, 1998.
- [MM96] A.B. METHA, K.T. MULLEN, Temporal mechanisms underlying flicker detection and identification for red-green and achromatic stimuli, *Journal of the Optical Society of America*, 13(10):1969–1980, 1996.
- [MS74] J.L. MANNOS, D.J. SAKRISON, The effects of a visual fidelity criterion on the encoding of images, *IEEE Trans, On Information Theory*, IT-20(4): 525-536, July 1974.
- [Mul85] K.T. MULLEN, The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings, *Journal of Physiology*, 359: 381-400,1985.
- [NR99] M. NADENAU, J. REICHEL. Compression of color images with wavelets under consideration of the hvs. In *SPIE Human Vision and Electronic Imaging*, volume 3644, pp. 237-242, San Jose, CA, 1999.
- [NRK03] M. J. NADENAU, J. REICHEL, M. KUNT, Wavelet-based color image compression: Exploing the contrast sensitivity function, *IEEE Transactions on Image Processing*, 12(1): 58-70, 2003.
- [OT86] A.L. OLZAK, J.P. THOMAS, Seeing spatial patterns. In L. Kaufman R. Boff and J.P. Thomas editors, editors, *Handbook of Perception and Human Performance*, volume 1, chapter 7. John Wiley & Sons, 1986.
- [PL62] J.J. DE PALMA, E.M. LOWRY, Sine-wave response of the visual system. ii sine-wave and square-wave contrast sensitivity, *Journal of the Optical Society of America*, 52(3): 328-335, 1962.
- [PR83] D. A. POLLEN, S.F. RONNER, Visual cortical neurons as localized spatial frequency filters, *IEEE Transactions on Systems, Man, and Cybernetics*, 13(5): 907-916, 1983.
- [pro00] JJ2000 PROJECT, JJ2000 : A JAVA implementation of JPEG 2000, white paper, Technical report, 2000.
- [PW93] A.B. POIRSON, B.A. WANDELL, Appearance of colored patterns: patterncolor separability, *Optics and Image Science*, 10(12): 2458–2470, 1993.
- [PW96] A.B. POIRSON, B.A. WANDELL, Pattern-color separable pathways predict sensitivity to simple colored patterns, *Vision-Research*, 36(4): 515-526, 1996.
- [SB82] B. SAKITT, H.B. BARLOW, A model for the economical encoding of the visual image in cerebral cortex, *Biological Cybernetics*, 50: 43-97, 1982.
- [SCE01] A. SKODRAS, C. CHRISTOPOULOS, T. EBRAHIMI, The JPEG 2000 still image compression standard. *IEEE Signal Processing Magazine*, 18(5), pp. 36-58, 2001.
- [Tau03] D.S. TAUBMAN, Kakadu survey documentation, Technical report, Kakadu Web Site, <http://kakadusoftware.com/>, 2003.
- [TM02] D.S. TAUBMAN, M.W. MARCELLIN, *JPEG2000 image compression fundamentals, standards and practice*. Kluwer Academic Publishers, 2002.
- [Wan95] B.A. Wandell, *Foundations of Vision*, Sinauer Associates, Inc., Sunderland MA, 1995.
- [WYSV97] A.B. WATSON, G.Y. YANG, J.A. SOLOMON, J. VILLASENOR. Visual thresholds for wavelet quantization error. In *SPIE Human Vision and Electronic Imaging*, volume 2657, 1997.
- [ZDL00] W. ZENG, S. DALY, S. LEI, Point-wise extended visual masking for JPEG2000 image compression. In *IEEE International Conference on Image Processing*, Vancouver, Canada, Sept 2000.
- [ZG81] E. ZRENNER, P. GOURAS, Characteristics of the blue sensitive cone mechanism in primate retinal ganglion cells, *Vision Research*, 21: 1605-1609, 1981.

Annexe

Des exemples de facteurs de pondération visuelle pour la compression JPEG2000 d'images couleur, obtenue avec la méthode

proposée dans cet article, sont présentés dans les tableaux 1 et 2. Ces facteurs sont calculés dans l'espace $YCbCr$ (Y représente la composante 1, Cb la composante 2 et Cr la composante 3), pour une distance d'observation de 85 et 114cm respectivement et une résolution d'écran de 1024×768 pixels.

Tableau 1. Fichier contenant des facteurs de pondération visuelle pour la compression JPEG2000, dans l'espace couleur $YCbCr$, pour des distances d'observation : a) de 85 cm et b) de 114 cm.

<pre> # # File holding the CSF weights # for each component and each subband. # Component 1: 1 0.37455931 0.36494891 0.36282431 1 0.04541019 0.03290261 0.03597109 1 0.99921101 0.99925283 0.99934435 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 Component 2: 1 0.17890396 0.18316656 0.18569195 1 0.41238412 0.40853942 0.41026656 1 0.61635217 0.62775187 0.62196691 1 0.77297945 0.78587782 0.78153883 1 0.88916248 0.89354522 0.88963015 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 Component 3: 1 0.19465322 0.19189999 0.20204259 1 0.42672868 0.42315904 0.42922911 1 0.61282996 0.62140225 0.61871203 1 0.77435471 0.78079066 0.77220772 1 0.88722389 0.88362862 0.88608932 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 </pre>	<pre> # # File holding the CSF weights # for each component and each subband. # Component 1: 1 0.18497085 0.17910707 0.17777632 1 0.45025489 0.43475749 0.4386358 1 0.97947817 0.98062315 0.98077934 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 Component 2: 1 0.10991105 0.11293475 0.11503608 1 0.31714218 0.31385165 0.31538797 1 0.53133327 0.54408675 0.53767324 1 0.7111605 0.72677609 0.72148072 1 0.85237315 0.85794222 0.85299424 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 Component 3: 1 0.12410514 0.12200413 0.12998783 1 0.33409998 0.33102423 0.33667455 1 0.53009166 0.53944237 0.53664925 1 0.71472519 0.72229575 0.71214845 1 0.85076985 0.84634833 0.84936342 1 1.00000000 1.00000000 1.00000000 1 1.00000000 1.00000000 1.00000000 </pre>
--	--

(a)

(b)



Adrian **Stoica**

Adrian Stoica a reçu son diplôme d'ingénieur en électronique et télécommunications de l'Université polytechnique de Bucarest en 2000 et son DEA Transmission d'information : Informatique, Image et Automatique de l'Université de Poitiers en 2001. Il a reçu son Doctorat en traitement d'images de l'Université de Poitiers en décembre 2004. Son travail s'est concentré sur les modèles du système visuel humain, la compression d'images couleur et l'évaluation de la qualité, avec application à la norme de compression JPEG2000.



Mohamed-Chaker **Larabi**

Mohamed-Chaker Larabi a reçu son DEA en automatique industrielle de l'Université de Lyon en 1999 et son Doctorat en électronique de l'Université de Poitiers en 2002. Il est actuellement Maître de Conférences au sein de cette même Université et est membre du laboratoire SIC. Ses intérêts de recherche sont concentrés sur l'évaluation de la qualité, les modèles d'apparence couleur, la compression d'images (JPEG2000) et l'indexation de vidéo par le contenu.



Christine **Fernandez-Maloigne**

Christine Fernandez-Maloigne a reçu son DEA et son Doctorat de l'Université Technologique de Compiègne respectivement en 1986 et 1989. Elle a passé son HDR à l'Université des Sciences et Technologie de Lille en 1995. Elle est actuellement Professeure à l'Université de Poitiers et responsable de l'équipe ICONES, Images COuleur dyNamiques Et Statique du laboratoire de SIC. Son travail est concentré sur le traitement d'images couleur, les modèles d'apparence couleur et l'évaluation de qualité.



