

Indice de qualité des images avec référence par apprentissage statistique*

C. Charrier

O. Lézoray

G. Lebrun

Universtié de Caen Basse-Normandie, GREYC UMR CNRS 6072

{christophe.charrier,olivier.lezoray,gilles.lebrun}@unicaen.fr

Résumé

L'utilisation d'un processus de classification et de régression basé sur les Séparateurs à Vaste Marge permet de plagier le fonctionnement du SVH qui utilise plusieurs échelles interne d'évaluation de qualité selon le principe du modèle perception-conscience défini par Freud. L'application des SVM dans un cadre multi-classe implique la définition d'une stratégie de combinaison des résultats réalisée sous raisonnement incertain. Ceci permet de prendre en compte les incertitudes des classificateurs impliqués. Une fois le processus de classification exécuté, une régression par SVM est effectuée afin de passer d'une représentation vectorielle de l'image à un score de qualité. Une fonction de régression est construite pour chacune des classes de qualité de façon à garantir une complexité faible de ladite fonction.

Mots clefs

Classification, Indice de qualité, Régression, Apprentissage statistique

1 Introduction

Les schémas de compression avec perte d'information des images couleur induisent indubitablement des dégradations dont la visibilité est plus ou moins élevée selon leur importance et leur voisinage proche (effet de masquage, ou, son dual, effet de facilitation). L'évaluation de la qualité des images compressées est importante et peut être basée sur l'observation humaine. Les résultats ainsi obtenus, permettent de déterminer des seuils de sensibilité de l'observateur humain. L'un des inconvénients majeurs de ce type de mesure, est une mise en œuvre particulièrement longue ainsi que l'utilisation d'un protocole contraignant [1]. L'objectif est alors de disposer d'une mesure automatique de qualité, plus rapide, et, *a priori*, sans intervention directe de l'être humain dans le processus de prédiction du score de qualité d'une image. Quelques travaux d'estimation de la qualité basés sur le développement d'un expert informatique existent [2, 3, 4]. Toutes ces approches reposent sur l'utilisation de réseaux de neurones de type perceptron multicouche. Plus récemment, Narwaria [5] a proposé une méthode d'estimation de la qualité utilisant une régression à base de vecteurs support (*SVR-Support*

Vector Regression). Les vecteurs de caractéristiques utilisés sont ceux obtenus après application d'une décomposition en valeurs singulières de la matrice image sur toute la dynamique de la qualité. Or, la sensibilité du SVH aux dégradations n'évolue pas de manière linéaire.

Dès lors, l'idée développée est d'utiliser un processus de classification multi-classe de la qualité des images [6] auquel sera joint un processus de régression pour chacune des classes de manière à obtenir une prédiction du score final de qualité, tel que présenté dans la figure 1. Il sera fait référence à cette méthode sous l'acronyme *MLIQM-Machine Learning based Image Quality Metric*.

2 Création et manipulation du vecteur de qualité : le choix des attributs

D'après le domaine de recherche sur la perception, il existe une hypothèse selon laquelle le Système Visuel Humain (SVH) a évolué au cours du temps pour s'adapter à son environnement, notamment en développant une adaptation aux statistiques des scènes naturelles (SSN) qui l'entoure. Afin de prendre en compte cette hypothèse, les premiers critères sélectionnés sont ceux utilisés pour définir la métrique MS-SSIM [7]. On retrouve ainsi le critère de distorsion de luminance $l(I, J)$, de contraste $c(I, J)$ et un critère de comparaison de structure $s(I, J)$ entre une image originale I et sa version dégradée J . Contrairement à la définition de l'indice MS-SSIM original dans lequel le critère $l(I, J)$ était uniquement calculé sur le dernier niveau de résolution, chacun des trois critères de distorsion est calculé à travers les cinq échelles de résolution. Chacun des critères ainsi obtenu sera utilisé en tant que caractéristique à part entière dans le vecteur de qualité. Ainsi, à ce stade, le vecteur de qualité est constitué de 15 caractéristiques.

La sensibilité du SVH aux variations fréquentielles est établie depuis de nombreuses années et a permis de formaliser diverses propriétés fonctionnelles pertinentes du SVH, telles la sensibilité au contraste, les effets de masquage, etc. Plusieurs modèles estimant la visibilité des erreurs en intégrant ces divers formalismes ont été développés depuis lors. Tous ces modèles reposent sur une décomposition multi-canal du signal, chacun des canaux étant sélectif en orientation et en fréquences spatiales de manière à prendre en compte la sensibilité du SVH aux fréquences spatiales.

*Ce travail a été financé par le projet ANR #ANR-08-SECU-007-04.

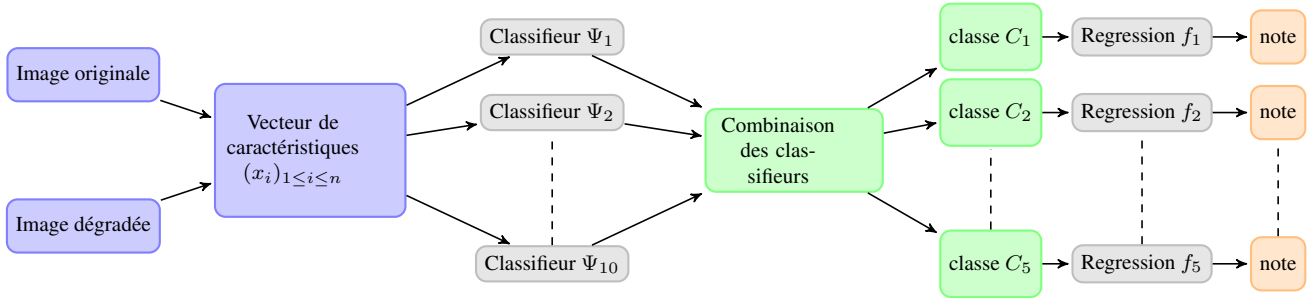


Figure 1 – Synopsis de la méthode d'évaluation de la qualité basée sur une approche d'apprentissage statistique, nommée MLIQM–Machine Learning based Image Quality Metric.

Les diverses décompositions disponibles diffèrent principalement en nombre de canal (bande radiale), en orientation et en largeur de bande [8, 9].

Notre choix s'est porté sur la transformation en pyramide orientée (*steerable pyramid transform*) [10] de manière à mesurer les effets de masquage. Une décomposition en trois niveaux et quatre orientations (0, 45, 90 et 135 degrés) suivie d'un filtrage passe-bas isotrope a été utilisée. Soit $a(x, y, f, \theta)$ le coefficient obtenu après application de la transformation précédemment décrite à la position (x, y) dans la bande fréquentielle f d'orientation θ . De manière à obtenir une mesure d'énergie locale, le coefficient $a(x, y, f, \theta)$ est normalisé comme suit :

$$r(x, y, f, \theta) = k \frac{(a(x, y, f, \theta))^2}{\sum_{\phi \in \{0, 45, 90, 135\}} (a(x, y, f, \phi))^2 + \sigma^2}, \quad (1)$$

Cette étape de normalisation est utile pour discriminer les différences de contraste le long de bande de faible dynamique. Ceci requiert l'utilisation de nombreuses bandes pour déterminer les variations de contraste. L'erreur de contraste entre l'image originale $r_0(x, y, f, \theta)$ et sa version dégradée $r_1(x, y, f, \theta)$ pour chaque bande fréquentielle f d'orientation θ est définie par :

$$\Delta r(f, \theta) = \left\| \sum_{x, y} r_0(x, y, f, \theta) - r_1(x, y, f, \theta) \right\|^2 \quad (2)$$

Ainsi, nous disposons de 12 valeurs supplémentaires qui viennent compléter le vecteur des caractéristiques.

Le vecteur final S contient alors 27 attributs $(s_k)_{1 \leq k \leq 27}$ de qualité, tant spatial que fréquentiel.

3 Le processus de classification du vecteur de qualité

La théorie de l'Apprentissage Statistique de Vapnik et de Chervonenkis [11] a conduit au développement d'une classe d'algorithmes connus sous le nom de SVM (Support Vector Machines). Ils permettent de réaliser des estimations en classification (et en régression). Une des originalités de la méthode est de produire une fonction de décision qui n'utilise qu'un sous-ensemble de la base d'apprentissage. Les éléments de ce sous-ensemble sont nommés

Vecteurs de Support (SV). En outre, le choix de cette technique de classification a été motivé par ses fortes capacités en généralisation de données de grande dimension.

Soit une base d'apprentissage $A = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_k, y_k)\}$ composée de k couples (vecteur d'attributs, label) avec $x_i \in \mathcal{R}^n$ et $y_i \in \{-1, +1\}$. L'algorithme des SVM projette les vecteurs x_i dans un espace de travail \mathbf{H} à partir d'une fonction non linéaire $\phi : \mathcal{R}^n \rightarrow \mathbf{H}$. L'hyperplan optimal de séparation des deux classes dans l'espace \mathbf{H} est ensuite recherché. Cet hyperplan (\mathbf{w}, b) matérialise la frontière de séparation entre les deux classes. La classe y d'un nouvel exemple \mathbf{x} est définie par :

$$y = \text{sign}(\mathbf{w} \cdot \phi(\mathbf{x}) + b) \quad (3)$$

L'hyperplan est optimal s'il maximise la distance qui le sépare des exemples dont il est le plus proche. Cette distance est usuellement appelée marge du classificateur. Il a été démontré [11] que maximiser cette marge correspond à maximiser le pouvoir généralisateur du classificateur. En choisissant une fonction noyau $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$, il a également été montré que la résolution du problème dual correspondant produit une fonction de décision de la forme

$$f(x) = \sum_{x_i \in SV} \alpha_i y_i K(x_i, x) + b \quad (4)$$

où $\{\alpha_i\}$ et b correspondent à la solution optimale du problème dual.

3.1 La sélection du modèle

Le choix de la fonction noyau est critique dans la construction d'un expert informatique. La fonction RBF (*Radial Basis Function*) est très souvent utilisée comme fonction noyau ; la principale raison étant qu'elle fonctionne comme une mesure de similarité entre deux exemples à classer. Puisque aucune connaissance *a priori* existe sur l'importance d'un attribut s par rapport aux autres, la fonction classique RBF a été modifiée de manière à modéliser les modulations d'importance des attributs (s_k) du vecteur d'attributs S telle que :

$$K_{\beta}(s_i, s_j) = \exp\left(-\sum_{k=1}^n \beta_k (s_i^k - s_j^k)^2 / r^2\right) \quad (5)$$

avec $s_{k,i}$ le $k^{\text{ième}}$ attribut de l'image i . β_k , qui est valeur booléenne, permet de modéliser la prise en compte ou non de la caractéristique k . Lorsque les valeurs de β_k ne sont pas fixées par des *a priori*, elles sont déterminées en utilisant le paradigme de sélection d'attributs. La qualité d'un sous-ensemble d'attributs utilisé pour l'apprentissage des SVM binaires est mesurée par le taux de reconnaissance.

3.2 Les règles de combinaison

Les SVM étant des classificateurs binaires, ils ne peuvent pas être directement appliqués pour définir un schéma de classification multiclass. Il convient d'effectuer une décomposition en plusieurs classificateurs binaires et de combiner les résultats obtenus par chacun des classificateurs binaires pour aboutir à la décision finale. Le schéma de décomposition un-contre-un (UCU) a été utilisé dans notre étude.

La décision finale est prise à partir de celle obtenue pour chacune des fonctions de décision. Plusieurs stratégies de combinaison peuvent être adoptées afin d'obtenir la décision finale. Le critère du vote majoritaire est celui qui est le plus couramment utilisé. Cependant, l'application d'un tel choix peut se révéler drastique. En effet, la classification d'un exemple peut être incertain lorsqu'il se situe à la frontière de deux classes. Dans ce cas, il conviendrait de prendre la décision finale en intégrant cette incertitude.

Parmi tous les outils qui s'offrent à nous dans ce domaine, nous décidons d'utiliser la théorie de l'évidence (aussi appelée théorie de Dempster-Shafer ou théorie des fonctions de croyance) [12], qui permet d'une part de traiter des informations incertaines, et d'autre part de combiner des informations provenant de plusieurs sources. Ce cadre théorique permettra de traiter l'incertitude relative à chacun des classificateurs binaires qui peuvent être considérés comme des sources d'information imprécises et incertaines. Conceptuellement, la décision finale est prise en fonction de la confiance que l'on accorde au classificateur binaire. Cette confiance peut revêtir plusieurs aspects : un taux de reconnaissance, une probabilité de vraisemblance, une probabilité *a posteriori*, etc. Cependant, un SVM ne permet pas d'aboutir directement à de telles mesures de confiance. Notre choix s'est porté sur le calcul d'une probabilité *a posteriori* de classification. Afin d'affecter une probabilité *a posteriori* au classificateur (ce qui peut se traduire par une confiance que l'on accorde à la décision prise par un classificateur), une fonction paramétrique de type fonction logistique est utilisée [13] :

$$p(\mathbf{x}_i = 1 | \Psi_i) = \frac{1}{1 + \exp(a\Psi_i + b)}, \quad (6)$$

Les paramètres a et b sont estimés en utilisant la procédure du maximum de vraisemblance. Ainsi, la confiance que l'on accorde à la classification d'un exemple \mathbf{x}_i par le classificateur Ψ_i est définie par la probabilité conditionnelle $p(\mathbf{x}_i = 1 | \Psi_i)$. Reste alors à combiner les résultats ainsi obtenus afin de prendre la décision finale d'apparte-

nance à l'une des cinq classes de qualité, selon le formalisme de la théorie de l'évidence.

4 Combinaison des classificateurs sous raisonnement incertain

4.1 Éléments de théorie de l'information

Soit $\Omega = \{\omega_1, \dots, \omega_N\}$, l'ensemble des N classes finales possibles pour un pixel, appelé cadre de discernement. Dans notre cadre applicatif, $N = 5$ et Ω correspond aux cinq classes $\{\omega_E, \omega_B, \omega_{MO}, \omega_{MA}, \omega_{TM}\}$, respectivement pour les classes de qualité excellente, bonne, moyenne, mauvaise et très mauvaise. Au lieu de restreindre ces mesures à cet ensemble (comme le ferait la théorie des probabilités contrainte par son axiome d'additivité), la théorie de l'évidence s'étend sur l'ensemble puissance de Ω , noté 2^Ω , l'ensemble des 2^N sous-ensembles de Ω . On définit alors une fonction de masse m qui représente la croyance que l'on accorde aux différents états du système, à un instant donné. Cette fonction est également connue en tant que masse de croyance élémentaire. Elle est définie de 2^Ω dans $[0, 1]$ et vérifie $\sum_{A \subseteq \Omega} m(A) = 1$ et $m(\emptyset) = 0$ où $m(A)$ représente la croyance que l'on place dans la proposition A . Il est à noter que A représente soit un singleton ω_n , soit une disjonction d'hypothèses. Ici, $m(A)$ quantifie la croyance que l'on accorde au fait que la classe cherchée appartienne au sous-ensemble A de Ω (et à aucun autre sous-ensemble de A). Les propositions A telles que $m(A) > 0$ sont appelées éléments focaux. De plus, dans le cas où l'ensemble des hypothèses est exhaustif et exclusif, la masse de l'ensemble vide \emptyset est nulle.

Deux fonctions de masse initiales m_1 et m_2 représentant les informations respectives de deux sources différentes, peuvent être combinées selon la règle de Dempster [14].

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K}, \quad \forall A \in \Omega, \quad A \neq \emptyset. \quad (7)$$

K est appelé le facteur de conflit et représente le désaccord entre les deux sources. Il correspond à la masse de l'ensemble vide $K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$.

Après combinaison, il reste à prendre une décision quant à l'élément le plus « probable » de Ω . Plusieurs règles de décision sont alors possibles, mais une des plus utilisées reste celle du maximum de probabilité pignistique $\text{BetP}(\omega)$. Cette règle de décision présentée par SMETS [15] utilise la transformation pignistique, qui permet de répartir uniformément la masse associée à un sous-ensemble de Ω sur chacun de ses éléments :

$$\text{BetP}(\omega) = \sum_{\omega \in A \subseteq \Omega} \frac{m(A)}{|A|}, \quad \forall \omega \in \Omega \quad (8)$$

$|A|$ étant le cardinal de A . La décision va alors vers l'élément de Ω dont cette valeur est la plus grande :

$$\omega^* = \text{Arg} \{ \max_{\omega \in \Omega} [\text{BetP}(\omega)] \}. \quad (9)$$

4.2 Construction des fonctions de masse

L'une des difficultés réside dans le choix de la modélisation des fonctions de masse. Parmi les modélisations existantes, celle proposée par DENÈUX [16] a été utilisée car elle est basée sur la notion de distance aux voisins et permet d'intégrer divers critères de voisinage. Ainsi, la masse $m(\{\omega_n\})$ est définie comme une fonction décroissante de la distance d entre le vecteur \mathbf{x}_i à classifier et le barycentre de la classe ω_n :

$$\begin{cases} m(\omega_n) = \alpha \exp(-\gamma_n d^2) \\ m(\Omega) = 1 - m(\omega_n) \end{cases} \quad (10)$$

où $0 < \alpha < 1$ correspond à la probabilité *a posteriori* associée au classificateur binaire Ψ_l ; $\gamma_n > 0$ dépend de la classe ω_n et est calculé par minimisation d'un critère d'erreur basé sur l'algorithme SEM (*Stochastic Expectation Maximization*).

Au lieu de se restreindre à une seule classe, la théorie de l'évidence permet d'agréger dans un même élément focal plusieurs classes candidates selon un critère de similarité. Ainsi, un élément à classifier est associé à un ensemble de classes, qui est au plus l'ensemble Ω . Dès lors, la contrainte d'affectation a été relâchée de manière à construire l'ensemble A tel que :

$$A = \{\omega_l \in \Omega / \max(p_l) - \delta_l \leq p_l \leq \max(p_l)\} \quad (11)$$

où $l \in \{1, \dots, 5\}$ et δ_l est une constante ad-hoc liée au classificateur utilisé.

Dans ce cas, les classes pour lesquelles la probabilité *a posteriori* associée appartient à l'intervalle $\max(p_l) - \delta_l \leq p_l \leq \max(p_l)$ sont considérées comme candidates pour le processus de classification durant l'étape de fusion.

5 Schéma de régression

Même si l'estimation de la qualité par un observateur humain est plus en adéquation avec un processus de classification qu'avec un processus de notation scalaire, force est de constater que la nécessité d'aboutir à une note de qualité est indiscutable. Ceci est inhérent à l'existence d'un ordre total dans l'espace \mathbb{R} permettant un classement des valeurs par ordre croissant (ou décroissant).

Dans ce cas, en considérant la base d'apprentissage A telle que définie dans la section 3, l'idée est d'estimer une fonction f qui présente au plus une déviation maximale ϵ traduisant la dépendance entre un exemple \mathbf{x}_i et la classe d'affectation y_i et qui soit la plus régulière possible. C'est ainsi que l'utilisation des SVM dans un contexte de régression est connue sous le terme « ϵ -SV régression ». Dans ce cas, les fonctions de coûts L_ϵ utilisées sont linéaires ou quadratiques et ne prennent en compte que les déviations $|y_i - f(\mathbf{x}_i)| > \epsilon$. Cette fonction de coût est communément appelée fonction ϵ -insensible et est définie par

$$L_\epsilon(x, y) = \begin{cases} 0 & \text{si } |y - f(\mathbf{x})| \leq \epsilon \\ |y - f(\mathbf{x})| - \epsilon & \text{sinon} \end{cases} \quad (12)$$

Les motivations relatives à l'utilisation d'un schéma ϵ -SV régression sont :

1. la solution en régression ne fait intervenir qu'un nombre limité d'exemples \mathbf{x}_i , ce qui permet de faire une sélection des exemples pertinents pour chacune des cinq classes de qualité,
2. la forme de la solution dépend du noyau utilisé. Dans ce cas, l'exploration de plusieurs espaces d'hypothèses est possible de manière à ne retenir que celui dans lequel les taux de corrélation avec les valeurs DMOS sont maximum.

De manière à rester cohérent avec l'échelle de qualité de l'ITU, un intervalle numérique est assigné à chacune des cinq classes de qualité. La dynamique de l'intervalle global est $[0;5]$, ce qui induit une longueur d'intervalle d'une unité pour chacune des classes. Dès lors, la classe de « très mauvaise qualité » est associée à l'intervalle numérique $[0;1]$, la classe « mauvaise qualité » est associée à l'intervalle numérique $]1;2]$ et ainsi de suite jusqu'à la classe de qualité « excellente » qui est associée à l'intervalle $]4;5]$. Ainsi, aucun chevauchement des scores n'est possible entre les classes.

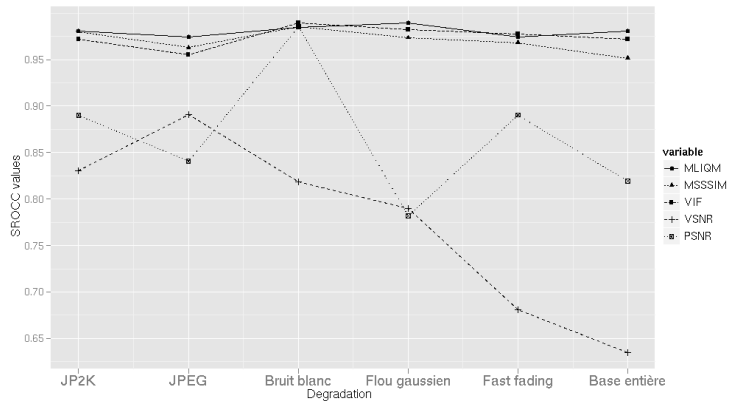


Figure 2 – Obtained correlation coefficients between the predicted DMOS values and the subjective DMOS scores considering LIVE database test set.

6 Mesure des performances du schéma général

6.1 Apparatus

De manière à évaluer les performances du schéma développé, deux bases d'images sont utilisées : la base LIVE [18] et la base TID2008 [19]. Une base d'apprentissage (TrainC1) issue de la base LIVE est construite. Elle est composée de 12 images de référence et de toutes les versions dégradées associées. De plus, afin de procéder à l'apprentissage des fonctions de régression f_1, f_2, \dots, f_5 , cinq bases d'apprentissage (TrainR1, TrainR2, ..., TrainR5) sont créées ; une pour chaque classe de qualité. Les images non utilisées de

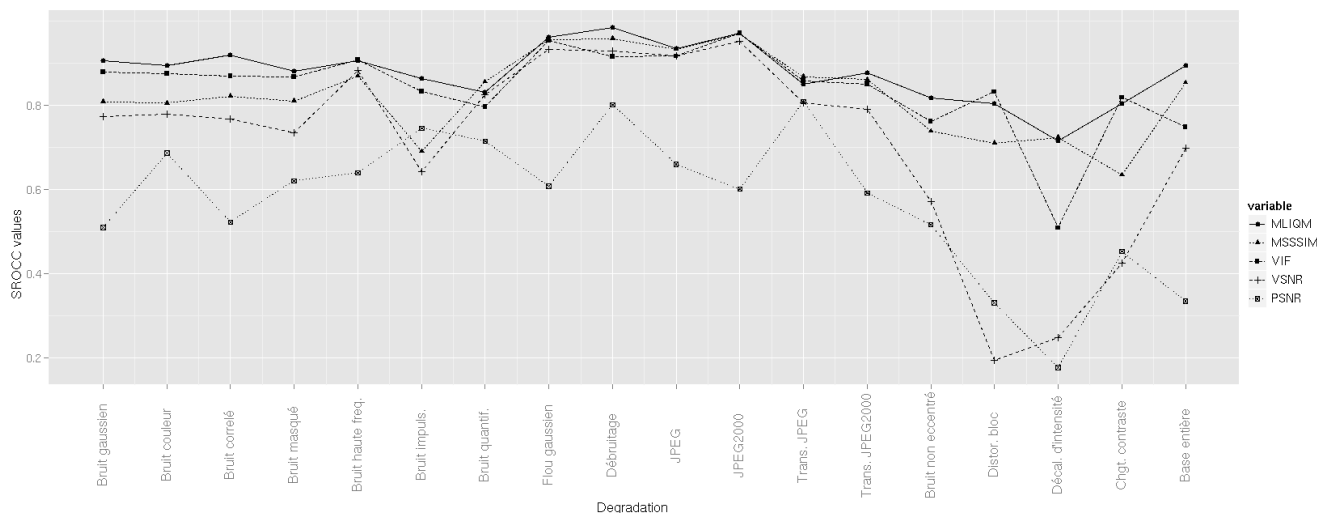


Figure 3 – Obtained Spearman rank order correlation coefficient (SROCC) between the predicted DMOS values and the subjective DMOS scores considering TID2008 database as test set. The type of degradations are described in table ??.

la base LIVE constituent la base d'apprentissage `TestC1`. De même, la base TID2008 servira de base de test.

Les paramètres du schéma de classification multi-classe basé sur les SVM ainsi que des cinq schémas de régression sont déterminés par une technique de validation croisée à 10 parties. Qui plus est, une technique de bootstrap à 999 réplifications est appliquée pour valider les résultats. Afin de comparer l'approche développée, quatre techniques d'évaluation de la qualité des images avec référence sont utilisées : il s'agit de l'indice MS-SSIM [20], de la mesure VSNR [21], de l'indice VIF [22] et de la mesure du PSNR. Toutes ces techniques d'évaluation ont la particularité commune d'être appliquées sur la composante achromatique des images. Les valeurs quantitatives des performances sont évaluées à l'aide de la mesure de corrélation de Spearman (SROCC).

6.2 Résultats

La figure 2 présente les résultats obtenus lors du calcul du coefficient de corrélation SROCC entre les métriques d'évaluation de la qualité et les valeurs subjectives DMOS de la base de test `TestC1`. En utilisant l'approche développée, les résultats montrent que les corrélations sont plus élevées lorsque la base entière est considérée. Pour la plupart des sous-ensembles de la base de test (*i.e.*, par type de dégradation), la méthode développée permet une amélioration de la corrélation avec les scores des observateurs humains. Même si les améliorations induites ne sont pas toutes statistiquement significatives (au sens du test de Fisher), la consistance observée dans l'augmentation des valeurs de corrélation peut-être interprétée comme un indicateur de validité de l'approche développée. Une seconde interprétation des résultats concerne les attributs utilisés. Étant donné que la qualité des attributs revêt une importance somme toute particulière dans les performances des techniques de classification et/ou de régression par appren-

tissage statistique, les résultats obtenus tendent à démontrer (de manière empirique) que les 27 attributs utilisés dans cette étude sont pertinents. Néanmoins, on peut se demander si l'utilisation d'une base d'apprentissage extraite de la base LIVE pour régler les différents paramètres de nos SVM ne vient pas quelque peu biaiser les résultats.

La figure 3 présente les résultats du calcul du coefficient de corrélation SROCC pour les cinq méthodes d'évaluation de la qualité avec les valeurs subjectives DMOS en considérant la base de test TID2008 (Il est à noter qu'aucune nouvelle phase d'apprentissage n'a été effectuée, et que les paramètres obtenus suite à la phase d'entraînement du modèle sur les ensembles d'apprentissage sont conservés.) La méthode proposée permet d'obtenir des taux de corrélation plus élevés qu'avec les quatre méthodes concurrentes, lorsque la base d'images complète est considérée. À l'exception des dégradations « Bruit Haute Freq. », « Bruit quantif. », « Trans. JPEG », « Distor. de bloc », « Décal d'intensité » et « Chgt. contraste », le schéma MLIQM est le mieux corrélé avec le jugement humain.

Si l'on considère la dégradation « Bruit Haute Freq. », la différence de corrélation n'est pas statistiquement significative. La dégradation « Bruit quantif. » (qui est une distorsion de quantification) peut être interprétée comme une perte de couleur pouvant induire d'importantes déformations structurelles sur l'image lors d'une forte quantification. Dans ce cas, les informations capturées par les attributs dédiés aux déformations structurelles sont importantes, et ceci est parfaitement mesuré dans le cas de l'indice de qualité MS-SSIM. Dans notre cas, ces informations se trouvent en compétition avec les 22 attributs restant ce qui a pour effet d'en atténuer le poids. Néanmoins, la différence de corrélation est faible.

Pour les dégradations orientées compression (sauf la dégradation « Trans. JPEG » qui concerne des erreurs de transmission JPEG), le schéma développé permet une meilleure

corrélation des scores prédits avec le jugement humain. Cependant, une faible différence de corrélation est observée pour les dégradations « Trans. JPEG » et « Distor. de bloc ». Pour cette dernière, les erreurs aboutissent à l'apparition de blocs de taille et d'intensité différentes, ce qui peut être considéré comme des erreurs de transmission, puisque le bloc reconstruit n'est pas identique à celui d'origine. Les dégradations « Décal d'intensité » et « Chgt. contraste » concernent des changements d'intensité et de contraste. Elles ne peuvent pas être considérées uniquement comme des processus de dégradation de la qualité, mais comme des changements de l'aspect naturel des images. Dans ce cas, il se peut que ces changements ne soient pas perçus comme des dégradations mais comme une amélioration des images.

6.3 Conclusion

L'approche basée sur l'utilisation d'un processus de classification suivi d'une étape de régression permettant d'aboutir à une note de qualité des images semble pertinente au regard des résultats obtenus pour deux bases d'images test. La stratégie de classification peut alors être modélisée par un classificateur de grande dimension puisque de nombreux détails peuvent venir influencer la prise de décision finale. En outre, l'incertitude liée aux processus de classification peut-être prise en compte dans le cadre de la combinaison des informations sous raisonnement incertain. L'aptitude des êtres humains à réaliser des opérations d'ordonnement peut-être modélisée par une fonction de décision de type fonction de régression non-linéaire.

Références

- [1] ITU-R Recommendation BT.500-11. Methodology for the subjective assessment of the quality of television pictures. Rapport technique, International Telecommunication Union, Geneva, Switzerland, 2002.
- [2] A. Bouzerdoun, A. Havstad, et A. Beghdadi. Image quality assessment using a neural network approach. Dans *Fourth IEEE Inter. Symp. on Signal Proc. and Information Tech.*, pages 330–333, 2004.
- [3] P. Gastaldo, R. Zunino, I. Heynderickx, et E. Vicario. Objective quality assessment of displayed images by using neural networks. *Signal Processing : Image Communication*, 20 :643–661, 2005.
- [4] R. V. Babu, S. Suresh, et A. Perki. No-reference JPEG image quality assessment using GAP-RBF. *Signal Processing*, 87(6) :1493–1503, 2007.
- [5] M. Narwaria et W. Lin. Objective image quality assessment based on support vector regression. *IEEE Transactions on Neural Networks*, 21(3) :515–519, 2010.
- [6] C. Charrier, G. Lebrun, et O. Lezoray. Image quality assessment with manifold and machine learning. Dans *SPIE, Image Quality and System Performance VI*, volume 7242, San-Jose, California, Janvier 2009.
- [7] Z. Wang et A. C. Bovik. A universal quality index. *IEEE Signal Processing Letters*, 9(3) :81–84, 2002.
- [8] A. B. Watson. The cortex transform : Rapid computation of simulated neural images. *Computer Vis. Graphics and image proces.*, 39 :311–327, 1987.
- [9] S. Daly. A visual model for optimizing the design of image processing algorithm. Dans *ICIP*, volume 2, pages 16–20, 1994.
- [10] E. P. Simoncelli et W. T. Freeman. The steerable pyramid : a flexible architecture for multi-scale derivative computation. Dans *ICIP*, pages 444–447, Washington, DC, 1995.
- [11] V. N. Vapnik. *Statistical Learning Theory*. Wiley, New York, 1998.
- [12] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.
- [13] J. Platt. Probabilities for SV machines. Dans D. Schuurmans P. J. Bartlett, B. Schölkopf et A. J. Smola, éditeurs, *Advances in Large-Margin Classifiers*, Neural Information Processing Systems. MIT Press, 2000.
- [14] A. Dempster. Upper and Lower Probabilities Induced by Multivalued Mapping. *Ann. Math. Statist.*, 38 :325–339, 1967.
- [15] P. Smets. Constructing the pignistic probability function in a context of uncertainty. *Uncertainty in Artificial Intelligence*, 5 :29–39, 1990. Elsevier Science Publishers.
- [16] T. Denoeux. A k-nearest neighbor classification rule based on dempster-shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(5) :804–813, 1995.
- [17] Laboratory for Image & Video Engineering, University of Texas (Austin). LIVE Image Quality Assessment Database. <http://live.ece.utexas.edu/research/Quality>, 2002.
- [18] N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian et J. Astola, et F. Battisti. Color image database for evaluation of image quality metrics. Dans *International Workshop on Multimedia Signal Processing*, pages 403–408, Australia, Octobre 2008.
- [19] Z. Wang, E. P. Simoncelli, et A. C. Bovik. Multi-scale structural similarity for image quality assessment. Dans *IEEE Asilomar Conference on Signals, Systems, and Computers*, pages 1398–1402, 2003.
- [20] D. M. Chandler et S. S. Hemami. VSNR : A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16(9) :2284–2298, 2007.
- [21] H. R. Sheikh et A. C. Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2) :430–444, Février 2006.