

Différents types de dégradations

Effets de bloc



La principale source d'erreur lors de la compression JPEG est ce qu'on appelle l'effet de blocs. Visuellement, cette distorsion se manifeste généralement au niveau des frontières entre blocs et apparaît comme des contours verticaux et horizontaux dont la visibilité dépend fortement de la distribution spatiale du signal image. En effet, tous les blocs sont encodés indépendamment les uns des autres. Il peut donc arriver qu'à la frontière entre deux blocs, il y ait une discontinuité facilement perceptible par l'oeil humain.

Effet d'oscillations parasites (ringing effect)



Cette dégradation est due en général à l'étape de quantification ou de décimation des coefficients hautes fréquences. Elle se manifeste sous forme d'oscillations au voisinage des régions à fort contraste et est souvent définie comme un bruit autour de ces régions. Ce sont les ondelettes dont le support croise le bord d'un objet qui créent ce type d'artefact.

Flou



Effet esthétique qui donne à voir un contour imprécis. Ce type d'erreur qui apparaît principalement dans la compression JPEG et JPEG2000 d'où la perte de netteté de l'image. Cela se caractérise par une image plus floue, dont les bords des objets sont plus diffus.

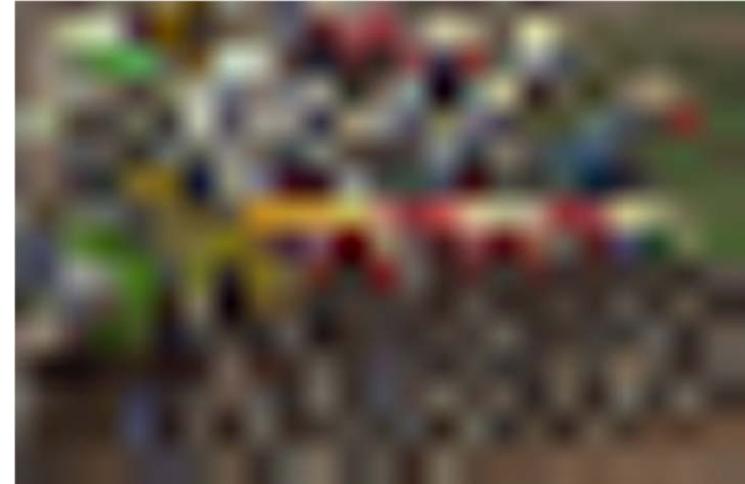
En plus, le flou apparaît dans l'image de différentes manières à savoir

Le flou au niveau de l'acquisition

Le mouvement



La perte d'informations



Le flou volontaire



Bruit

On peut définir le bruit comme étant une dégradation dans l'image, provoquée par une perturbation externe. Généralement, on peut savoir les types d'erreurs à attendre, et donc le type de bruit sur l'image, d'où nous pouvons choisir la méthode la plus adaptée pour réduire les effets.

Bruit poivre et sel

On l'appelle aussi le bruit impulsionnel, le bruit de grenaille, ou le bruit binaire. Cette dégradation peut être causée par de fortes perturbations soudaines dans le signal d'image. Son apparence est éparpillée au hasard en pixels blancs ou noirs (ou les deux) sur l'image. Par contre, ce bruit est obtenu en ajoutant n pixels blancs et n pixels noirs aléatoirement dans une image. On le caractérise souvent par le pourcentage de pixels remplacés.



Bruit Gaussien

Le bruit gaussien est une forme idéalisée de bruit blanc, qui est provoqué par des fluctuations aléatoires dans le signal. Nous pouvons observer le bruit blanc en regardant une télévision qui est mal accordée à un canal particulier. Si l'image est représentée comme I , et le bruit gaussien par N , alors nous pouvons modeler une image bruyante en additionnant simplement les deux : $F=I+N$



Évaluation subjective

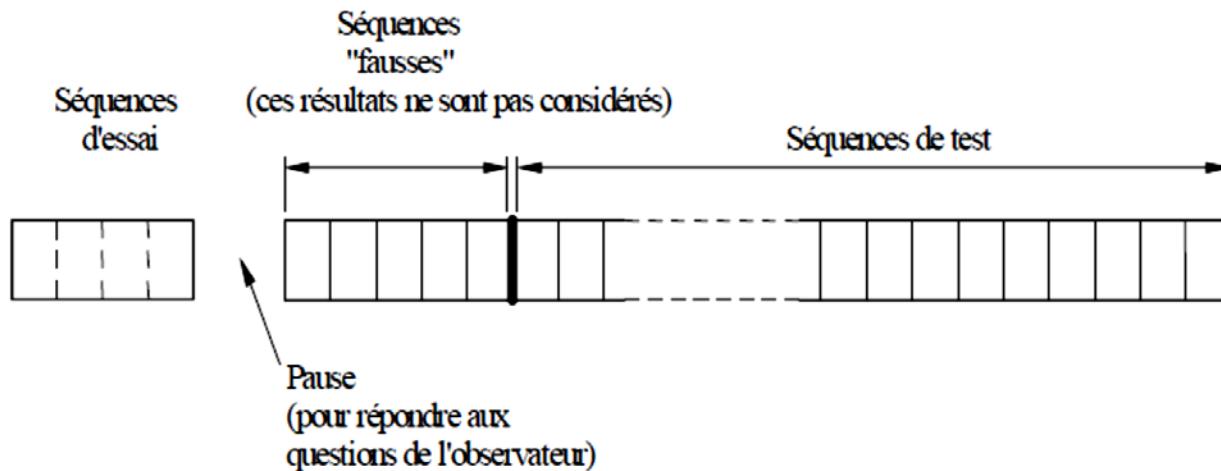
La mesure de la qualité visuelle d'images, appelée classiquement "méthode subjective d'évaluation de la qualité d'images", l'évaluation par les observateurs humains c'est le jugement le plus able. Un groupe de personnes (expert et non-expert) décide de la qualité d'une image.

Ce dernier est amené à juger la qualité des images qui lui sont présentées selon une grille d'évaluation à plusieurs niveaux d'appréciation. A l'issu de ces tests, une note subjective est obtenue.

Facteurs influents

Pour assurer la fiabilité de l'évaluation subjective, il est essentiel de prendre en compte certains facteurs qui peuvent influencer énormément le jugement des observateurs, les plus pertinents sont :

1. *Distance d'observation*
2. *Conditions de visualisation*
3. *Écran*
4. *Le choix des images*
5. *Les observateurs*
6. *Facteurs psychologiques*



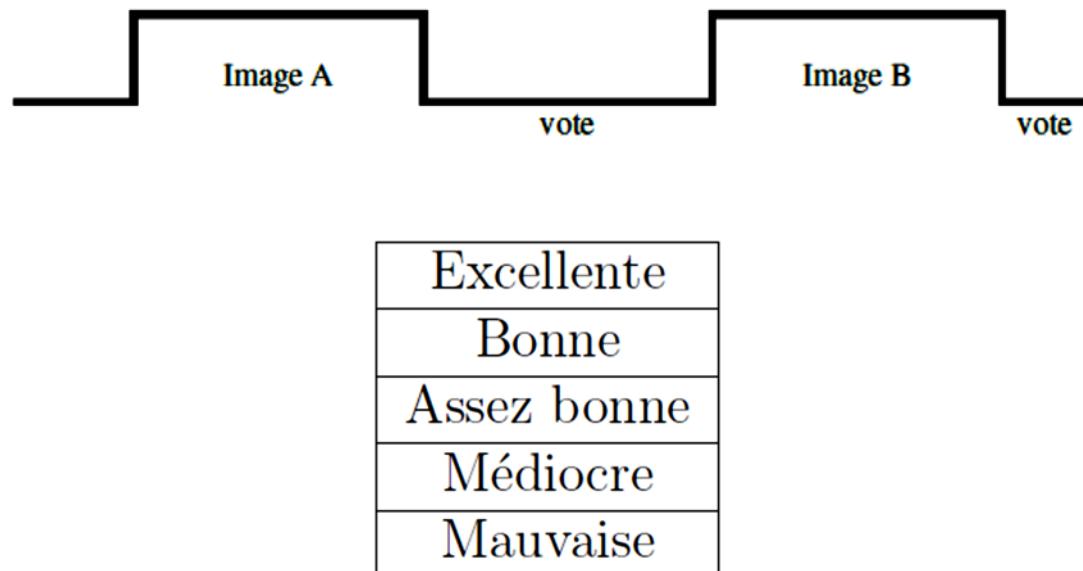
Méthodes catégorielles

Les méthodes catégorielles d'évaluation de la qualité d'une image (ou vidéo) offrent à l'observateur une échelle de catégories : celui-ci doit attribuer à l'image (ou la vidéo) présentée une note de qualité située dans la catégorie qui lui paraît la plus proche de son jugement personnel. Cette échelle peut être explicitée sous la forme de vocabulaire qualificatif (adjectifs) allant de "mauvaise qualité" à "excellente qualité" par exemple ou bien sous une forme numérique, allant de 0 à 100. Dans ce dernier cas, l'échelle est dite continue.

Selon le mode de présentation des images (ou vidéos) à juger, nous pouvons classer les méthodes catégorielles en deux groupes : méthodes à simple stimulus et méthodes à double stimuli.

Méthodes à simple stimulus

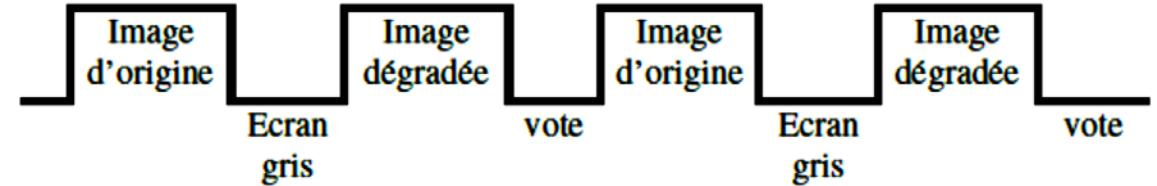
Ces méthodes consistent à présenter à l'observateur l'image (ou vidéo) dont la qualité est à évaluer : il lui attribue une note de qualité puis après une période dite de latence lui sera présentée l'image (ou vidéo) suivante et ainsi de suite. L'échelle de catégories peut être élargie ou réduite selon la dynamique des dégradations à évaluer en se conformant toujours aux normes



Méthodes a double stimulus

La méthode a double stimulus est appelée "Double Stimuli Continuous Quality Scale (DSCQS)". Ce protocole de test présente à l'observateur l'image à évaluer et la version originale correspondante. Il lui est demandé d'évaluer la fidélité de l'image par rapport à sa version originale. La similarité entre les deux images ou la conservation de l'information représentent la fidélité.

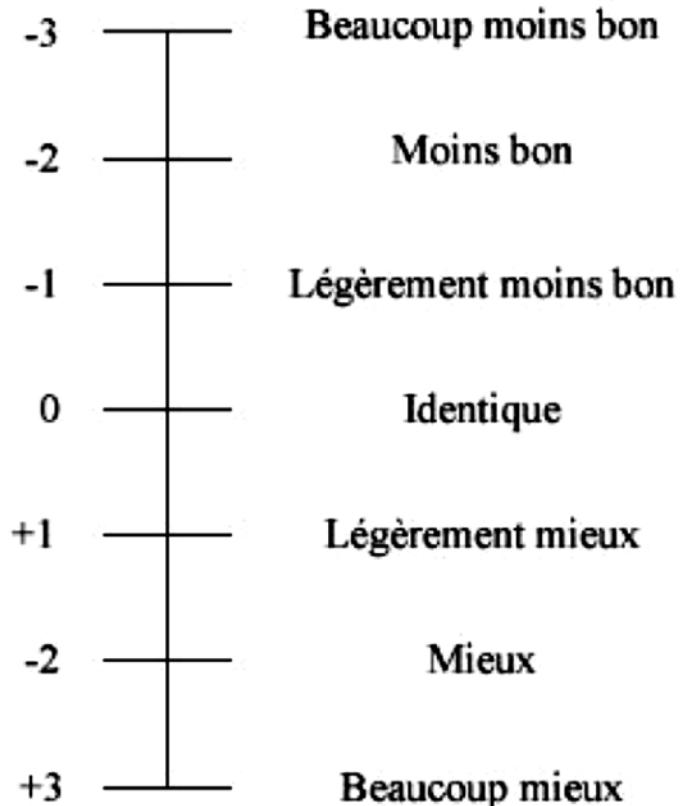
Généralement, l'image originale est présentée en premier suivie d'un écran gris, puis de l'image à évaluer, suivie d'un second écran gris permettant à l'observateur de donner sa note. Toutefois, il existe d'autres procédés pour présenter les images (image dégradée / image originale) ou (image dégradée / image dégradée / image originale / image originale). Il est toutefois important d'imposer un temps de présentation identique pour les deux stimuli.



Note	Qualité	Niveau de dégradation
5	Excellente	Imperceptible
4	Bonne	Perceptible mais non gênante
3	Assez bonne	Légèrement gênante
2	Médiocre	Gênante
1	Mauvaise	Très gênante

Méthodes comparatives

Les méthodes comparatives d'évaluation de la qualité d'une image (ou vidéo) consistent à noter la différence entre deux versions d'une même image (ou vidéo) à l'aide d'une échelle : les deux images (ou vidéos) sont montrées en même temps et l'observateur juge laquelle est de meilleure qualité.



MOS : Score d'Opinion Moyen

Lorsque les tests se terminent, on cherche à trouver la moyenne de la qualité. Une note est attribuée à chaque image en combinant les notes de tous les observateurs. Cette note appelée "Mean Opinion Score (MOS)" est donnée par la formule suivante :

$$MOS(i) = \frac{1}{N} \sum_{j=1}^N Note_i(j)$$

où N est le nombre d'observateurs et $Note_i(j)$ est la note de l'observateur j pour la dégradation i de l'image.

DMOS : Difference mean opinion score

Au lieu d'appliquer directement les résultats d'évaluation, les mesures modernes de qualité d'image utilisent les différences de qualité entre les images. Le DMOS est définie comme la différence entre les notes de qualité brutes des images de référence et de test. Le DMOS est calculé à l'aide du équation suivante :

$$d_{i,j} = r_{i,ref}(j) - r_{i,j}$$

Évaluation objective de la qualité des images fixes

Limites des métriques subjectives

- Ils sont longs et coûteux. Cela est dû au fait que les résultats subjectifs sont obtenus par des expériences avec de nombreux observateurs.
- Ils ne peuvent pas être incorporés dans des applications en temps réel telles que la compression et la transmission d'images.
- Leurs résultats dépendent fortement des conditions physiques et de l'état émotionnel des observateurs. De plus, d'autres facteurs tels que le dispositif d'affichage et les conditions d'éclairage affectent les résultats de ces expériences.

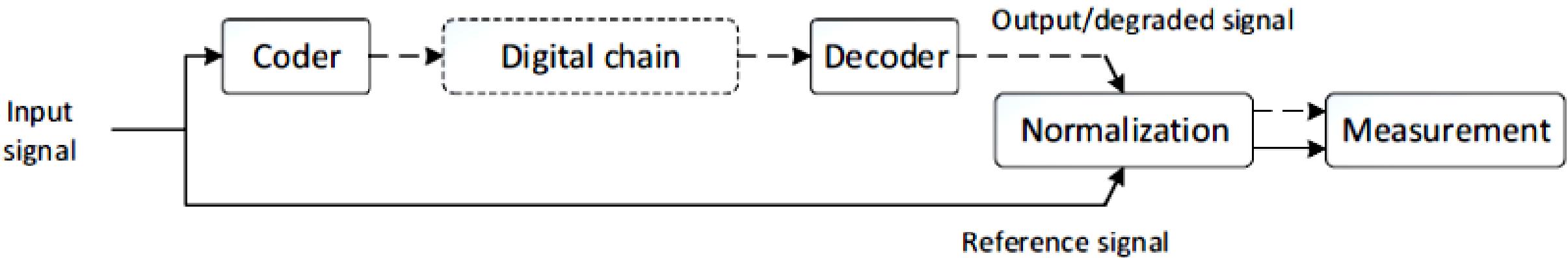
Le but de l'évaluation objectives de la qualité est de concevoir des modèles mathématiques capables de prédire la qualité d'une image avec précision et automatiquement. L'idéal est que la méthode objective doit être capable d'imiter les prévisions de qualité d'un observateur humain. Les méthodes IQA ont une grande variété applications:

- Ils peuvent être utilisés pour surveiller la qualité de l'image dans les systèmes de contrôle de la qualité. Par exemple, les systèmes d'acquisition d'images peuvent utiliser une métrique IQA objective pour surveiller et s'adapter automatiquement afin d'obtenir la meilleure Qualité d'image.
- Ils peuvent être utilisés pour évaluer les algorithmes de traitement d'image. Par Exemple, si un certain nombre d'algorithmes d'amélioration d'image sont disponibles, une métrique IQA objective peut être utilisée pour choisir l'algorithme qui fournit des images de meilleure qualité.
- Ils peuvent être utilisés pour optimiser les systèmes de transmission et de traitement d'image . Par exemple, dans un réseau de communication visuelle, une métrique IQA objective Peut être utilisée pour optimiser les algorithmes de pré-filtrage et d'attribution de bit dans la partie codage et de post-filtrage et de reconstruction dans la partie décodage.

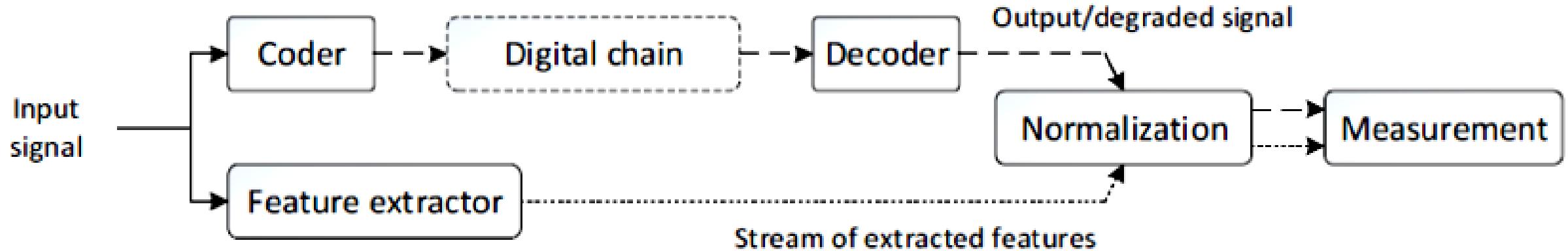
Catégories des métriques IQA

En se basant sur la présence du signal original, on peut classer les métrique IQA en trois catégories

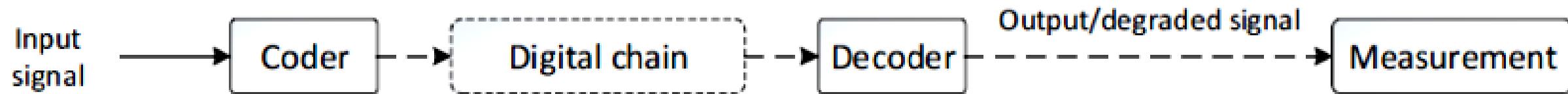
Métriques à référence complète



Métriques à référence réduite



Métriques sans référence



Métriques à référence complète

Métriques simples(sans prendre en compte le HVS)

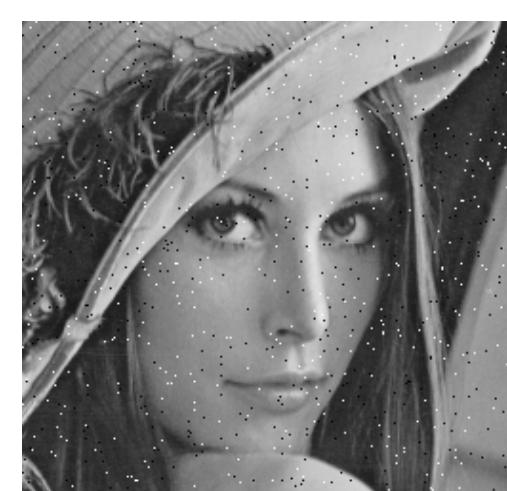
PSNR (Peak Signal to Noise Ratio)

$$PSNR = 10 \log_{10} \left(\frac{I_{\max}^2}{MSE} \right)$$

MAX représente la valeur maximal que peut prendre un pixel, le MSE (Mean Square Error) peut être défini par la formule suivante

$$MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (I(m, n) - \hat{I}(m, n))^2$$

Limite du PSNR

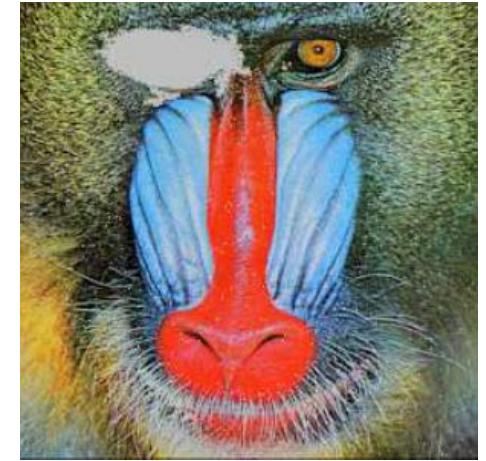




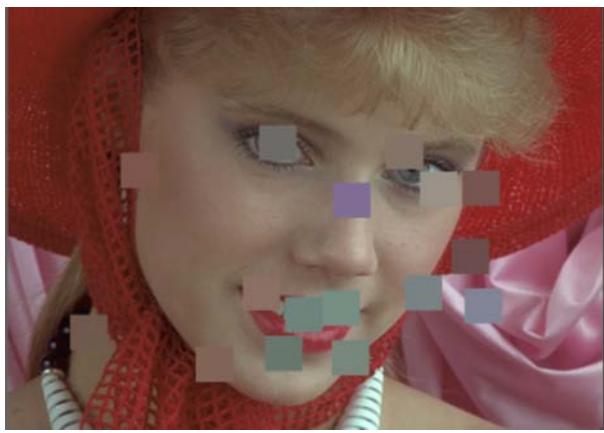
21,51 dB



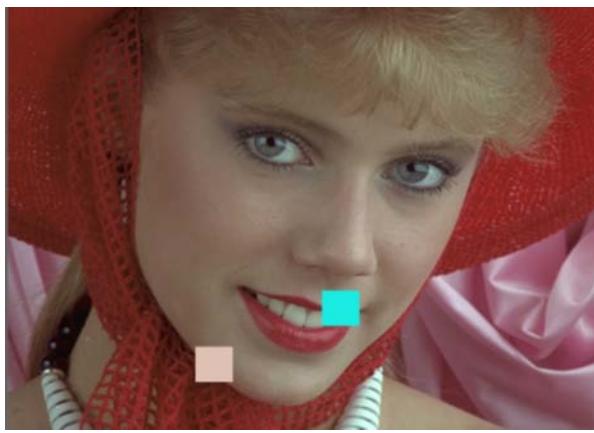
27,22 dB



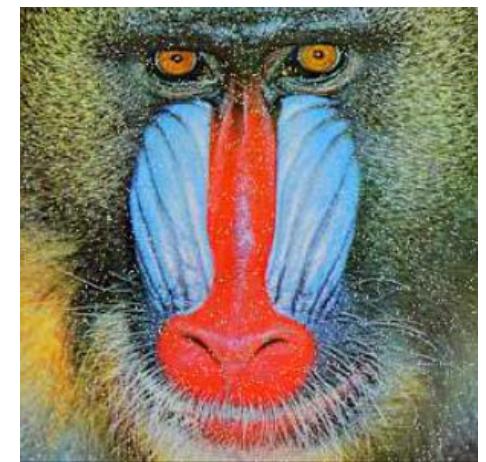
11.06 dB



26,79 dB



26,11 dB

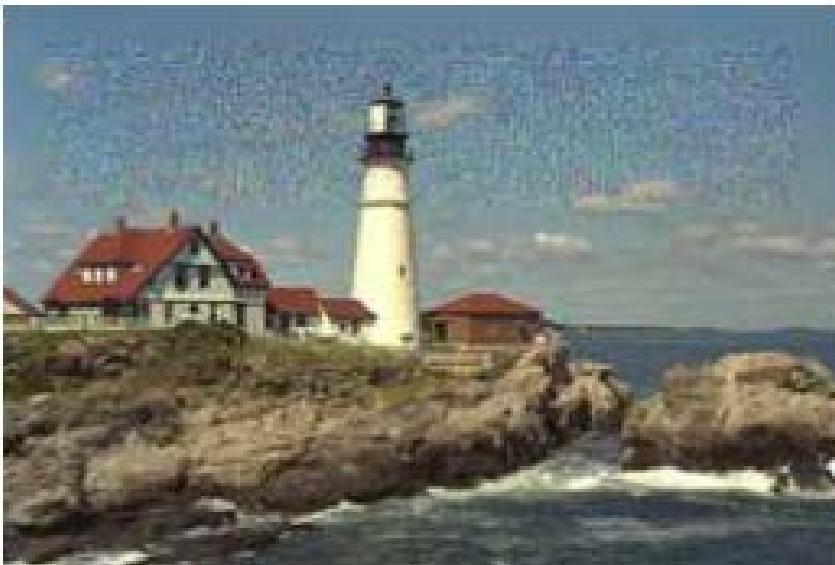


11.06 dB

Métriques avec prise en compte des caractéristiques SVH

La plupart des modèles HVS dans le traitement d'image utilisent trois propriétés fondamentales de la vision humaine:

- ❖ la sensibilité de fréquence détermine la sensibilité de l'œil à diverses fréquences spatiales,
- ❖ la sensibilité à la luminance mesure l'effet du seuil de détection du bruit Sur un fond constant,
- ❖ l'effet de masquage détermine la visibilité d'un signal en présence d'un autre signal.



Indice de similarité structurelle (SSIM)

SSIM est une mesure de similarité entre deux images numériques. Elle a été développée pour mesurer la qualité visuelle d'une image déformée, par rapport à l'image originale. L'idée de SSIM est de mesurer la similarité de structure entre les deux images, plutôt qu'une différence pixel à pixel comme le fait par exemple le PSNR. L'hypothèse sous-jacente est que l'œil humain est plus sensible aux changements dans la structure de l'image.

La métrique SSIM est calculée sur plusieurs fenêtres d'une image. On dénote x et y l'image originale et l'image déformée respectivement.

La similarité compare la luminance, le contraste et structure entre chaque couple de fenêtres.

La luminance est estimée par la mesure de l'intensité moyenne de chaque fenêtre:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i$$

N : le nombre de pixels de chaque fenêtre.

xi: l'intensité d'un pixel.

Le contraste pour chaque fenêtre est mesuré par:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}}$$

La similarité est déterminée par la perte de corrélation entre les deux fenêtres:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)$$

La fonction de comparaison de luminance notée par $I(x, y)$ est une fonction de x et y .

Le contraste entre deux fenêtres noté par $c(x, y)$, comparent les variances de x et y .

La troisième fonction compare les structures des deux fenêtres données par la fonction $s(x, y)$ en fonction de deux fenêtres normalisés.

Finalement la fonction mesurant la similarité, est une fonction de $I(x, y)$, $c(x, y)$ et de $s(x, y)$.

$$S(x, y) = f(I(x, y), c(x, y), s(x, y))$$

Wang et Al ont proposé une fonction de comparaison de la luminance de la forme :

$$l(x, y) = \frac{2\mu_x\mu_y + (K_1 L)^2}{\mu_x^2 + \mu_y^2 + (K_1 L)^2}$$

On remarque que cette équation est conforme à la loi de Weber, largement utilisée pour la modélisation de l'adaptation à la lumière dans le SVH (système visuel humain). Selon la loi de weber, la variation de luminance est proportionnelle à la luminance de fond. En d'autres termes, le SVH est sensible aux variations relatives de luminance entre les deux signaux ou encore entre les deux images.

La fonction de comparaison de contraste prend une forme similaire :

$$c(x, y) = \frac{2\sigma_x\sigma_y + (K_2L)^2}{\sigma_x^2 + \sigma_y^2 + (K_2L)^2}$$

La comparaison de structure est réalisée par la corrélation entre les deux vecteurs après soustraction de la luminance et normalisation par la variance. Notons que la corrélation entre les deux vecteurs est une mesure simple et effective de la similarité structurelle. Alors la fonction de comparaison structurelle est donnée par :

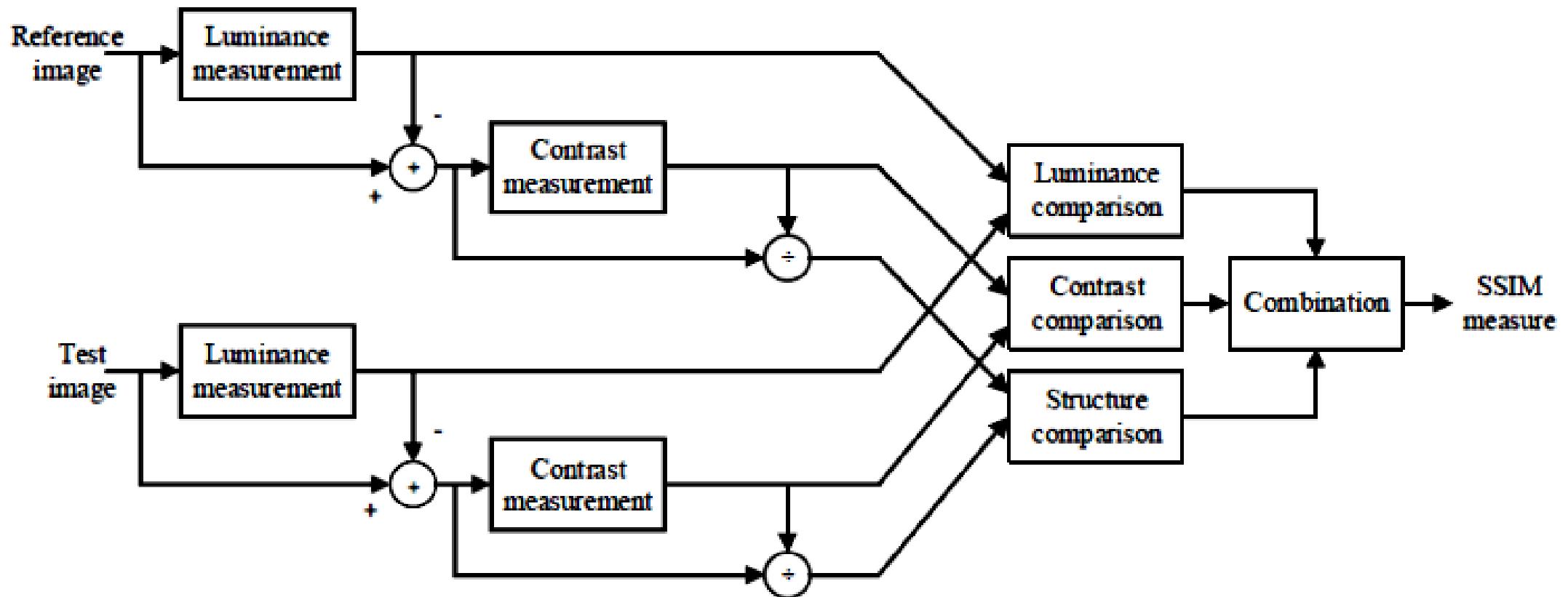
$$s(x, y) = \frac{2\sigma_{xy} + (K_2L)^2}{2\sigma_x\sigma_y + (K_2L)^2}$$

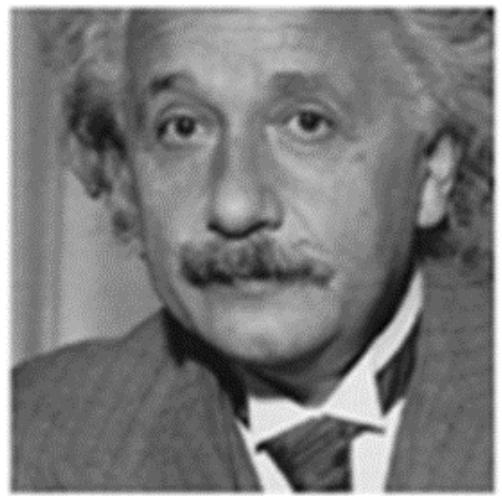
Finalement, la mesure de similarité résultante des trois comparaisons est donnée par une expression simplifiée de l'index de similarité structurelle entre x et y :

$$SSIM(x, y) = I(x, y) \cdot C(x, y) \cdot S(x, y)$$

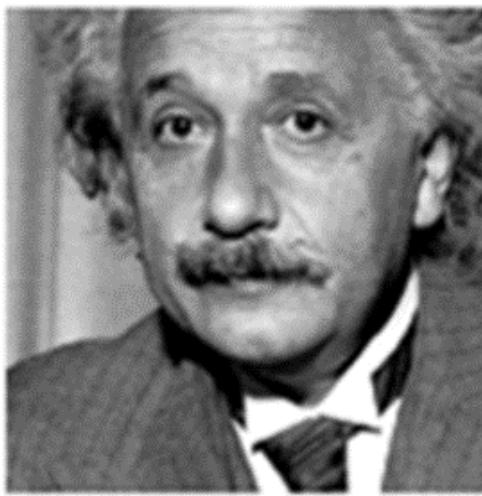
où encore sous une forme simplifiée:

$$SSIM(x, y) = \frac{2\mu_x\mu_y + (K_1L)^2}{\mu_x^2 + \mu_y^2 + (K_1L)^2} * \frac{2\sigma_{xy} + (K_2L)^2}{\sigma_x^2 + \sigma_y^2 + (K_2L)^2}$$

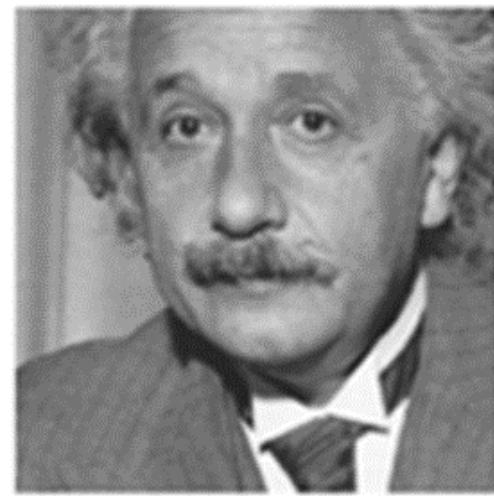




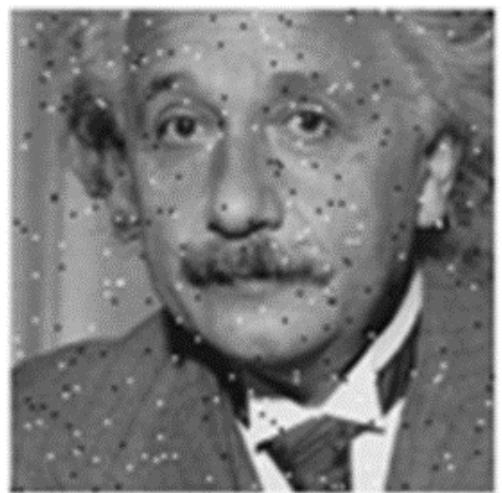
(a) MSE=0, SSIM=1
CW-SSIM=1



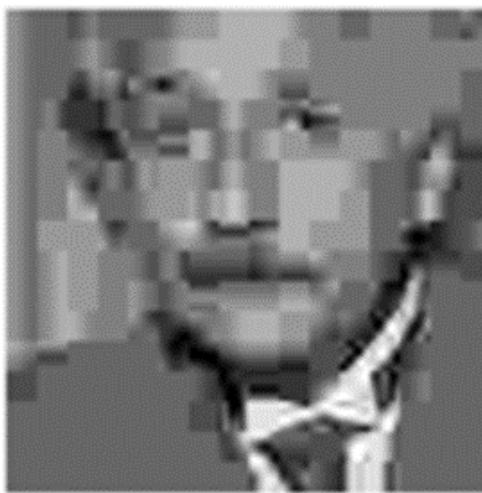
(b) MSE=306, SSIM=0.928
CW-SSIM=0.938



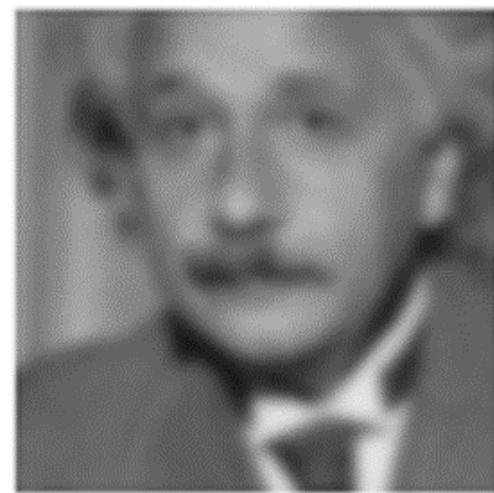
(c) MSE=309, SSIM=0.987
CW-SSIM=1.000



(e) MSE=313, SSIM=0.730
CW-SSIM=0.811



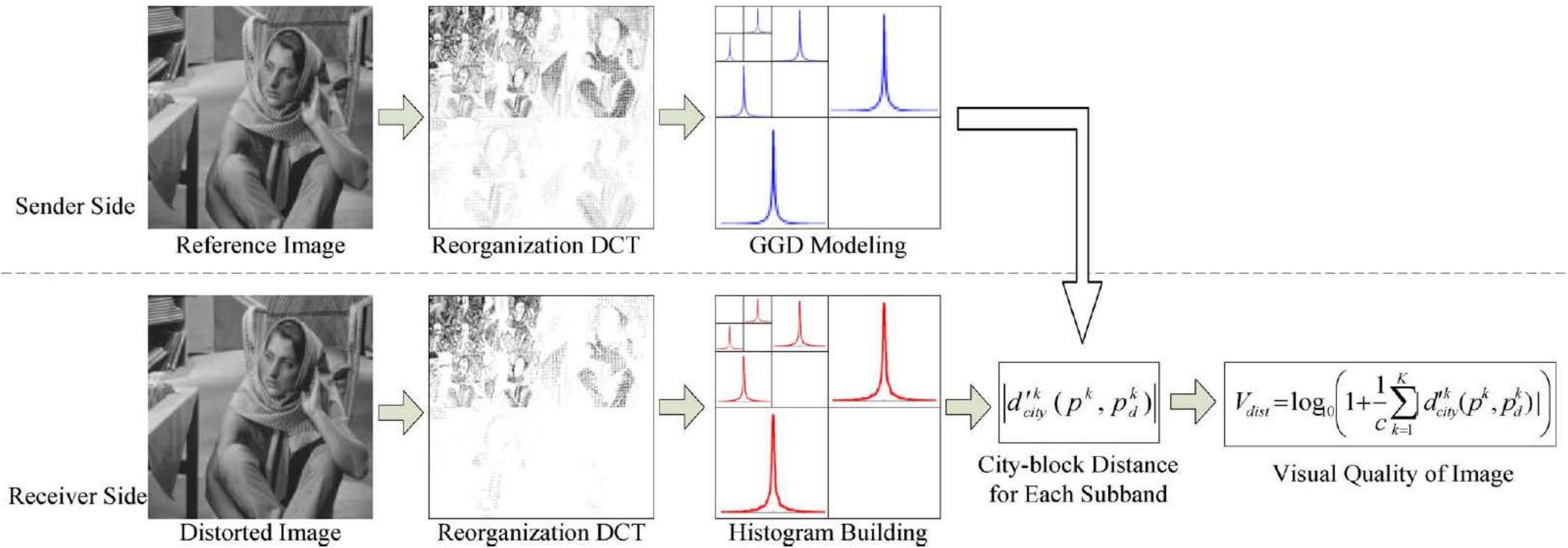
(f) MSE=309, SSIM=0.580
CW-SSIM=0.633



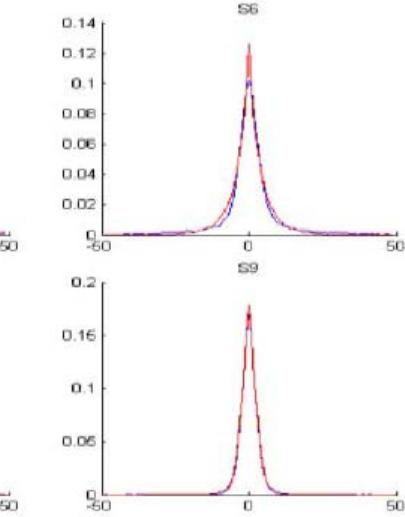
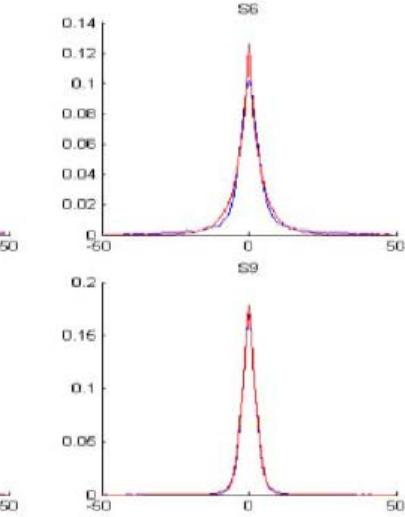
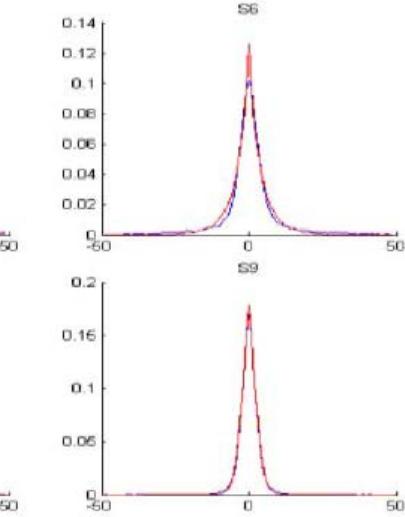
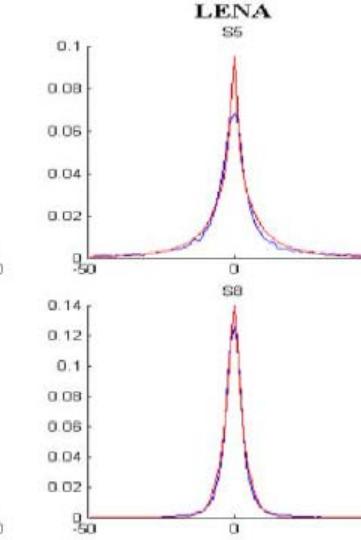
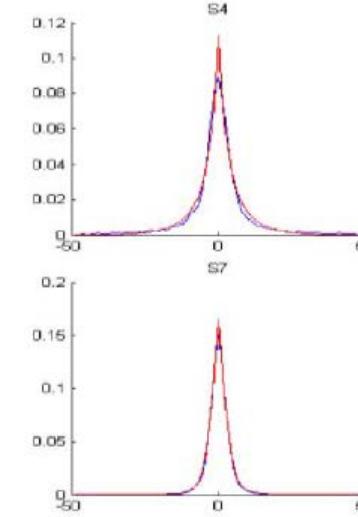
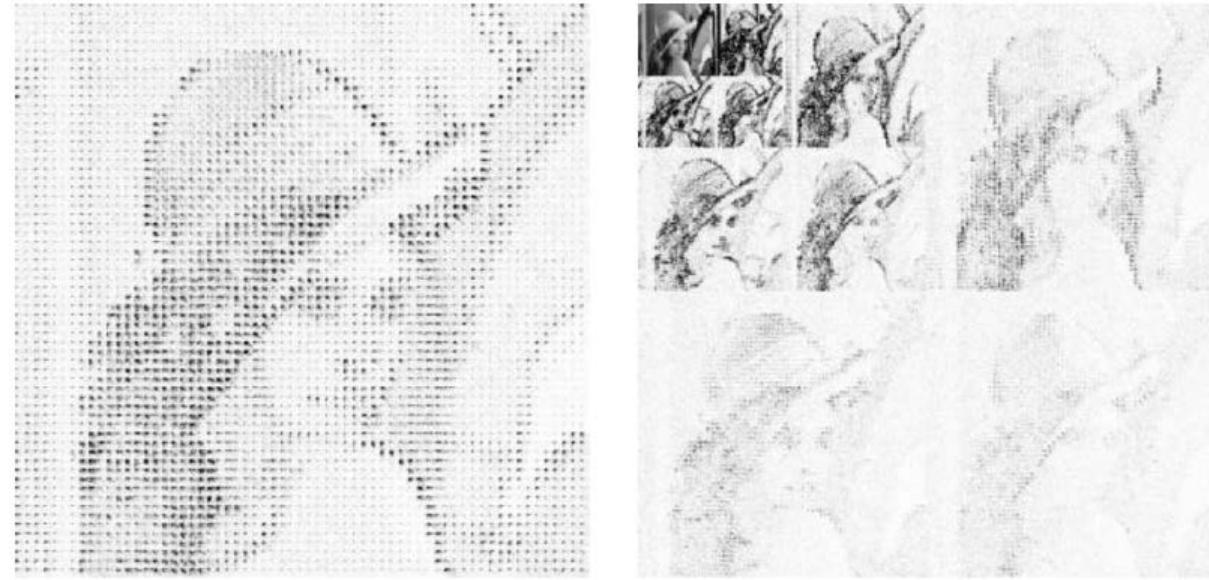
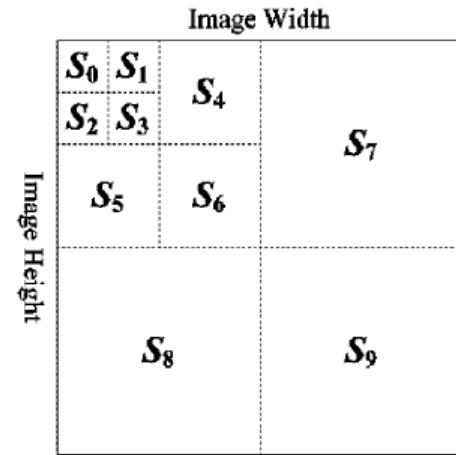
(g) MSE=308, SSIM=0.641
CW-SSIM=0.603

Métrique de qualité d'image à référence réduite

Le schéma fonctionnel de la métrique à référence réduite proposés est illustré à la figure suivante . Dans le côté de l'expéditeur, la réorganisation du DCT est d'abord appliquée à l'image de référence . Par la suite, la GGD est utilisée pour modéliser la distribution de chaque sous-bande DCT réorganisée. Du côté des récepteurs, la même réorganisation du DCT est appliquée à l'image déformée. Ensuite, l'histogramme de chaque sous-bande DCT réorganisée est construit. En se référant aux caractéristiques GGD extraites de l'image de référence, la distance city-block est utilisée pour représenter la distance de l'histogramme entre les sous-bandes du DCT organisées. Enfin, l'indice de qualité visuelle de l'image est obtenue en regroupant les distances ensemble.



0	1	4		
2	3		7	
5	6			
8		9		



$$d_{city}(r, q) = \sum_{i=1}^L |r(i) - q(i)|$$

Permet de calculer la distance city entre deux distribution (histogramme)

$$d'_{city}(p, p_d) = d_{city}(p_{\alpha, \beta}, p_d) - d_{city}(p_{\alpha, \beta}, p)$$

Calculer la distance entre chaque sous bande (image originale et dégradée), avec introduction des erreurs d'estimation

$$\begin{aligned} V_{dist} &= \log_{10} \left(1 + \frac{1}{c} \sum_{k=1}^K |d'_{city}(p^k, p_d^k)| \right) \\ &= \log_{10} \left(c + \sum_{k=1}^K |d'_{city}(p^k, p_d^k)| \right) - \log_{10}(c) \end{aligned}$$

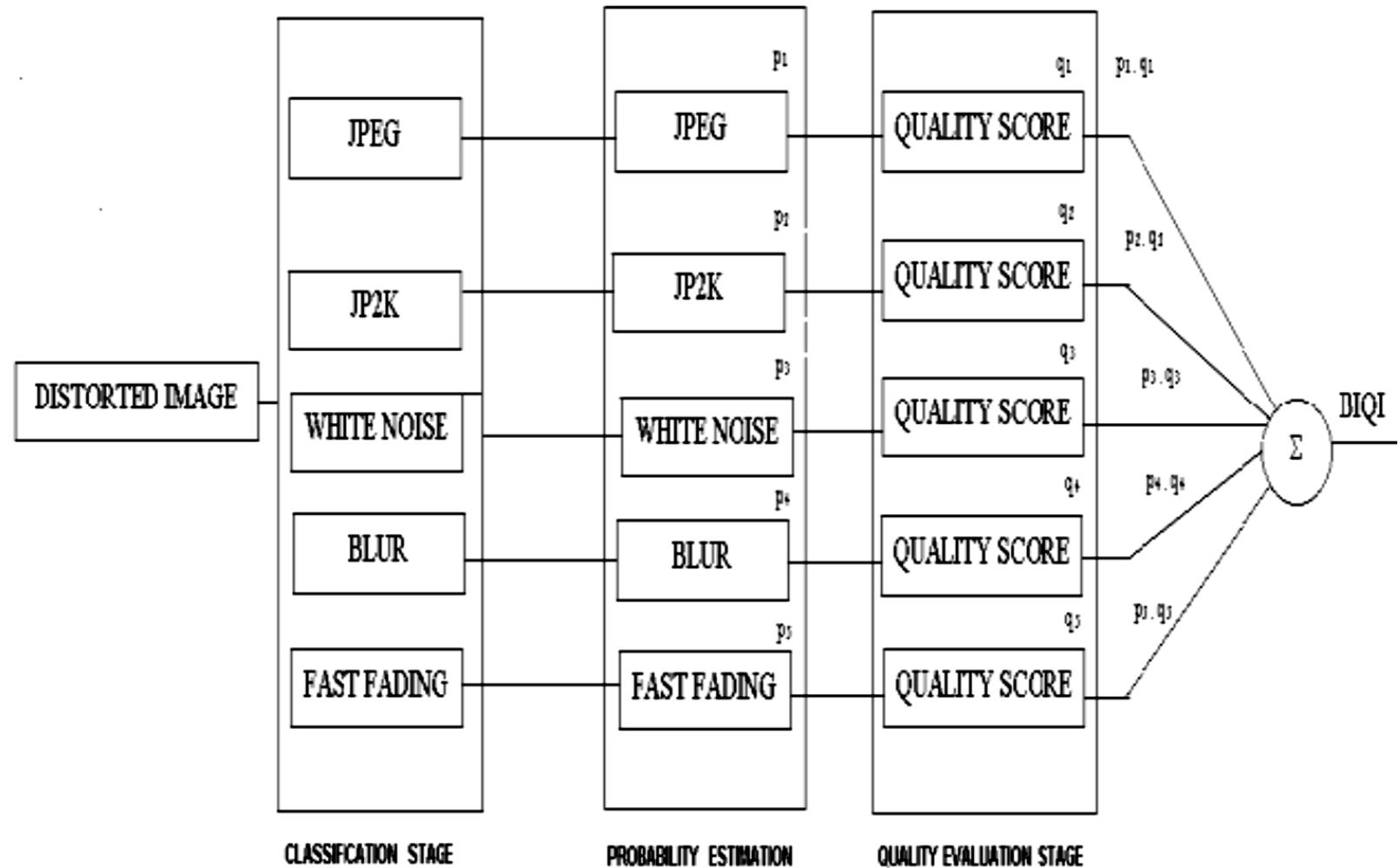
Combinaison des distances obtenues dans les six bande utilisées pour avoir une estimation globale de qualité

$$p_{\alpha, \beta}(x) = \frac{\alpha}{2\beta \Gamma(\frac{1}{\alpha})} \exp \left\{ - \left(\frac{|x|}{\beta} \right)^{\alpha} \right\}$$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$

La fonction GGD et la fonction Gamma

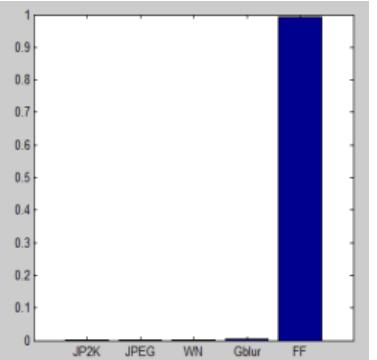
Métrique de qualité d'image sans référence



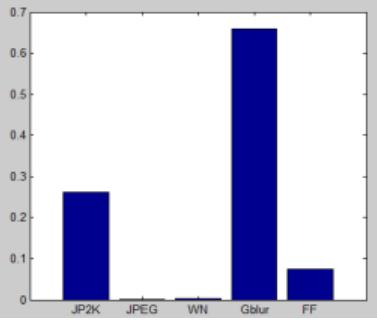
$$BIQI = \sum_{i=1}^5 p_i q_i$$



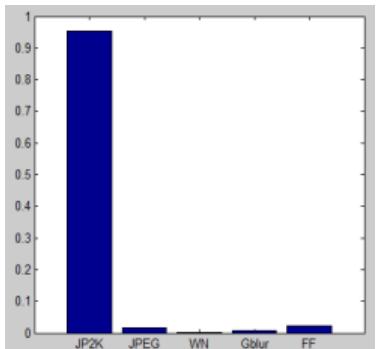
Distorted Image:FF



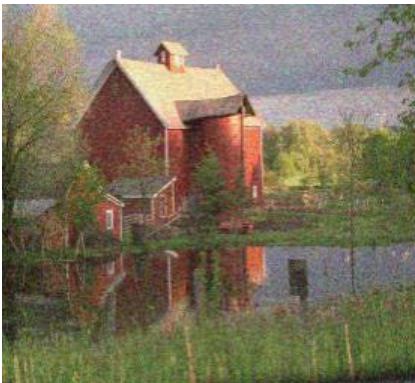
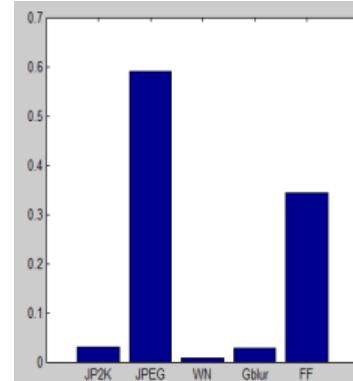
Distorted Image:Gblur



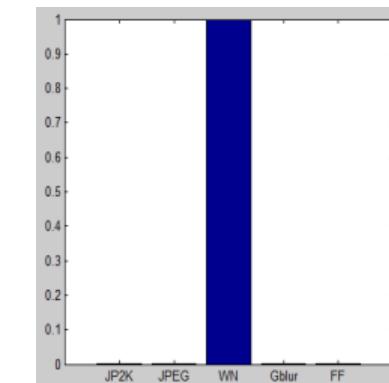
Distorted Image: JP2K



Distorted Image:JPEG



Distorted Image: WN



Évaluation des performances des métriques de qualité

La performance des mesures objectives est évaluée en ce qui concerne leur capacité d'estimer l'évaluation subjective de la qualité visuelle en trois aspects:

- **précision de la prédition (prediction accuracy)**: la capacité de prévoir les estimations de qualité subjectives avec une erreur faible,
- **monotonie de la prévision (prediction monotonicity)** : le degré auquel les prédictions du modèle sont conformes aux importances relatives d'estimations de la qualité subjective,
- **Cohérence de la prévision (prediction consistency)**: le degré auquel le modèle maintient la précision de la prédition sur toute la gamme de la séquences de test d'image, à savoir, que sa réponse est robuste par rapport à une variété de déficience d'image.

Coefficient de corrélation linéaire (précision)

Le coefficient de corrélation linéaire CC exprime la dépendance linéaire entre les mesures objectives M_{obj} et les notes subjectives moyennes MOS (Mean Opinion Score). Il est donné par :

$$CC = \frac{1}{NbJK} \frac{\sum_{j=1}^J \sum_{k=1}^{K(j)} (MOS_{jk} - \overline{MOS})(Mobj_{jk} - \overline{Mobj})}{\sqrt{\sigma_{MOS}^2 \cdot \sigma_{Mobj}^2}}$$

avec :

$$NbJK = \sum_{j=1}^J K(j)$$

le nombre d'images (ou vidéos) testées (on suppose que J versions originales ont subi chacune $K(j)$ dégradations différentes)

- $\overline{MOS} = \frac{1}{NbJK} \sum_{j=1}^J \sum_{k=1}^K (j) MOS_{jk}$ le MOS moyen
- $\overline{Mobj} = \frac{1}{NbJK} \sum_{j=1}^J \sum_{k=1}^K (j) Mobj_{jk}$ la mesure objective moyenne
- $\sigma_{MOS}^2 = \frac{1}{NbJK - 1} \sum_{j=1}^J \sum_{k=1}^K (j) (MOS_{jk} - \overline{MOS})^2$ la variance des MOS
- $\sigma_{Mobj}^2 = \frac{1}{NbJK - 1} \sum_{j=1}^J \sum_{k=1}^K (j) (Mobj_{jk} - \overline{Mobj})^2$ la variance des mesures objectives

La valeur du coefficient CC est toujours comprise entre -1 et 1. Une valeur proche de 1 ou de -1 indique une bonne corrélation.

Coefficient de corrélation de rang (monotonie)

Le coefficient de corrélation de rang CCR décrit combien les mesures objectives et subjectives de qualité, après remise en ordre propre à chaque type, suivent un ordre similaire c'est-à-dire évoluent dans le même sens. Il est donné par la relation :

$$CCR = 1 - \frac{6 \cdot \sum_{j=1}^J \sum_{k=1}^{K(j)} d_{jk}^2}{NbJK^3 - NbJK}$$

avec d_{jk} la différence de classement de l'image (ou vidéo) j ayant subi la dégradation k entre M_{obj} et MOS. Une valeur de CCR proche de 1 montre que le classement des images (ou vidéos) par le critère objectif de qualité est similaire à celui fait par les observateurs humains.

Erreur de prédiction de la qualité

Écart-type de l'erreur de prédiction de la qualité

C'est la mesure de la différence entre MOS et Mobj pour une même image (ou vidéo). L'écart-type est donné par :

$$\sigma(\epsilon_{MOS}) = \sqrt{\frac{\sum_{j=1}^J \sum_{k=1}^{K(j)} (MOS_{jk} - Mobj_{jk})^2}{NbJK - 1}}$$