

La qualité d'image

1.1 Introduction

Le traitement d'images est une discipline des mathématiques appliquées qui étudie les images numériques et leurs transformations, pour améliorer leur qualité ou d'en extraire de l'information.

L'idée de l'évaluation de la qualité d'image a commencé par l'invention des premiers instruments optiques, le télescope optique et le microscope. Ce concept appartient encore aux jours de la photographie, 1860-1930, cependant le développement de la télévision, 1935-1955, avance avec l'image numérique jusqu'à, aujourd'hui.

Dans ce chapitre nous étudions les différents types de dégradations, et nous allons parler de deux manières de mesure de la qualité des images : subjective et objective.

1.2 Image et qualité

Actuellement, l'évaluation de la qualité d'image est très demandée dans diverses applications de communications multimédias. Différents traitements de codage et de compression subis par les images, ainsi que l'impact des réseaux de communications qui sont traversés durant la transmission des données multimédias, introduisent des distorsions dans l'image et altèrent sa qualité.

La qualité d'image pour la plupart est déterminée par la résolution et un équilibre de couleurs qui se nomme habituellement contraste. La résolution se rapporte à la quantité de détails dans une image. Habituellement plus il y a de détails dans une image, plus la résolution est haute.[14]

1.3 Différents types de dégradations

L'image numérique peut subir plusieurs dégradations ou distorsions, nous allons présenter les différents types de dégradation et donner des exemples, dont les plus connues sont le bruit, l'effet de bloc, l'effet de flou et l'effet de sonnerie.

1.3.1 Effet de bloc

L'effet de bloc est la conséquence d'un découpage de l'image en blocs et leur traitement de façon indépendante. C'est généralement le principal source d'erreur lors de la compression JPEG. Il se manifeste généralement au niveau des frontières entre les blocs et apparaît comme des contours verticaux et horizontaux dont la visibilité dépend fortement de la distribution spatiale du signal image [3]. En effet, tous les blocs sont encodés indépendamment les uns des autres. Il fait créer une discontinuité facilement perceptible par l'œil humain.

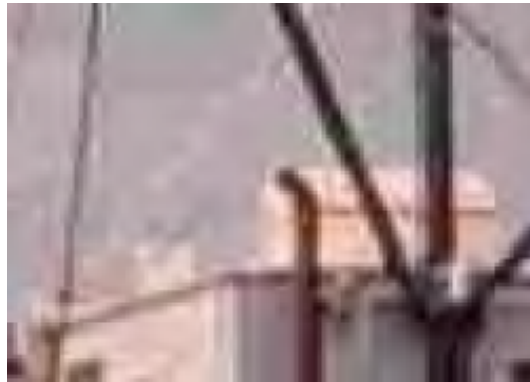


FIGURE 1.1 – Exemple d'effet de bloc

1.3.2 Effet de sonnerier "ringing"

Effet de ringing, Communément appelé 'Gibbs' est associé à l'apparition des oscillations proximités des régions à fort contraste. Cet artefact est la conséquence directe de la quantification des coefficients hautes fréquences. Sa visibilité est accrue le long des contours à fort contraste proches des zones homogènes et dépend fortement de la valeur du contraste [13].



FIGURE 1.2 – Exemple de l'effet de ringing

1.3.3 Le flou

Mauvaise quantification où les coefficients de basses fréquences sont favorisés par rapport aux coefficients en hautes fréquences, toutes les informations liées au détail sont donc

atténuées ou supprimées, même si le contenu d'image reste compréhensible mais les bords des objets ou les hautes fréquences représentant les contours sont réduites ce qui cause un effet de flou.



FIGURE 1.3 – Exemple d'un image flou

Ce type d'erreur qui apparaît principalement dans la compression JPEG et JPEG2000 d'où la perte de netteté de l'image. Il existe plusieurs types de flou dans les images, dont :

- ✓ **Le flou au niveau de l'acquisition** : Le flou peut provenir d'une erreur de manipulation du matériel d'acquisition telle qu'une mauvaise focalisation ou un bougé.
- ✓ **Le mouvement** : Un objet qui se déplace à grande vitesse peut produire un effet de flou sur l'image. Ce type de flou est aussi lié aux limitations du matériel de capture.

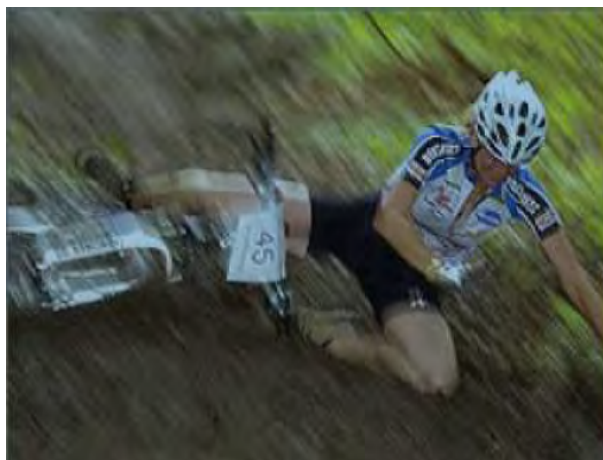


FIGURE 1.4 – Exemple de flou mouvement

- ✓ **La perte d'informations** : La perte d'information le long des canaux de transmission est aussi une cause d'apparition de flou dans l'image.



FIGURE 1.5 – Exemple de flou perte d'informations

- ✓ **Le flou volontaire** : Il peut aussi s'agir d'un flou volontairement ajouté à des fins artistiques, permettant ainsi de faire ressortir certaines zones de l'image au détriment d'autres [14].



FIGURE 1.6 – Exemple de flou volontaire

1.3.4 Bruit

On peut définir le bruit comme étant un signal qui dégrade la qualité d'une image, provoqué par une perturbation externe. dans le cas de la transmission d'une image par voie électronique(via le satellite ou la transmission sans fil, ou par le câble réseau) on peut s'attendre à des erreurs se produisant dans le signal de l'image. Ces erreurs peuvent dégrader la qualité de l'image reçue, généralement, on peut savoir les types d'erreurs à attendre, et donc le type de bruit sur l'image. On va présenter quatre types différents de bruit, et leurs effets sur la qualité d'une image [15].

- ✓ **Bruit blanc (Gaussien)** : Les principales sources de bruit gaussien dans les images numériques apparaissent lors de l'acquisition. Par exemple le bruit du capteur du à un mauvais éclairage et / ou transmission, comme le bruit de circuit électronique.
- ✓ **Bruit de speckle** : Le bruit de granularité cohérente, ou "speckle", constaté la première fois en 1960, lors d'expérience d'illumination d'objets avec une source de lumière cohérente, le laser.

- ✓ **Bruit de quantification** : Le bruit provoqué par la quantification des pixels d'une image détectée en un certain nombre de niveaux discrets est appelé bruit de quantification. Il a une distribution approximativement uniforme.
- ✓ :
- ✓ **Bruit poivre et sel** : On l'appelle aussi le bruit impulsif, le bruit de grenaille ou le bruit binaire. Cette dégradation peut être causée par de fortes perturbations apparaissant soudainement dans le signal d'image.

1.4 Evolution des techniques de mesure

Pourquoi ? Les techniques d'évaluation de la qualité d'image sont nécessaires pour :

- Pour l'optimisation, où on maximise la qualité à un coût donné (exp : Compresser une image avec un taux donné tout en conservant une certaine qualité)
- Pour l'analyse comparative entre différentes solutions de rechange (exp : un concepteur de dispositif médical de formation image peut vouloir décider entre deux dispositifs alternatifs de rayons X quel est celui qui donne les meilleurs résultats)
- La surveillance de la qualité dans des applications en temps réel (exp : la surveillance de la dégradation dans la qualité d'image envoyée dans un réseau)

1.4.1 Evaluation 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 fiable. 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'issue de ces tests, une note subjective appelée MOS (Mean Opinion Score) est obtenue.

Il existe deux modes d'évaluation subjective : l'évaluation avec référence et l'évaluation sans référence. Le premier mode, fournit aux observateurs des images aux références standards qui agissent en tant que calibrages dans le jugement subjectif.

Le second mode, les observateurs doivent juger la qualité d'image sans utiliser l'image de référence. Pendant la réalisation de ces tests, certains facteurs doivent être pris en charge affectant effectivement les jugements des observateurs. Dans ce qui suit on présente ces facteurs :

1.4.1.1 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 :

Distance d'observation La distance d'observation affecte la visibilité d'un stimulus. Cette distance allant de 4 à 6 fois la hauteur de l'image de projection et qu'il faut conserver tout au long du test.

Conditions de visualisation L'éclairage de la salle ainsi que l'environnement (couleur de fond) jouent également un rôle important dans la perception des stimuli. Un éclairage élevé risque d'éblouir l'observateur et ainsi altérer sa perception. De même, la couleur de fond de la salle a tendance à affecter l'apparence de certaines couleurs.

Écran Il est nécessaire de calibrer le support d'affichage qu'est l'écran afin qu'il fonctionne dans des conditions optimales. Sans cette calibration, les couleurs affichées peuvent être différentes du stimulus d'origine [13].

Le choix des images Un panel d'images de contenus visuels assez variés et le plus représentatif possible des scènes, liées à l'application visée, doit être utilisé. Une palette d'images avec un contenu visuel assez varié doit être utilisée, les scènes associées à l'application cible doivent être représentées autant que possible.

Facteurs psychologiques Pour minimiser les effets liés à l'état psychologique de l'observateur, il convient de débiter les tests par une séance d'initiation permettant à l'observateur de comprendre mieux le test. Il est aussi important d'expliquer clairement l'objectif du test et le protocole à respecter sans pour autant influencer le jugement des participants. De plus, il est préférable de mener de nombreux tests de courte durée (environ 10min) et variés plutôt que de contraindre l'observateur à des séances longues (environ 30min) et fatigantes qui risquent de l'amener à un jugement rapide et aberrant.

Les observateurs Le choix des observateurs doit être établi en fonction de l'application visée par les tests et le plus représentatif possible de la population moyenne concernée. L'observateur doit avoir ainsi une acuité visuelle de 10/10 pour les deux yeux avec ou sans correction. Un nombre suffisant de participants doivent participer aux tests. Il est généralement admis que les tests sont statistiquement valides si le nombre de participants est supérieur ou égal à 16.[13]

Ainsi, pour minimiser les erreurs de jugements, tous ces paramètres doivent être pris en considération et notés soigneusement. Notons que l'évaluation subjective de la qualité d'image dépend d'autres facteurs encore mal maîtrisés tel que l'humeur, l'âge, la culture, le niveau intellectuel. Ici, nous nous sommes limités aux facteurs les plus influents [17].

1.4.1.2 Protocoles suivis

L'union internationale des télécommunications (ITU) propose essentiellement trois grandes familles communes d'évaluation subjective [18] :

- Tests à simple stimulus : l'observateur doit évaluer la qualité d'une image sans aucune information relative à l'image originale (sans référence).
- Tests à doubles stimulus : l'observateur doit évaluer le degré de similarité entre deux images (avec référence)
- Tests comparatifs : l'observateur compare deux ou plusieurs stimuli.

Simple stimulus La méthode à simple stimulus, appelée Single Stimulus Continuous Quality Evaluation, SSCQE, Ce protocole de test se présente comme un test où les différentes images à évaluer sont montrées les unes après les autres avec un temps mort entre chaque image permettant d'évaluer la qualité de l'image qui vient d'être montrée.

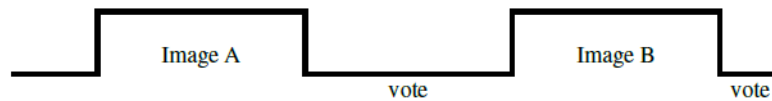


FIGURE 1.7 – Méthodes à simple stimulus

L'observateur visualise les images sans aucune information relative à l'image originale. De ce fait, un seul stimulus à la fois est jugé dans ce type de protocole, d'où son nom «protocole à simple stimulus».

Excellente
Bonne
Assez bonne
Médiocre
Mauvaise

TABLE 1.1 – Echelle d'évaluation de qualité d'image.

On peut aussi utiliser une échelle de qualité d'image comportant 5, 6, 7 ou 100 niveaux de qualité d'image.

double stimulus La méthode à 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.

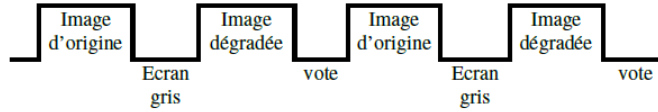


FIGURE 1.8 – Méthodes à double stimulus

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.

Ce protocole utilise également différentes échelles d'évaluation où chacune correspond à un degré de dégradation par rapport à l'image originale. La table représente une échelle à 5 niveaux avec les dégradations correspondant à chaque niveau.

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

TABLE 1.2 – Échelle représentative d'un protocole double stimulus.

Comparatif L'objectif des méthodes comparatives est de comparer deux ou plusieurs images selon leur qualité. La différence majeure réside dans le fait que l'on souhaite ici quantifier la perception des différences entre des stimuli tableau suivant :

Note	Qualité
5	Moins bonne
4	Identique
3	Légèrement mieux
2	Mieux
1	Beaucoup mieux

TABLE 1.3 – Échelle représentative d'un protocole comparatif.

1.4.1.3 MOS "Mean Opinion Score"

Une fois tous les tests subjectifs effectués, une note finale par image appelée MOS (Mean Opinion Score) est obtenue en combinant les différentes notes données par les divers observateurs. Cette note représente la moyenne des scores donnés par les observateurs pour chaque image. Elle est donnée par la relation suivante :

$$MOS(i) = \frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} Note_i(j) \quad (1.1)$$

où N_{obs} est le nombre total de participants et $Note_i(j)$ la note affectée à l'image i par l'observateur j .

Afin de prendre en compte d'éventuelles erreurs dues aux estimations des observateurs, un intervalle de confiance d'environ 95% est généralement utilisé autour du MOS. Cet intervalle est donné par :

$$[MOS(i) - e_j, MOS(i) + e_j] \quad (1.2)$$

où

$$e_j = 1.96 \cdot \sigma_j \quad (1.3)$$

et

$$\sigma_j = \sqrt{\frac{1}{N_{obs} - 1} \sum_{\omega=1}^{N_{obs}} (Note_i(k) - MOS(k))^2} \quad (1.4)$$

Une analyse plus approfondie permet de détecter et de rejeter d'éventuelles erreurs. dans l'évaluation subjective il y a plusieurs obstacles. Par exemple, le nombre d'observateurs est très coûteux financièrement (pour assurer la confiance statistique le nombre d'observateurs doit être plus grand) donc le temps d'exécution sera augmenté sans oublier que les résultats influencés par les conditions expérimentales et la relation entre le jugement de l'humain et ses croyances.

1.4.2 Evaluation objective :

L'évaluation objective contrairement à l'évaluation subjective est reproductible et ne nécessite pas la présence de human. Cette technique d'évaluation permet de mesurer automatiquement la qualité des images. Elle permet de calculer à l'aide d'un algorithme mathématique une grandeur spécifiant la qualité globale ou locale de l'image. Les recherches dans le domaine de l'évaluation de la qualité de l'image augmentent ces dernières années, en particulier dans le domaine de la mesure de qualité objective. Différents types de critères ont été définis à cet effet et de nouvelles idées sont proposées pour des mesures objectives. [21] [13]

Les méthodes d'évaluation objective de la qualité peuvent être classées en trois catégories :

- ✓ **Les méthodes avec référence** : l'algorithme a un accès à une version parfaite de l'image avec laquelle il peut comparer la version dégradée (dégradée).
- ✓ **Les méthodes avec référence réduite (reduced reference)** : l'image de référence est disponible partiellement (une solution intermédiaire).
- ✓ **Les méthodes sans référence** : l'algorithme n'a accès qu'à l'image dégradée et doit estimer la qualité de l'image sans connaissance de la version parfaite.

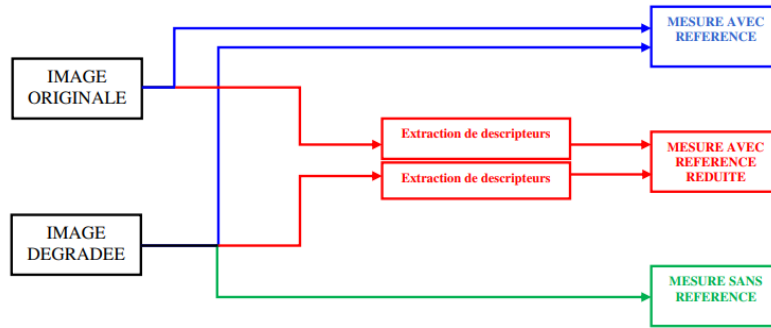


FIGURE 1.9 – Qualité d'image objective : les différentes approches

1.4.2.1 Evaluation d'images fixes avec référence complète

dans les quelles l'algorithme a un accès à une version parfaite de l'image avec les quelles il peut comparer la version dégradée. La version parfaite vient généralement d'un dispositif de haute qualité d'acquisition, ensuite elle est dégradée par la compression et les erreurs de transmission.

Approches basée sur l'erreur quadratique moyenne (MSE) : Pour déterminer le rapport ressemblance il faut faire une comparaison entre l'image dégradée I et l'image parfaite I' par la mesure de l'erreur quadratique moyenne entre les pixels de ces deux images [19] :

Cette mesure n'est rien d'autre que la moyenne quadratique du signal erreur ou distorsion. Elle est donnée par :

$$MSE = \frac{1}{M \times N} \sum_{k=m-1}^M \sum_{k=n-1}^N (I'(m, n) - I(m, n)) \quad (1.5)$$

avec $(M \times N)$ c'est la taille de l'image, et $I(m, n)$ et $I'(m, n)$ sont respectivement les amplitudes de pixels sur les images parfaite et dégradée. Il est vraisemblable que l'œil tient beaucoup plus compte des erreurs à grandes amplitudes, ce qui favorise la mesure quadratique [14].

Le rapport crête signal sur bruit (PSNR) Cette mesure permet de quantifier la fidélité qui existe entre deux images [19] (le rapport entre la puissance maximale possible d'un signal (amplitude de pixel) et la puissance du bruit), elle est une fonction de MSE :

$$PSNR = 10 \log_{10} \left(\frac{I_{max}^2}{MSE} \right) \quad (1.6)$$

Pour les images couleur il faut calculer le PSNR sur chacun des trois plans colorimétriques puis faire la moyenne, dans une image à niveau de gris, I_{max} désigne la luminance maximale possible. Une valeur de PSNR infinie correspond à une image non dégradée (MSE converge vers le zéro), et cette valeur décroît en fonction de la dégradation. Le PSNR lie donc le MSE à l'énergie maximale de l'image.

Le rapport signal sur bruit (SNR) : Cette mesure lie le MSE à l'énergie moyenne de l'image :

$$SNR = 10 \log_{10} \left(\frac{\frac{1}{N} \sum I^2}{MSE} \right) \quad (1.7)$$

N : la taille de l'image

I : les amplitudes de pixels .

Approches Structurelles

L'index de similarité structurelle (SSIM) le système visuel humain est fortement influencé par les structures présentes dans la scène (figure 1.13) donc l'idée est de mesurer la similarité structurelle entre les deux images (parfaite et dégradée) [20] :

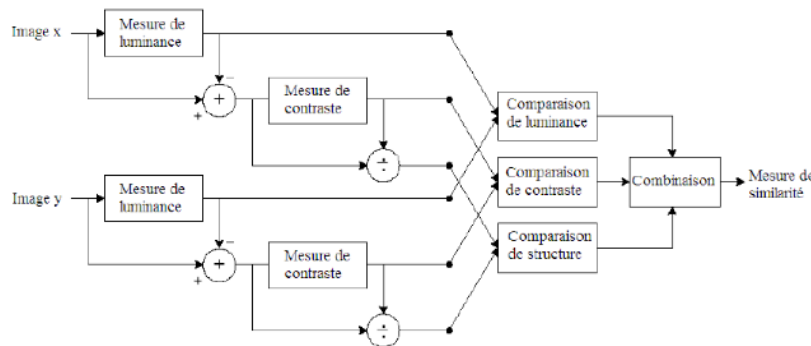


FIGURE 1.10 – Diagramme du système de mesure de similarité structurelle (SSIM)

La version de base de SSIM est décrite, où l'information structurelle est recueillie par une comparaison de la luminance, le contraste et la structure. soit x et y Deux signaux d'image non négatifs de longueur N .

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 :

$$\mu_x = \frac{1}{N} \sum_1^N x_i \quad (1.8)$$

N : le nombre de pixels de chaque fenêtre.

x_i : l'intensité d'un pixel.

Le contraste pour chaque fenêtre est mesuré par :

$$\sigma_x = \left(\frac{1}{N-1} \sum_1^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}} \quad (1.9)$$

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

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (1.10)$$

La fonction de comparaison de luminance notée par $l(x, y)$ 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 $l(x, y)$, $c(x, y)$ et de $s(x, y)$. Elle est de la forme :

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (1.11)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1.12)$$

où μ_x (resp. μ_y) est l'intensité moyenne de X (resp. Y), σ_x (resp. σ_y) est l'écart-type des intensités de X (resp. Y) et σ_{xy} est la covariance entre les intensités de X et Y. $C_1 = (k_1d)^2$, $C_2 = (k_2d)^2$ sont deux petites constantes positives nécessaires pour stabiliser la division. d est l'étendue des intensités et $k_1 = 0.01$, $k_2 = 0.03$ par défaut.

Complexe Wavelet SSIM (CW-SSIM) L'indice CW-SSIM est une extension de la méthode SSIM au domaine des ondelettes complexes. Le principal inconvénient de l'algorithme SSIM du domaine spatial est qu'il est très sensible à la translation, à la mise à l'échelle

et à la rotation des images. C'était la motivation initiale pour étendre l'idée de SSIM., proposé dans [1] pour mesurer la qualité des images qui ont une distorsion géométrique telle que la rotation, le redimensionnement. Comme le SSIM, la valeur maximale de CW-SSIM est 1 qui indique que les deux images sont parfaitement identiques.

L'objectif est de concevoir une mesure insensible aux distorsions géométriques "non structurées" de l'image qui sont généralement causées par des facteurs de nuisance, tels que le mouvement relatif du dispositif d'acquisition d'image, plutôt que par un changement réel de la structure des objets. Nous considérons des ondelettes complexes symétriques dont les "ondelettes mères" peuvent être écrites comme une modulation d'un filtre passe-bas $\omega(\mu) = g(\mu)e^{j\omega\mu}$, où $\omega(\mu)$ se trouve la fréquence centrale de la modulé, et est un filtre passe-bas, et $g(\mu)$ est une fonction symétrique qui varie lentement. La famille des ondelettes est constituée de versions dilatées/contractées et traduites de l'ondelette mère

$$\omega_{s,p}(\mu) = \frac{1}{\sqrt{s}} \omega\left(\frac{\mu-p}{s}\right) = \frac{1}{\sqrt{s}} g\left(\frac{\mu-p}{s}\right) e^{j\omega s(\mu-p)/s} \quad (1.13)$$

$$CW - SSIM(C_x, C_y) = \left(\frac{2 \sum_{i=1}^N |C_{x,i}| |C_{y,i}| + K}{\sum_{i=1}^N |C_{x,i}|^2 + |C_{y,i}|^2 + K} \right) \left(\frac{2 \sum_{i=1}^N |C_{x,i} C_{y,i}^*| + K}{2 \sum_{i=1}^N |C_{x,i} C_{y,i}^*| + K} \right) \quad (1.14)$$

Où C_x est la transformation en ondelettes complexe du signal x et C_y est la transformation en ondelettes complexe du signal y . De plus, K est un petit nombre positif utilisé aux fins de la stabilité de la fonction.

RMSE Est utilisé pour mesurer la déviation sage de pixel entre deux entités, décrite comme suit :

$$RSME = \left[\sum_{n=1}^N \sum_{m=1}^M \frac{I(m,n) - I_n(m,n)}{MN} \right]^{\frac{1}{2}} \quad (1.15)$$

Où I et I_N sont l'image original et l'image bruité et $M \times N$ est la taille d'image.

1.4.2.2 Evaluation d'images fixes avec référence réduite

Les méthodes avec référence réduit est une mixtes entre les méthodes avec référence complete et les méthodes sans référence. Seules des caractéristiques de l'image de référence sont envoyées au récepteur comme montre la figure 1.14

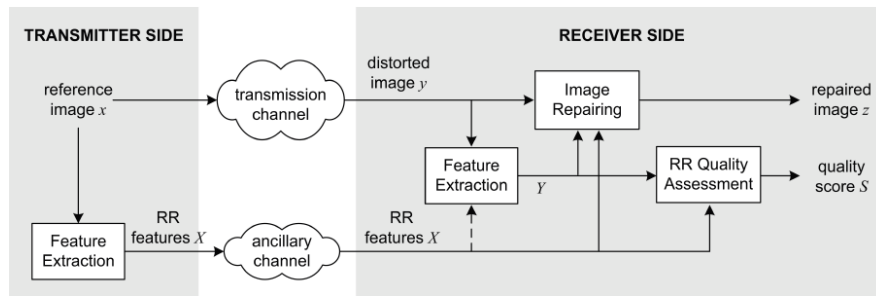


FIGURE 1.11 – Structure d'évaluation de qualité avec référence réduite.

A cause de facteur de temps et le manque des données parfait, on a besoin de développer des méthodes basées sur l'extraction partielle de l'information (Seules des caractéristiques de l'image référence sont envoyées au récepteur).

1.4.2.2.1 L'extraction des caractéristiques

la selection aléatoire : l'extraction des caractéristique est tout simplement aléatoire, le procédé choisit un nombre fini des pixels au hasard ,après l'extraction des vecteurs et la compression d'image référence il remplace les pixels non-selections par des zéros dans les deux images(référence et dégrader),et enfin c'est l'application d'une métrique connue entre les deux images

Le DCT : c'est évident que la selection aléatoires est très loin d'être idéal donc pour l'améliorer ce procédé transforme l'image vers le DCT, et avec la prédiction des régions les plus probables des déficiences on peut limiter les effets de la election aleatoire

1.4.2.3 Evaluation d'images fixes sans référence

L'évaluation de la qualité d'images sans référence (SR) est le type le plus important dans l'évaluation de la qualité parce qu'il est l'exemple idéal du processus de jugement humain. Les méthodes sans référence visent à atteindre les objectifs en utilisant uniquement des images de test pour estimer la qualité de l'image perceptuelle.

Ces méthodes sont nettement plus complexes que les méthodes avec référence et avec référence réduit en raison de l'absence d'informations sur les images source.

Certaines des premiers algorithmes d'évaluation sans référence ont été conçus pour un spécifique type d'application, par exemple, dans une application permettant de mesurer les pertes de qualité dans les images compressées. Les approches d'évaluation sans référence existantes pourraient être subdivisées en deux grandes catégories.

Approches basées sur les statistiques de scènes naturelles (NSS) L'idée principale de ces approches est que les images de scènes naturelles présentent des régularités statistiques affectés par la présence de distorsions, et ces changements statistiques dans la distorsion

des images peuvent être mesurées pour évaluer la qualité de l'image. Parmi les méthodes proposées :

BIQI Cette approche a été présentée par Moorthy et Bovik [23], ils estiment la qualité à partir d'une transformée en ondelettes utilisant la base d'ondelettes de Daubechies 9/7 [16]. La transformation est effectuée sur trois échelles et trois orientations. Le coefficient de sous-bande de la transformation est paramétré à l'aide d'une distribution gaussienne généralisée (DGG).

Il y a trois paramètres de la distribution : la moyenne, la variance et le paramètre de forme. Mais deux paramètres seulement sont utilisés qui sont la variance et le paramètre de forme car la moyenne dans la base des ondelettes est égale à zéro. Par conséquent, trois échelles, trois orientations et deux paramètres génèrent 18 caractéristiques pour l'image ($3 \times 3 \times 2$). Après avoir extrait les caractéristiques, l'image testée est classée par un classificateur dans une des cinq classes représentant cinq distorsions différentes (jpeg, jpeg2000, bruit blanc, erreur de lecture rapide et ou). Ce classificateur ne donne pas seulement une classification absolue, mais aussi en donnant le montant de chaque distorsion présente dans l'image.

Approches basées sur l'apprentissage Dans ces approches les caractéristiques sont dérivées pendant le processus d'apprentissage lui-même. Ces approches extraient d'abord les caractéristiques des données d'entrée par le biais d'apprentissage, qui sont ensuite mises en commun pour produire une visibilité globale de la distorsion, puis converties en un score de qualité perceptuelle par ajustement à un modèle de régression. Les approches basées sur l'apprentissage des caractéristiques sont plus efficaces que les approches basées sur NSS en raison de leurs capacités d'apprendre automatiquement de meilleures fonctionnalités à partir de pixels d'image brutes. Parmi les méthodes proposées :

Les auteurs de [24] ont proposé une technique d'évaluation de qualité d'image basée sur un réseau neuronal convolutif (CNN) qui peut prédire avec précision la qualité des distorsions images par rapport à la perception humaine. Cette technique apprend et prédit la qualité de l'image sur des régions locales au lieu de l'image entière. La première étape dans la technique proposée consiste à effectuer une normalisation de contraste de l'image en niveau de gris dont le but est d'avoir des patches d'image non chevauchantes. Ensuite, le CNN estime un score de qualité pour chaque patch et le score

nale sera obtenu en faisant une moyenne de tous les scores estimés.

Cornia Dans [25] l'auteur présenté un objectif général efficace évaluation de la qualité d'image (IQA) sans référence (NR) cadre basé sur l'apprentissage des fonctionnalités non supervisé. est une nouvelle stratégie basé sur l'utilisation des morceaux d'image brute comme information caractéristique dans la phase d'apprentissage au lieu des caractéristiques basées sur la connaissance préalable des différences entre les images altérées et les images originales.

Les approches précédentes pour ce problème reposent généralement sur caractéristiques artisanales soigneusement conçues sur la base des connaissances antérieures. La méthode CORNIA comprend les étapes suivantes : extraction locale des caractéristiques, construction de livre de code (codebook), codage par affectation souple, regroupement maximal et régression linéaire.

1.4.2.4 Paramètres d'analyse comparative

Les différentes approches d'évaluation de qualité sans référence fournissent différents scores de qualité. Donc, pour comparer la performance de ces approches il faut avoir un système de comparaison, les paramètres utilisés pour faire la comparaison sont cités ci-dessous :

PCC est utilisé pour mesurer la dépendance entre les variables. Sa valeur est entre $(-1, +1)$ où la valeur proche de $+1$ indique que les deux variables ont une corrélation positive et la valeur proche de -1 indique que les deux variables ont une corrélation négative. Une très faible ou un zéro indique que les deux variables ne sont pas corrélées. L'équation de PCC entre X , Y est décrite comme suit :

$$\rho = \frac{\text{Covariance}(X, Y)}{\sigma_x \sigma_y} \quad (1.16)$$

où X est l'évaluation objective et Y est l'évaluation subjective

Le PCC est la covariance entre le score SSIM1 et SSIM2, rapportée au produit de leurs écart-types

SCC Fournit la relation entre deux variables classées. Sa valeur est entre $(-1, +1)$ avec la même interprétation que celle de PCC. Le SCC est calculé comme suit :

$$\rho = 1 - \frac{6 * \sum d^2}{n(n^2 - 1)} \quad (1.17)$$

Où d est la différence dans les rangs de deux variables X et Y , $d_i = x_i - y_i$. Et n est le nombre de paires de valeurs.

Outlier's Ratio Est défini comme le pourcentage de la sortie de l'algorithme qui dépasse le double de l'écart type des scores subjectifs. S'il y a des images " I " et que " S_i " est le score de qualité subjective de cette image, le score subjectif moyen (S_m) est calculé comme suit :

$$S_m = \frac{1}{I} \sum_{i=1}^I S_i \quad (1.18)$$

L'écart type des scores subjectifs (σ_s) est calculé comme suit :

$$\sigma_s = \left[\frac{1}{2} \sum_{i=1}^I (S_i - S_m)^2 \right]^{\frac{1}{2}} \quad (1.19)$$

Supposons qu'il existe des images "P" avec un score de qualité objective individuel (O_i) tel que :

$$|O_i - S_i| > 2 * \sigma_s \quad (1.20)$$

Alors le ratio des valeurs aberrantes est :

$$\text{Outlier'sRatio} = \frac{P}{I} \quad (1.21)$$

Le ratio augmente si le résultat obtenu par l'algorithme est totalement différent du résultat obtenu par l'évaluation subjective.

1.4.3 Mesures inspirées du Système Visuel Humain

Dans cette partie, on va voir, quelques approches de système visuel humain (SVH). Différentes caractéristiques du SVH ont été identifiées et utilisées dans le développement de nombreux algorithmes de traitement d'images et de vidéos et d'évaluation de la qualité.

Avec des observateurs humains comme "récepteur" final de l'image, il était important pour comprendre le processus fonctionnel du système visuel humain (SVH). Malheureusement, comprendre la façon dont l'information est recueillie par l'œil humain est pas suffisante, car une grande partie de la perception visuelle se fait dans le cerveau. Définir la manière Les informations saisies sont traitées et jouent un rôle majeur dans le développement le traitement des images et des vidéos et les algorithmes de prédiction de la qualité. Les tests des propriétés du HVS ont été effectués par le biais d'expériences physiologiques et psychophysiques

1.4.3.1 Caractéristiques du Système Visuel Humain

Fovéation Elles atteignent leur maximum à la fovéa et diminuent rapidement avec la distance à la fovéa. Un résultat normal est que toutes les fois qu'un observateur humain fixe un point dans son environnement, la région autour du point fixé est résolue avec la résolution spatiale la plus élevée, alors que la résolution diminue loin de la distance du point de fixation.

l'adaptation à la lumière SVH fonctionne sur un éventail de valeurs d'intensité de la lumière, enjambant plusieurs ordres de grandeur à partir d'une nuit éclairée par la lune à un jour ensoleillé lumineux.

Fonctions de sensibilité de contraste La fonction de sensibilité de contraste (CSF) modélise la variation de la sensibilité du SVH aux différentes fréquences spatiales et temporelles qui sont présentes dans le stimulus visuel, c'est-à-dire déterminer l'existence de « bords » entre surfaces homogènes.

L'effet de masquage dans SVH Le masquage peut être défini comme l'interaction entre différents stimuli et l'effet d'un stimulus sur la perception d'un autre dans la même scène visuelle. Des études menées au SVH indiquent que les deux stimuli sont vus, tant qu'ils dépassent tous deux leur seuil de détection. Cependant, un stimulus (masqueur) peut modifier le seuil de détection d'un autre stimulus (masqué), ce qui rend sa détection plus difficile pour la SVH.

1.4.3.2 Méthodes basées sur SVH

Algorithme de Daly Il vise à calculer une probabilité de la carte de détection entre l'image référence et l'image déformée [43].

Algorithme de Lubin Il estime une probabilité de détection des différences entre les versions déformée et originale [44].

Métrique de Watson Elle est basée sur la transformée DCT (Transformée en Cosinus Discrète). Elle est utilisée généralement dans la compression d'image [45].

1.5 Conclusion

Dans ce chapitre, nous avons présenté des définitions de base de la qualité d'image, des différents types de dégradation pouvant atteindre une image comme l'effet de blocs qui se manifestant généralement au niveau des frontibres entre les blocs. Nous avons aussi parlé de d'autres types de dégradations tel que le bruit, le flou etc.

Ensuite nous avons présenté les différentes méthodes d'évaluation de la qualité d'image qui sont divisées en deux familles : méthodes subjectives et méthodes objectives

En premier lieu la notion de qualité et de fidélité d'images. L'évaluation subjective a ensuite été présentée. Les différents protocoles y ont été décrits ainsi que les principaux facteurs influençant l'appréciation humaine. Le MOS (Mean Opinion Score) obtenu à partir de la compilation des notes subjectives a ensuite été défini.

Ainsi, la notion d'évaluation objective a été introduite. Nous avons dans un premier temps défini les différentes catégories de mesures de qualité, à savoir les mesures avec référence, avec référence réduite et sans référence, en soulignant leurs différences et leurs approches.

Dans le chapitre suivant, nous allons présenter la méthode de transformation ondelettes en arbre-doublé (dual-tree complex wavelet transform).