

**République Algérienne Démocratique et Populaire**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**



**Université de Mentouri – Constantine**  
**Faculté des sciences de l'ingénieur**  
**Département d'informatique**

Année : 2012  
N° d'ordre : 143/TS/2012  
N° de série : 09/inf/2012

# THESE

Pour l'obtention du diplôme de Docteur en Sciences  
Spécialité : Informatique

Titre :

## **Evaluation de la qualité d'image**

**Présentée par : AHMED SEGHIR Zianou**

Soutenue le 24/06/2012

Devant le jury composé de :

<b>Président : Z. BOUFAÏDA</b>	Professeur Université Mentouri de Constantine
<b>Rapporteur : F. HACHOUF</b>	Professeur Université Mentouri de Constantine
<b>Examineur : A. BELHAJD AISSA</b>	Professeur USTHB, Alger
<b>Examineur : MC. BATOUCHE</b>	Professeur Université Mentouri de Constantine
<b>Examineur : A.MOUSSAOUI</b>	Maître conférence Université Farhat Abès de Sétif
<b>Invité : S.CHIKHI</b>	Professeur Université Mentouri de Constantine

# Remerciements

Je remercie tout d'abord le bon dieu pour m'avoir donné le courage et la santé pour accomplir ce travail.

Mes vifs remerciements accompagnés de toute ma gratitude vont ensuite à mon promoteur **HACHOUF FELLA** Professeur à l'université de Constantine, qui a dirigé cette thèse, qui m'a amicalement conseillé tout au long de ce travail et qui m'a proposé un cadre de travail très favorable. Je la remercie d'avoir consacré beaucoup de son temps pour les nombreuses relectures de mon document et les articles publiés. Ses remarques m'ont permis de faire progresser ce travail.

Je remercie la présidente de jury madame **Z. BOUFAÏDA**, professeur à l'université de Constantine qui nous a fait l'honneur de présider le jury.

Je remercie également madame **A. BELHADJ AISSA**, professeur à USTHB, Alger pour m'avoir fait l'honneur d'évaluer mon travail.

Je remercie aussi monsieur **MC. BATOCHE**, professeur à l'université de Constantine, monsieur **A.MOUSSAOUI**, maître de conférence à l'université de Farhat Abès de Sétif et monsieur **S.CHIKHI** professeur à l'université de Constantine pour l'intérêt qu'ils ont bien voulu porter à ce travail en acceptant d'être examinateurs.

Enfin, que tous ceux qui nous ont aidé et encouragé de près ou de loin dans la concrétisation de ce projet, trouvent ici notre gratitude et nos sincères remerciements.

## **Résumé**

Dans ce travail nous avons proposé une nouvelle mesure objective avec référence complète. Celle-ci est basée sur l'information contour et région et les pixels déformés et déplacés. Elle permet de comparer deux images originale et déformée. L'intérêt de cette mesure est : d'une part, elle détermine automatiquement l'erreur à l'intérieur de l'image déformée, et d'autre part le regroupement de toutes les mesures locales entre deux images donne l'évaluation globale de l'image qui comporte l'ensemble des mesures. Une étude comparative avec des mesures existant dans la littérature est effectuée. En utilisant une base d'images reconnue dans la littérature, cette comparaison nous a conduit à tirer la conclusion que la méthode proposée donne des résultats satisfaisants par apport aux autres méthodes.

Mots-clés : Evaluation de la qualité d'image, information contour et région, pixels déformés, pixels déplacés.

## **Abstract**

Image quality assessment (IQA) plays a major role in a broad range of applications. Evaluating the adequacy of a quality image for a given application (image or video processing applications) is a requisite. However, objective quality assessment is far from being a solved problem. In this work, we have addressed the problem of assessment of images quality. An approach based on edge-region information, distorted and displaced pixels have been developed for image quality assessment. First, the distorted and original images are divided into blocks of  $11 \times 11$  pixels, and then we compute distorted pixels and displacement which can be used to compute the global error. Simulation results with well known methods show that the performance of the proposed technique is found to be better than previously metrics which do not use the pixel displacement in calculation. Thus, the results indicate that this global method can be helpful in the assessment and comparison of quality of images. In the conducted experiments, the proposed IQA method has showed the aptitude to evaluate the image quality according to what a human observer would do.

**Keywords:** Image Quality, Distorted Pixels, Region information, Displaced Pixels.

## التلخيص

يلعب تقييم جودة الصورة دورا رئيسيا في مختلف التطبيقات. الهدف الرئيسي من مختلف الخوارزميات ذات الطابع التقييمي للصور هو تقييم مدى نوعية الصورة لتطبيق معين (تطبيقات معالجة الصور أو الفيديو). ومع ذلك، لا تعتبر تقييم جودة الصورة معضلة قد طوي صيتها فائيا. نطمح في هذا العمل المتواضع إلى حل بعض المشاكل، و ذلك باقتراح حل بسيط. تنتسب الخوارزمية المطروحة إلى فئة الخوارزميات ذات الصورة المرجعية. في البداية نقوم بتجزئ الصورة الأصلية و الصورة ذات النوعية الرديئة إلى كتل من حجم  $11 \times 11$  بكسل. يليها حساب مدى انتقال و تشوه بكسل. بينت نتائج المحاكاة للخوارزمية المقترحة و ذلك مقارنته بخوارزميات معروفة، أن أداء التقنية المقترحة قد أعطى نتائج أفضل. زد على ذلك، تشير النتائج إلى أن هذه الخوارزمية يمكنها أن تكون مفيدة في تقييم ومقارنة جودة الصور. في التجارب التي أجريت، أظهرت إمكانية الخوارزمية المقترحة لتقييم جودة الصورة وفقا لمراقبة الإنسان.

**الكلمات الرئيسية:** جودة الصورة، تشوه بكسل ، معلومات المنطقة، إزاحة بكسل.

# Table des matières

<b>Table des figures</b> .....	iv
<b>Liste des tableaux</b> .....	vii
<b>Liste des abréviations</b> .....	viii
<b>Introduction générale</b> .....	01
<b>1 Chapitre 1 : Qualité d'images et différents types de dégradations</b> .....	04
1.1 Introduction .....	04
1.2 Définitions .....	04
1.2.1 Fidélité .....	05
1.2.2 Qualité d'images .....	05
1.3 Motivation pour l'évaluation de la qualité d'image .....	05
1.4 Compression .....	06
1.4.1 Taux de Compression .....	07
1.4.2 Codage d'images fixes : standards JPEG .....	08
1.4.2.1 Norme JPEG .....	08
1.4.2.2 Norme JPEG 2000 .....	11
1.5 Différents types de dégradations .....	14
1.5.1 Effets de bloc .....	14
1.5.2 Effet d'oscillations parasites (ringing effect) .....	14
1.5.3 Flou .....	15
1.5.4 Bruit .....	17
1.5.4.1 Différents types de bruit .....	18
1.6 Restauration et suppression des bruits .....	22
1.7 Conclusion .....	22
<b>2 Chapitre 2 : Mesure de qualité : état de l'art</b> .....	23
2.1 Introduction .....	23
2.2 Evaluation subjective .....	23
2.2.1 Facteurs influants .....	24
2.2.2 Protocoles .....	25
2.2.2.1 Discussion .....	27
2.2.2.2 Analyse des résultats .....	27
2.3 Bases d'images .....	28

2.3.1	Base CSIQ .....	28
2.3.2	Base IVC .....	33
2.3.3	Base A57 .....	34
2.3.4	Base Touyama .....	37
2.3.5	Base LIVE .....	39
2.3.6	Base TID 2008 .....	40
2.4	Evaluation objective .....	42
2.4.1	Méthodes avec référence complète .....	43
2.4.1.1	Erreur quadratique moyenne (MSE) .....	43
2.4.1.2	Rapport crête signal sur bruit (PSNR) .....	43
2.4.1.3	Rapport signal sur bruit (SNR) .....	43
2.4.1.4	Indice de similarité structurelle (SSIM) .....	44
2.4.1.5	Similitude structurale à base de gradient .....	46
2.4.1.6	Indice de la région d'intérêt ( <i>VroiWQI</i> ) .....	47
2.5	Mesures inspirées du Système Visuel Humain .....	48
2.5.1	Système visuel humain .....	48
2.5.1.1	Œil humain.....	48
2.5.1.2	Rôle du cerveau.....	50
2.5.2	Caractéristiques psychophysiques de SVH .....	51
2.5.2.1	Fovéation.....	51
2.5.2.2	Phénomène d'adaptation à la lumière.....	51
2.5.2.3	Fonctions de sensibilité de contraste .....	52
2.5.2.4	Effet de masquage.....	52
2.5.3	Méthodes basées sur SVH .....	52
2.5.3.1	Algorithme de Daly .....	52
2.5.3.2	Algorithme de Lubin .....	53
2.5.3.3	Métrique de Watson .....	53
2.6	Conclusion .....	53
<b>3</b>	<b>Chapitre 3: Méthode proposée</b> .....	<b>54</b>
3.1	Introduction .....	54
3.2	Etat de l'art .....	55
3.3	Travaux précédents .....	56
3.4	Méthode proposée (ERDDM).....	57

3.4.1	Mesure de pixels déformés et déplacés .....	60
3.4.2	Régions visuelles d'intérêt .....	68
3.4.3	Information contour .....	68
3.4.4	Erreur globale .....	70
3.4.5	Discussion .....	71
3.5	Conclusion .....	73
<b>4</b>	<b>Chapitre 4 : Résultats et analyses</b> .....	<b>74</b>
4.1	Introduction .....	74
4.2	Base d'images Live .....	74
4.2.1	Choix des images .....	74
4.2.2	Type de déformations .....	75
4.2.3	Détails de l'expérience .....	76
	4.2.3.1 Méthodologie de test .....	76
	4.2.3.2 Configuration d'équipement .....	76
4.2.4	Traitement des points bruts .....	78
	4.2.4.1 Détection de valeurs aberrantes et rejet d'observateur .....	78
	4.2.4.2 Scores de MOS .....	79
	4.2.4.3 Scores de DMOS .....	79
4.3	Méthodologie pour l'évaluation de la performance .....	79
4.3.1	Métriques relatives à la précision de prédiction d'une méthode .....	81
4.3.2	Métriques relatives à la monotonie de prédiction d'une méthode ...	81
4.3.3	Métriques relatives à la cohérence de prédiction d'une méthode .....	82
4.4	Analyse des résultats .....	83
4.4.1	Discussion .....	102
4.5	Conclusion .....	103
<b>5</b>	<b>Conclusion générale</b> .....	<b>104</b>
	<b>Bibliographie</b> .....	<b>106</b>



# Table des figures

<b>Figure 1.1</b> : Schéma bloc de la compression .....	07
<b>Figure 1.2</b> : Etapes de compression de la norme JPEG .....	08
<b>Figure 1.3</b> : Parcours en zig-zag d'un bloc de $8 \times 8$ pixels issu de la <i>DCT</i> .....	10
<b>Figure 1.4</b> : Arbre binaire de Huffman .....	11
<b>Figure 1.5</b> : Diagramme de la chaîne de codage de l'algorithme JPEG2000 .....	11
<b>Figure 1.6</b> : Effet de bloc .....	14
<b>Figure 1.7</b> : Effet de ringing .....	15
<b>Figure 1.8</b> : Flou .....	15
<b>Figure 1.9</b> : Exemple de flou de mouvement .....	16
<b>Figure 1.10</b> : Flou dû à une perte d'information .....	17
<b>Figure 1.11</b> : Flou artistique .....	17
<b>Figure 1.12</b> : (a) image originale (b) image avec le bruit poivre et sel .....	19
<b>Figure 1.13</b> : Distribution de bruit .....	20
<b>Figure 1.14</b> : Images avec le bruit gaussien .....	20
<b>Figure 1.15</b> : Bruit de speckle .....	21
<b>Figure 2.1</b> : Méthode à double stimulus .....	25
<b>Figure 2.2</b> : Méthode à simple stimulus .....	26
<b>Figure 2.3</b> : Animaux .....	29
<b>Figure 2.4</b> : Paysage .....	29
<b>Figure 2.5</b> : Personnages .....	29
<b>Figure 2.6</b> : Plantes .....	30
<b>Figure 2.7</b> : Paysages urbains .....	30
<b>Figure 2.8</b> : Compression JPEG .....	31
<b>Figure 2.9</b> : Compression JPEC 2000 .....	31
<b>Figure 2.10</b> : Décroissances globales de contraste (Global contrast decrements ) ...	31
<b>Figure 2.11</b> :Flou gaussien .....	31
<b>Figure 2.12</b> : Bruit blanc gaussien additif .....	32
<b>Figure 2.13</b> : Bruit rose gaussien additif .....	32
<b>Figure 2.14</b> : Images originales de la base d'images IVC .....	33
<b>Figure 2.15</b> : Image dégradée par la compression JPEG .....	33
<b>Figure 2.16</b> : Image dégradée par la compression JPEG2000.....	34

<b>Figure 2.17 :</b> Image dégradée par le codage de LAR .....	34
<b>Figure 2.18 :</b> Flou Gaussien .....	34
<b>Figure 2.19 :</b> Images originales .....	35
<b>Figure 2.20:</b> Quantification des sous-bandes LH de DWT avec 5 niveaux .....	35
<b>Figure 2.21:</b> Compression JPEG .....	35
<b>Figure 2.22:</b> Compression JPEG 2000 .....	36
<b>Figure 2.23:</b> Compression JPEG 2000+DCQ .....	36
<b>Figure 2.24:</b> Flou Gaussien .....	36
<b>Figure 2.25:</b> Bruit blanc gaussien .....	36
<b>Figure 2.26:</b> Images originales .....	37
<b>Figure 2.27:</b> Compression JPEG.....	38
<b>Figure 2.28:</b> Compression JPEG2000 .....	38
<b>Figure 2.29:</b> Images originales .....	40
<b>Figure 2.30:</b> Schéma de l'œil et système visuel humain .....	49
<b>Figure 3.1</b> (a) image originale (b) image déformée (c) pixel central $X_{23}$ de taille $3 \times 3$ de l'image originale (d) pixel central $X'_{23}$ de taille $3 \times 3$ de l'image déformée .....	62
<b>Figure 3.2 :</b> Image et ses centres de la gravité: (a) image originale ( $M \times N$ ), (b) image déformée ( $M \times N$ ).....	64
<b>Figure 3.3 :</b> (a) Image originale (b) La carte des régions d'intérêt (c) L'information de contour .....	69
<b>Figure 3.4 :</b> Organigramme calculant <i>ERDDM</i> .....	72
<b>Figure 4.1 :</b> Images de références .....	75
<b>Figure 4.2 :</b> Images échantillons et leurs versions déformées .....	76
<b>Figure 4.3 :</b> Interface pour l'étude subjective .....	77
<b>Figure 4.4:</b> Graphe de <i>DMOS</i> contre <i>ERDDM</i> pour toute la base de données (a) <i>ERDDM</i> ( $8 \times 8$ ) (b) <i>ERDDM</i> ( $11 \times 11$ ) (c) <i>ERDDM</i> ( $15 \times 15$ ) .....	85
<b>Figure 4.5 :</b> Graphe de la variation de la performance d' <i>ERDDM</i> selon les métrique : (a) CC, (b) MAE et (c) RMS.....	86
<b>Figure 4.6:</b> (a) Image originale (b) l'information de contour de(a) (c) l'information de contour de l'image déformée (d) carte des régions d'intérêt de (a) .....	87
<b>Figure 4.7:</b> Comparaison des images "lighthouse" et "building2". (a) Image	

originale de “lighthouse”. (b) Image floue de “lighthouse”(c) Image originale de “building2”. (d) Image floue de “building2”.

.....	89
<b>Figure 4.8.</b> Coefficient de corrélation linéaire de Pearson ( <i>CC</i> ) .....	90
<b>Figure 4.9:</b> Racine carrée de la moyenne des erreurs de prédiction ( <i>RMS</i> ) .....	91
<b>Figure 4.10:</b> Coefficient de corrélation des rangs de Spearman ( <i>ROCC</i> ) .....	93
<b>Figure 4.11:</b> Erreur absolue maximale de prédiction ( <i>MAE</i> ) .....	94
<b>Figure 4.12:</b> Rapport des valeurs aberrantes ( <i>OR%</i> ).....	95
<b>Figure 4.13:</b> Graphe de <i>DMOS</i> avec les prévisions de modèles de flou gaussien, bruit blanc, erreurs de transmission dans JPEG2000 (Fastfading). (a) <i>ERDDM</i> (b) <i>VroiWQI</i> (c) <i>GSSIM</i> (d) <i>PSNR</i> .....	96
<b>Figure 4.14:</b> Graphe de <i>DMOS</i> avec les prévisions de modèles de flou gaussien, bruit blanc, Erreurs de transmission dans JPEG2000 (Fastfading). (a) <i>MSSIM</i> (b) <i>ERDM</i> (c) <i>ERDMSSIM</i> .....	97
<b>Figure 4.15:</b> Graphe de <i>DMOS</i> avec les prévisions de modèles de JPEG2000, JPEG. (a) <i>ERDDM</i> (b) <i>VroiWQI</i> (c) <i>GSSIM</i> (d) <i>PSNR</i> .....	98
<b>Figure 4.16:</b> Graphe de <i>DMOS</i> avec les prévisions de modèles de JPEG2000, JPEG. (a) <i>MSSIM</i> (b) <i>ERDM</i> (c) <i>ERDMSSIM</i> .....	99
<b>Figure 4.17:</b> Graphe de prévisions de qualité par les six méthodes après la régression non linéaire: (a) <i>PSNR</i> (b) <i>MSSIM</i> (c) <i>WroiWQI</i> (d) <i>GSSIM</i> .....	100
<b>Figure 4.18:</b> Graphe de prévisions de qualité par les six méthodes après la régression non linéaire:(a) <i>ERDMSSIM</i> (b) <i>ERDM</i> (c) <i>ERDDM</i> .....	101

# Liste des tableaux

<b>Tableau 1.1</b> : Code de Huffman .....	11
<b>Tableau 2.1</b> : Exemple d'échelle de notation .....	26
<b>Tableau 2.2</b> : Echelle d'évaluation de qualité d'image .....	27
<b>Tableau 2.3</b> : Paramètres et les conditions de teste subjectives .....	39
<b>Tableau 2.4</b> : Types de déformations .....	41
<b>Tableau 2.5</b> : Sommaire des bases d'images avec l'année de la publication et le nombre d'images pour chaque base .....	41
<b>Tableau 4.1</b> : Sessions d'évaluation subjectives: nombre d'images en chaque session et le nombre d'observateurs participant à chaque session. Les images de référence ont été incluses dans chaque session.....	78
<b>Tableau 4.2</b> : Comparaison de la performance de méthodes d'évaluation de la qualité d'image.....	86
<b>Tableau 4.3</b> : Coefficient de corrélation linéaire de Pearson ( <i>CC</i> ).....	90
<b>Tableau 4.4</b> : Racine carrée de la moyenne des erreurs de prédiction ( <i>RMS</i> ) ...	91
<b>Tableau 4.5</b> : Coefficient de corrélation des rangs de Spearman ( <i>ROCC</i> ) .....	92
<b>Tableau 4.6</b> : Erreur absolue maximale de prédiction ( <i>MAE</i> ) .....	93
<b>Tableau 4.7</b> : Rapport des valeurs aberrantes ( <i>OR%</i> ) .....	94

# Liste des abréviations

DSCQS : Echelle continue de la qualité sur stimulus double (Double Stimuli Continuous Quality Scale).

DSIS : Echelle de dégradation sur stimulus double (Double-Stimulus Impairment scale).

SSCQE : Evaluation continue de la qualité sur stimulus unique (Single-Stimulus Continuous Quality Evaluation).

MOS : Score Optimal Moyen (Mean Opinion Score).

DMOS: Différence de Score Optimal Moyen (Difference Mean Opinion Score).

MSE : Erreur quadratique moyenne (Mean Square Error).

PSNR : Rapport crête signal sur bruit (Peak Signal to Noise Ratio).

SNR : Rapport signal sur bruit (Signal to Noise Ratio).

SSIM : Indice de similarité structurelle (Structural Similarity Index).

GSSIM : Similitude structurale à base de Gradient (Gradient-based Structural Similarity).

*VroiWQI* : Indice de la région d'intérêt (Visual region of interest Weighted Quality Index).

SVH : Système Visuel Humain.

CSF : Fonction de Sensibilité de Contraste.

PSF : Vision Fovéale et Périphérique.

UQI : Métrique de qualité universelle (The Universal Quality Metric).

MSSIM : Indice de similarité structurelle moyenne (The Mean Structural Similarity Index).

*DM*: Mesure de différence.

*ODM*: Mesure globale de Pixel déformés entre les deux blocs.

*ERDDM*: Mesure de contour-région, pixels déformés et déplacés.

*VQEG* : Groupe d'expert de la qualité de vidéo (Video Quality Expert Group).

*CC* : Coefficient de corrélation linéaire de Pearson (Pearson linear correlation coefficient).

*ROCC* : Coefficient de corrélation des rangs de Spearman (Spearman rank-order correlations coefficient).

*OR* : Rapport des valeurs aberrantes (Outlier ratio).

*RMS*: Racine carrée de la moyenne des erreurs de prédiction ou erreur quadratique moyenne de prédiction (Root mean square prediction error).

*MAE* : Erreur absolue maximum de prédiction (Maximum absolute prediction error).

*DMOS<sub>p</sub>* : Différence prévus de Score Optimal Moyen (the predicted Difference Mean Opinion Score).

*DCT* : Transformée en Cosinus Discrète (Discrete Cosine Transform).

*ERDMSSIM*: Similarité structurale de contour, région et pixels déformés (the edge-region and distorted pixels structural similarity).

*ERDM*: Mesure de contour -région et pixels déformés avec pixels déplacés (the edge-region and distorted pixels with pixel Displaced structural similarity).

# Introduction générale

- **Contexte de la thèse**

Les images ont un rôle important dans les médias, mais leur utilisation dépend de sa qualité. Les images numériques sont affectées par une grande variété de déformations pendant l'acquisition, le traitement, la compression, le stockage, la transmission et la reproduction.

Il existe deux façons de mesurer la qualité d'une image. La première consiste à mener des tests subjectifs. Ces tests exigent un équipement approprié et des protocoles normalisés pour permettre des coopérations, échanges et comparaisons des résultats fournis par différents laboratoires. Elle est effectuée par les observateurs humains. Le groupe de travail inclut des observateurs « non experts » et des observateurs « experts ». Un observateur non expert focalise son attention sur la vue globale, mais un observateur qualifié peut se concentrer sur les détails. Les résultats de ces tests représentent la référence dans l'évaluation de la qualité. Toutefois ces tests sont lourds à mettre en œuvre, chers et surtout très longs et ne constituent donc pas une solution pratique pour les différents opérateurs. Pour éviter un tel inconvénient, les métriques perceptuelles (méthodes objectives) représentant la deuxième alternative, ont pour objectif de définir des mesures de qualité qui soient fortement corrélées aux notes de qualité qu'aurait donnée un ensemble d'observateurs. On distingue alors trois grandes familles de métriques objectives :

**Mesures avec référence :** Les mesures avec référence sont utilisées pour estimer la qualité entre deux images, l'originale et sa version dégradée. Généralement, aucune information liée au type de dégradation n'est donnée. Elles sont donc supposées universelles. La majorité des métriques proposées dans la littérature sont avec référence.

**Mesures sans référence** : Orientées pour une dégradation spécifique, les mesures sans référence sont les plus attrayantes d'un point de vue applicatif.

En effet, l'image dégradée est la seule donnée nécessaire. Cependant, le type de distorsion est en général supposé connu.

**Mesures avec référence réduite** : Les mesures avec référence réduite ont été développées afin de pallier aux contraintes liées aux métriques avec et sans référence. Ce type de mesure nécessite uniquement certaines caractéristiques de l'image originale. De plus, aucune information à priori n'est généralement exploitée. Elles restent les méthodes les moins développées.

- **Problématique**

Mesurer la qualité d'une image soulève plusieurs questions. D'abord, la qualité représente aujourd'hui une des clés du développement des applications et des services multimédias. En effet, la bande passante disponible étant limitée est chère. Les opérateurs sont donc amenés à fixer d'abord les ressources disponibles dans le transport et la qualité souhaitée à la réception avant de choisir les algorithmes et les standards de compression dont ils ont besoin. Pour les applications TV, les télédiffuseurs choisissent un débit cible et fixent, en fonction de leur expérience, un certain nombre de paramètres pour l'encodeur. La conséquence de ces choix au niveau de la réception se traduit par une qualité perçue par l'observateur qui peut varier considérablement d'un contenu à un autre.

Le besoin d'un algorithme automatique est l'évaluation objective et l'analyse des images est de faire le rapport de leur qualité sans intervention humaine. Le but est la conception d'une méthode dont son évaluation est proche du jugement humain. Cette méthode peut éliminer le besoin d'une étude coûteuse de l'évaluation subjective. Le problème de la conception d'une méthode d'évaluation objective avec référence a été abordé dans cette thèse.

- **Organisation de la thèse**

La thèse est organisée en quatre chapitres.

**Le premier chapitre** introduit des définitions relatives à la qualité d'image. Ainsi que les différents types de déformation, à savoir les effets de bloc, le flou, le bruit.

**Le deuxième chapitre** est l'état de l'art des méthodes d'évaluation de la qualité d'image. Des définitions relatives à la qualité d'images sont introduites. Ensuite, les différents protocoles nécessaires pour effectuer une évaluation subjective sont présentés. En plus, les six bases d'image les plus utilisées dans le domaine de la qualité d'image sont décrites, suivi par



quelques méthodes d'évaluation objectives telle que PSNR, MSSIM, etc. le chapitre se termine par une introduction au système visuel humain (SVH), ses caractéristiques et quelques méthodes basée sur SVH.

**Le troisième chapitre** décrit la méthode d'évaluation de la qualité d'images proposée basées sur l'information contour, région, pixels déformés et déplacés. Cette dernière s'intègre dans les méthodes objectives d'évaluation de la qualité d'images avec référence. Celle-ci est une amélioration de la méthode MSSIM (Mean Structural Similarity Index) [4] en introduisant le concept du déplacement des pixels ainsi que ses intensités afin de calculer le taux de déformation de l'image de test par rapport à l'image originale.

**Le quatrième chapitre** décrit la mise en œuvre notre méthode proposée de l'évaluation de la qualité d'images en utilisant base d'images existant sur le web : la base d'images Live. Dans un premier temps nous présentons la base d'images Live en expliquant les différentes étapes à suivre afin qu'elle soit en service. Ensuite, on suit la procédure standard d'évaluation des performances utilisées par le groupe d'expert de la qualité de vidéo (*VQEG*). Enfin une étude comparative intégrant un ensemble des méthodes d'évaluation existant dans la littérature. Suivie d'une discussion et d'une analyse des résultats obtenus.

Enfin, la thèse se termine par une conclusion générale qui récapitule les travaux réalisés et propose quelques visions pour les travaux futurs.

# Chapitre 1

## La qualité d'images et différents types de dégradations

### 1.1 Introduction

L'idée de l'évaluation de la qualité de l'image a commencé par l'invention des premiers instruments optiques, le télescope optique et le microscope (1600-1620) (Galilée était une image principale dans ces deux inventions). Ce concept apparaît encore aux jours de la photographie, 1860-1930, pendant le développement de la télévision, 1935-1955, et continue avec l'image numérique jusqu'à aujourd'hui.

Dans ce chapitre, nous étudions les différents types de dégradations les plus répandues pouvant affecter une image, à savoir les effets de bloc, bruit, la compression JPEG, JPEG2000 et le flou. Généralement, le flou apparaît principalement dans la compression JPEG et JPEG2000 d'où la perte de netteté de l'image. De plus, le bruit est une dégradation dans l'image, provoquée par une perturbation externe. Le principe des effets de bloc est sa manifestation au niveau des frontières entre blocs et son apparition comme des contours verticaux et horizontaux dont la visibilité dépend fortement de la distribution spatiale de l'image.

### 1.2 Définitions

Dans cette partie on va donner quelques définitions relatives à l'évaluation de la qualité d'image.

### 1.2.1 Fidélité

La fidélité, comme son nom l'indique, permet principalement de savoir si la reproduction de l'image est fidèle ou non à l'originale. Dans ce cas, on met en place une mesure qui calcule la distance entre les deux images. Cette distance symbolise numériquement l'écart qu'il peut y avoir entre les deux reproductions de l'image [52].

### 1.2.2 Qualité d'images

La qualité d'image est un terme général employé par l'observateur pour évaluer la qualité inhérente de l'image. Elle est principalement subjective, signifiant que chaque observateur évalue la qualité d'image différemment. Certains peuvent penser qu'une image est d'assez haute qualité, mais on peut avoir un individu avec des niveaux plus élevés qui peut indiquer que l'image manque de qualité.

Avec des images numériques, beaucoup de personnes peuvent évaluer et classer des images par catégorie basées sur les bits de données ou les pixels. Par exemple, si on a deux images numériques, une d'un méga de pixels (1 million de pixels) et une image différente ayant 5 méga pixels de d'information dans l'image, la plupart des personnes déclareraient que l'image avec plus de pixels a la qualité plus élevée. Et ce serait vrai dans un bon nombre de circonstances. Cependant, si l'image de 5 méga pixels était prise dans l'obscurité complète et l'image de 1 méga pixels était prise en utilisant la lumière régulière, la plupart des personnes conviendraient que bien qu'une image puisse avoir plus de méga pixels ne signifie pas nécessairement qu'elle a une meilleure qualité d'image.

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.

## 1.3 Motivation pour l'évaluation de la qualité d'image

Maintenant que l'on a expliqué l'importance de la qualité des images et l'équilibre associé de coût-qualité, la question évidente qui se pose est pourquoi on doit mesurer la qualité. La réponse est simple et pourrait être illustrée par quelques exemples. Si un concepteur conçoit une télévision à objectif de qualité élevé, et veut savoir à quoi va ressembler la courbe de rapport qualité-coût, il a besoin évidemment d'un mécanisme pour mesurer la qualité de la vidéo et le rendement quand sa

conception fonctionne dans certaine configuration coûtant une certaine somme. Dans un autre scénario, 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. Il a besoin aussi d'une méthode de comparer scientifiquement la qualité des deux systèmes. Une autre situation peut être celle d'un entrepreneur, qui vend le contenu d'un divertissement par le câble, le satellite, ou l'Internet, et veut surveiller la qualité du contenu qui est fourni à ses clients.

Les algorithmes d'évaluation de la qualité sont nécessaires pour trois types d'applications:

1. Pour l'optimisation, où on maximise la qualité à un coût donné.
2. Pour l'analyse comparative entre différentes solutions de rechange.
3. Pour la surveillance de la qualité dans des applications en temps réel.

#### **1.4 Compression**

L'utilisation de données sous leurs formes numériques ne serait pas possible aujourd'hui sans la compression [65-69] préalable de celles-ci. Et ceci pour plusieurs raisons :

- Les capacités de stockages des utilisateurs, même si elles ne cessent d'augmenter, ne sont pas infinies. Dans ce cas les données sont compressées avant d'être stockées sur une unité de stockage. En cas de besoin, des données archivées seront restaurées à l'aide d'une procédure de décompression.
- La durée des transmissions de ces données numériques est conditionnée par le débit du réseau qui est utilisé et qui est parfois relativement faible. L'expéditeur compresse les données avant de les transmettre et le receveur les décompresse à la réception. Dans le domaine de la transmission d'image ou de séquences d'image, ces dernières sont lourdes dans leurs représentations en bits surtout dans le cas d'une image volumineuse. Ce qui nécessite une compression avant l'expédition pour un gain de coût de transmission.

La compression de données est l'opération consistant à lire un ensemble d'informations et les transformer en code. Dans le cas de la compression d'image, ces données ne sont qu'une suite de symboles binaires représentant une image ou une

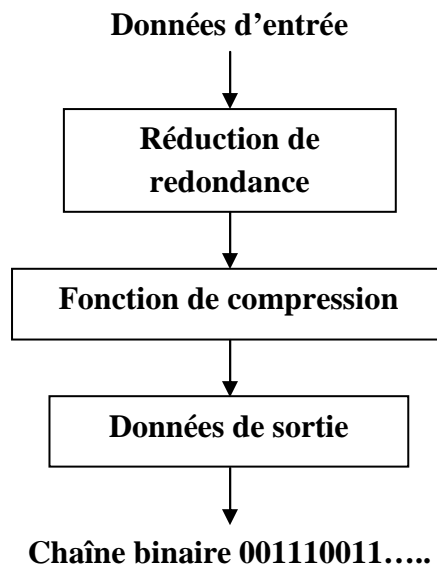
séquence d'image. Le schéma général de la compression est donné dans la figure 1.1.

En fait, la compression de données est efficace si la taille de l'information résultante est plus petite que la taille de données originales.

$$D_s = F(D_e) \begin{cases} D_s: \text{données de sortie} \\ D_e: \text{données d'entrée} \\ F: \text{fonction de compression} \end{cases}$$

Donc la compression vise à remplacer une suite de bits par une autre suite de bits moins longue pour réduire la taille physique d'un bloc de données. D'autre part, la décompression consiste à retrouver l'information de départ.

En plus, la compression des données numériques permet donc de diminuer la taille de stockages et de rendre possible leurs transports sur des réseaux de communications (Internet, GSM, câble, TV satellite).



**Figure 1.1** : Schéma bloc de la compression

### 1.4.1 Taux de Compression

On calcule souvent le taux de compression qui détermine le degré de compression obtenu : [69]

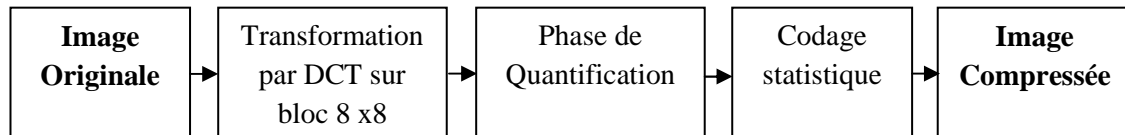
$$Taux = \frac{\text{Nombre de bits dans l'image originale}}{\text{Nombre de bits dans l'image compressée}} \quad (1.1)$$

Pour la même méthode de compression et un même taux de compression appliqué sur des images différentes, la qualité obtenue peut très bien varier d'une image à l'autre. Les propriétés statistiques des images originales jouent un rôle prépondérant dans les résultats obtenus. Par exemple, avec une image contenant beaucoup de redondances, il est facile d'obtenir des taux élevés [70].

## 1.4.2 Codage d'images fixes : standards JPEG

### 1.4.2.1 Norme JPEG

La norme JPEG [72] (*Joint Photographic Experts Group*) permet la compression avec perte d'information. Elle a été établie en 1991. La figure 1.2 introduit les étapes essentielles utilisées dans la compression de la norme JPEG.



**Figure 1.2 :** Etapes de compression de la norme JPEG

#### a. Découpage

Dans cette étape on découpe l'image originale en blocs de taille  $8 \times 8$  pixels. Le partitionnement de l'image permet de diminuer le temps de calcul dans les opérations qui viennent.

#### b. Transformée en cosinus discrète(DCT)

La transformée en cosinus discrète, aussi appelée *Discrete Cosine Transform (DCT)*, est à la base de la transformation utilisée dans la norme *JPEG*.

On applique la *DCT* sur chacun des blocs de taille  $8 \times 8$  pixels précédemment définis. Elle donne à nouveau un bloc de  $8 \times 8$  pixels.

La définition de la *DCT* est donnée par la formule suivante :

$$DCT(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) \cdot \cos\left(\frac{(2x+1)i\pi}{2N}\right) \cdot \cos\left(\frac{(2y+1)j\pi}{2N}\right) \quad (1.2)$$

où

- $DCT(i, j)$  représente la valeur de la *DCT* au point de coordonnées  $(i, j)$  dans le bloc résultat de  $8 \times 8$  pixels.

- $N$  représente le nombre de pixels dans la largeur d'un bloc, dans la norme  $JPEG N = 8$ .
- $pixel(x, y)$  représente la valeur du pixel de coordonnées  $(x, y)$  dans le bloc de l'image originale de  $8 \times 8$  pixels.
- $C(x) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } x = 0 \\ 1 & \text{sinon} \end{cases}$

La  $DCT$  inverse sur chacun des blocs  $8 \times 8$  est définie par :

$$I(i, j) = \frac{1}{\sqrt{2N}} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} C(i).C(j).DCT(x, y) . \cos\left(\frac{(2x+1)i\pi}{2N}\right) . \cos\left(\frac{(2y+1)j\pi}{2N}\right) \quad (1.3)$$

### c. Quantification

Dans le but d'atteindre des taux de compression intéressants, on permet un certain niveau de perte d'information sur l'image. Cette phase s'appelle la quantification. Elle permet d'éliminer certaines valeurs d'un signal dont l'influence est faible pour la reconstruction de celui-ci.

La quantification consiste à diviser chaque bloc de pixels par une matrice de même taille (la table de quantification). Les valeurs du bloc qui sont inférieures aux valeurs correspondantes dans la table de quantification sont donc mises à zéro. Elle s'applique sur un bloc de la  $DCT$  de telle sorte que :

$$DCT_{quantifiee} = \text{arrondi}\left(\frac{DCT(x, y)}{Q(x, y)}\right) \quad (1.4)$$

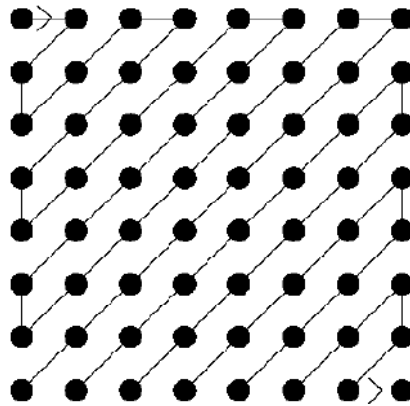
### d. Codage RLE et Huffman

Ce codage transforme les valeurs numériques en bits. Dans la norme  $JPEG$ , il se décompose en trois phases :

1. le parcours en zig-zag de la matrice  $DCT$
2. l'encodage RLE
3. l'encodage VLE

- **Parcours en zig-zag**

Une fois qu'on a quantifié la matrice  $DCT$ , on va la linéariser en la parcourant en zig-zag (voir figure 1.3) afin d'obtenir un maximum de zéros à la fin de la séquence.



**Figure 1.3 :** Parcours en zig-zag d'un bloc de  $8 \times 8$  pixels issu de la *DCT*

▪ **Codage R.L.E (Run Length Encoding )**

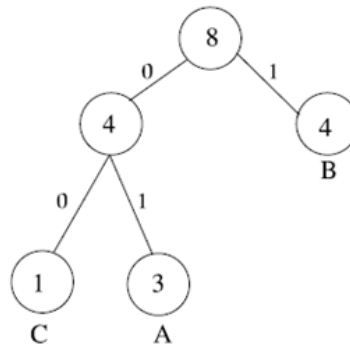
Ce type de code [71] permet de coder les éléments du message par sa valeur de répétition. Considérons le message "AAAAAABBBBBCCC", le code RLC correspondant est "6A5B3C", ce qui permet d'obtenir un gain de  $(14-6)/14$ , soit 57%.

▪ **Codage VLE (Variable Length Encoding)**

Cette méthode d'encodage consiste à coder les valeurs qui ont une probabilité forte par un petit nombre de bits : le code de Huffman [71] en est la meilleure illustration. Ce type de codage consiste à coder les symboles par une représentation de bits à longueur variable. Les symboles ayant la probabilité d'apparition forte sont codés avec des chaînes de bits plus courtes, tandis que les symboles dont la probabilité d'apparition est faible sont codés par des chaînes plus longues. Le code d'un symbole ne doit pas être le préfixe d'un autre code. Cette propriété est admise afin que la reconnaissance soit possible. Pour représenter le codage de Huffman, on utilise l'arbre binaire.

Soit un message à coder "ABBBBAAC". La fréquence d'apparition ainsi que le code Huffman correspondant sont donnés dans le tableau 1.1 et représentés par la figure 1.4.





**Figure 1.4 :** Arbre binaire de Huffman

**Tableau 1.1 :** Code de Huffman.

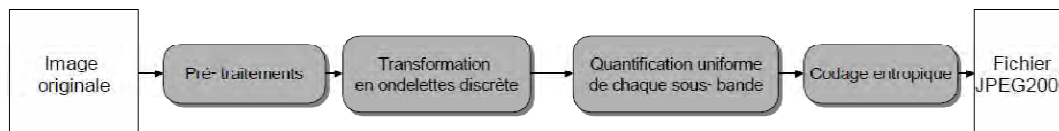
Symbole	A	B	C
Fréquence d'apparition	3	4	1
Code de Huffman	01	1	00

#### 1.4.2.2 Norme JPEG 2000

JPEG2000 [71] est une norme de compression d'image fixe. Introduite en mars 1997 par l'ISO/IEC, elle est devenue le standard en décembre 2000. L'objectif de la norme est de compléter les performances du standard JPEG mais sans la remplacer.

Par rapport à JPEG, à la place de la transformée en cosinus discrète (DCT), le standard JPEG2000 utilise la transformée en ondelettes discrète (DWT). La norme JPEG2000 emploie un codage entropique basé sur le codage arithmétique par plan de bits (EBCOT) au lieu du codage de Huffman. Ce dernier permet l'accès aléatoire, la flexibilité et la résistance aux erreurs.

La compression et la décompression d'une image au format JPEG2000 se passent en quatre étapes : Transformation, Quantification, Codage plus une étape de pré-traitement de l'image qui à pour but de rendre l'opération de codage plus efficace. Ces différentes étapes sont illustrées sur la figure 1.5 [70].



**Figure 1.5 :** Diagramme de la chaîne de codage de l'algorithme JPEG2000.

### **a. Pré-filtrage de l'image**

Cette étape comporte une transformation de l'espace des couleurs RGB en l'espace  $YCbCr$ . Cette transformation permet de coder les couleurs avec un nombre de bits inférieur.

### **b. Transformée en ondelettes discrète**

La transformation en ondelettes discrète provient de l'analyse multi-résolution qui a été développée par Stéphane Mallat et Yves Meyer. Comme son nom l'indique, le but de cette théorie est de décomposer un signal suivant différentes résolutions. On procède ainsi à une dé-corrélation de l'information qu'il contient. L'analyse multi-résolution a le même effet qu'un microscope aux pouvoirs de grossissement variables. Les basses résolutions représentent la forme grossière du signal tandis que les hautes résolutions encodent les détails du signal.

L'avantage de la représentation multi-résolution réside dans la dualité contenu-fréquence. Contrairement à la transformée de Fourier qui projette le signal dans l'espace des fréquences, l'analyse multi-résolution représente le signal conjointement dans son espace réel et dans son domaine fréquentiel. Pour des signaux 2D comme les images, des propriétés topologiques (orientations, agencement du contenu) sont ainsi conservées après la transformation multi-résolution. Le codeur qui est basé sur une telle transformation peut ainsi prendre en compte la redondance spatiale et fréquentielle de l'image. Il est ainsi plus efficace.

### **c. Quantification des sous-bandes**

L'algorithme de compression JPEG2000 adopte une quantification linéaire pour chaque sous-bande. Le pas de quantification utilisé est cependant beaucoup plus faible pour les sous-bandes basses fréquences, qui contiennent beaucoup d'information, que pour les sous-bandes hautes fréquences.

### **d. Codage des sous-bandes**

Le codage de chaque sous-bande s'effectue par plan de bits. Cette technique de codage est effectuée en trois étapes :

- une étape qui code les bits non nuls isolés,
- une étape qui code les bits dont le bit parent est non nul,
- une étape qui code le reste des bits non nuls.

### e. Codage par régions

L'algorithme JPEG2000 permet également de coder différentes régions de l'image avec des qualités différentes. Cette fonctionnalité est mise en œuvre en augmentant le nombre de bits de poids forts des coefficients appartenant à une zone d'intérêt. De part l'orientation du codage, ces coefficients seront alors codés en priorité. Il existe deux manières différentes de coder une région d'intérêt :

- il est possible d'insérer dans le fichier compressé, un fichier codant la forme de la région d'intérêt.
- il est également possible de doubler le nombre de bits appartenant aux zones d'intérêts : on passe de N bits classiquement utilisés à 2N bits. Ces coefficients seront codés en premier lieu et il sera facile de montrer qu'ils appartiennent à une zone d'intérêt.

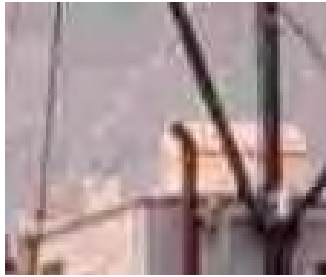
### ❖ Avantages des ondelettes (JPEG 2000) par rapport à la DCT (JPEG) [83]

- La compression DCT du format JPEG analyse l'image par bloc 8 par 8 pixels ce qui produit un effet de mosaïque (les limites des blocs sont visibles à fort taux de compression).
- La compression par ondelettes ne présente pas cet effet de mosaïque indésirable. Il est donc possible de compresser des images avec un taux de compression élevé tout en conservant une bonne qualité picturale.
- Les blocs JPEG 8x8 sont quantifiés indépendamment les uns des autres ce qui ne permet pas de réduire les redondances au-delà d'un bloc. Au contraire, la compression par ondelettes est une méthode globale sur toute l'image. Cet avantage se traduit par une efficacité encore plus importante sur les grosses images. Une image de 50 MO peut être réduite à 1 MO.
- *discutions* : La compression d'image peut engendrer du flou. En effet, le principe de la compression irréversible est de retenir les composantes les plus énergétiques et visibles. Les détails qui se situent en général dans les hautes fréquences sont sacrifiés conduisant ainsi à une perte de finesse et de netteté qui se traduit par un flou dû à cet effet de filtrage passe-bas.

## 1.5 Différents types de dégradations

### 1.5.1 Effets de bloc

La principale source d'erreur lors de la compression JPEG est ce qu'on appelle l'effet de blocs [66]. 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'œil humain. Par exemple, considérons une zone de l'image avec un dégradé subtil de couleurs, comme dans l'image suivante (voir figure 1.6).



**Figure 1.6 :** Effet de bloc

Localement, l'image est approximativement constante et elle sera donc seulement représentée par un ou quelques coefficients non-nuls après la quantification de chaque bloc. Après la dé-quantification, on aura un choix limité de valeurs pour le coefficient principal indiquant l'intensité moyenne du bloc. Il pourrait alors arriver qu'une petite différence entre deux blocs engendre une grosse différence entre la valeur de ces coefficients, d'où l'apparition de l'effet de blocs.

### 1.5.2 Effet d'oscillations parasites (ringing effect)

Cette dégradation est due en général à l'étape de quantification ou de décimation des coefficients hautes fréquences [51]. 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. En effet, pour décrire le changement brusque dans l'image, on prend des ondelettes d'une résolution élevée. Lors de la quantification des coefficients, l'approximation de ces coefficients se répercute sur les ondelettes qui

font apparaître leurs oscillations sur l'image reconstruite. Comme c'était le cas pour l'image de la figure 1.7 on observe ce phénomène.



**Figure 1.7 :** Effet de ringing

### 1.5.3 Flou

Effet esthétique qui donne à voir un contour imprécis (voir figure 1.8). 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 [51]:

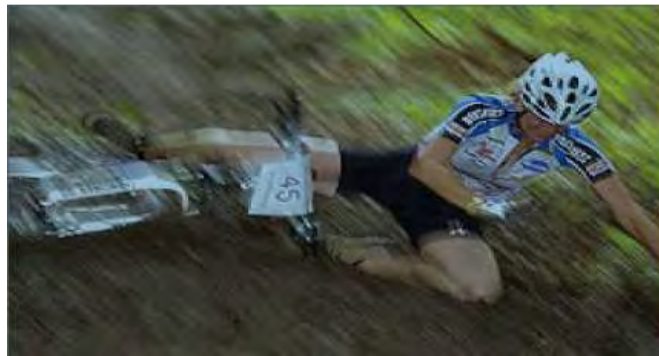


**Figure 1.8 :** Image flou

- **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 flou de défocalisation est lié à la notion de profondeur de champ qui n'est rien d'autre que l'espace d'excursion du plan de focalisation permettant d'avoir une image nette. En effet, lors d'une prise de vue, la focalisation sur un plan donné entraîne inévitablement un flou au niveau des objets se trouvant en dehors du champ de netteté. Le flou de bougé est dû essentiellement à la limitation de la vitesse d'acquisition du système de capture et de la focale. Ce flou apparaît

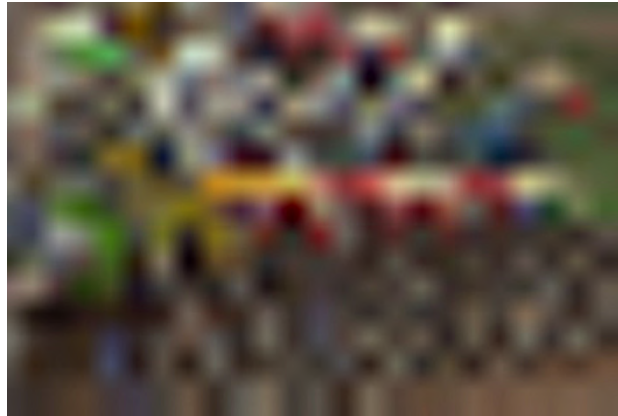
surtout lors de l'utilisation du zoom. Ce défaut se manifeste par un effet de trainée. Cela est dû au fait que chaque élément de la scène sensibilise plusieurs pixels de la caméra. On peut limiter cet effet indésirable en utilisant des stabilisateurs optiques.

- **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 d'acquisition d'image. On modélise ce flou par une opération de convolution d'une fonction porte orientée dans le sens du déplacement et dont l'étendue ou support n'est rien d'autre que le produit  $V.T$  où  $V$  est la vitesse de déplacement et  $T$  la durée d'exposition. Un exemple de flou de mouvement est illustré par la figure 1.9.



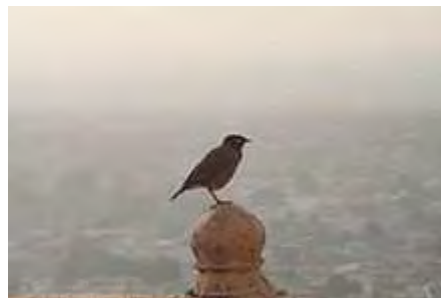
**Figure 1.9** : Exemple de flou de mouvement.

- **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. La figure 1.10 présente un exemple d'image dont la perte de données est à l'origine du flou.



**Figure 1.10 :** Flou dû à une perte d'information.

- **Les traitements :** Certains traitements visant à atténuer quelques distorsions tels que le filtrage du bruit peuvent engendrer du flou.
- **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. Un exemple d'image contenant un flou ajouté volontairement est présentée figure 1.11. Il est à noter qu'il semble difficile de différencier un flou artistique volontaire d'un flou indésirable.



**Figure 1.11:** Flou artistique.

#### 1.5.4 Bruit

On peut définir le bruit [13,73] comme étant une dégradation dans l'image, provoquée par une perturbation externe. Si une image est envoyée par voie électronique d'un endroit à un autre, 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 apparaissent sur l'image de sortie de différentes manières en fonction du type de la perturbation dans le signal. 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. Le filtrage d'une image corrompue par le bruit est donc un domaine important de la restauration d'image.

Soit  $f(.)$  une image. L'image sans bruit est notée  $g(.)$ , et le composant de bruit est noté  $q(.)$ . La décomposition la plus courante de l'image  $f(.)$  est *additive* :

$$f(.) = g(.) + q(.) \quad (1.5)$$

Par exemple, le bruit gaussien est habituellement considéré comme un composant additif. Le deuxième le plus courant est *multiplicatif* :

$$f(.) = g(.) q(.) \quad (1.6)$$

Un exemple d'un bruit souvent modelé en tant que multiplicatif est le speckle (voir paragraphe *bruit speckle*).

#### 1.5.4.1 Différents types de bruit

On va présenter quatre types différents de bruit, et comment ils apparaissent sur une image.

##### a. 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 (voir figure 1.12). 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.





(a)

(b)

**Figure 1.12 :** (a) image originale (b) image avec le bruit poivre et sel

Un exemple où le bruit poivre et sel surgit, est dans la transmission des images à travers des canaux numériques bruyants. Chaque pixel est quantifié par  $B$  bits. La valeur du pixel peut être écrite comme  $X = \sum_{i=0}^{B-1} b_i 2^i$ .

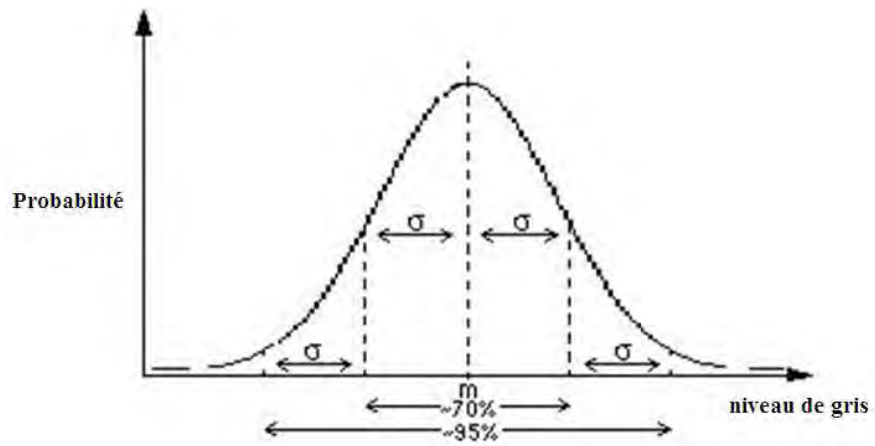
On suppose que le canal est symétrique binaire avec une probabilité de croisement  $\epsilon$ . Alors chaque bit est renversé avec la probabilité  $\epsilon$ . On appelle la valeur reçue  $Y$ .

$$Pr(|X - Y| = 2^i) = \epsilon \quad (1.7)$$

### b. Bruit Gaussien

Le bruit gaussien (voir figure 1.13) 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 modéliser une image bruyante en additionnant simplement les deux :

$$F = I + N \quad (1.8)$$



**Figure 1.13:** Distribution de bruit

L'ajout d'un bruit blanc gaussien sur une image modifie donc la valeur de chaque pixel (voir figure 1.14).



**Figure 1.14 :** Images avec le bruit gaussien

La fonction de densité du bruit gaussien,  $q$ , avec la moyenne  $\mu$  et la variance  $\sigma^2$  est

$$p_q(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (1.9)$$

Pour  $-\infty < x < +\infty$ . On note que la gamme des valeurs de  $x$  où la fonction de densité de probabilité est différente de zéro, dans les directions positives et négatives. Mais, si on considère une image comme une carte d'intensité, alors les valeurs doivent être non négatives. Dans la pratique, la gamme des valeurs du bruit gaussien est limitée à approximativement  $\mp 3\sigma$  et la densité gaussienne est un modèle utile et précis pour beaucoup de processus.

### c. bruit de speckle

Le phénomène de speckle ; "chatolement" dans les images radar, ou tavelure (en astronomie) ou granularité (en optique) et *speckle* en anglais fut constaté la première fois en 1960, lors d'expérience d'illumination d'objets avec une source de lumière cohérente, le laser. La granularité détectée n'avait pas une relation simple avec les propriétés macroscopiques de l'objet. Tandis que le bruit gaussien peut être modelé par des valeurs aléatoires sur une image; le bruit de speckle (voir figure 1.15) peut être modelé par des valeurs aléatoires multipliées par les valeurs de Pixels, par conséquent il s'appelle également le bruit multiplicatif. Le bruit de speckle est un problème important dans quelques applications de radar et d'échographie.



**Figure 1.15:** Bruit de speckle

## 1.6 Restauration et suppression des bruits

Le filtrage d'une image est une tâche essentielle dans le traitement d'image. Il existe une multitude de méthodes pour la suppression de bruit dans l'image ou dans un ensemble de données. La propriété importante de ces méthodes est qu'elle devrait complètement enlever le bruit aussi loin que possible dans l'image ainsi que la conservation de ses contours. Traditionnellement, il y a deux types de modèles c.-à-d. les modèles linéaires et les modèles non-linéaires.

Généralement, les modèles linéaires sont les plus employés. Les avantages des modèles linéaires est la vitesse et leurs limitations. Ils ne peuvent pas préserver les contours des images d'une façon efficace c.-à-d. les contours, qui sont identifiés comme discontinuités dans l'image sont masqués (présentés comme une région floue). D'autre part, les modèles non linéaires peuvent présenter les contours d'une manière bien meilleure que les modèles linéaires [67].

## 1.7 Conclusion

Dans ce chapitre, nous avons présenté des notions de base de la qualité d'image. Ainsi que, des définitions relatives à la qualité d'image. Ensuite, les différents types de déformations pouvant atteindre une image pendant la transmission, le traitement, la compression, etc. ont été exposés. Les normes JPEG et JPEG2000 ont été introduites d'où JPEG 2000. Les images compressées par la norme JPEG comportent un type de distorsion caractérisé par l'effet de blocs, se manifestant généralement au niveau des frontières entre blocs. Il apparaît comme des contours verticaux et horizontaux. D'autre type de déformations ont été présentées à savoir le flou et le bruit.

Dans le chapitre suivant, on va introduire un état de l'art concernant les méthodes d'évaluation de la qualité d'image.

# Chapitre 2

## Mesure de qualité : état de l'art

### 2.1 Introduction

Dans le chapitre qui précède un panorama sur les différents types de dégradations qui changent la qualité d'une image est présenté telles que : le bruit, l'effet de bloc, le flou etc. Afin de mesurer la qualité des images dégradées, deux manières sont exploitées. Méthodes subjectives et méthodes objectives. Plusieurs mesures sont développées dans la première catégorie impliquant l'observateur humain dont l'attribution d'une note à une image dégradée. La validation des résultats obtenus par une méthode objective et sa conformément à l'évaluation visuelle nécessite l'implication des procédures ainsi que des standards d'évaluation subjective de la qualité des images. Ces standards fixent plusieurs contraintes pour mener à bien l'évaluation. A cette fin plusieurs bases d'images sont mise en œuvre qui sont développées par des Laboratoires. Ces bases sont utilisées dans la vérification des méthodes objectives.

### 2.2 Evaluation subjective

Actuellement, le jugement le plus fiable de l'évaluation de la qualité d'images est l'évaluation subjective par les observateurs humains. Typiquement, le groupe d'observateurs inclut des observateurs « non experts » et des observateurs « experts ». Une visionneuse non experte peut prêter plus d'attention sur la vue globale, mais une visionneuse qualifiée peut se concentrer sur les détails.

Il y a habituellement deux sortes d'évaluation subjective : l'évaluation avec référence et l'évaluation sans référence. Dans le premier cas, on fournit aux observateurs des images standards de référence qui agissent en tant que calibrages dans le jugement subjectif. Dans le second cas, 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 [51].

### 2.2.1 Facteurs influants

Certains facteurs sont pris en compte pour s'assurer de la fiabilité de l'évaluation subjective. Ces derniers peuvent influencer le jugement des observateurs, les plus pertinents sont :

- **Distance d'observation**

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

- **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 [51].

- **Conditions de visualisation**

L'environnement et l'éclairage de la salle affecte effectivement la perception des stimuli (séquence de test). Un éclairage élevé risque de troubler l'observateur et ainsi pourrir sa perception. En revanche, la couleur de fond de la salle affecte l'apparence de certaines couleurs.

- **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.

- **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 mieux appréhender 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. [51]

- **Observateurs**

Il est recommandé d'avoir un panel d'observateurs le plus large possible, au moins quinze individus. Ils peuvent être experts ou novices, en ce qui concerne le thème de la campagne d'évaluation. Avant chaque séance, les observateurs seront sélectionnés pour leur acuité visuelle normale ou rendue normale par correction et leur vision



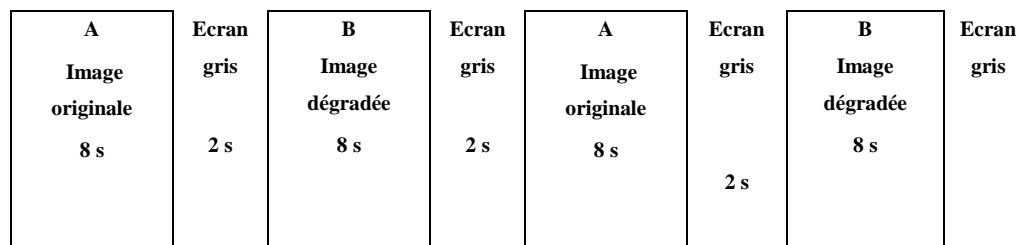
normale des couleurs. L'observateur doit avoir ainsi une acuité visuelle de 10/10 pour les deux yeux avec ou sans correction. Le panel d'observateurs choisi dépend à la fois des types de traitement à évaluer et de la nature de l'étude à mener. Il faut donc prendre en compte, lors de la sélection de ce panel, plusieurs critères comme l'âge mais aussi l'origine socioculturelle, professionnelle, etc. Ainsi, la campagne d'évaluation peut être menée sur un panel du genre "femmes européennes de 30 à 40 ans ou hommes asiatiques de plus de 60 ans". [52]

### 2.2.2 Protocoles

Il y a essentiellement trois grandes familles communes d'évaluation subjective définies par l'ITU [53] : Echelle continue de la qualité sur stimulus double (DSCQS), Echelle de dégradation sur stimulus double (DSIS) et Evaluation continue de la qualité sur stimulus unique (SSCQE).

- **Echelle continue de la qualité sur stimulus double (DSCQS)**

La méthode à double stimulus [53] est appelée "Double Stimuli Continuous Quality Scale (DSCQS)". Dans cette méthode, chaque essai se compose d'une paire de stimulus : un stimulus de la référence, et l'autre du test. Le stimulus de test est habituellement la référence après avoir subi un certain type de traitement. Les deux stimuli sont chacun présentés deux fois dans un test, d'une façon alternative, avec l'ordre du choix aléatoirement pour chaque épreuve. Des signaux audio sont employés pour indiquer quand une épreuve commence, quand un nouveau stimulus commence, quand voter, et ce qui est le nombre de tests courants dans la séquence de test qui compose une session de tests (voir figure 2.1).



**Figure 2.1 : Méthode à double stimulus**

Le tableau 2.1 donne un exemple d'une échelle de mesure à 5 niveaux de notation utilisée pour évaluer la qualité d'une image.

**Tableau 2.1** : Exemple d'échelle de notation

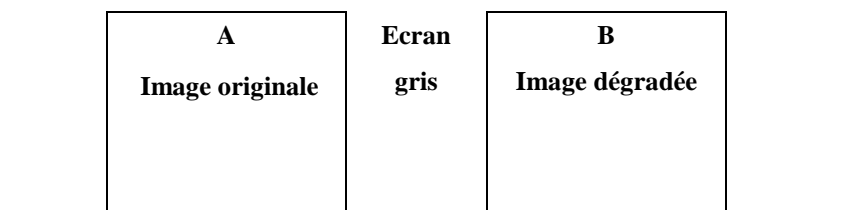
Echelle à cinq notes		
Qualité		Dégradation
Excellente	5	Imperceptible
Bonne	4	Perceptible mais non gênante
Assez bonne	3	Légèrement gênante
Médiocre	2	Gênante
Mauvaise	1	Très gênante

▪ **Echelle de dégradation sur stimulus double (DSIS)**

Comme dans la méthode de DSCQS, chaque essai se compose d'une paire de stimulus: la référence et l'essai. Cependant, dans la méthode de Double-Stimulus Impairment scale (DSIS) [53], les deux stimulus sont toujours présentés dans le même ordre : la référence est toujours la première, suivi du test. Dans la méthode de DSIS, les observateurs comparent les deux stimuli dans un essai et évaluent la dégradation du stimulus d'essai en ce qui concerne la référence, en utilisant une échelle de dégradation de cinq niveaux. Ainsi, seulement une voix est faite pour chaque essai de DSIS.

▪ **Evaluation continue de la qualité sur stimulus unique (SSCQE)**

Au lieu de voir des paires de courtes séquences séparées, les observateurs observent un programme de la durée 20-30 minutes en général qui a été traité par le système du test; la référence n'est pas montrée (Single-Stimulus Continuous Quality Evaluation, SSCQE). En utilisant un glisseur, les observateurs évaluent continuellement la qualité instantanément perçue sur l'échelle de DSCQS du mauvais à l'excellent (voir figure 2.2).



**Figure 2.2** : Méthode à simple stimulus.



Une échelle de notation utilisée pour évaluer la qualité d'une image est donnée dans le tableau 2.2. Ces échelles peuvent varier de 1 à 5 jusqu'à 100.

**Tableau 2.2** : Echelle d'évaluation de qualité d'image.

Excellente
Bonne
Assez bonne
Médiocre
Mauvaise

### 2.2.2.1 Discussion

Dans la procédure subjective d'évaluation, il y a plusieurs points qui doivent être soulignés. Tout d'abord, le nombre d'observateurs doit être assez grand. Pour assurer la confiance statistique en expériences de qualité d'image. En second lieu, les résultats du test subjectif sont influencés par les conditions expérimentales, ce qui signifie que les conditions expérimentales de visionnement doivent être conçues pour unir les conditions du visionnement dans la pratique aussi étroitement que possible. En outre, il y a un certain nombre d'autres éléments procéduraux à noter, comme : le choix des observateurs, les conditions de référence, les sources de signal pour les scènes d'essai, la synchronisation de la présentation des diverses scènes de test, le choix d'une gamme des scènes de test.

### 2.2.2.2 Analyse des résultats

Pour chaque paramètre de test, la moyenne et l'intervalle de confiance à 95% de la distribution statistique des notes doivent être calculés (pour plus de détails voir le chapitre 4).

#### a) 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 :

$$MOS(i) = \frac{1}{N} \sum_{j=1}^N Note_i(j) \quad (2.1)$$

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.

### b) Intervalle de confiance

Afin de réduire l'impact d'éventuelles erreurs, un intervalle de confiance est souvent associé à chaque note  $MOS$ . Il est généralement fixé à 95% et est donné par la relation suivante :

$$[MOS(i) - e_j, MOS(i) + e_j] \quad (2.2)$$

Où

$$e_j = 1.95 \frac{\sigma_j}{\sqrt{N}} \quad (2.3)$$

Et l'écart-type pour chaque présentation  $\sigma_j$ , est donné par :

$$\sigma_j = \sqrt{\frac{\sum_{k=1}^N (Note_i(k) - MOS(k))^2}{N - 1}} \quad (2.4)$$

Une analyse plus approfondie permet de détecter et de rejeter d'éventuelles erreurs.

## 2.3 Bases d'images

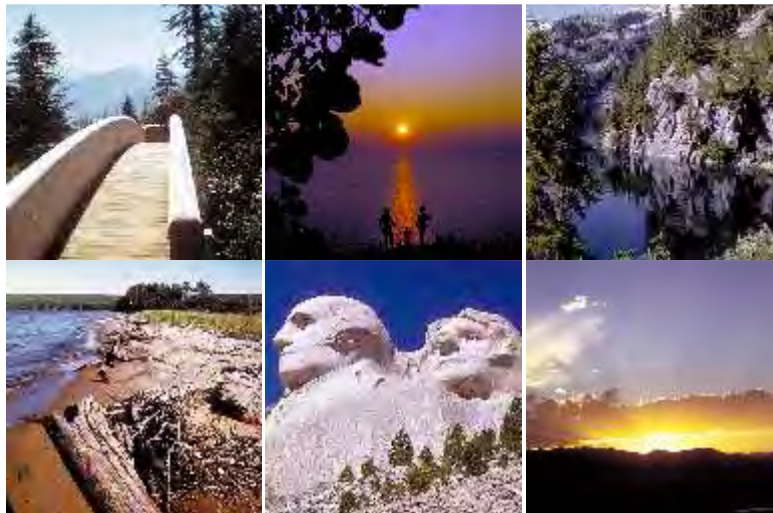
Plusieurs bases d'images ont été proposées dans le domaine de la qualité d'images. Dans cette partie nous présentons les six bases d'images les plus utilisées dans la littérature. Pour chacune d'elles, nous présentons les différents types d'image ainsi que les protocoles utilisés dans l'évaluation subjective.

### 2.3.1 Base CSIQ

La base d'images de CSIQ est une base d'images du laboratoire CPIQL (Computational Perception and Image Quality Lab) [47]. Elle comporte 30 images originales (voir figure 2.3-7) à partir desquelles sont générées des images dégradées en utilisant six types de dégradations.



**Figure 2.3: Animaux**



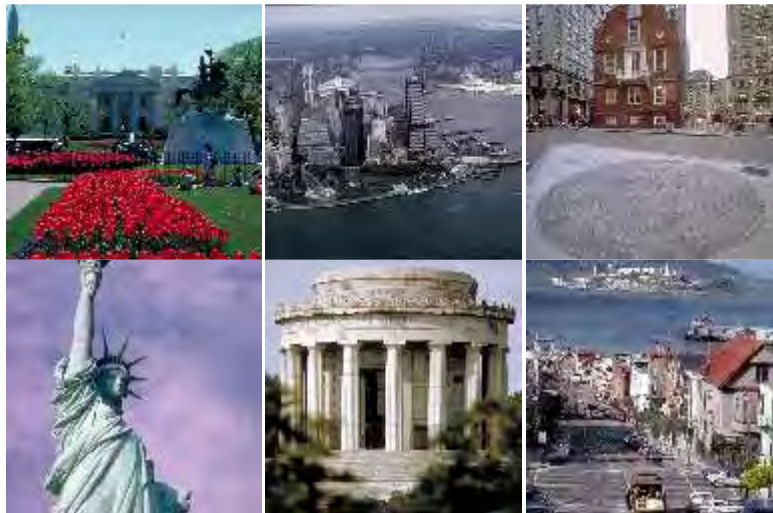
**Figure 2.4: Paysage**



**Figure 2.5 : Personnages**



**Figure 2.6 : Plantes**



**Figure 2.7 : Paysages urbains**

Les dégradations utilisées dans CSIQ sont: JPEG, JPEG2000, décroissances globales de contraste (global contrast decrements), bruit gaussien rose additif (additive pink Gaussian noise), bruit gaussien blanc additif (additive white Gaussian noise), et flou gaussien (Gaussian blurring). Alors, CSIQ est constitué de 866 d'images dégradées étant le résultat de la déformation des images originales par les différents types de dégradations précédents. Des exemples de ces types de dégradation sont montrés ci-dessous (voir figure 2.8-13):

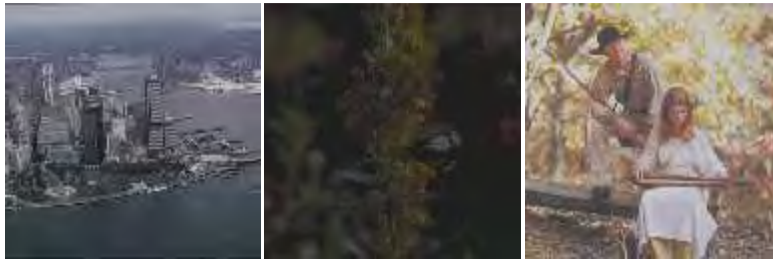




**Figure 2.8 :** Compression JPEG



**Figure 2.9:** Compression JPEC 2000



**Figure 2.10:** Décroissances globales de contraste



**Figure 2.11:** Flou gaussien



**Figure 2.12:** Bruit blanc gaussien additif



**Figure 2.13:** Bruit rose gaussien additif

Les images de CSIQ sont subjectivement évaluées en se basant sur un déplacement linéaire des images. Quatre écrans d'affichage à cristaux liquides (LCD) à la résolution de  $1920 \times 1200$  étaient calibrés en suivant l'espace de couleur RVB et placés côte à côte avec la distance égale de visionnement à l'observateur. Des observateurs ont été chargés de garder la distance de visionnement stable approximativement de 80 cm.

Toutes les versions dégradées d'une image originale ont été regardées simultanément à travers l'écran (l'écran= la rangée de moniteurs). Chaque observateur a placé ces images dans l'écran, dans lequel la distance horizontale entre deux images reflète la qualité perçue de ces images. Une image placée près des autres signifie que : l'observateur pense que leur qualité n'est pas beaucoup différente.

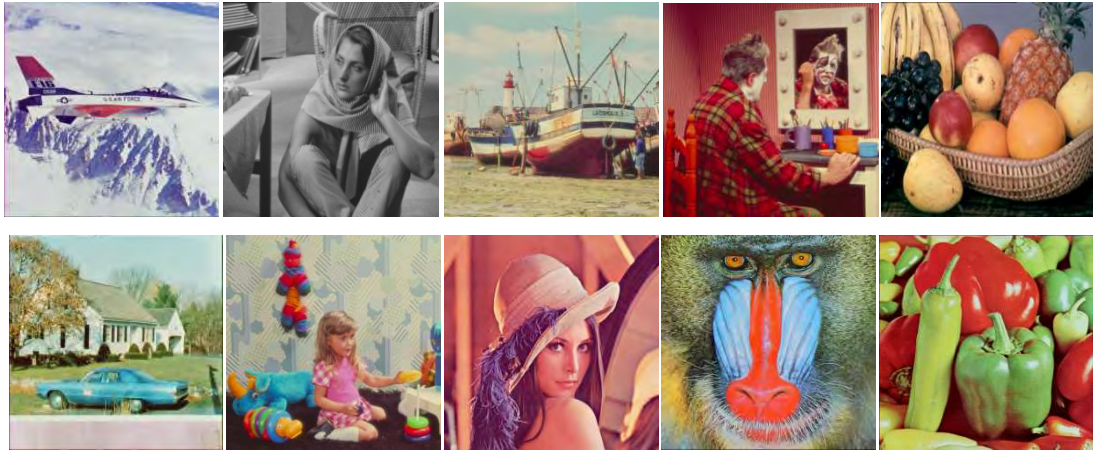
D'autre part, une place d'une image éloignée à la gauche des autres signifie que l'observateur pense que l'ancienne image est beaucoup plus mauvaise en qualité en comparant à la dernière image. Comme étape finale, des estimations d'image sont réalignées.

Dans cette expérience, 35 observateurs ont participé. Les observateurs sont femmes et des hommes avec l'acuité visuelle normale ou corrigée à la normale. Les âges des observateurs s'étendent de 21 à 35. De façon générale la base d'images contient 5000

estimations subjectives qu'elles soient rapportées sous forme de DMOS (Différence Mean Opinion Score). Le détail du calcul de DMOS est donné dans le chapitre 4.

### 2.3.2 Base IVC

La base d'images IVC [44] comprend 10 images originales qui sont utilisées afin d'engendrer 235 images déformées (voir figure 2.14).



**Figure 2.14 :** Images originales de la base d'images IVC

Les quatre types de déformations utilisées sont: JPEG, JPEG2000, Codage de LAR (LAR coding) et Flou Gaussien (voir figure 2.15-18). Ces algorithmes ont l'avantage de produire davantage de types de différentes dégradations. Des évaluations subjectives ont été faites où la distance de visionnement est égale à 6 fois la taille d'écran (image) en utilisant la méthode DSIS (Double Stimulus Impairment Scale) avec une échelle de 5 niveaux et 15 observateurs.



**Figure 2.15 :** Image dégradée par la compression JPEG





**Figure 2.16 :** Image dégradée par la compression JPEG2000



**Figure 2.17 :** Image dégradée par le codage de LAR



**Figure 2.18 :** Flou Gaussien

Des déformations pour chaque traitement et chaque image ont été optimisées afin de couvrir uniformément l'échelle subjective. Cette base de données a été employée dans [48].

### 2.3.3 Base A57

Trois images, *cheval*, *port*, et *bébé*, obtenues à partir de la base d'images de Kodak ont servi comme images originales [45]. Ainsi, un total de 57 images a été examiné: trois images originales et 54 images dégradées. Les images numériques sont des images niveau-de-gris de taille 512×512. La figure 2.19 présente les trois images originales.





**Figure 2.19 :** Images originales

Ces images ont été dégradées par six types de dégradations (voir figure 2.20-25):

1. La quantification des sous-bandes LH de DWT avec 5 niveaux.
2. Bruit blanc gaussien additif.
3. Compression JPEG.
4. Compression JPEG2000.
5. Compression JPEG2000 avec l'algorithme de la quantification à base le contraste dynamique (DCQ).
6. Flou gaussien.



**Figure 2.20:** Quantification des sous-bandes LH de DWT avec 5 niveaux



**Figure 2.21:** Compression JPEG



**Figure 2.22:** Compression JPEG 2000



**Figure 2.23:** Compression JPEG 2000+DCQ



**Figure 2.24:** Flou Gaussien



**Figure 2.25:** Bruit blanc gaussien

Une expérience psychophysique de gradation a été exécutée sur une variété d'images dégradées pour obtenir des estimations subjectives de la fidélité visuelle (base d'images A57) ; la métrique a été alors appliquée à ces images, et alors les résultats prévus ont été comparés aux résultats subjectifs réels.

Il est important de noter qu'en raison du nombre limité d'images et du nombre limité des observateurs humains, la base de données A57 est de fiabilité statistique limitée.

### 2.3.4 Base Touyama

Les expériences subjectives dans la base Touyama [43] ont été menées sur des images couleur RVB de 24 bit/pixel. Dans ces tests, un certain nombre d'observateurs humains ont été invités à assigner à chaque image des points indiquant leur évaluation de la qualité. On compte 98 images de taille 768×512 pixels dans la base d'images pour chaque groupe JPEG et JPEG2000 dont, 14 sont des images originales (voir la figure 2.26) dans chaque groupe. Le reste des images sont des images JPEG et JPEG2000 (voir figure 2.27-28) (c.-à-d. 84 images compressées dans chaque groupe).



**Figure 2.26:** Images originales



Les six échelles de qualité (15, 20, 27, 37, 55 et 79) ont été choisies pour l'encodeur JPEG et six rapports de compression (12, 24, 32, 48, 72 et 96) ont été choisis pour l'encodeur JPEG2000. Seize observateurs non-experts ont évalué la base d'images, la plupart d'entre eux sont des étudiants. Chaque observateur a regardé les images aléatoirement. Les observateurs ont été invités à fournir leur perception de la qualité sur une échelle discrète de qualité qui a été divisée en cinq et identifiée par la valeur numérique des adjectifs (Mauvaise = 1, Médiocre = 2, Assez bonne = 3, bonne = 4, et excellente = 5) dans les conditions de test de l'ITU [53]. Les seize notes de chaque image ont été ramenées à une moyenne pour obtenir une note « Mean Opinion Score » (MOS) finale de l'image avec la fiabilité d'observateur de l'intervalle de confiance de 95%. Les paramètres et les conditions de test subjectives sont donnés dans le tableau 2.3:



**Figure 2.27:** Compression JPEG



**Figure 2.28:** Compression JPEG2000

Afin de calculer l'intervalle de confiance 95% (IC), la formule suivante est utilisée [43]:

$$IC = \text{moyen}(\text{MOSs}) \pm 2.13(\sigma(\text{MOSs})/\sqrt{N}) \quad (2.5)$$

Où  $N$  = le nombre total d'observateurs, dans ce cas  $N = 16$ .

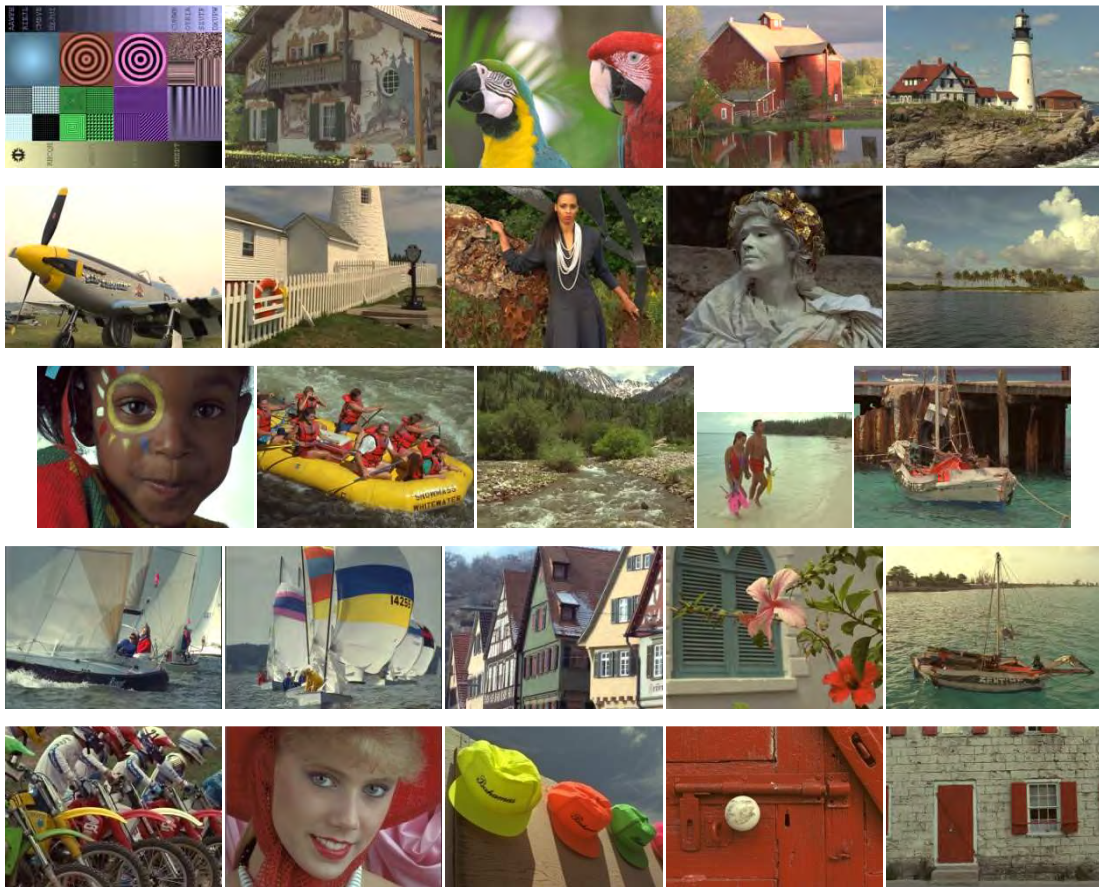
**Tableau 2.3** : Paramètres et les conditions de test subjectifs

Méthode	Single Stimulus (Adjectival categorical judgment)
Les échelles de l'Evaluation	5 Grades (Echelles d'adjectifs)
Images	24-bits/pixel RGB color (768x512)
Nombre d'images de référence	14
Encodeur	JPEG et JPEG2000
Logiciel (codec de JPEG)	cjpeg software <a href="ftp://ftp.uu.net/graphics/jpeg/jpegsrc.v6b.tar.gz">ftp://ftp.uu.net/graphics/jpeg/jpegsrc.v6b.tar.gz</a>
Logiciel (codec de JPEG2000)	JasPer version 1.701.0 <a href="http://www.ece.uvic.ca/~mdadams/jasper/">http://www.ece.uvic.ca/~mdadams/jasper/</a>
Paramètres de codage	6 (échelles de qualité: 15, 20, 27, 37, 55 et 79 pour JPEG et le rapport compression: 12, 24, 32, 48, 72 et 96 pour JPEG2000)
Observateurs	16 (Non expert, étudiants d'université )
Affichage	CRT 17-puce (1024x768)
Distance de visionnement	4 T (T: Taille d'image (largeur de l'image))
Illumination de pièce	Basse

### 2.3.5 Base LIVE

C'est la base d'images (the Live Image Quality Assess Database Release2) [3] du laboratoire de l'ingénierie d'image et de vidéo de l'université du Texas à Austin. Des détails sur cette base d'image sont trouvés dans le quatrième chapitre.

### 2.3.6 Base TID 2008



**Figure 2.29:** Images originales

TID2008 [46] contient 1700 images de tests (25 images de référence-voir figure 2.29, des images dégradées par 17 types de déformations pour chaque image de référence (voir tableau 2.4), avec 4 niveaux différents (100 images dégradées) pour chaque type de déformation). Le MOS de cette base d'images a été obtenu à partir de plus de 800 expériences. Pendant ces tests, les observateurs de trois pays (Finlande, Italie, et Ukraine) ont effectué environ 256000 jugements de qualité. Le MOS obtenu peut être utilisé pour tester efficacement les anciennes métriques visuelles de qualité aussi bien dans la conception de la nouvelle métrique.

**Tableau 2.4 :** Types de déformations

N°	Type de déformation (quatre niveaux pour chaque déformation)
1	Bruit gaussien
2	Bruit couleur.
3	Bruit corrélé
4	Bruit masqué
5	Bruit haute fréquence
6	Bruit impulsionnel
7	Bruit quantification
8	Flou gaussien
9	Débruitage
10	Compression JPEG
11	Compression JPEG2000
12	Transmission de JPEG
13	Transmission de JPEG2000
14	Bruit non écentré
15	Distortions de bloc
16	Décalage d'intensité
17	Changement de contraste

Dans la suite une table récapitulative des bases d'images existant dans la littérature.

**Tableau 2.5 :** Sommaire des bases d'images avec l'année de la publication et le nombre d'images pour chaque base.

Base d'images	Université/pays	images
TOYAMA(2000)	Université de Toyama, Japon	196
LIVE(2005)	Université du Texas à Austin	982
IVC(2006)	Université de Nantes, France	160
A57(2007)	Université De Cornell, Ithaca, New York, Etats-Unis	54
TID2008(2008)	Université de Tampere de technologie, Finlande	1700
CSIQ(2009)	Université de l'Etat De l'Oklahoma, Etats-Unis	866

## 2.4 Evaluation objective

En un mot, la mesure de la qualité objective (par opposition à l'évaluation subjective de qualité par les observateurs humains) cherche à déterminer la qualité des images algorithmiquement. Le but de la recherche de l'évaluation de la qualité objective est de concevoir des algorithmes dont la prévision de la qualité est en accord avec l'évaluation subjective des observateurs humains.

Les méthodes d'évaluation de la qualité d'image peuvent être classifiées dans trois larges catégories :

- **Les méthodes avec référence complète**, dans lesquelles l'algorithme a un accès à une version parfaite de l'image avec lesquelles il peut comparer la version dégradée. La version parfaite vient généralement d'un dispositif de haute qualité d'acquisition, après elle est dégradée par la compression et les erreurs de transmission.
- **Les méthodes sans référence**, dans lesquelles l'algorithme a un accès qu'au signal déformé et doit estimer la qualité du signal sans connaissance de la version parfaite. Puisque les méthodes sans référence n'exigent aucune information de référence, elles peuvent être employées dans n'importe quelle application où la mesure de qualité est exigée.
- **Méthodes avec référence réduites**, dans lesquelles l'information partielle concernant la version parfaite est disponible. Un côté canal existe par lequel quelque information concernant la référence peut être rendue disponible à l'algorithme d'évaluation de la qualité. Les algorithmes réduits de référence emploient cette information de référence partielle pour juger la qualité du signal déformé.

Les mesures quantitatives les plus utilisées sont : l'erreur quadratique moyenne (MSE), le rapport crête signal sur bruit (Peak Signal to Noise Ratio, PSNR), le rapport signal sur bruit (Signal to Noise Ratio :SNR) etc..



## 2.4.1 Méthodes avec référence complète

### 2.4.1.1 Erreur quadratique moyenne (MSE)

L'image dégradée  $\hat{I}$  est toujours comparée à l'originale  $I$  pour déterminer son rapport de ressemblance. Ce critère est le plus utilisé. Il est basé sur la mesure de l'erreur quadratique moyenne (MSE) calculée entre les pixels originaux et dégradés:

$$MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (I(m,n) - \hat{I}(m,n))^2 \quad (2.6)$$

Où  $(M \times N)$  est la taille de l'image, et  $I_p$  et  $\hat{I}_p$  sont respectivement les amplitudes des pixels sur les images originale et dégradée. Il est vraisemblable que l'œil tienne beaucoup plus compte des erreurs à grandes amplitudes, ce qui favorise la mesure quadratique.

### 2.4.1.2 Rapport crête signal sur bruit (PSNR)

Au lieu de mesurer la distorsion, cette valeur (Peak Signal to Noise Ratio, PSNR) mesure la fidélité, puisqu'elle est proportionnelle à la qualité. Tout de même, elle est une fonction de MSE ; sa définition et son utilisation proviennent du domaine du traitement de signal:

$$PSNR = 10 \log_{10} \left( \frac{I_{\max}^2}{MSE} \right) \quad (2.7)$$

Pour une image à niveau de gris,  $I_{\max}$  désigne la luminance maximale possible. Une valeur de PSNR infini correspond à une image non dégradée. Et cette valeur décroît en fonction de la dégradation. Le PSNR relie donc le MSE à l'énergie maximale de l'image.

### 2.4.1.3 Rapport signal sur bruit (SNR)

On utilise parfois une autre variante du rapport signal sur bruit (Signal to Noise Ratio : SNR), qui relie le MSE à l'énergie moyenne de l'image :

$$SNR = 10 \log_{10} \left( \frac{\frac{1}{N} \sum I^2}{MSE} \right) \quad (2.8)$$

#### 2.4.1.4 Indice de similarité structurelle (SSIM)

SSIM [4] 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:

$$\mu_x = \frac{1}{N} \sum_1^N x_i \quad (2.9)$$

$N$  : le nombre de pixels de chaque fenêtre.

$x_i$ : l'intensité d'un pixel.

Le contraste pour chaque fenêtre est mesuré par:

$$\sigma_x = \left( \frac{1}{N-1} \sum_1^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}} \quad (2.10)$$

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

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (2.11)$$

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 :

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (2.12)$$

Wang et Al ont proposé dans [4] une fonction de comparaison de la luminance de la forme :

$$l(x, y) = \frac{2\mu_x\mu_y + (K_1L)^2}{\mu_x^2 + \mu_y^2 + (K_1L)^2} \quad (2.13)$$

$K_1$  est une constante de très faible valeur. La constante  $(k_1L)^2$  évite l'instabilité de la fonction de comparaison lorsque  $\mu x^2 + \mu y^2$  sont très proches de zéro. Dans le cas des images à niveau de gris, la constante est  $L = 255$  (dynamique du capteur).

On remarque que l'équation (2.13) 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 :

$$c(x, y) = \frac{2\sigma_x\sigma_y + (K_2L)^2}{\sigma_x^2 + \sigma_y^2 + (K_2L)^2} \quad (2.14)$$

où  $K_2$  est une constante de très faible valeur.

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 :

$$s(x, y) = \frac{2\sigma_{xy} + (K_2L)^2}{2\sigma_x\sigma_y + (K_2L)^2} \quad (2.15)$$

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

$$SSIM(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y)$$

où encore sous une forme simplifiée:

$$SSIM(x, y) = \frac{2\mu_x\mu_y + (K_1L)^2}{\mu_x^2 + \mu_y^2 + (K_1L)^2} * \frac{2\sigma_{xy} + (K_2L)^2}{\sigma_x^2 + \sigma_y^2 + (K_2L)^2} \quad (2.16)$$

L'index de similarité est exploité pour un choix approprié des constantes  $K_1$  et  $K_2$  pour généraliser l'évaluation à toute l'image.

#### 2.4.1.5 Similitude structurale à base de gradient

Un algorithme amélioré de SSIM a été développé (la similitude structurale à base de Gradient (GSSIM)) [8], qui compare l'information de contour entre les blocs de l'image déformée et l'originale, et remplace la comparaison de contraste  $c(x, y)$  et la comparaison de structure  $s(x, y)$  par la comparaison de contraste à base gradient  $c_g(x, y)$  et la comparaison de la structure  $s_g(x, y)$  respectivement. Il y a un certain nombre de manières d'obtenir l'information de contour, telle que l'algorithme simple de détection de contour.

Deux cartes de gradient  $X'$  et  $Y'$  représentant respectivement l'image originale et déformée sont générées en utilisant l'opérateur de Sobel. En plus les blocs des images  $X'$  et  $Y'$  sont donnés par  $x'$  et  $y'$ . La comparaison de contraste à base de gradient  $c_g(x, y)$  et la comparaison de structure  $s_g(x, y)$  sont décrites par :

$$c_g(x, y) = \frac{2\sigma_{x'}\sigma_{y'} + C_2}{\sigma_{x'}^2 + \sigma_{y'}^2 + C_2} \quad (2.17)$$

$$s_g(x, y) = \frac{\sigma_{x'y'} + C_3}{\sigma_{x'}\sigma_{y'} + C_3} \quad (2.18)$$

Où  $\sigma_{x'}$  et  $\sigma_{y'}$  sont l'écart type des blocs  $x'$  et  $y'$  respectivement,  $\sigma_{x'y'}$  est la covariance des blocs  $x'$  et  $y'$ , et  $C_2, C_3$  sont des petites constantes évitant la dénomination de la valeur zéro. La similitude structurale à base de Gradient (GSSIM) est décrite comme suit :

$$GSSIM(x, y) = [l(x, y)]^\alpha \cdot [c_g(x, y)]^\beta \cdot [s_g(x, y)]^\gamma \quad (2.19)$$

La similitude globale de structure d'image est calculée comme la moyenne de toutes les sous-images (blocs) de GSSIM, les coefficients de pondérations ( $\alpha, \beta$  et  $\gamma$ ) sont fixées à la valeur '1'.

$$MGSSIM(X, Y) = \frac{1}{M} \sum_{i=1}^{Eb} GSSIM(x_i, y_i) \quad (2.20)$$

Où  $X$  et  $Y$  sont l'image originale déformée respectivement,  $M$  est le nombre de des blocs.

#### 2.4.1.6 Indice de la région d'intérêt (*VroiWQI*)

Les régions d'intérêt dans l'image originale sont modélisées par l'entropie  $e = -\sum_{i=1}^k p(H_i) \log_2 p(H_i)$  ( $H_i$  est une variable aléatoire indiquant l'intensité,  $p(H_i)$  est l'histogramme des niveaux d'intensité dans une région,  $K$  est le nombre de niveaux possibles d'intensité) pour chaque bloc de taille  $8 \times 8$  de l'image originale. En plus, l'entropie normalisée de chaque bloc est collectée dans une matrice  $E$  :

$$E = \begin{pmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{pmatrix} \quad (2.21)$$

D'abord l'image originale et l'image déformée sont divisées en bloc non chevauchant de taille  $8 \times 8$ . L'index de la qualité pour chaque bloc est calculé par la formule de l'entropie. Tous les indices (**L'index de similarité structurelle (SSIM)**) de l'image peuvent être collectés sous forme d'une matrice *SSIM*, où chaque élément  $q_{ij}$  représente la mesure de la distorsion entre un bloc de l'image originale et son bloc correspondant de l'image déformée :

$$SSIM = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ q_{m1} & q_{m2} & \cdots & q_{mn} \end{pmatrix} \quad (2.22)$$

Ensuite, la matrice  $E$  est obtenue à partir de l'image originale et l'index *SSIM* est obtenu des blocs de l'image originale et déformée. L'équation de *VroiWQI* [16] est donnée comme suit:

$$VroiWQI = \frac{\sum_{l=1}^M \sum_{j=1}^N [E] * [SSIM]}{\sum_{l=1}^M \sum_{j=1}^N [E]} \quad (2.23)$$

Où  $M$ ,  $N$  sont la longueur et la largeur des images respectivement,  $E$  est la matrice de l'entropie normalisée,  $SSIM$  est la matrice de l'index de similarité structurelle.

## 2.5 Mesures inspirées du Système Visuel Humain

Dans cette partie, on va décrire, quelques approches relatives au système visuel humain (SVH). Des notions de bases sont introduites dans la section suivante.

### 2.5.1 Système visuel humain

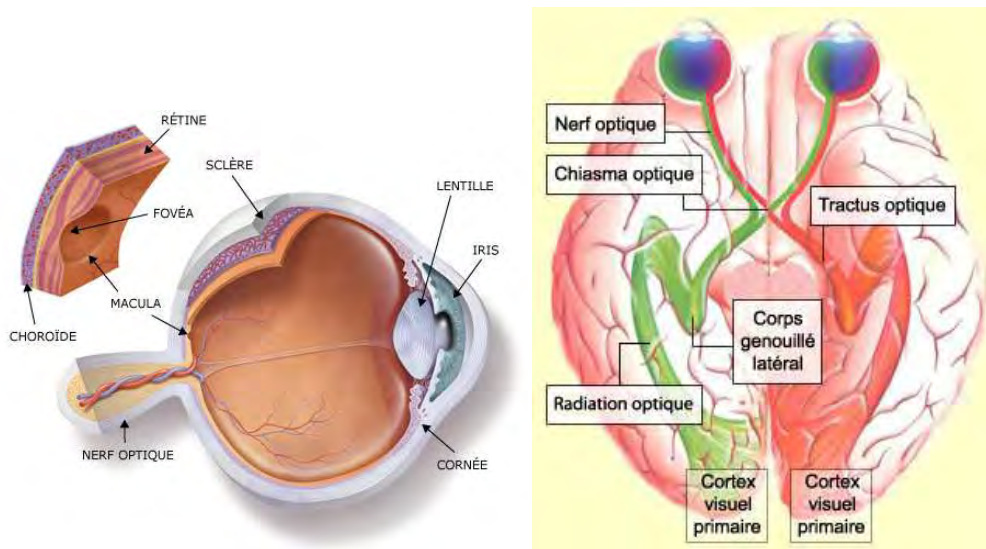
Plusieurs méthodes d'évaluation de la qualité d'image sont inspirées à partir des propriétés du système visuel humain (SVH). Nous allons décrire dans cette partie le SVH en premier. Nous commençons avec la structure physique qui est assez bien établi et qui peut nous aider à comprendre les caractéristiques assez complexe du comportement perceptif. Ensuite, nous présenterons quelques caractéristiques du système visuel humain. Enfin, nous introduisons quelques travaux relatifs au SVH [82].

#### 2.5.1.1 Œil humain

L'**œil** possède plusieurs parties assurant chacune une fonction différente (voir figure 2.30).

La **sclérotique** est la partie blanche et rigide qui forme le globe oculaire. Elle remplit un rôle de protection de l'œil car elle est formée d'un tissu dense.

La **cornée** remplace la sclérotique du côté antérieur. Elle est transparente et permet ainsi la pénétration des rayons lumineux dans le globe oculaire. La cornée est la principale lentille de l'œil. Elle assure en effet près de 80% de la réfraction. Plus interne que ces enveloppes protectrices se trouvent la **tunique vasculaire**, elle est composée de trois parties : la **choroïde**, le **corps ciliaire** et l'**iris**.



**Figure 2.30:** Schéma de l'œil et système visuel humain

La **choroïde** est une membrane fortement vascularisée. Elle recouvre les trois cinquièmes postérieurs du globe oculaire. Elle nourrit les **photorécepteurs** de la **rétine** grâce à ses nombreux vaisseaux sanguins.

Du côté antérieur, l'**iris** remplace la choroïde. C'est la partie colorée de l'œil. L'iris est composé de muscles lisses, il s'ouvre ou se ferme pour changer la taille de la **pupille** (son ouverture centrale) afin de moduler la quantité de lumière entrant dans l'œil.

Le **corps ciliaire** modifie la taille du **crystallin** grâce à ses muscles lisses. Le cristallin est une lentille molle composée de fines couches superposées, il filtre la lumière. La mise au point et la formation d'une image nette de l'objet sur la rétine dépend de la forme du cristallin.

La **rétine** est l'endroit où l'énergie lumineuse est transformée en activité nerveuse. Elle est sensible à la lumière grâce à ses deux types de **photorécepteurs** : les cônes et les bâtonnets.

Les **cônes** sont environ 5 à 7 millions dans la **fovéa** (une région qui forme une dépression au centre de la rétine et où la perception du plus petit détail possible est à son maximum). Les cônes sont peu sensibles à la lumière mais ils permettent la vision des couleurs et nous assurent une capacité à distinguer les détails. Cela est dû au fait que chaque cône de la fovéa transmet son information à plusieurs fibres du nerf

optique. Les cônes servent donc à la vision de jour. La taille des cônes varie de 420 nm à 565 nm. Les cônes contiennent un pigment appelé **rhodopsine**, qui est décomposé et décoloré à la lumière. Ce procédé de **décomposition** crée un potentiel électrique qui transforme l'énergie de la lumière en **impulsion nerveuse**. Celle-ci est transmise au cerveau par l'intermédiaire du **nerf optique**. Il existe trois types de cônes

- Cône contenant de la cyanopsine : sensible au bleu.
- Cône contenant de la chloropsine : sensible au vert.
- Cône contenant de l'erythropsine : sensible au rouge.

Les **bâtonnets** (Rods) sont beaucoup plus nombreux que les cônes, il y en a environ 130 millions. La taille des bâtonnets est d'environ 0.06 mm de long et 0.25 mm d'épaisseur. Ils sont répartis tout autour de la fovéa. Les bâtonnets sont très sensibles à la lumière et ils servent à la vision de nuit en raison de leur capacité à percevoir de faibles lueurs dans l'obscurité. Contrairement aux cônes, ils ont une faible sensibilité aux couleurs et aux détails car des dizaines de bâtonnets ne sont liés qu'à une seule fibre du nerf optique.

### 2.5.1.2 Rôle du cerveau

L'œil n'est qu'un ensemble de capteurs recueillant des informations mais ne pouvant pas les interpréter. Il transforme ces informations en messages nerveux transmis au cerveau (voir figure 2.30).

Le message nerveux est transmis par l'œil au cerveau via le nerf optique. Les nerfs optiques vont des yeux jusqu'à l'arrière du cerveau, dans le cortex visuel primaire (ou aire V1) de chaque hémisphère, celui-ci ayant un rôle majeur dans la vision des êtres humains. Les nerfs optiques de chaque œil se rejoignent au niveau du chiasma optique. Le chiasma permet la répartition de l'information visuelle. Le champ visuel de l'œil est divisé en deux, chaque hémisphère perçoit une des deux moitiés. Au niveau du chiasma optique, il y a inversion, l'hémi-champ gauche va dans l'hémisphère droit et vice-versa. Grâce à cela, le cerveau est beaucoup moins vulnérable car si un des nerfs est lésé, les deux hémisphères pourront recevoir des informations provenant des deux yeux. Le cortex visuel occupe le lobe occipital du cerveau (la nuque). C'est lui qui traite les informations provenant des yeux et qui



constitue l'image que nous percevons après le traitement. Les informations visuelles arrivent dans le cortex visuel primaire (ou aire V1). Celles qui viennent du centre du champ visuel sont traitées par la partie occipitale de l'aire V1. Après cela, les informations sont séparées en deux voies distinctes au delà de l'aire V2 :

- La voie ventrale : dans laquelle les informations de forme et de couleurs sont analysées afin d'obtenir une reconnaissance des objets vus.
- La voie dorsale : dans laquelle les informations liées au mouvement et aux positions dans l'espace sont analysées de façon à permettre les mouvements en direction ou non des objets.

## **2.5.2 Caractéristiques du Système Visuel Humain**

### **2.5.2.1 Fovéation**

Les densités des cellules de cônes et les cellules de ganglions dans la rétine ne sont pas uniformes. 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. La vision à haute résolution due à la fixation par l'observateur sur une région s'appelle la vision fovéale, alors que la vision progressivement inférieure de résolution s'appelle la vision périphérique. La plupart des modèles d'évaluation de la qualité d'image fonctionnent avec la vision fovéale. Les modèles peuvent également échantillonner en deuxième fois l'image avec la densité de prélèvement des récepteurs dans la fovéa afin de fournir une meilleure approximation du SVH aussi bien que fournir un calibrage plus robuste du modèle [54, 55 , 14].

### **2.5.2.2 Phénomène d'adaptation à la lumière**

Le 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. Il fait face à une gamme si étendue par un phénomène connu sous le nom d'adaptation à la lumière, qui fonctionne en contrôlant la quantité de lumière entrant dans l'œil par la pupille, aussi bien que des mécanismes d'adaptation dans les cellules rétiniennes qui ajustent le gain des neurones post-récepteur dans la

réfine. Le résultat est que la réfine chiffre le contraste du stimulus visuel au lieu des intensités de la lumière absolue. Le phénomène qui maintient la sensibilité de contraste du SVH sur un éventail d'intensité de la lumière de fond est connu comme loi de Weber [14,56, 57].

### **2.5.2.3 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 [14,56-59].

### **2.5.2.4 Effet de masquage**

Masque/facilitation se rapporte au fait que la présence d'un composant d'image (le signal) diminuera/augmentera la visibilité d'un autre composant d'image (le masqueur). Le masque réduit généralement la visibilité du signal en comparaison du cas de l'absence du masque. Habituellement, l'effet de masque est plus fort quand le masque et le signal ont le même contenu de fréquence. La plupart des méthodes d'évaluation de qualité incorporent un modèle de masque, alors que certains incorporent également la facilitation [14, 56, 57, 60, 61].

## **2.5.3 Méthodes basées sur SVH**

Dans cette partie, nous présentons quelques métriques d'évaluation de la qualité d'images qui sont basées sur le paradigme de sensibilité à l'erreur du SVH [14,56, 57].

### **2.5.3.1 Algorithme de Daly**

Il vise à calculer une probabilité de la carte de détection entre l'image référence et l'image déformée [5]. La valeur à chaque point dans la carte est la probabilité qu'un observateur humain percevra une différence entre la référence et les images déformée à ce point. La référence et les images déformées (exprimées en valeurs de luminance) subissent une série de processus : le filtrage de CSF, la décomposition de canal, calcul de contraste, modélisation d'effet de masquage et la probabilité de calcul de la détection. La transformée en Cortex est employée pour la décomposition de canal, qui transforme le signal d'image en cinq niveaux spatiaux suivis de six niveaux d'orientation, menant à un total de 31 sous-images (y compris la

bande de base). Pour chaque canal, une carte d'altitude de seuil est calculée à partir du contraste dans ce canal. Une fonction psychométrique est employée pour convertir des forces d'erreur en carte de probabilité de détection pour chaque canal.

### **2.5.3.2 Algorithme de Lubin**

Il estime une probabilité de détection des différences entre les versions déformée et originale. Un flou Gaussien est appliqué pour modéliser la fovéation du système optique d'œil. Une pyramide laplacienne est employée pour décomposer les images en sept bandes, suivie par des calculs de contraste de chaque bande [54,55].

### **2.5.3.3 Métrique de Watson**

Elle est basée sur la transformée DCT (Transformée en Cosinus Discrète)  $8 \times 8$ . Elle est utilisée généralement dans la compression d'image et vidéo. À la différence des modèles ci-dessus, cette méthode divise le spectre dans 64 sous-bandes uniformes (8 dans chaque dimension cartésienne) [62].

## **2.6 Conclusion**

Dans ce chapitre, nous avons présenté un état de l'art sur les méthodes d'évaluation de la qualité d'image, ainsi que, leurs définitions. Ensuite, les différents protocoles nécessaires pour effectuer une évaluation subjective sont présentés. Les méthodes d'évaluation de la qualité d'images se divisent en deux grandes familles ; les méthodes subjectives basées sur l'observateur humain. Ces méthodes sont coûteuses au sens du matériel et du temps. En plus, l'implication des observateurs non experts dans la procédure d'évaluation pour affecter le score donné à une image. Les deuxièmes techniques sont les méthodes objectives. Où l'introduction de l'observateur humain n'est pas nécessaire. Les résultats obtenus par ces méthodes doivent être conformes à l'évaluation subjective. Quelques méthodes d'évaluation objectives sont présentées dans ce chapitre telles que PSNR, MSSIM, etc. Le chapitre se termine par une introduction au système visuel humain. Ses caractéristiques et quelques méthodes basées sur SVH. Dans le chapitre suivant, la méthode d'évaluation proposée intégrée dans les méthodes d'évaluation objective est introduite.

# Chapitre 3

## Méthode proposée

### 3.1 Introduction

Au fil des années, plusieurs chercheurs ont cherché des nouvelles mesures de qualité, mieux adaptée à la perception humaine [9-11]. Il est important d'évaluer la performance de ces mesures dans un contexte comparatif et en analysant leurs pouvoirs et faiblesses.

Comme nous l'avons vu dans le chapitre précédent, les méthodes d'évaluation de la qualité d'images [7, 14-15, 35-37] peuvent être divisées en méthodes subjectives et méthodes objectives [12-14]. La première approche de l'évaluation de la qualité d'image est la méthode de test subjectif basée sur plusieurs observateurs humains estimant la qualité d'image. Cette dernière exige un temps d'exécution considérable en impliquant un grand ensemble d'observateurs. La deuxième approche est la mesure d'évaluation objective [24-29, 30, 31-34, 38] basée sur des calculs mathématique. Ces types de méthodes sont simples et rapides puisque les observateurs humains ne sont pas demandés. L'une des propriétés des mesures d'évaluations de la qualité d'image est la production des résultats qui sont en accord avec l'évaluation visuelle. En outre, quelques méthodes objectives utilisant le système visuel humain (SVH) [5, 6] sont proposées. Ces modèles de SVH présentent une partie de la vision humaine (telles que la résolution spatiale, le mouvement temporaire, résolution de couleur...). La majorité de ces méthodes réclament l'image de test et l'image de référence afin de déterminer la différence perceptuelle entre elle.

La plupart des méthodes d'évaluation de la qualité d'image objective existant dans la littérature sont des mesures à base d'erreur. Ces techniques utilisent les métriques à base de différences de pixels telles que : la moyenne de l'erreur absolue (MAE), l'erreur quadratique moyenne ( $MSE$ ), la racine de  $MSE$ , rapport crête signal sur bruit (PSNR) et rapport signal sur bruit (SNR). Cependant, nombreux de ces algorithmes ne peuvent pas bien corrélérer avec la qualité perçue [1].

Les méthodes d'évaluation de la qualité proposées ont suivi une stratégie de modification de la mesure  $MSE$  afin que les erreurs donnent une valeur conforme à la qualité visuelle [4, 15].

Plusieurs chercheurs ont montré que la fonction principale de l'œil humain est d'extraire les informations de contours ou structures du champ de vision, en plus SVH est très adapté à ce but [15]. Ce dernier est sensible à l'information contour et région. En effet, les images naturelles sont très texturées, et plus précisément les échantillons des signaux sont fortement corrélés. Alors que, toute sorte d'artefact crée des distorsions de pixels. Par conséquent, dans ce travail, nous avons développé une nouvelle mesure de la qualité d'évaluation basée sur l'information contour-région et les pixels déformés et déplacés.

### 3.2 Etat de l'art

Plusieurs travaux de recherche évaluent la qualité d'image de manières différentes. Dans cette partie nous introduisons les limitations des mesures présentées dans le chapitre précédent.

$MSE$  calcule l'erreur quadratique moyenne des pixels de l'image dégradée et de référence, ainsi que les quantités de PSNR liées. Ces deux mesures sont simples à implémenter et à calculer. Mais elles sont mal corrélées à la qualité visuelle perçue [1]. La métrique de qualité universelle ( $UQI$ ) [2] est basée sur la comparaison de la structures des images déformée et référence et présente la déformation d'image comme une combinaison de trois facteurs : la perte de la corrélation, la déformation de la luminance, et la déformation du contraste. Cependant, les mesures  $d'UQI$  [2] donnent une importance égale pour toutes les déformations dans chaque région de l'image.

Dans [4], Zhou Wang a proposé une nouvelle idée de l'indice de similarité structurelle (SSIM) pour la mesure de la qualité d'image basée sur l'hypothèse que le SVH est très adapté pour l'extraction de l'information structurale du champ

d'observation. L'indice de similarité structurel moyenne (MSSIM) comporte trois parties: comparaison de la luminance  $l(x, y)$ , comparaison du contraste  $c(x, y)$  et la comparaison de la structure  $s(x, y)$ . Toutefois, MSSIM a été détecté pour être moins compétitif lorsqu'il est utilisé pour évaluer les images floues et bruitées [8].

Il est évident que les contours et les régions des images ont une importance dans la perception de la vision. *MSSIM* ne tient pas compte des facteurs comme l'importance visuelle des caractéristiques de l'image. Quelques chercheurs ont étudié la possibilité d'améliorer les performances de *MSSIM* par accorder une importance des poids visuel aux valeurs de *MSSIM*.

Dans [8], les auteurs ont proposé une méthode améliorée de SSIM appelée similitude structurale à base de Gradient (GSSIM) en se basant sur l'information de contour comme l'information de structure de l'image la plus importante. L'opérateur de Sobel est utilisé pour extraire l'information de contour en raison de sa simplicité[4]. Cet algorithme est très intéressant dans l'utilisation de l'information de contour avec MSSIM mais il ne peut pas être utilisé pour mesurer la qualité de l'information de région.

Dans [16], un indice de la région d'intérêt est présenté. L'indice de la région d'intérêt ne prend pas en compte l'information de contour dans le processus de l'évaluation de la qualité d'image.

Motivés par ces études, nous proposons dans ce travail, d'adopter une approche de bas niveau pour le problème. Par conséquent, nous cherchons à évaluer la qualité des images (référence et déformée) avec la différence des régions et des contours. Plus précisément, nous utilisons l'opérateur de Sobel pour trouver l'information contour et l'opérateur entropie pour produire les régions visuelles de la référence et de l'image déformée. Par ailleurs, les mesures de pixels déplacés et déformés sont utilisées dans la méthode proposée. L'information contour-région, pixels déformés et déplacés (ERDDM- edge-region and pixel deformed with pixel displacement measure) [40] qui n'exploite pas une connaissance a priori sur l'image déformée et des types d'artefacts sont introduites.

### 3.3 Travaux précédents

Des instances précédentes de cette méthode sont faites dans [17-22, 39, 41, 42] et des résultats prometteurs sur des tests simples sont accomplis. Premièrement [21], les images déformées et originales sont divisées en blocs de  $8 \times 8$ , ensuite le déplacement

de position et la différence de couleur sont calculés, ce qui peut être utilisé pour calculer l'erreur entière. Deuxièmement [22], l'image déformée et l'image originale sont transformées dans l'espace couleur  $L^*a^*b^*$ . Puis, nous suivons les mêmes étapes que les travaux antérieurs. Troisièmement [20], nous proposons une amélioration de la moyenne de l'indice de similarité structurelle (MSSIM) en introduisant la position et l'intensité à l'intérieur de l'erreur. Quatrièmement [19], l'information de contour et le déplacement des pixels sont utilisés avec la méthode MSSIM. Cinquièmement [18,41], la région d'intérêt avec le déplacement et la déformation de pixels sont utilisés pour améliorer MSSIM. Sixièmement [17], l'information de contour et de la région avec des pixels déformés sont introduits pour couvrir certaines limites des travaux antérieurs. La mesure proposée est calculée en utilisant la base d'Images LIVE [3], et les résultats trouvés ont été promoteurs. Dans [39], nous avons amélioré la mesure proposée dans [17] en introduisant à l'intérieur la mesure du pixel déplacé. Les résultats sont capables d'imiter la perception visuelle sur la base d'Images LIVE.

### 3.4 Méthode proposée (ERDDM)

Afin d'évaluer la qualité d'une image déformée en utilisant une méthode d'évaluation avec référence complète, la version déformée de cette dernière est comparée avec la version complète (référence ou originale)

Avant d'introduire notre algorithme quelques notions doivent être présentées. L'image originale et l'image déformée sont respectivement notées par  $Original(m, n)$  et  $Distorted(m, n)$ . Ces deux images ont une taille de  $M \times N$  pixels. En plus, elles sont partitionnées en blocs chevauchant de taille  $11 \times 11$ , où la région chevauchante est d'un seul pixel. La méthode d'évaluation de la qualité est calculée dans les blocs chevauchant de taille  $11 \times 11$  menant à une carte de qualité de l'image. La valeur globale de la mesure de la qualité est la moyenne de la carte de la qualité.

Le choix du chevauchement de taille  $11 \times 11$  est motivé. Par le fait que la subdivision de l'image en blocs chevauchant de taille  $8 \times 8$  montre souvent des problèmes indésirables comme l'effet de bloque [7]. Ainsi, ce genre de problème est évité et disparaît, quand une partition supérieure à  $8 \times 8$  pixels est employée. En outre, la performance de la mesure développée est optimale quand le chevauchement  $11 \times 11$  est choisi.

Dans le chapitre 4, une étude comparative entre les résultats de la mesure proposée est effectuée (justifiant la partitionnement de l'image en blocs de taille  $11 \times 11$ ). À cette fin, la performance de la mesure est calculée en utilisant trois types de chevauchement de blocs,  $8 \times 8$ ,  $11 \times 11$  et  $15 \times 15$ .

Par conséquent, toutes les variables utilisées pour la mesure proposée sont définies comme suit:

*db*: bloc chevauchant de taille  $11 \times 11$  pixels de l'image déformée.

*ob*: bloc chevauchant de taille  $11 \times 11$  pixels de l'image originale.

*DM*: mesure de différence.

*M<sub>f<sub>i</sub></sub>*

*X<sub>f<sub>ij</sub></sub>*

*M<sub>i</sub>*: la moyenne des pixels des blocs de taille  $3 \times 3$  de l'image originale.

*X<sub>ij</sub>*: l'intensité des pixels dans l'image originale.

*ODM*: La mesure globale des pixels déformés entre les deux blocs *ob* et *db*.

*sdb*: sous-blocs chevauchant de taille  $3 \times 3$  de l'image déformée.

*sob*: sous-blocs chevauchant de taille  $3 \times 3$  de l'image originale.

*Gdb*: ensemble de tous les centres de la gravité de *db*; dans la figure 3.2 *Gdb* inclut vingt centre de la gravité.

*Gob*: ensemble de tous les centres de gravité de *ob*; dans la figure 3.2 *Gob* inclut vingt centre de la gravité.

*G<sub>i</sub>*: centre de gravité appartenant au *Gob* de l'image originale.



$v_i(x_i, y_i)$ : pixel du centre de gravité de  $Gdb$ .

$d((x_i, y_i), (x'_i, y'_i))$ : Distance Euclidienne entre le centre de gravité des deux pixels.

$diff(v_i(x_i, y_i), Gob)$ : la distance entre le pixel du centre de gravité ' $v_i(x_i, y_i)$ ' et  $Gob$ .

$Dp$ : mesure du déplacement du pixel du centre de gravité de  $Gdb$  de l'image déformée.

$DPDM$ : Mesure globale des pixels déformés et déplacés entre deux blocs  $ob$  et  $db$ .

$K$ : nombre possible de niveaux d'intensité.

$p(H_i)$ : l'histogramme d'une région.

$H_i$ : variable aléatoire indiquant l'intensité.

$e$ : entropie du bloc  $ob$ .

$s_R(ob, db)$ : Mesure de la comparaison de la structure d'une région.

$Xorg$ : représente la carte de gradient de l'image originale.

$Ydist$ : représente la carte de gradient de l'image déformée.

$ob'$ : bloc (sa taille est  $11 \times 11$  pixels) de  $Xorg$ .

$db'$ : bloc (sa taille est  $11 \times 11$  pixels) de  $Ydist$ .

$s_E(ob', db')$ : représente la mesure de l'information de contour.

$ERDDM$ : mesure de contour-région, pixels déformés et déplacés.

$E_b$ : nombre total des blocs.

### 3.4.1 Mesure de pixels déformés et déplacés

Les déformations des pixels sont estimées en utilisant les caractéristiques suivantes. D'abord, la mesure de différence ( $DM$ ) d'un pixel central est estimée pour les pixels du voisinage  $3 \times 3$ . Ensuite, elle est généralisée pour tous les pixels centraux disponibles dans l'image. Les valeurs de  $DM$  sont alors ramenées à une moyenne du bloc de chevauchement de taille  $3 \times 3$  (voir figure 3.1).

Soient  $X_{23}$  et  $Xf_{23}$  les pixels centraux de voisinage  $3 \times 3$  dans l'image originale et déformée respectivement comme montré dans la figure 3.1 (c)(d); soient  $Mf_i$  la moyenne des blocs de taille  $3 \times 3$  pixels de l'image déformée et  $Xf_{ij}$  son intensité.

Tous les pixels centraux disponibles dans les images (originale et déformée) sont situés dans la ligne grasse comme montré dans la figure 3.1 (a)(b).

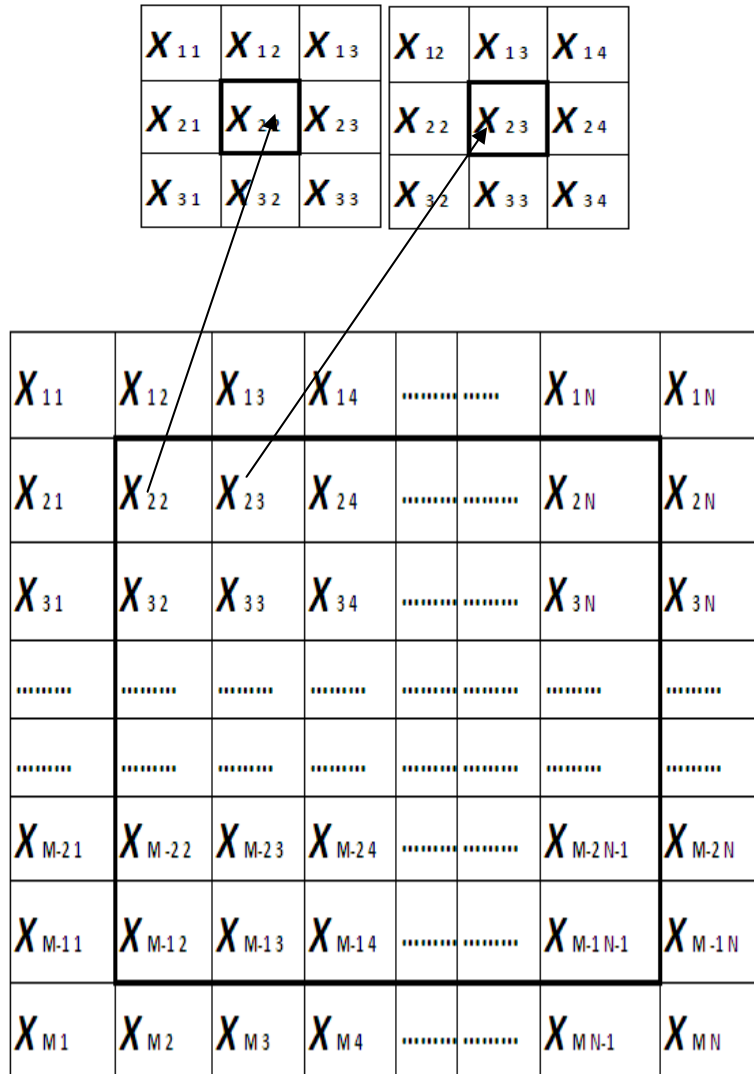
La moyenne peut être estimée comme suit :

$$Mf_i = \frac{1}{9} \sum_{i=1}^3 \sum_{j=1}^3 Xf_{ij} \quad (3.1)$$

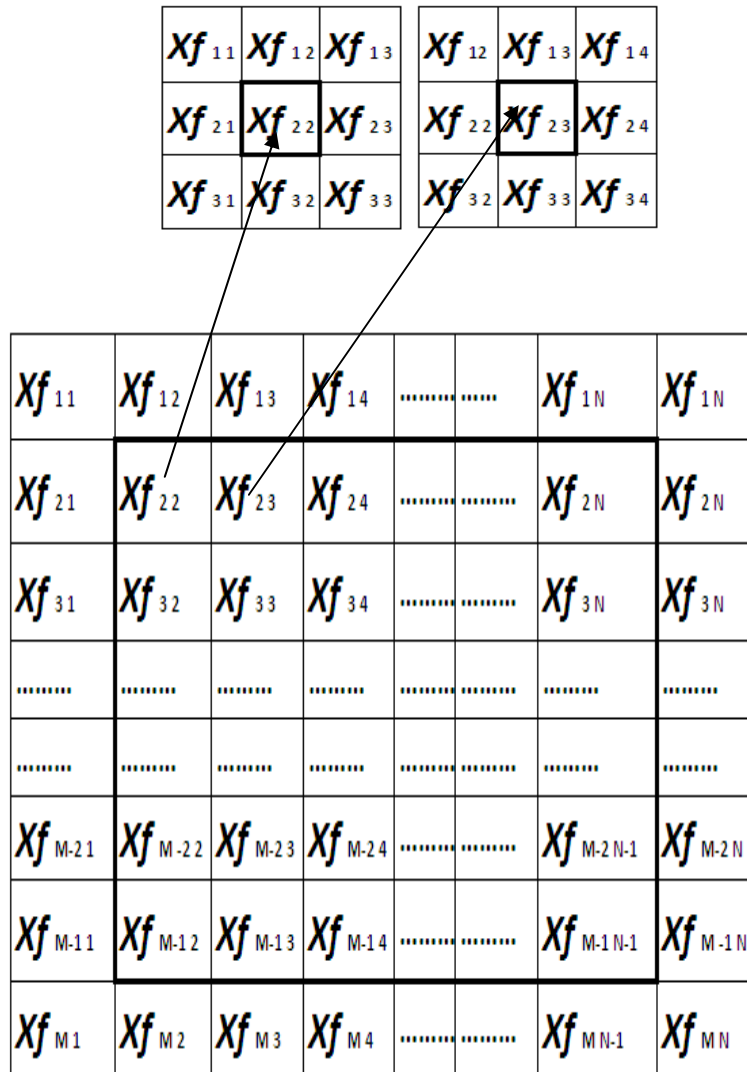
La même procédure est appliquée sur l'image originale. Celle-ci a également une moyenne  $M_i$  et  $X_{ij}$  son intensité. Pour chaque bloc chevauchant de taille  $3 \times 3$  de l'image originale, la moyenne est calculée comme suit :

$$M_i = \frac{1}{9} \sum_{i=1}^3 \sum_{j=1}^3 X_{ij} \quad (3.2)$$

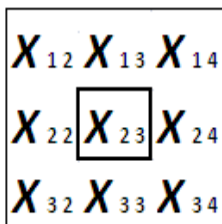
La mesure de différence ( $DM$ ) entre les deux blocs chevauchant de taille  $3 \times 3$  de l'image originale et déformée peut être estimée comme suit:



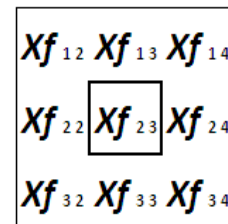
(a) image originale



(b) Image déformée



(c)



(d)

**Figure 3.1** (a) image originale (b) image déformée (c) pixel central  $X_{23}$  de taille  $3 \times 3$  de l'image originale (d) pixel central  $Xf_{23}$  de taille  $3 \times 3$  de l'image déformée.

$$DM = \left( \frac{1}{9} \sum_{i=1}^3 \sum_{j=1}^3 |(Mf_i - X_{ij})^2 - (M_i - Xf_{ij})^2| \right)^{\frac{1}{2}} \quad (3.3)$$

La mesure déformation globale de pixels  $ODM$  entre deux blocs  $ob$  et  $db$  est définie par:

$$ODM = \frac{1}{20} \sum_{i=1}^{20} DM_i \quad (3.4)$$

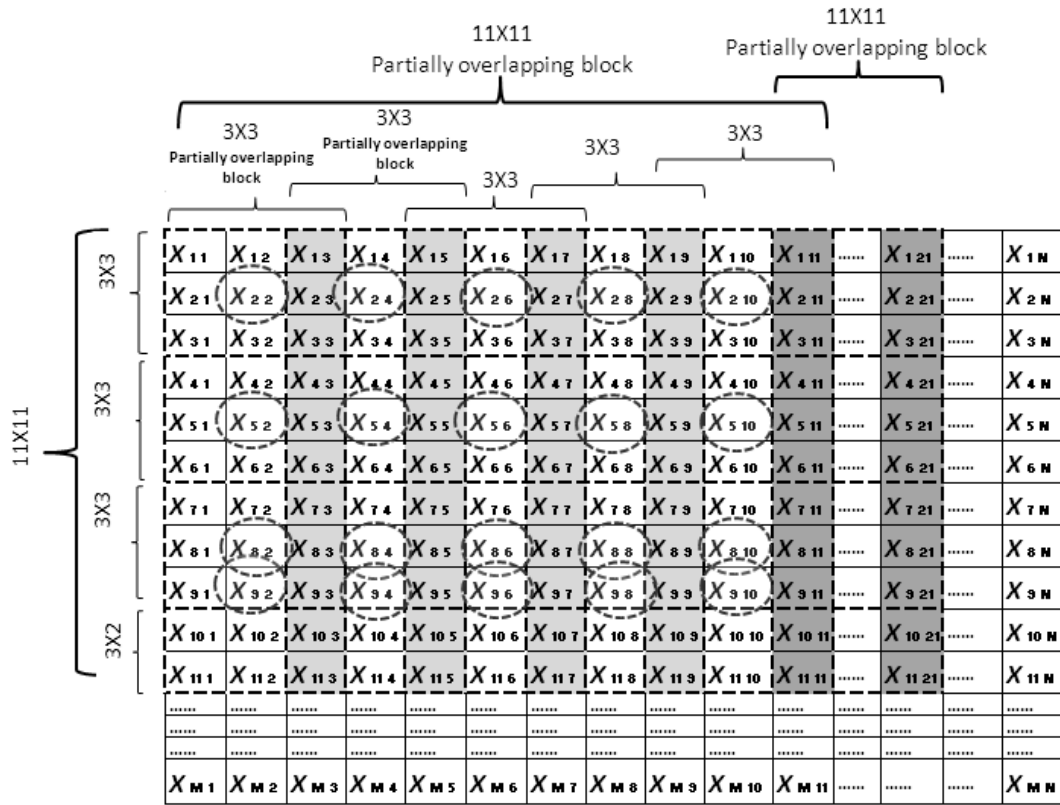
Où 20 dénote le nombre de blocs de taille  $3 \times 3$  pixels dans le bloc du  $11 \times 11$  pixels. Tout  $ODM$  peut être rangé dans la matrice  $ODM_{global} = [ODM_{i,j}]$ .

Afin de refléter les différences entre les deux images au niveau local, nous avons calculé des centres de gravité des images originale et déformée sur la base des blocs  $11 \times 11$  pixels. Dans cette étape, la mesure du déplacement de pixels sera calculée. Ce critère fonctionne en trois étapes :

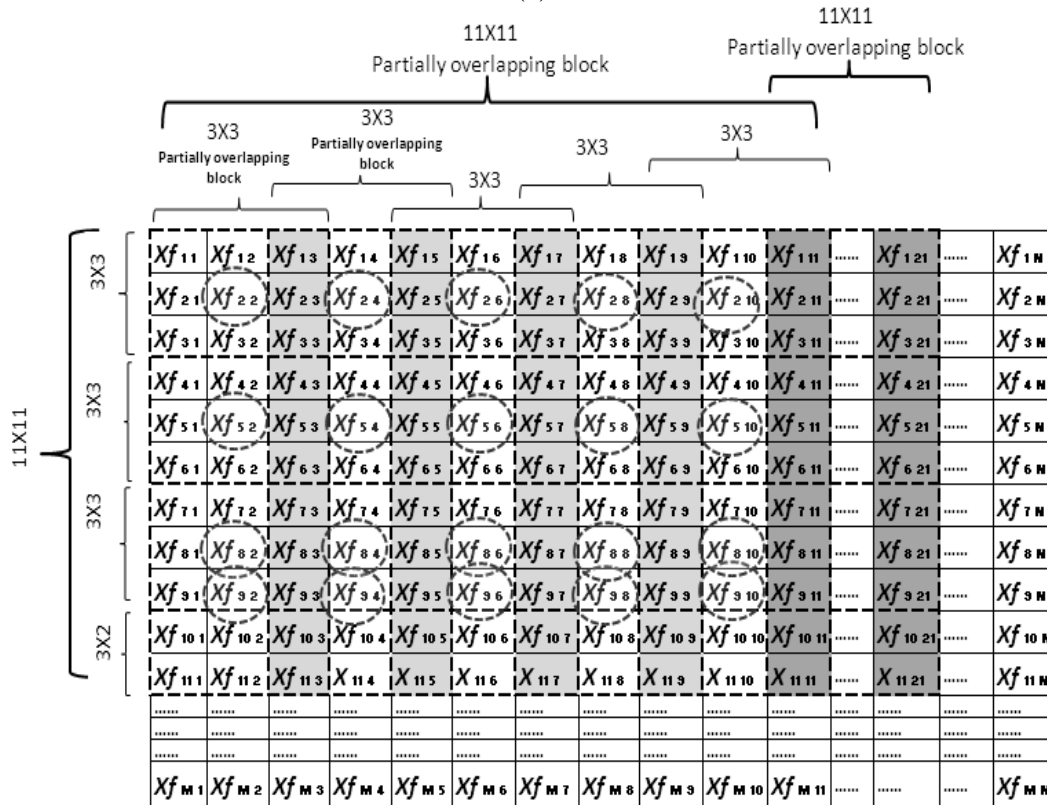
1. La première étape consiste à identifier tous les centres de gravité des sous-blocs (pixels de voisinage  $2 \times 3$  et pixels de voisinage  $3 \times 3$ ) de  $db$  et  $ob$ . En outre, ces blocs ( $db$  et  $ob$ ) sont divisés en sous-blocs ( $sdb$  et  $sob$ ) chevauchant de taille  $3 \times 3$  de l'image déformée et originale respectivement.
2. Dans la deuxième étape, la distance entre un centre gravité de  $Gdb$  et tous les centres de gravité de  $Gob$  de l'image originale est mesurée. Tous les centres de gravité disponibles des images (originale et déformée) sont situés dans les cercles pointillés comme il est montré dans figure 3.2.
3. La troisième étape calcule tous les déplacements des pixels des centres de gravité liée à  $Gdb$  de l'image déformée.

Cette étape est basée sur le calcul du centre de gravité de sous-blocs défini par des moments d'ordre  $p$  et  $q$  :

$$mm_{p,q} = \sum_{xx=begin1}^{end1} \sum_{yy=begin2}^{end2} xx^p yy^q \quad (3.5)$$



(a)



(b)

**Figure 3.2 :** Image et ses centres de la gravité: (a) image originale ( $M \times N$ ), (b) image déformée ( $M \times N$ )

Où  $end1$  et  $end2$  sont la longueur et la largeur d'un sous-bloc;  $begin1$  et  $begin2$  sont les positions du premier pixel dans le sous-bloc; par exemple  $begin1=1$  et  $begin=1$  sont la position du premier pixel ( $X_{11}$ ) du sous-bloc de taille  $3 \times 3$  (voir figure 3.2).

Le centre de gravité est donné par :

$$G = \left( \frac{mm_{10}}{mm_{00}}, \frac{mm_{01}}{mm_{00}} \right) \quad (3.6)$$

Où  $mm_{00}$  est la surface du sous-bloc,  $mm_{10}$  est la somme des positions des pixels  $xx$  qui appartiennent au sous-bloc,  $mm_{01}$  sont la somme de positions des pixels  $yy$  qui appartiennent au sous-bloc.

Ce critère calcule la distance entre les centres de gravités de l'image déformée et l'image originale. Il prend le centre de gravité de l'image déformée de  $Gdb$  et calcule le déplacement du pixel dans  $Gob$  de l'image originale.

On définit ce critère pour chaque centre de gravité ' $v_i(x_i, y_i)$ ' de  $Gdb$  en utilisant l'équation (3.7) :

$$diff(v_i(x_i, y_i), Gob) = \begin{cases} 0 & \text{if } x_i = x'_i \text{ and } y_i = y'_i \text{ and } f(v_i) = f(G_i) \\ \min(position, color) & \text{if } x_i = x'_i \text{ and } y_i = y'_i \\ & \text{and } f(v_i) \neq f(G_i) \end{cases} \quad (3.7)$$

Où :

$$position = \min \left( d \left( (x_i, y_i), (x'_1, y'_1) \right), \dots, d \left( (x_i, y_i), (x'_{20}, y'_{20}) \right) \right) :$$

si  $x_i = x'_i$  et  $y_i = y'_i$ ;  $d((x_i, y_i), (x'_i, y'_i))$  est supprimé de la *position*,

$$color = \min(|f(v_i) - f(G_1)|, \dots, |f(v_i) - f(G_{20})|) \quad ,$$

$$Gob = \{G_1(x'_1, y'_1), \dots, G_i(x'_i, y'_i), \dots, G_{20}(x'_{20}, y'_{20})\},$$

$G_i$  : centre de gravité appartenant à  $Gob$  de l'image originale.

$$Gdb = \{v_1(x_1, y_1), \dots, v_i(x_i, y_i), \dots, v_{20}(x_{20}, y_{20})\}, v_i(x_i, y_i) \in Gdb.$$

$f$  : peut être l'intensité de pixel du centre du gravité ou tout autre attribut,

$d((x_i, y_i), (x'_i, y'_i))$  est la distance Euclidienne entre les deux centre de gravité.

$x_i$  et  $y_i$  : les positions de centre de gravité de ' $v_i$ ' dans  $Gdb$ .

$x'_i$  et  $y'_i$  : les positions de centre de gravité de  $G_i$  dans  $Gob$ .

$diff(v_i(x_i, y_i), Gob)$  est la distance entre le centre de gravité ' $v_i(x_i, y_i)$ ' et  $Gob$ .

Pour illustrer cette formule, l'exemple suivant l'explique (voir figure 3.2).

Soit  $X_{f_2 2}$  un pixel du centre de gravité dont le voisinage  $3 \times 3$  pixels dans l'image déformée comme il est montré sur la figure 3.2 (b); ses positions sont  $x_I=2$  et  $y_I=2$  et son intensité est  $f(X_{f_2 2})=40$ .

Soient  $Gob=\{X_{2 2}, X_{2 4}, X_{2 6}, X_{2 8}, X_{2 10}, X_{5 2}, X_{5 4}, X_{5 6}, X_{5 8}, X_{5 10}, X_{8 2}, X_{8 4}, X_{8 6}, X_{8 8}, X_{8 10}, X_{9 2}, X_{9 4}, X_{9 6}, X_{9 8}, X_{9 10}\}$  les pixels représentant centre de gravité des pixels de voisinage  $11 \times 11$  pixels (voir figure 3.2 (a)) dans l'image originale, et leurs intensités sont  $f(X_{2 2})=52, f(X_{2 4})=35, f(X_{2 6})=42, f(X_{2 8})=55, f(X_{2 10})=25, f(X_{5 2})=16, f(X_{5 4})=20, f(X_{5 6})=88, f(X_{5 8})=45, f(X_{5 10})=42, f(X_{8 2})=77, f(X_{8 4})=14, f(X_{8 6})=13, f(X_{8 8})=55, f(X_{8 10})=22, f(X_{9 2})=45, f(X_{9 4})=45, f(X_{9 6})=49, f(X_{9 8})=54$  et  $f(X_{9 10})=12$ . Nous voyons que  $f(X_{f_2 2}) \neq f(X_{2 2})$  alors, nous prenons la deuxième partie de la formule pour calculer la mesure,

Et

$$diff(X_{f_2 2}, Gob) = \min(\text{position}, \text{color})$$

Où

$$\begin{aligned} \text{position} = \min & \left[ [(2-2)^2 + (4-2)^2]^{1/2}, \right. \\ & [(2-2)^2 + (6-2)^2]^{1/2}, \\ & [(2-2)^2 + (8-2)^2]^{1/2}, \\ & [(2-2)^2 + (10-2)^2]^{1/2}, \\ & [(5-2)^2 + (2-2)^2]^{1/2}, \\ & [(5-2)^2 + (4-2)^2]^{1/2}, \\ & [(5-2)^2 + (6-2)^2]^{1/2}, \\ & [(5-2)^2 + (8-2)^2]^{1/2}, \\ & [(5-2)^2 + (10-2)^2]^{1/2}, \\ & [(8-2)^2 + (2-2)^2]^{1/2}, \\ & [(8-2)^2 + (4-2)^2]^{1/2}, \\ & [(8-2)^2 + (6-2)^2]^{1/2}, \\ & [(8-2)^2 + (8-2)^2]^{1/2}, \\ & [(8-2)^2 + (10-2)^2]^{1/2}, \\ & [(9-2)^2 + (2-2)^2]^{1/2}, \\ & [(9-2)^2 + (4-2)^2]^{1/2}, \\ & [(9-2)^2 + (6-2)^2]^{1/2}, \\ & [(9-2)^2 + (8-2)^2]^{1/2}, \\ & \left. [(9-2)^2 + (10-2)^2]^{1/2} \right] \end{aligned}$$

$$\text{et } \text{color} = \min[|f(X_{2 2}) - f(X_{f_2 2})|, |f(X_{2 4}) - f(X_{f_2 2})|,$$



$$\begin{aligned}
 & |f(X_{2\ 6}) - f(X_{f_2\ 2})|, |f(X_{2\ 8}) - f(X_{f_2\ 2})|, \\
 & |f(X_{2\ 10}) - f(X_{f_2\ 2})|, |f(X_{5\ 2}) - f(X_{f_2\ 2})|, \\
 & |f(X_{5\ 4}) - f(X_{f_2\ 2})|, |f(X_{5\ 6}) - f(X_{f_2\ 2})|, \\
 & |f(X_{5\ 8}) - f(X_{f_2\ 2})|, |f(X_{5\ 10}) - f(X_{f_2\ 2})|, \\
 & |f(X_{8\ 2}) - f(X_{f_2\ 2})|, |f(X_{8\ 4}) - f(X_{f_2\ 2})|, \\
 & |f(X_{8\ 6}) - f(X_{f_2\ 2})|, |f(X_{8\ 8}) - f(X_{f_2\ 2})|, \\
 & |f(X_{8\ 10}) - f(X_{f_2\ 2})|, |f(X_{9\ 2}) - f(X_{f_2\ 2})|, \\
 & |f(X_{9\ 4}) - f(X_{f_2\ 2})|, |f(X_{9\ 6}) - f(X_{f_2\ 2})|, \\
 & |f(X_{9\ 8}) - f(X_{f_2\ 2})|, |f(X_{9\ 10}) - f(X_{f_2\ 2})|, \\
 & |f(X_{5\ 5}) - f(X_{f_2\ 2})|, |f(X_{8\ 5}) - f(X_{f_2\ 2})|, \\
 & |f(X_{9\ 5}) - f(X_{f_2\ 2})|, |f(X_{2\ 8}) - f(X_{f_2\ 2})|, \\
 & |f(X_{5\ 8}) - f(X_{f_2\ 2})|, |f(X_{8\ 8}) - f(X_{f_2\ 2})|, \\
 & |f(X_{9\ 8}) - f(X_{f_2\ 2})|, |f(X_{10\ 10}) - f(X_{f_2\ 2})|, \\
 & |f(X_{8\ 5}) - f(X_{f_2\ 2})|, |f(X_{5\ 9}) - f(X_{f_2\ 2})|, \\
 & |f(X_{2\ 8}) - f(X_{f_2\ 2})|.
 \end{aligned}$$

Alors,  $diff(X_{f_2\ 2}, Gob) = 2$ ,  $color = 2$  et  $position = 2$ . D'autre part, si  $f(X_{f_2\ 2}) = f(X_{2\ 2})$ ;  $diff(X_{f_2\ 2}, Gob) = 0$ ,  $color = 0$  et  $position = 0$ .

La formule suivante est employée pour calculer la mesure du déplacement du pixel centre de gravité liée à  $Gdb$  de l'image déformée:

$$Dp = \frac{1}{20} \sum_{i=1}^{20} diff(v_i(x_i, y_i), Gob) \quad (3.8)$$

Où 20 est le nombre de centres de gravité de  $Gdb$  et  $v_i(x_i, y_i)$  est le centre de gravité de  $Gdb$ .

Tous les  $Dp$  sont rangés dans la matrice:  $Dp_{global} = [Dp_{i,j}]$ .

La mesure globale des pixels déformés et déplacés ( $DPDM$ ) entre les deux blocs  $ob$  et  $db$  est obtenue en combinant les deux résultats définis précédemment, et est donnée par:

$$DPDM(ob, db) = Max(ODM, Dp) \quad (3.9)$$

### 3.4.2 Régions visuelles d'intérêt

Les régions d'intérêt ont été généralement employées dans des applications de communication visuelles dans un effort de fournir des images de bonne qualité.

Dans cette étape, la même fonction développée dans [16] a été employée où les régions dans l'image originale sont modelisées en calculant l'entropie "e" pour chaque bloc de 11×11 Pixels de l'image originale, en utilisant l'équation (3.10).  $H_i$  est une variable aléatoire indiquant l'intensité,  $p(H_i)$  est l'histogramme des niveaux d'intensité dans une région,  $K$  est le nombre de niveaux possibles d'intensité.  $K$  change de 0 à 255 pour des images niveau de gris.

$$e = - \sum_{i=1}^k p(H_i) \log_2 p(H_i) \quad (3.10)$$

Tous  $e$  sont rangées dans la matrice  $E = [e_{i,j}]$ , où chaque élément  $e_{i,j}$  est l'entropie d'un bloc.

Dans cette étape, la mesure de comparaison de la structure des régions est donnée par :

$$s_R(ob, db) = e * s(ob, db) \quad (3.11)$$

Où  $s(ob, db) = (\sigma_{ob\ db} + C_3) / (\sigma_{ob} \sigma_{db} + C_3)$ ,  $C_3 = C_2/2$  où  $C_2 = (K_2 L_2)^2$ ,  $K_2 \ll 1$ ,  $L_2$  est la dynamique du capteur,  $\sigma_{ob}$  et  $\sigma_{db}$  sont l'écart type des blocs  $ob$  et  $db$  respectivement,  $\sigma_{ob\ db}$  est la covariance de bloc  $ob$  et  $db$ .

La figure 3.3 montre l'image originale "maison" (768 × 512) (figure 3.3 (a)) et sa carte normalisée d'entropie  $E$  (figure 3.3 (b)), avec une valeur '1' indiquant la région d'intérêt le plus élevé représenté en tant que blanche, et une valeur '0' indiquant la région de plus bas intérêt, représentée en tant que noir. Les autres niveaux de l'image changeant du blanc au noir indiquent les régions visuelles avec le niveau descendant d'intérêt.

### 3.4.3 Information contour

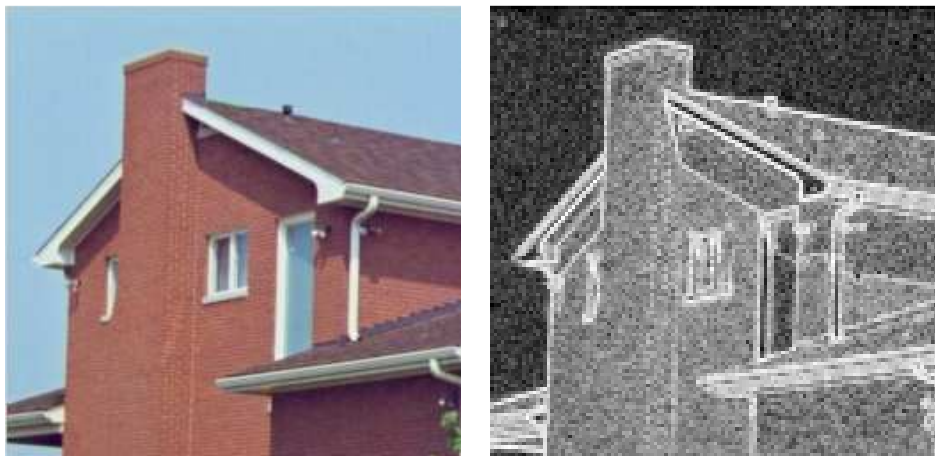
Selon le paradigme texture-contour des images, les contours forment la plus instructive partie de l'image. Par exemple, dans la perception du contenu de scène par SVH (système visuel humain), le contour joue un rôle principal. D'une manière semblable, les mesures de vision par machine se fondent souvent sur des cartes d'attributs obtenues à partir des contours. Ainsi, la performance de cette tâche dans la vision par des humains ou des machines est fortement liée à la qualité des contours et

d'autres attributs bidimensionnels tels que les coins.

Dans [8], une évaluation améliorée de la qualité appelée *GSSIM* est présentée se basant sur l'information de contour comme information la plus importante.

Les opérateurs de Sobel sont utilisés pour obtenir l'information de contour due à sa simplicité. *Xorg* et *Ydist* représentent la carte des gradients de l'image originale et déformée respectivement.

Soient *ob'* and *db'* des blocs (de tailles sont  $11 \times 11$  pixels) de *Xorg* et de *Ydist* respectivement.



(a)

(b)



(c)

**Figure 3.3 :** (a) Image originale (b) La carte des régions d'intérêt (c) L'information de contour

Cette métrique est employée dans la détermination de l'information de contour où la comparaison de contraste  $c(ob, db)$  et la comparaison de structure  $s(ob, db)$  dans [7] sont remplacées par la comparaison du contraste à base de gradient  $c_g(ob', db') = (2\sigma_{ob'}\sigma_{db'} + C_2)/(\sigma_{ob'}^2 + \sigma_{db'}^2 + C_2)$  et la comparaison de structure  $s_g(ob', db')$  respectivement.

$s_E(ob', db')$ ; représente l'information de contour fournie par la formule:

$$s_E(ob', db') = \frac{\sigma_{ob'db'} + C_3}{\sigma_{ob'}\sigma_{db'} + C_3} \quad (3.12)$$

Où  $\sigma_{ob'}$  et  $\sigma_{db'}$  sont l'écart type des blocs  $ob'$  et  $db'$  respectivement,  $\sigma_{ob'db'}$  est la covariance des blocs  $ob'$  et  $db'$ .

La figure 3.3 montre l'image originale "maison" (figure 3.3 (a)) et sa transformée en utilisant des opérateurs de Sobel (figure 3.3 (c)).

#### 3.4.4 Erreur globale

Finalement, la mesure basée sur l'information contour-région, les pixels déformés et déplacés ( $ERDDM$ ) est décrite comme suit :

$$ERDDM(ob, db) = [l(ob, db)]^\alpha \cdot [c(ob, db)]^\beta \cdot [s_{ER}(ob, db)]^\gamma + \omega * [1 + DPDM(ob, db)]^{-1} \quad (3.13)$$

Où:  $s_{ER}(ob, db) = (e * s_E(ob', db') + s_R(ob, db))/2$  et  $\omega = e$ . Nous posons  $\alpha=\beta=\gamma=1$ .

La mesure globale de l'erreur de l'image originale et déformée est calculée comme la somme de  $ERDDM_i(ob_i, db_i)$ . La mesure normalisée est définie comme :

$$ERDDM(Original, distorted) = \frac{\sum_{i=1}^{Eb} ERDDM_i(ob_i, db_i)}{2 * \sum_{i=1}^{Eb} e_i} \quad (3.14)$$

Où  $Eb$  est le nombre total des blocs.

Cette mesure est plus proche de '1' quand l'image est proche de la référence. Dans la

mesure proposée, la comparaison de la structure  $s_E(ob', db')$  est employée pour calculer l'erreur liée à JPEG2000, JPEG et fast-fading. Tandis que  $c(ob, db)$  est remplacé dans l'équation (3.13) par  $c_g(ob', db')$  pour le bruit blanc (WN) et le flou gaussien (Gblur). La performance de la mesure proposée est améliorée en employant  $c_g(ob', db')$  dans le calcul de l'erreur de WN et de Gblur et diminuée dans l'autre cas. Ainsi, ces indications sont très importantes dans le calcul  $d'ERDDM$ . De plus,  $ERDDM$  est prévu pour des images de couleur.

### 3.4.5 Discussion

La mesure d'évaluation de la qualité d'image proposée compare une image de référence et une image dégradée en suivant ces étapes :

- Les mesures des Pixels déformés et déplacés sont données respectivement par la matrice ( $id_{global}$ ) ; matrice d'erreur ( $DPDM_{Global}$ ).
- On calcule l'entropie de régions.
- On calcule la mesure de l'information de contour.
- En fin on détermine la mesure de la qualité d'image en employant l'équation (3.14).

Un organigramme illustrant le calcul  $d'ERDDM$  est montré dans figure 3.4.

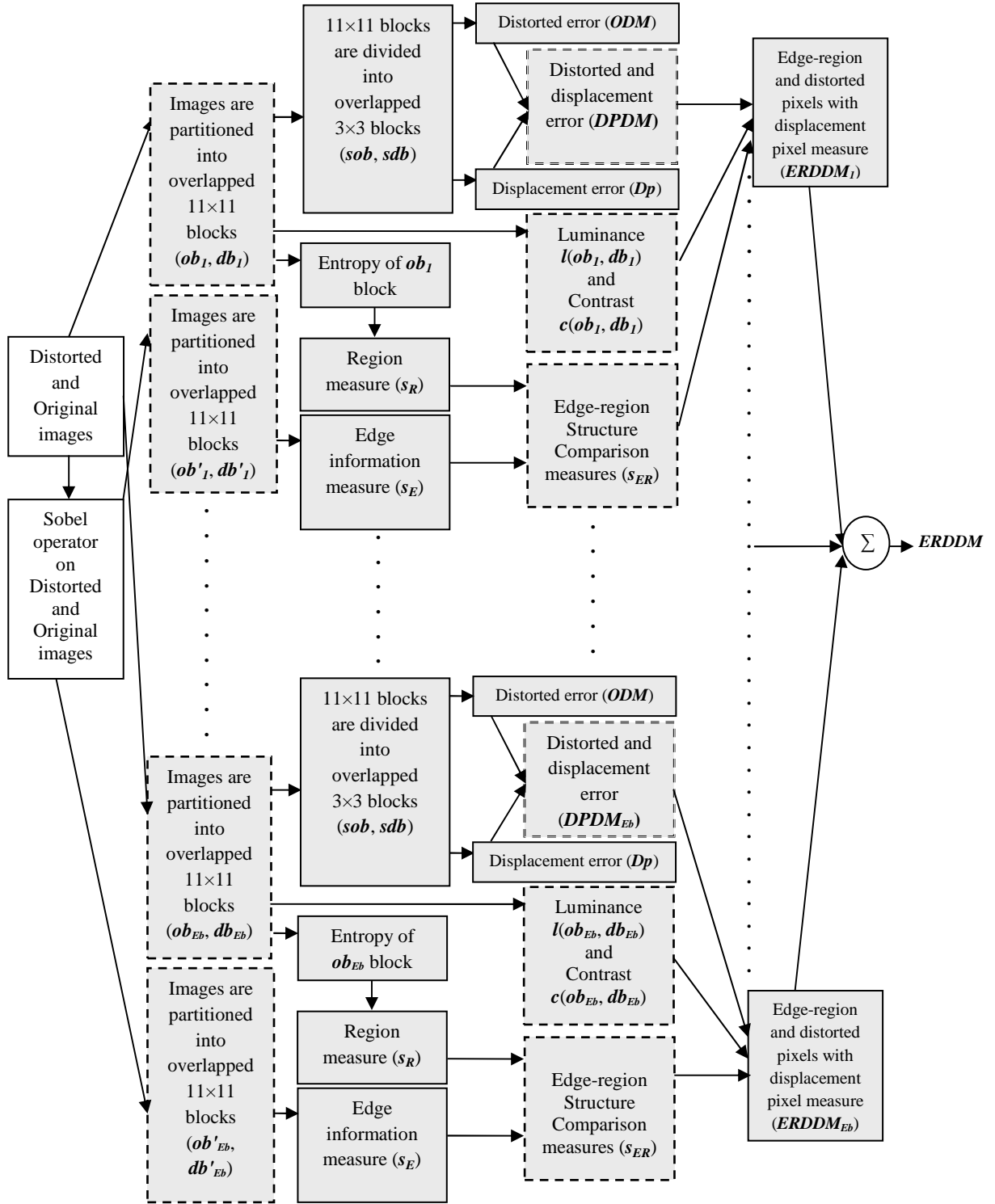


Figure 3.4 : Organigramme calculant ERDDM

### 3.5 Conclusion

Nous avons décrit dans ce chapitre la méthode d'évaluation de la qualité d'image proposée basée sur l'information contour, l'information région, les pixels déformés et déplacés. Cette dernière s'intègre dans les méthodes objectives d'évaluation de la qualité d'image avec référence. Celle-ci est une amélioration de la méthode MSSIM (Mean Structural Similarity Index) [4] en introduisant le concept de déplacement de pixels ainsi que son intensité afin de calculer le taux de déformation de l'image de test par rapport à l'image originale. En revanche, après la décomposition de l'image originale et de test en des blocs de taille  $11 \times 11$  dont la région de chevauchement est un pixel, on calcule localement la région d'intérêt, l'information contour ainsi que le déplacement et la déformation de pixels. L'introduction de la mesure des pixels déformés améliore les résultats obtenus.

Le chapitre suivant est une mise en œuvre de la mesure proposée ainsi qu'une étude comparative entre la méthode proposée et d'autres algorithmes existant dans la littérature.

# Chapitre 4

## Résultats et analyses

### 4.1 Introduction

Dans ce chapitre nous mettons en œuvre la méthode que nous proposons pour l'évaluation de la qualité d'image en utilisant la base d'images existant sur le web : la base d'images Live [3]. Dans un premier temps, la base d'images Live est introduite en expliquant les différentes étapes à suivre afin qu'elle soit exploitée. On suit la procédure standard d'évaluation des performances utilisées par le groupe d'expert de la qualité de vidéo (*VQEG*) [23]; le test de la phase II FR-TV, où cinq métriques principales d'évaluation sont employées (Métrique 1, Métrique 2, Métrique 3, Métrique 4 et Métrique 5). Enfin, une étude comparative intégrant un ensemble des méthodes d'évaluation existant dans la littérature, suivie d'une discussion et analyse des résultats obtenus sont effectuées.

### 4.2 Base d'images Live

#### 4.2.2 Choix des images

La conception des méthodes génériques d'évaluation de la qualité d'image exige un test avec un grand ensemble d'images ayant été déformées par une variété de dégradation. La base d'image Live a été construite à partir de vingt neuf images couleur de haute résolution ayant été rassemblées à partir l'internet et CD.

Ces images consistent en des images de visages, d'animaux, de scènes, objets synthétiques etc... La plupart des images ont la taille 768 × 512 pixels (voir figure 4.1).



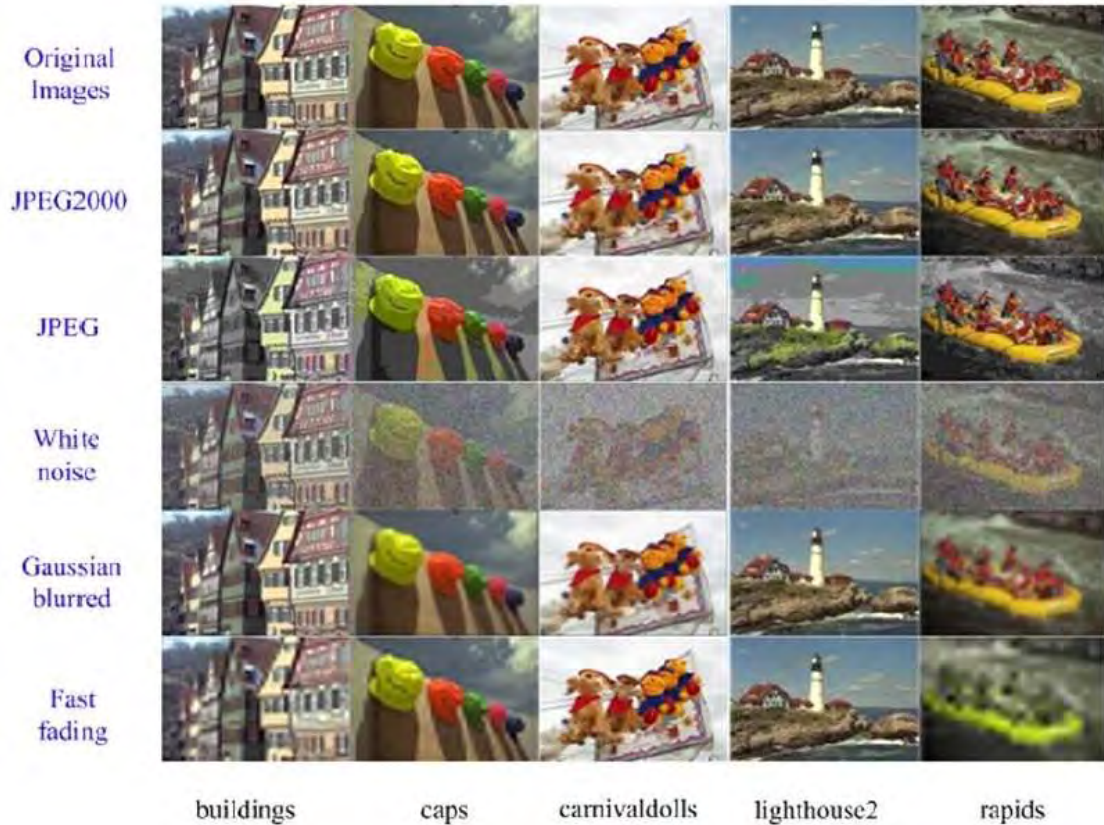


**Figure.4.1** : Images de références

### 4.2.3 Type de déformations

Cinq types de déformations ont été choisies afin de dégrader les images. Ces types peuvent se produire dans des applications réelles. Les types de déformation sont (voir figure 4.2): la compression sous format JPEG (JPEG\_1, PEG\_2: 233 images), la compression sous format JPEG2000 (JP2K\_1, JP2K\_2: 227 images), Bruit gaussien blanc additif dans les composantes RVB (WN: 174 images), flou gaussien dans les composantes RVB (Gblur: 174 images), et erreurs de transmission dans JPEG2000 (FF: 174 images). Ces déformations reflètent une large gamme de dégradations d'image, à partir du lissage, à la déformation structurée, et de bruit aléatoire. Le niveau de la déformation a été changé pour produire des images à une large gamme de qualité, à partir des niveaux imperceptibles aux niveaux élevés de dégradation qui entraveraient

de manière significative la compréhension cognitive du contenu de l'image. D'ailleurs, un total de 982 images, dont 779 sont des images dégradées et 203 sont des images de référence produites.



**Figure 4.2 :** Images échantillons et leurs versions déformées

#### 4.2.4 Détails de l'expérience

##### 4.2.4.1 Méthodologie de test

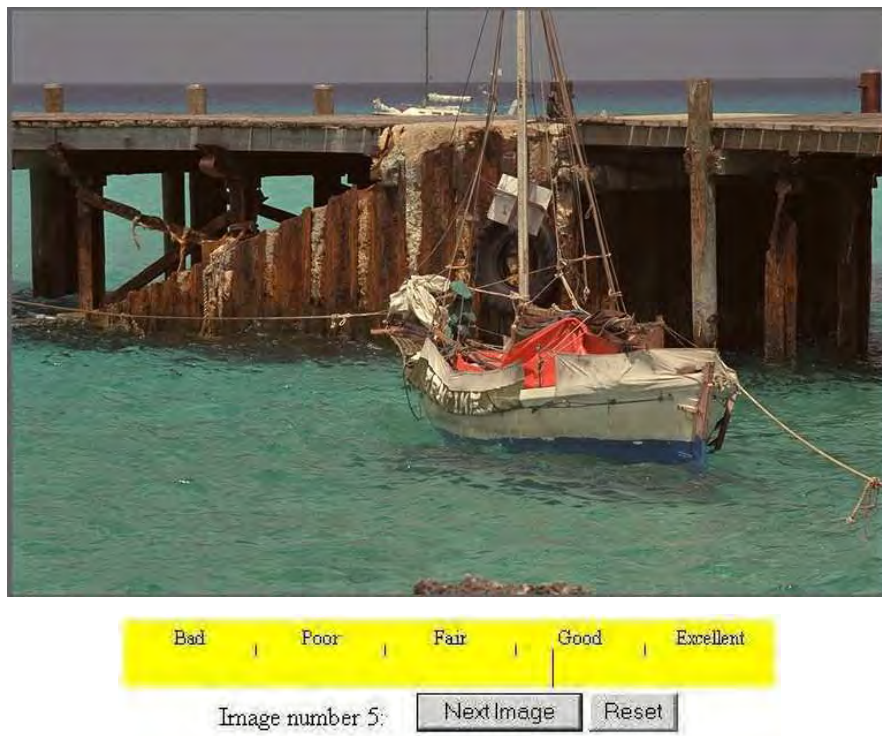
L'installation expérimentale employée est dérivée des recommandations normalisées selon les besoins et ressources spécifiques. En particulier, une méthodologie de simple stimulus [53] a été employée dans laquelle les images de référence sont également évaluées par les observateurs humains, dans la même session expérimentale comme les images de test. Comme les images de référence ont été également évaluées par les observateurs à chaque session, des points de différence (MOS) de qualité peuvent être dérivés pour toutes les images déformées.

##### 4.2.4.2 Configuration de l'équipement

Des postes de travail identiques contenant Microsoft Windows ont été utilisés dans les expériences. Une interface web montrant l'image à évaluer et un applet de Java



étant un glisseur pour assigner des points de qualité ont été employés pour imiter les conditions de jugement de la qualité de monde réel aussi étroitement que possible. Les postes de travail ont été placés dans un bureau avec un éclairage dans les normes. Les écrans de visualisation sont des écrans CRT ayant une taille 21 pouces et une résolution de  $1024 \times 768$  pixels. Bien que les écrans n'aient pas été calibrés, ils sont tous approximativement du même âge, et ils sont fixés à la même configuration d'affichage. Les observateurs ont été demandés de regarder les écrans à partir d'une distance de visionnement approximatif de 2-2.5 la taille de l'écran. Le logiciel inclut des attributs qui ne permettraient pas à des utilisateurs de sauter une image sans faire un jugement, et remettraient également le glisseur à zéro quand une nouvelle image a été montrée. La figure 4.3 montre l'interface utilisateur du logiciel d'évaluation de la qualité. Les expériences ont été effectuées en sept sessions : deux sessions pour JPEG2000, deux pour le JPEG, et un pour le bruit blanc, flou Gaussien, et les erreurs du canal. Chaque session inclut un ensemble complet des images de référence aléatoirement placé parmi les images dégradées. Le nombre d'images à chaque session est montré sur le tableau 4.1 (les images de référence ont été incluses dans chaque session).



**Figure 4.3 :** Interface pour l'étude subjective.

**Tableau 4.1:** Sessions d'évaluation subjectives: nombre d'images à chaque session et nombre d'observateurs participant à chaque session.

Session	Nombre d'images	Nombre d'observateurs
JPEG2000 #1	116	29
JPEG2000 #2	111	25
JPEG #1	116	20
JPEG #2	117	20
Bruit gaussien	174	23
flou gaussien	174	24
erreurs de transmission	174	20
Total	982	22.8 (la moyenne)

Les observateurs ont rapporté leurs jugements de qualité en traînant un glisseur sur une échelle de qualité. La balance de qualité est marquée numériquement. Elle a été divisée en cinq parties, qui ont été marquées avec des adjectifs : "mauvais", "pauvre", "juste", "bon", et "excellent". Après que l'observateur ait évalué l'image, la position du glisseur est convertie en points de qualité en traçant linéairement la valeur entière dans l'intervalle [1, 100]. De cette façon, les points de qualité bruts ont composé des nombres entiers dans la gamme 1 -100.

#### 4.2.5 Traitement des points bruts

Le traitement des points bruts a été fait pour calculer les MOS (mean opinion score) et les DMOS (difference mean opinion score) pour chaque image. MOS a été utilisé pour l'apprentissage et la validation des méthodes sans référence, alors que DMOS était employé pour former et tester les méthodes avec référence.

##### 4.2.5.1 Détection de valeurs aberrantes et rejet d'observateur

Un algorithme simple de détection de valeurs aberrantes (outliers) et de rejet d'observateur est choisi. Un point brut d'une image est considéré comme une valeur aberrante s'il est en dehors d'un intervalle de largeur  $\Delta$  des écarts-types sur le point (score) moyen (calculé à partir des points bruts de tous les observateurs) pour cette image. Et toutes les évaluations de la qualité d'un observateur sont rejetées ; si ses

évaluations aberrantes sont supérieures à un nombre  $R$ . Cet algorithme de rejet de valeurs aberrantes est exécuté deux fois. Un algorithme de minimisation est exécuté pour chaque type de déformation qui a changé  $\Delta$  et  $R$  pour réduire au minimum la largeur moyenne de l'intervalle de confiance de 95% pour chaque type de déformation. Trois observateurs de toutes les sessions sont rejetés.

#### 4.2.5.2 Scores de MOS

Afin de calculer les points de MOS, les points bruts (après l'élimination de valeurs aberrantes et le rejet d'observateurs)  $r_{ij}$  du  $i^{eme}$  observateur et  $j^{eme}$  image sont transformés en Z-points [49] :

$$z_{ij} = \frac{r_{ij} - \bar{r}_i}{\sigma_i} \quad (4.1)$$

Où  $\bar{r}_i$  est la moyenne des points bruts de toutes les images évaluées par l'observateur  $i$ , et  $\sigma_i$  est l'écart type. Les Z-points sont alors mesurés dans une gamme de 1-100 avant qu'ils soient transformés en une moyenne à travers des observateurs pour donner les points de MOS.

#### 4.2.5.3 Scores de DMOS

La procédure de calcul des scores de DMOS est semblable à celle des scores de MOS, sauf que les points bruts sont convertis la première fois en points de qualité de différence entre les images déformées et leurs images correspondantes de référence

$$d_{ij} = r_{iref(j)} - r_{ij} \quad (4.2)$$

Où  $r_{iref(j)}$  sont les points bruts de qualité assignés par l'observateur  $i$  à l'image de référence correspondant à l'image déformée  $j$ . Des valeurs aberrantes sont alors enlevées (un total de quatre observateurs sont rejetés dans toutes les sessions), et les Z-points sont calculés en utilisant  $d_{ij}$  avec (4.2). Ensuite, ils sont mesurés dans une gamme de 1-100 avant qu'ils soient transformés en une moyenne à travers des observateurs pour donner les points de DMOS.

### 4.3 Méthodologie pour l'évaluation de la performance

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 prédiction (prediction accuracy): la capacité de prévoir les estimations de qualité subjectives avec une erreur faible,
- monotonicité 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 prédiction 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.

Ces attributs (aspects) sont évalués par 5 métriques de performance indiquées dans le plan de test objectif, et sont discutés ci-dessous.

Les sorties du modèle objectif de la qualité d'image (l'estimation de la qualité visuelle,  $VQR$ ) devrait être mis en corrélation avec DMOS de l'observateur dans une manière prévisible et répétable. La relation entre les prédictions de chaque modèle et les évaluations subjectives est estimée en utilisant une régression non linéaire entre  $VQR$  et DMOS correspondant.

On a effectué un plan non linéaire entre les points objectifs et subjectifs [23]. Le plan non linéaire des cinq-paramètre ( $\theta_1, \theta_2, \theta_3, \theta_4$  et  $\theta_5$ ) est employé pour transformer l'ensemble des estimations de la qualité par les métriques objectives de qualité en un ensemble de valeurs prévues comme moyennes d'opinion de différence (scores de différence :  $DMOS$  : Difference Mean Opinion Score) dénoté  $DMOS_p$  (the predicted Difference Mean Opinion Score).

Le plan non-linéaire est choisi au-dessus du plan linéaire pour expliquer les caractéristiques non linéaires des points subjectifs aux extrémités des gammes d'essai. La fonction traçante est une fonction logistique avec le terme linéaire additif donné dans l'équation (4.3).

$$DMOS_p = \theta_1 \text{logistic}(\theta_2, (VQR - \theta_3)) + \theta_4 + \theta_5 \quad (4.3)$$

$$\text{logistic}(x, VQR) = \frac{1}{2} - \frac{1}{1 + \exp(VQR - x)} \quad (4.4)$$

Où  $VQR$  est l'estimation de la qualité des méthodes objective ( $ERDDM$ ,  $VroiWQI$ ,  $GSSIM$ ,  $MSSIM$  ou  $PSNR$ ) et  $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$  sont choisis pour le meilleur ajustement.

L'ajustement (optimisation des paramètres  $\theta$ ) est effectué par la fonction *fminsearch* du MATLAB.

Une fois la transformation non-linéaire est appliquée, les performances de prédiction de méthode objective sont ensuite évaluées en calculant plusieurs métriques sur les valeurs de *DMOS* et *DMOS<sub>p</sub>*. Le plan de test exige cinq métriques de la correspondance entre *VQR* et *DMOS*. En outre, il exige des contrôles de la qualité des données subjectives.

Afin de fournir des mesures quantitatives sur la performance de la méthode d'évaluation de la qualité d'image proposée *ERDDM*, on suit les procédures standard d'évaluation des performances utilisées par le groupe d'expert [23] de la qualité de la vidéo (*VQEG*) ; le test de la phase II FR-TV, où cinq métriques principales d'évaluation sont employées (Métrique 1, Métrique 2, Métrique 3, Métrique 4 et Métrique 5).

Les métriques suivantes sont employées dans le calcul entre les scores subjectifs et objectifs après un plan non-linéaire :

#### 4.3.3 Métriques relatives à la précision de la prédiction d'une méthode

- Métrique 1: Coefficient de corrélation linéaire de Pearson (*CC* : Pearson linear correlation coefficient) : C'est la covariance entre les scores subjectifs (*DMOS*) et objectifs (*DMOS<sub>p</sub>*), rapportée au produit de leurs écarts-types. Il fournit une évaluation de précision de la prédiction et il est défini par :

$$CC = \frac{\sum_{i=1}^N (DMOS(i) - \overline{DMOS})(DMOS_p(i) - \overline{DMOS_p})}{\sqrt{\sum (DMOS(i) - \overline{DMOS})^2} \sqrt{\sum (DMOS_p(i) - \overline{DMOS_p})^2}} \quad (4.5)$$

Où l'index *i* dénote l'échantillon d'image et *N* dénote le nombre d'échantillons.

#### 4.3.4 Métriques relatives à la monotonie de la prédiction d'une méthode

- Métrique 2: le coefficient de corrélation des rangs de Spearman (*ROCC* : Spearman rank-order correlations coefficient) : il compare simplement l'ordre dans lequel les valeurs apparaissent dans les 2 vecteurs sans faire d'hypothèse sur le type de fonction de dépendance, en calculant simplement un coefficient de corrélation de Pearson sur les rangs des modèles (*DMOS<sub>p</sub>*) et *DMOS*. Cette mesure de corrélation présume seulement un rapport monotone entre les deux quantités. Une vertu de ce type de corrélation est qu'il ne nécessite

pas l'hypothèse d'une forme fonctionnelle particulière dans la relation entre les données ( $DMOS$ ) et les prédictions ( $DMOS_p$ ). Elle est définie par:

$$ROCC = 1 - \frac{6 \sum (DMOS(i) - DMOS_p(i))^2}{N(N^2 - 1)} \quad (4.6)$$

Où 6 est une constante (elle est toujours employée dans la formule).

Cette formule est issue du coefficient de corrélation linéaire de Pearson, mais elle est appliquée aux rangs.

#### 4.3.5 Métriques relatives à la cohérence de la prédiction d'une méthode

- Métrique 3: Rapport des valeurs aberrantes (Outlier ratio :  $OR$ ). Cette métrique évalue la capacité d'un modèle objectif de donner des prévisions cohérentes et précises pour tous les types d'images et de ne pas échouer excessivement pour un sous-ensemble d'images, c.-à-d., la cohérence de prévision. Il est défini par :

$$OR = \frac{\text{(total number of outliers)}}{N} \quad (4.7)$$

Où une valeur aberrante est un point pour lequel :

$$|DMOS(i) - DMOS_p(i)| > 2 \times \sigma(DMOS(i))$$

$\sigma(DMOS(i))$ : représente l'écart type des différents points liés à l'échantillon de l'image  $i$ . Les différents scores sont approximativement distribués et donc deux fois la valeur de  $\sigma$  représente l'intervalle de confiance de 95%. Ainsi, la valeur de  $2 \times \sigma(DMOS(i))$  représente un bon seuil pour définir un point de valeur aberrante.

Plus d'information sur ces mesures peut être trouvée dans [23]. En plus de ces dernières, nous avons également calculé :

- Métrique 4: La racine carrée de la moyenne des erreurs de prédiction ou erreur quadratique moyenne de prédiction ( $RMS$ : Root mean square prediction error) : c'est la racine de la moyenne arithmétique des carrés des écarts entre les scores subjectifs ( $DMOS$ ) et objective ( $DMOS_p$ ). elle est définit par:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (DMOS(i) - DMOS_p(i))^2} \quad (4.8)$$



- Métrique 5: Erreur absolue maximale de prédiction (*MAE* : Maximum absolute prediction error) : c'est la moyenne arithmétique des valeurs absolues des écarts entre les scores subjectifs (*DMOS*) et objectifs (*DMOS<sub>p</sub>*). Elle est définie par :

$$MAE = \frac{1}{N} \sum_{i=1}^N |DMOS(i) - DMOS_p(i)| \quad (4.9)$$

Enfin, les métriques précédentes sont utilisées dans les sections qui viennent afin d'évaluer la performance de la méthode proposée ; leurs significations sont comme suit :

- Le coefficient de corrélation (*cc*) : il détermine le degré de corrélation entre chaque modèle et *DMOS*, une grande valeur de *cc* signifie une meilleure précision.
- le coefficient de corrélation des rangs de Spearman (*ROCC*): on le considère comme mesure de monotonie de prédiction, une valeur maximale de *ROCC* indique une meilleure monotonie de prédiction.
- le rapport des valeurs aberrantes (outlier ratio: *OR*), la racine carrée de la moyenne des erreurs (*RMS*) et l'erreur absolue maximale de prédiction (*MAE*) sont des mesures relatives à la cohérence de prédiction. Les plus petites valeurs de ces métriques signifient que la performance est meilleure.

#### 4.4 Analyse des résultats

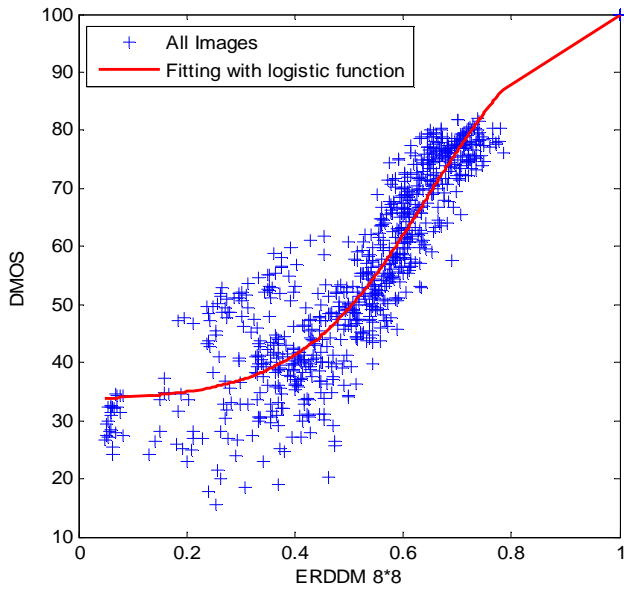
On va tester la performance de la méthode proposée dans cette partie. Ensuite, une étude comparative est effectuée en utilisant un ensemble des mesures reconnues dans la littérature d'évaluation de la qualité d'image à savoir *PSNR* et *MSSIM* etc...

Avant de comparer la méthode proposée avec *PSNR*, *WroiWQI*, *GSSIM*, *MSSIM*, *ERDMSSIM* (the edge-region and distorted pixels structural similarity) [17] et *ERDM* (the edge-region and distorted pixels with pixel displaced structural similarity) [39], une étude justifiant le choix de la fenêtre 11×11 est faite. Trois métriques sont calculées : Coefficient de corrélation linéaire de Pearson (*cc*), la racine carrée de la moyenne des erreurs (*RMS*) et l'erreur absolue maximale de prédiction (*MAE*).

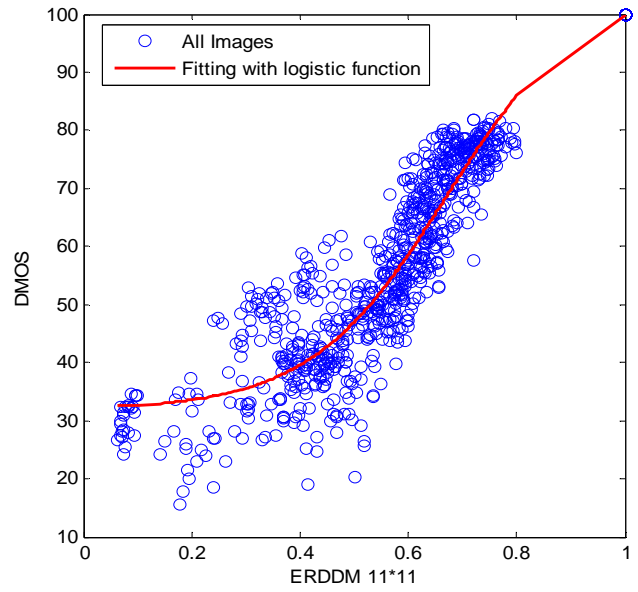
Durant les tests nous avons exécuté la méthode *ERDDM* sur la base d'image Live. Pour cela, la taille de partitionnement varie de la plus petite taille à la plus grande taille. A chaque exécution de la méthode, on calcule sa performance en utilisant les cinq métriques précédentes (*CC*, *RMS*,...) selon les procédures standard d'évaluation des performances utilisées par le groupe d'expert de la qualité vidéo (*VQEG*) ; le test

de la phase II FR-TV. Le temps d'exécution de la méthode augmente lorsque la taille de partitionnement est inférieure à  $11 \times 11$  et diminue quand la taille de division est plus grande. Donc, nous présentons dans cette partie la performance d'*ERDDM* avec trois types de partitionnement  $8 \times 8$ ,  $11 \times 11$ ,  $15 \times 15$ .

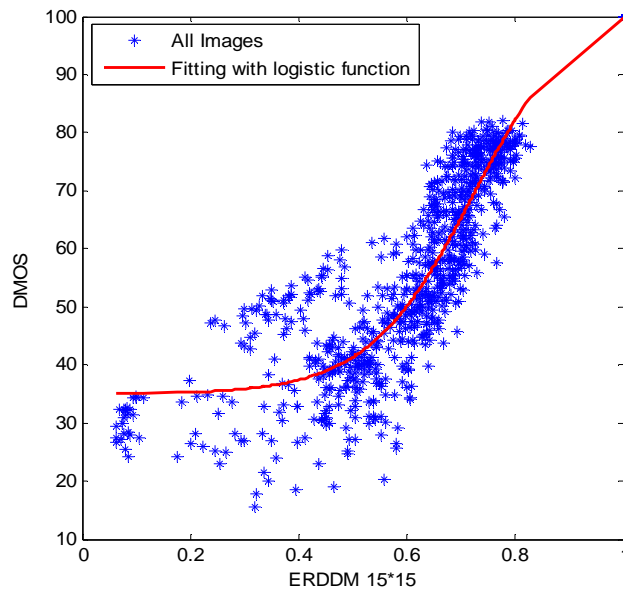
La figure 4.4 présente le graphe de dispersion de *DMOS* contre *ERDDM*. En plus, les graphes de dispersion sont des plans relatifs à la division  $8 \times 8$ ,  $11 \times 11$ ,  $15 \times 15$ . En examinant le tableau 4.2 et le graphe de dispersion (figure 4.4), on remarque que les métriques *CC*, *MAE* et *RMS* ont de bonnes valeurs lorsque la taille du bloc est  $11 \times 11$ . Comme nous l'avons vu dans la section précédente, les petites valeurs de *MAE*, *RMS* (voir figure 4.5) montrent les meilleures performances et une valeur de *CC* proche de 1 donne une bonne performance. La lecture de la figure 4.5 (a) montre qu'il existe un grand seuil pour le partitionnement  $11 \times 11$  où la performance change considérablement. Ainsi en observant la figure 4.5 (b) et (c), on découvre l'existence d'un petit seuil dans les deux graphes où la performance de la méthode change de la mauvaise vers la meilleur vers la mauvaise. Selon cette étude, la performance globale de la méthode *ERDDM* réduit quand la taille de la partition choisie est supérieure ou inférieure à la taille  $11 \times 11$ .



( a )

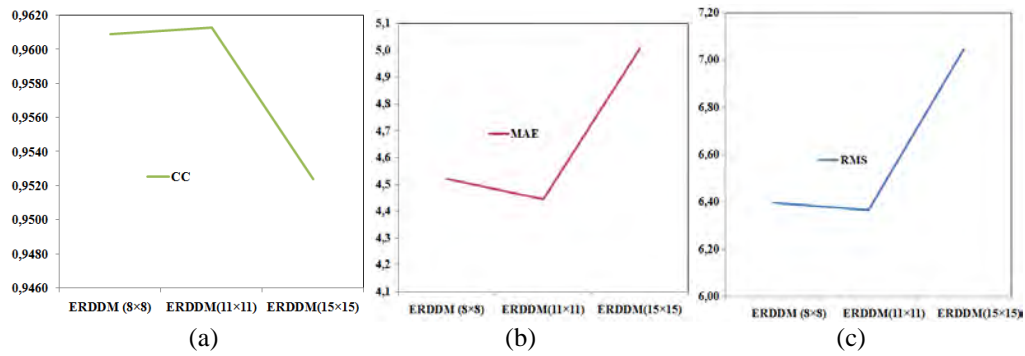


( b )



( c )

**Figure 4.4:** Graphe de *DMOS* contre *ERDDM* pour toute la base de données ( a ) *ERDDM* (8×8) ( b ) *ERDDM* (11×11) ( c ) *ERDDM* (15×15)



**Figure 4.5 :** Graphe de la variation de la performance d’ERDDM selon les métrique : (a) CC, (b) MAE et (c) RMS

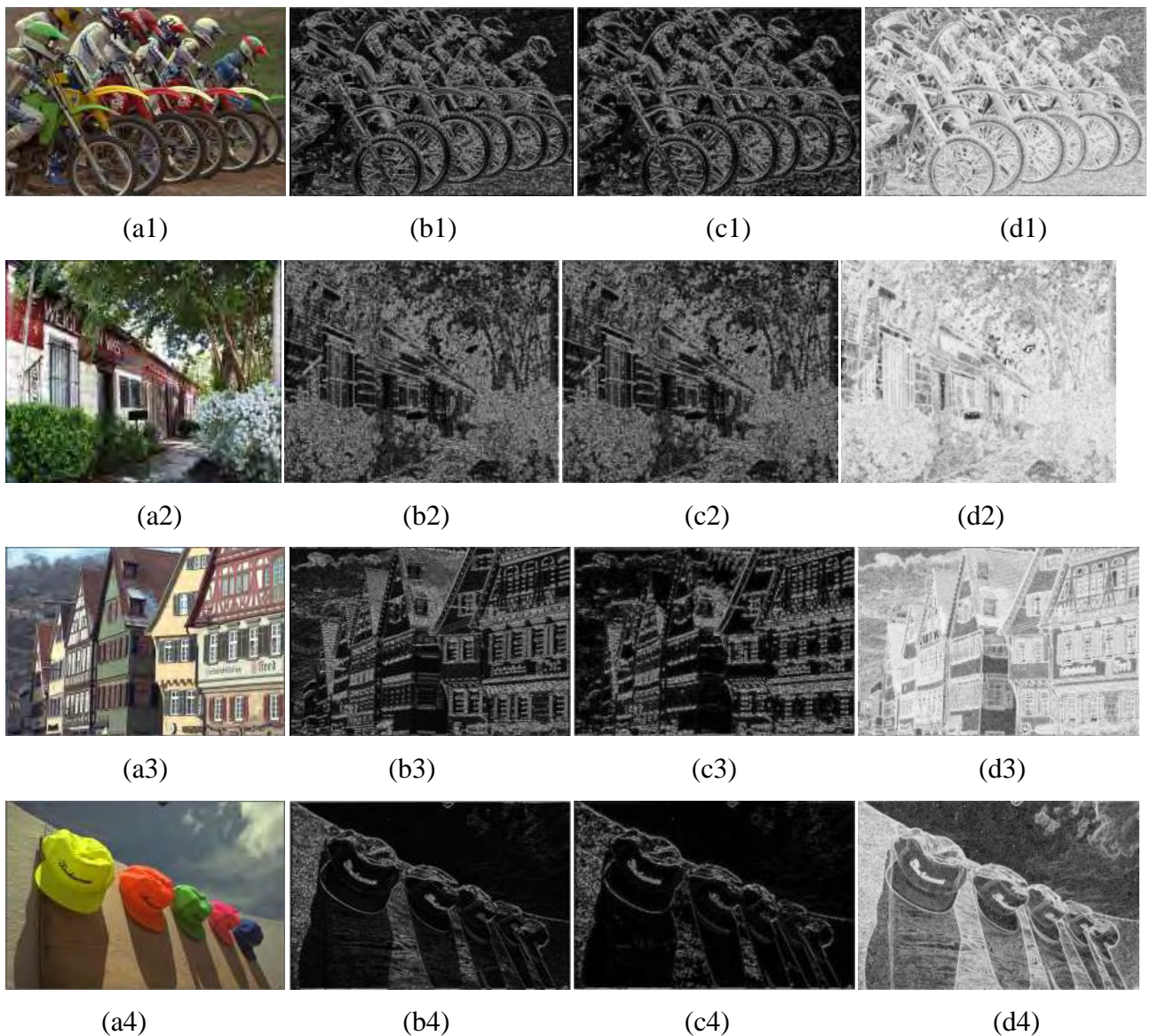
**Tableau 4.2 :** Comparaison de la performance de méthodes d’évaluation de la qualité d’image

Méthode	CC	MAE	RMS
<i>ERDDM (8×8)</i>	0.9609	4.5201	6.3978
<i>ERDDM(15×15)</i>	0.9524	5.0073	7.0454
<i>ERDDM(11×11)</i>	<b>0.9613</b>	<b>4.4454</b>	<b>6.3672</b>

De plus, la performance de la méthode proposée *ERDDM* a été vérifiée sur des images de différentes déformations de la base Live. La figure 4.6 montre les images originales avec leurs transformations. Dans cette figure nous avons présenté quelques échantillons des images de la base Live. La première colonne est consacrée aux images originales, la deuxième pour les transformations des images originales en utilisant l’opérateur de Sobel et la troisième colonne est réservée aux images déformées issues d’un traitement effectué par l’opérateur Sobel afin de mettre en évidence l’image contour qui sera utilisée dans les traitements postérieurs. En plus, l’application de l’opérateur d’entropie sur l’image originale permet de clarifier les régions d’intérêt de l’image. Cette dernière est présentée dans la dernière colonne.

Ainsi, nous avons comparé la méthode *ERDDM* à quelques mesures de l’évaluation de la qualité d’image à savoir, *PSNR*, *WroiWQI*, *GSSIM*, *MSSIM*, *ERDMSSIM* et *ERDM*. Les résultats que nous avons trouvés sur toutes les images déformées sont récapitulés dans les tableaux (tableau 4.3,..., tableau 4.7). Chaque tableau se compose de neuf colonnes et huit lignes. La première colonne est réservée aux méthodes d’évaluation de la qualité d’image ordonnées comme suit : *PSNR*, *WroiWQI*, *GSSIM*, *MSSIM*, *ERDMSSIM*, *ERDM* et *ERDDM*. La deuxième colonne jusqu’à la huitième

représentent les différentes déformations appliquée sur les images originales (la deuxième colonne-JP2K\_1 : compression sous format JPEG2000, la troisième colonne-JP2K\_2 : compression sous format JPEG2000, la quatrième colonne-JPEG\_1 : compression sous format JPEG, la cinquième colonne-JPEG\_2 : compression sous format JPEG, la sixième colonne-WN : Bruit gaussien blanc additif dans les composantes RVB, la septième colonne-Gblur : flou Gaussien dans les composantes RVB et la huitième colonne-FF : erreurs de transmission dans JPEG2000).



**Figure 4.6:** a) Image originale (b) l'information de contour de(a) (c) l'information de contour de l'image déformée (d) carte des régions d'intérêt de (a)

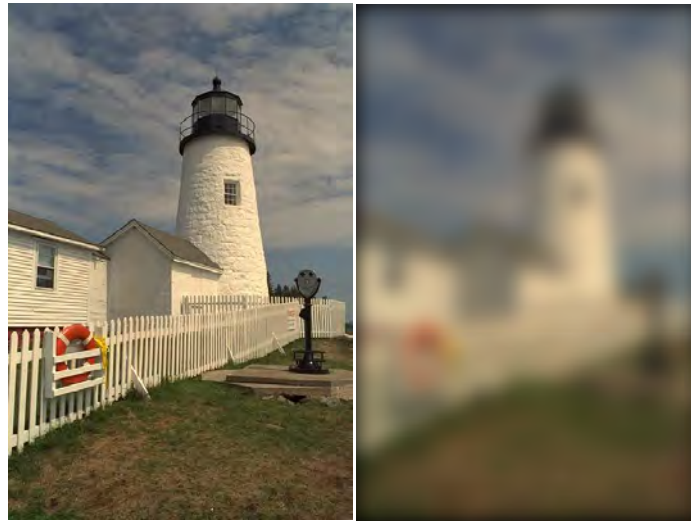
En plus, la dernière colonne (la neuvième) est la colonne de la mesure globale. Les résultats de chaque type de déformation relatifs à une mesure donnée sont montrés dans la première ligne jusqu'à la huitième.

En fait, cinq tableaux des résultats obtenus à partir de l'exécution des mesures sont trouvés dans les pages qui viennent. Ces tableaux sont ordonnés comme suit : **Tableau 4.3** : Coefficient de la corrélation linéaire de Pearson ( $CC$ ), **Tableau 4.4** : Erreur quadratique moyenne de prédiction ( $RMS$ ), **Tableau 4.5** : Coefficient de corrélation des rangs de Spearman ( $ROCC$ ), **Tableau 4.6** : Erreur absolue maximale de prédiction ( $MAE$ ) et **Tableau 4.7** : Rapport des valeurs aberrantes ( $OR\%$ ).

Nous remarquons que la méthode proposée  $ERDDM$  est parfaitement adaptée à l'évaluation de la qualité des images floues. Dans la suite, nous avons comparé les résultats obtenus par  $ERDDM$  avec  $MSSIM$  et  $PSNR$  sur un échantillon de deux images. Les images floues (b) et (d) dans la figure 4.7 ont presque la même valeur de  $PSNR$  (31.4895 et 31.3934 respectivement), mais leur qualité visuelle sont différentes, la qualité subjective de l'image floue du "phare" (figure 4.7 b où  $DMOS= 84.489$ ,  $PSNR= 31.4895$ ,  $MSSIM= 0.5062$ ,  $ERDDM= 0.1792$ ) est beaucoup plus mauvaise que l'image "building2" (figure 4.7 d), alors que les valeurs de  $MSSIM$  sont contraires à la qualité perceptuelle. L'image floue (figure 4.7 b) a une valeur plus élevée de  $MSSIM$  que l'image "building2" (figure 4.7 d où  $DMOS= 56.7104$ ,  $PSNR= 31.3934$ ,  $MSSIM= 0.448$ ,  $ERDDM= 0.376$ ).

D'ailleurs, une valeur de  $DMOS$  près de "100" signifie que l'image a une mauvaise qualité visuelle, alors qu'une valeur près de "0" indique qu'elle a une bonne qualité visuelle.





(a)

(b)



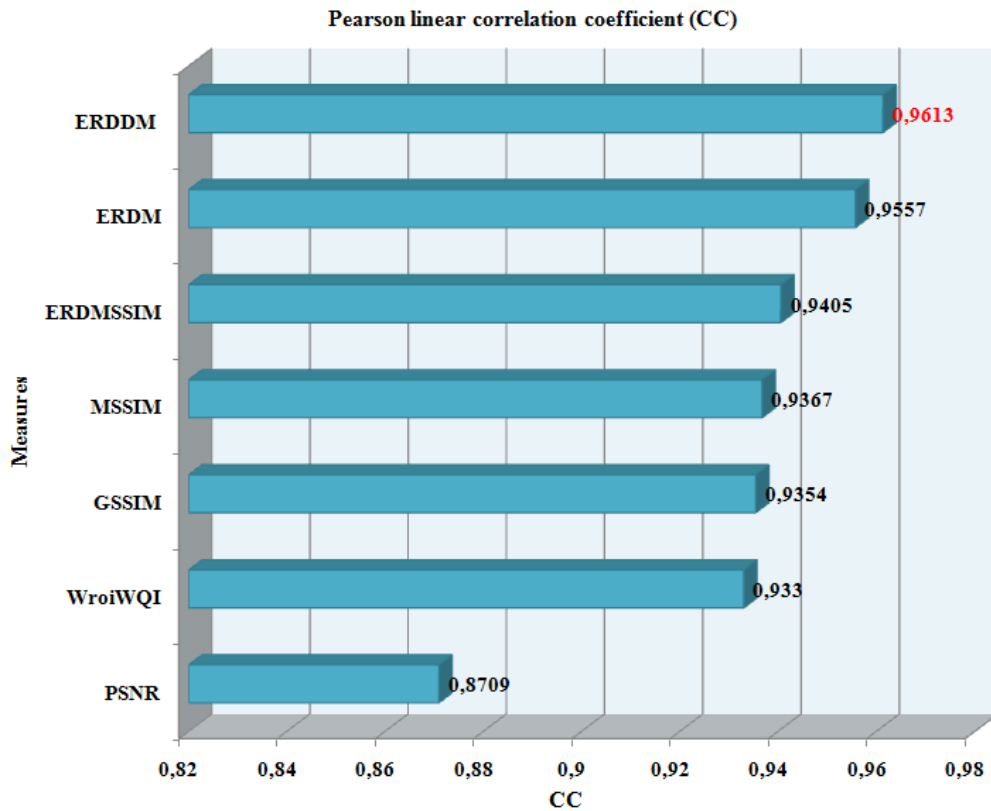
(c)

(d)

**Figure 4.7:** Comparaison des images “lighthouse” et “building2”. (a) Image originale de “lighthouse”. (b) Image floue de “lighthouse”(c) Image originale de “building2”. (d) Image floue de “building2”.

**Tableau 4.3:** Coefficient de corrélation linéaire de Pearson (*CC*)

Model	JP2K_1	JP2K_2	JPEG_1	JPEG_2	WN	Gblur	FF	All data
<i>PSNR</i>	0.9332	0.8740	0.8856	0.9167	0.9859	0.7834	0.8895	0.8709
<i>WroiWQI</i>	0.9721	0.9755	0.9715	0.9778	0.9902	0.9417	0.9620	0.9330
<i>GSSIM</i>	0.9252	0.9652	0.9156	0.9759	0.9666	0.9719	0.9660	0.9354
<i>MSSIM</i>	0.9761	0.9768	0.9732	0.9829	0.9844	0.9355	0.9675	0.9367
<i>ERDMSSIM</i>	0.9748	0.9840	0.9670	0.9875	0.9864	0.9584	0.9723	0.9405
<i>ERDM</i>	0.9596	0.9789	0.9496	0.9867	0.9818	0.9750	0.9775	0.9557
<i>ERDDM</i>	<b>0.9859</b>	<b>0.9936</b>	<b>0.9859</b>	<b>0.9932</b>	<b>0.9922</b>	<b>0.9798</b>	<b>0.9716</b>	<b>0.9613</b>



**Figure 4.8.** Coefficient de corrélation linéaire de Pearson (*CC*)

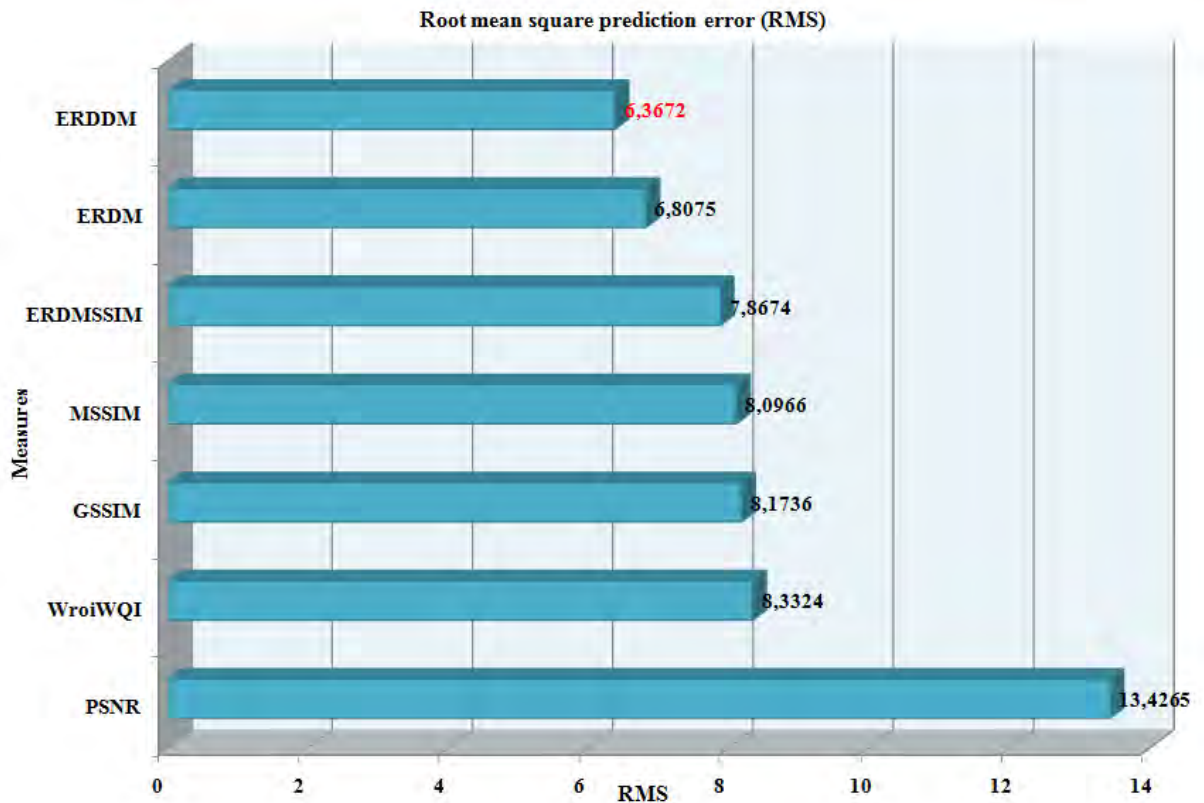
Comme nous l'avons vu dans la section précédente, une grande valeur du coefficient de corrélation (*cc*) signifie une meilleure précision. Nous remarquons d'après la figure 4.8 que la valeur globale du *CC* d'*ERDDM* est la plus grande valeur (**0.9613**) en comparant à celles des autres mesures. Ceci indique qu'*ERDDM* a la meilleure précision. La figure 4.8 donne un ordonnancement de la plus petite valeur en bas vers



la plus grande valeur en haut. Donc *PSNR* a la plus petite valeur 0.8709, *WroiWQI* a la valeur 0.9330, *GSSIM* a la valeur 0.9354, *MSSIM* a la valeur 0.9367, *ERDMSSIM* a la valeur 0.9405 et *ERDM* a la valeur 0.9557. En plus, en regardant le tableau 4.3, *ERDDM* a effectivement toujours les meilleures valeurs de *CC* selon la variation des déformations appliquées aux images originales.

**Tableau 4.4:** Racine carrée de la moyenne des erreurs de prédiction (*RMS*)

Model	JP2K_1	JP2K_2	JPEG_1	JPEG_2	WN	Gblur	FF	All data
<i>PSNR</i>	8.4858	12.7537	10.7793	13.5633	4.6689	11.4402	12.9716	13.4265
<i>WroiWQI</i>	5.7139	5.3829	5.8738	5.1227	3.0641	7.3216	6.0293	8.3324
<i>GSSIM</i>	9.2566	6.4016	9.6894	5.3252	5.6353	5.1223	5.7122	8.1736
<i>MSSIM</i>	5.2977	5.2360	5.5965	4.5070	4.0029	7.7146	5.5988	8.0966
<i>ERDMSSIM</i>	5.4526	4.3593	6.1820	3.8465	3.6071	6.207	5.1653	7.8674
<i>ERDM</i>	6.853	5.0011	7.551	3.9678	4.1962	4.8335	4.6603	6.8075
<i>ERDDM</i>	<b>4.0285</b>	<b>2.7714</b>	<b>4.0285</b>	<b>2.8530</b>	<b>2.7377</b>	<b>4.3465</b>	<b>5.2245</b>	<b>6.3672</b>



**Figure 4.9:** Racine carrée de la moyenne des erreurs de prédiction (*RMS*)

La racine carrée de la moyenne des erreurs (*RMS*) est une mesure relative à la cohérence de prédiction. La plus petite valeur signifie que la performance est meilleure. D'après la figure 4.9, nous remarquons qu'*ERDDM* a la plus petite valeur de *RMS* étant **6.3672** tandis que *PSNR* possède la plus mauvaise (13.4265). Ceci indique qu'*ERDDM* a la meilleure performance. La figure précédente est ordonnée de la plus grande valeur en bas (mauvaise performance) vers la plus petite valeur en haut (bonne performance). On remarque également à partir du tableau 4.4 que la méthode *ERDDM* a effectivement toujours les meilleures valeurs de *RMS*.

**Tableau 4.5:** Coefficient de corrélation des rangs de Spearman (*ROCC*)

Model	JP2K_1	JP2K_2	JPEG_1	JPEG_2	WN	Gblur	FF	All data
<i>PSNR</i>	0.9264	0.8549	0.8779	0.7699	0.9854	0.7823	0.8907	0.8755
<i>WroiWQI</i>	0.9854	0.9859	0.9846	0.9885	0.9981	0.9894	0.9928	0.9568
<i>GSSIM</i>	0.9618	0.9800	0.9581	0.9876	0.9937	0.9948	0.9935	0.9584
<i>MSSIM</i>	0.9875	0.9866	0.9860	0.9911	0.9968	0.9882	0.9938	0.9592
<i>ERDMSSIM</i>	0.9867	0.9907	0.9830	0.9935	0.9974	0.9924	0.9947	0.9615
<i>ERDM</i>	0.9791	0.9878	0.9746	0.9931	0.9965	0.9954	0.9957	0.9712
<i>ERDDM</i>	<b>0.9928</b>	<b>0.9963</b>	<b>0.9928</b>	<b>0.9964</b>	<b>0.9985</b>	<b>0.9963</b>	<b>0.9946</b>	<b>0.9748</b>

La figure 4.10 et le tableau 4.5 montrent le coefficient de corrélation des rangs de Spearman (*ROCC*). Cette mesure est considérée comme une mesure de la monotonie de prédiction. Une valeur maximale de *ROCC* indique une meilleure monotonie de prédiction. Nous remarquons d'après la figure 4.10 que la valeur globale *ROCC* d'*ERDDM* est la plus grande valeur (**0.9748**). Ceci indique qu'*ERDDM* a la meilleure monotonie de prédiction. La figure 4.10 est classée de la plus petite valeur en bas vers la plus grande valeur en haut. D'où *PSNR* a la plus petite valeur 0.8755, *WroiWQI* a la valeur 0.9568, *GSSIM* a la valeur 0.9584, *MSSIM* a la valeur 0.9592, *ERDMSSIM* a la valeur 0.9615 et *ERDM* a la valeur 0.9712. En examinant toutes les valeurs de *ROCC* existant dans le tableau 4.5 de tous les types de déformation à savoir JP2K\_1, JP2K\_2, JPEG\_1, JPEG\_2, WN, Gblur et FF, *ERDDM* a effectivement les meilleures valeurs de *ROCC*.

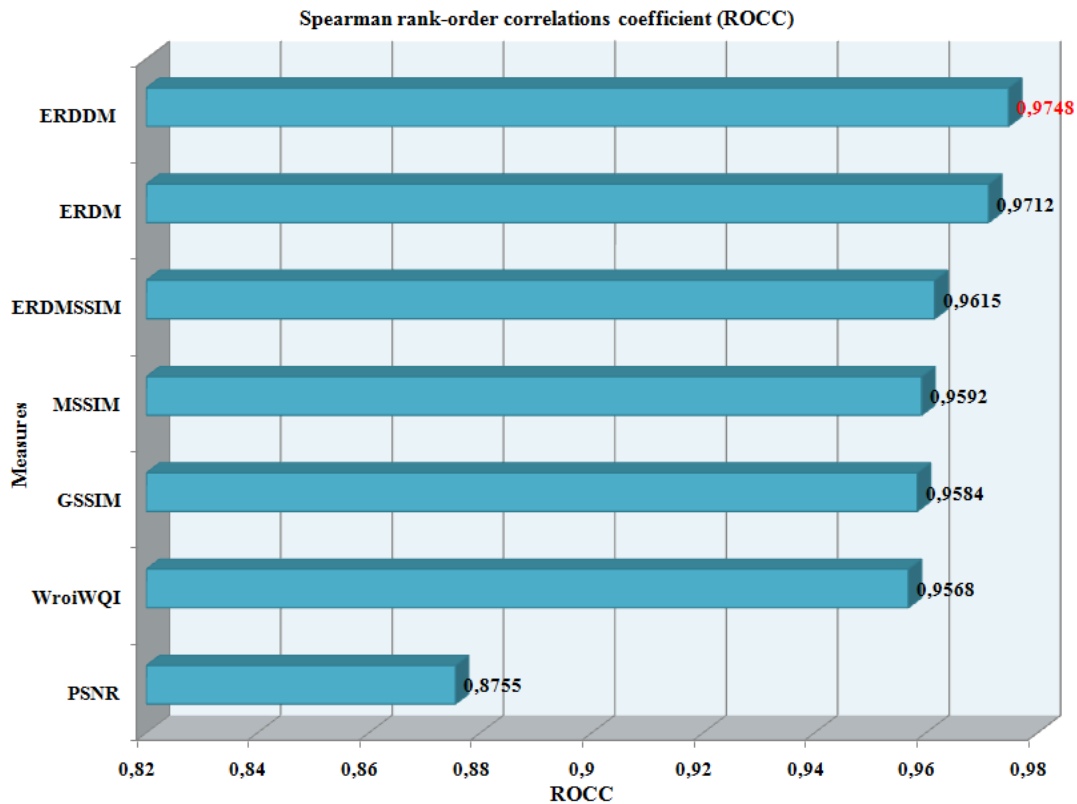
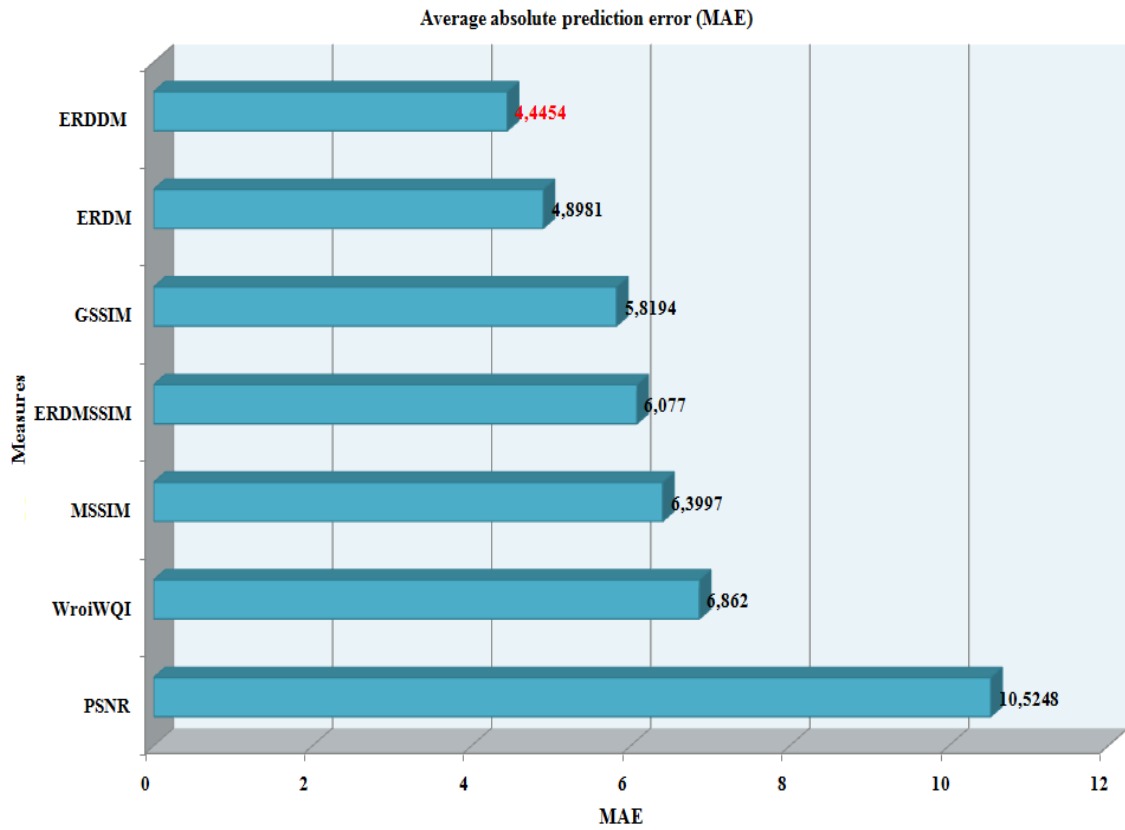


Figure 4.10: Coefficient de corrélation des rangs de Spearman (ROCC)

Tableau 4.6: Erreur absolue maximale de prédiction (MAE)

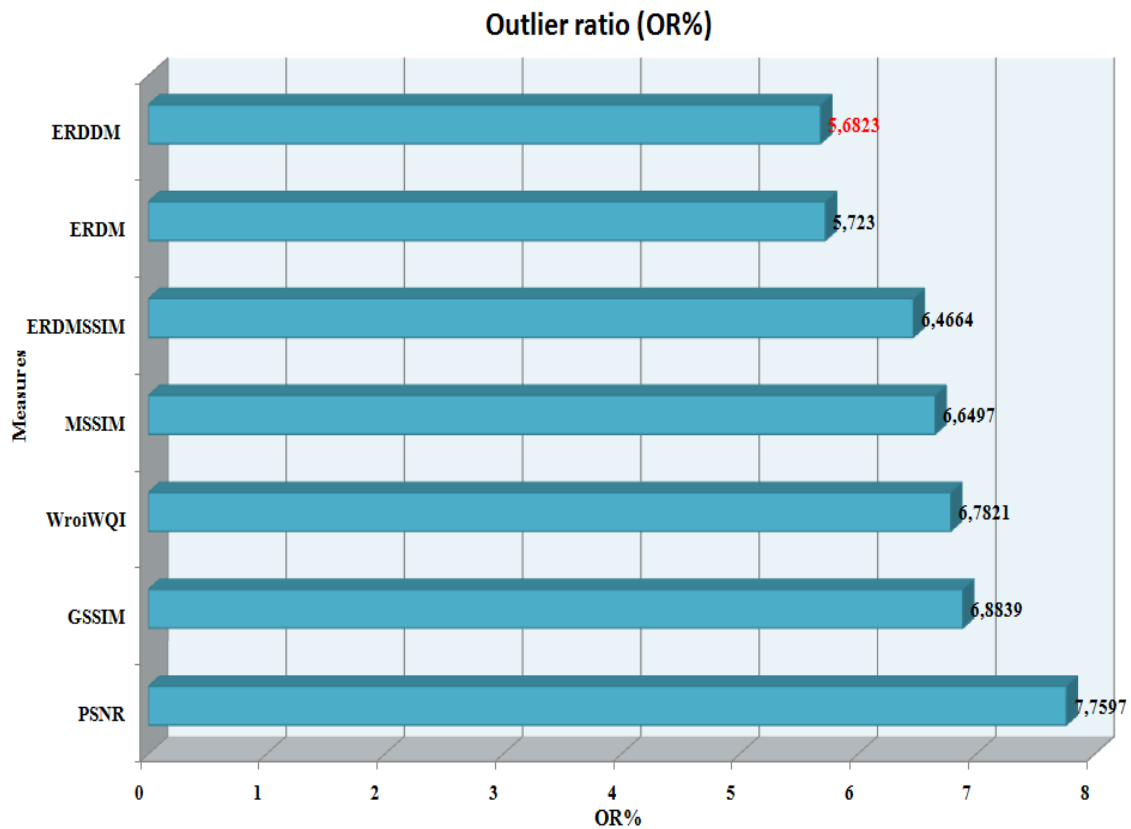
Model	JP2K_1	JP2K_2	JPEG_1	JPEG_2	WN	Gblur	FF	All data
<i>PSNR</i>	6.5033	9.9656	8.6900	10.013	3.4195	9.0550	9.9898	10.5248
<i>WroiWQI</i>	4.3480	4.0090	4.2327	3.9219	2.3080	5.3609	4.5646	6.8620
<i>GSSIM</i>	7.4196	4.4503	7.2624	3.8419	4.2076	3.8802	4.1853	5.8194
<i>MSSIM</i>	4.1134	3.7700	4.3231	3.2481	2.9122	5.6925	4.1862	6.3997
<i>ERDMSSIM</i>	4.3672	3.1846	5.0502	2.8498	2.7197	4.7522	3.8991	6.0770
<i>ERDM</i>	5.1974	3.421	5.8846	2.7788	3.1903	3.6616	3.5773	4.8981
<i>ERDDM</i>	<b>2.8756</b>	<b>1.9200</b>	<b>2.8756</b>	<b>1.9982</b>	<b>2.1228</b>	<b>3.1573</b>	<b>3.7065</b>	<b>4.4454</b>



**Figure 4.11:** Erreur absolue maximale de prédiction (*MAE*)

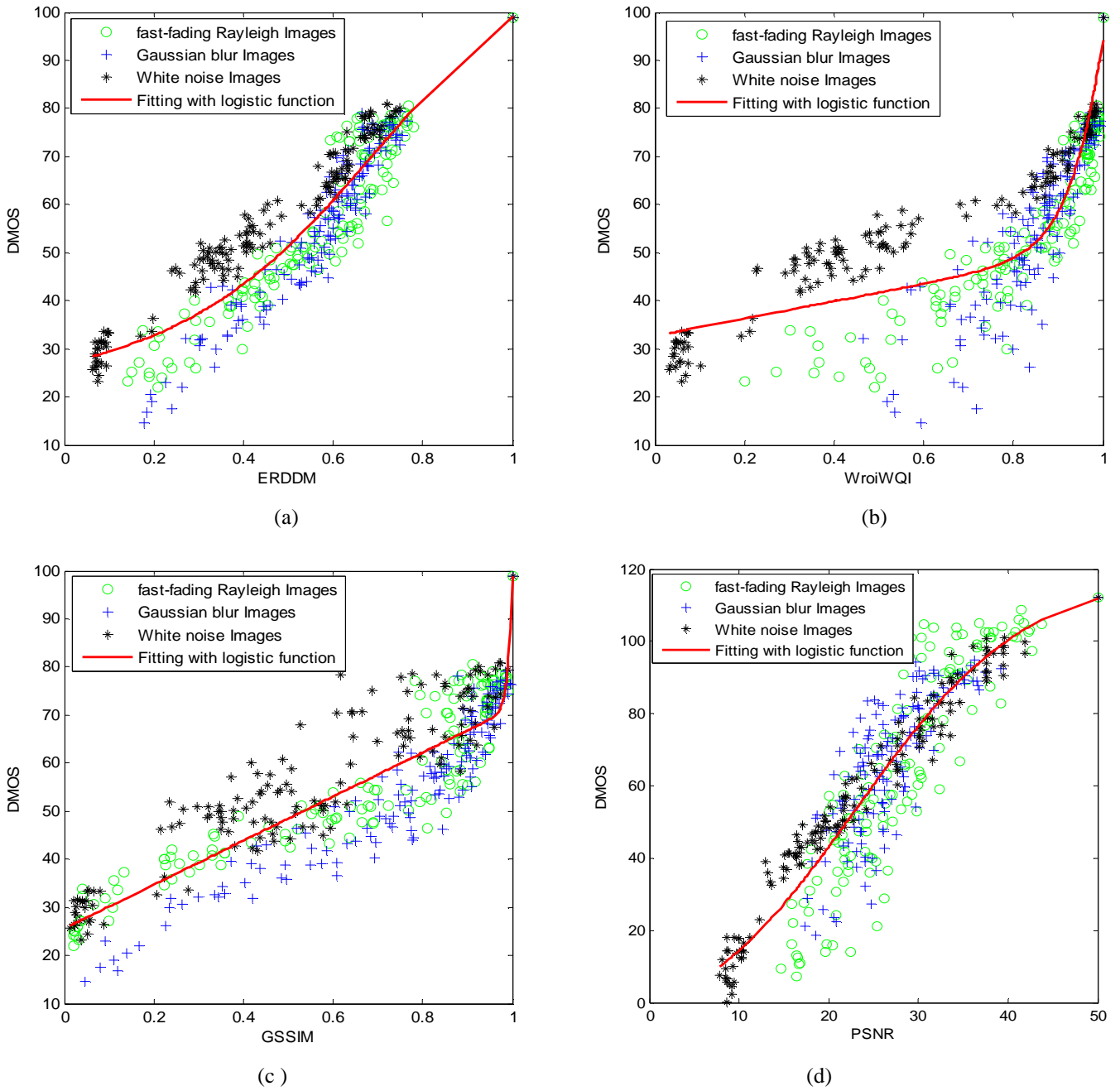
**Tableau 4.7:** Rapport des valeurs aberrantes (*OR%*)

Model	JP2K_1	JP2K_2	JPEG_1	JPEG_2	WN	Gblur	FF	All data
<i>PSNR</i>	6.6379	7.2072	6.8103	7.1795	4.1379	7.9310	6.7241	7.7597
<i>WroiWQI</i>	6.2069	5.7658	5.2586	4.9573	3.1609	6.6092	4.4828	6.7821
<i>GSSIM</i>	8.4483	6.5766	7.7586	6.0684	5.3448	5.7471	4.7701	6.8839
<i>MSSIM</i>	6.1207	5.6757	5.1724	4.7009	4.0230	6.6092	4.0230	6.6497
<i>ERDMSSIM</i>	6.1207	4.8649	5.4310	4.188	3.9655	6.3218	3.8506	6.4664
<i>ERDM</i>	7.2414	5.7658	6.8103	4.6154	4.5402	5.1149	3.8506	5.7230
<i>ERDDM</i>	<b>4.2241</b>	<b>3.5135</b>	<b>4.2241</b>	<b>3.2479</b>	<b>2.4713</b>	<b>4.8851</b>	<b>3.9655</b>	<b>5.6823</b>

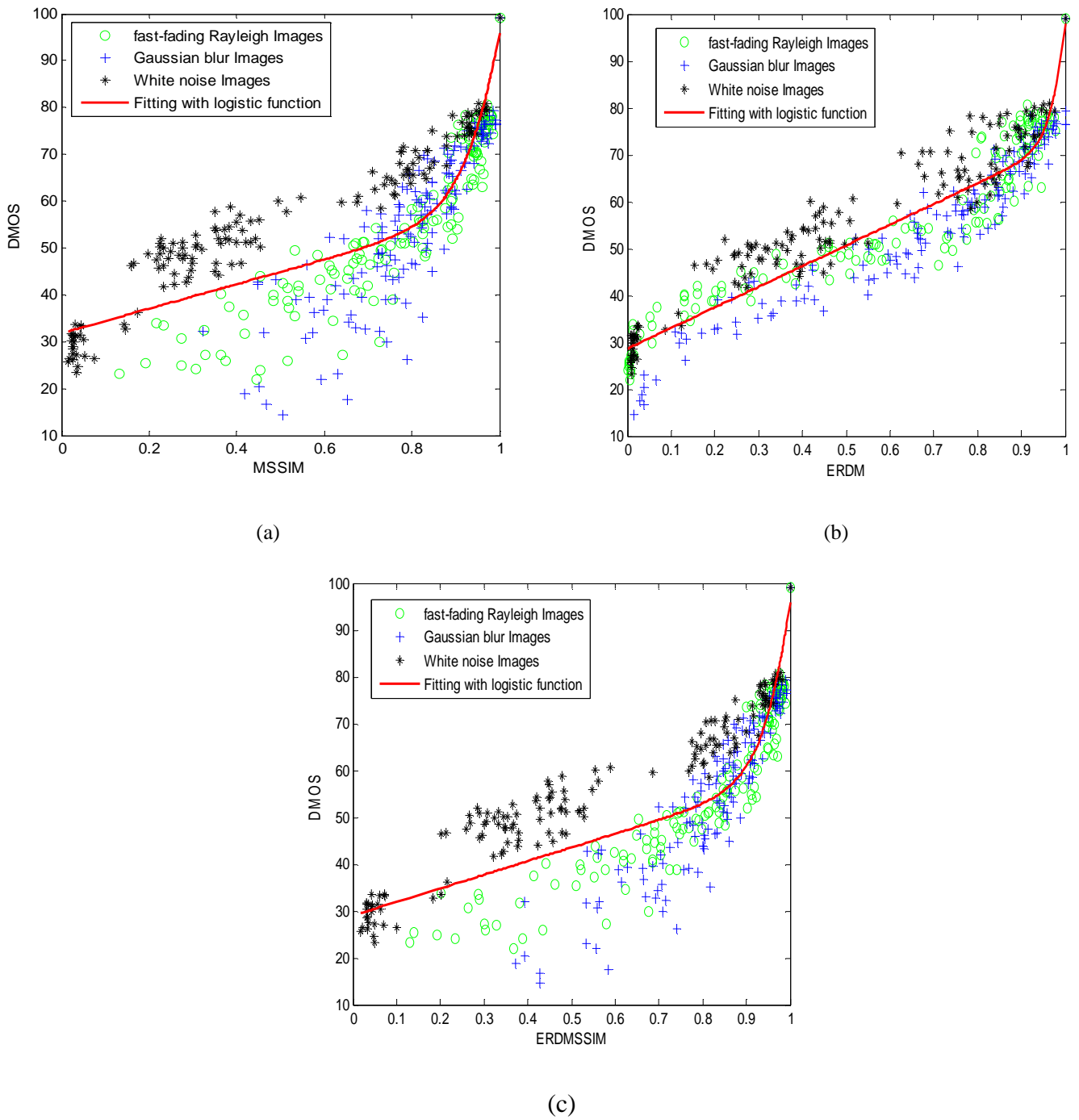


**Figure 4.12 :** Rapport des valeurs aberrantes (*OR*%)

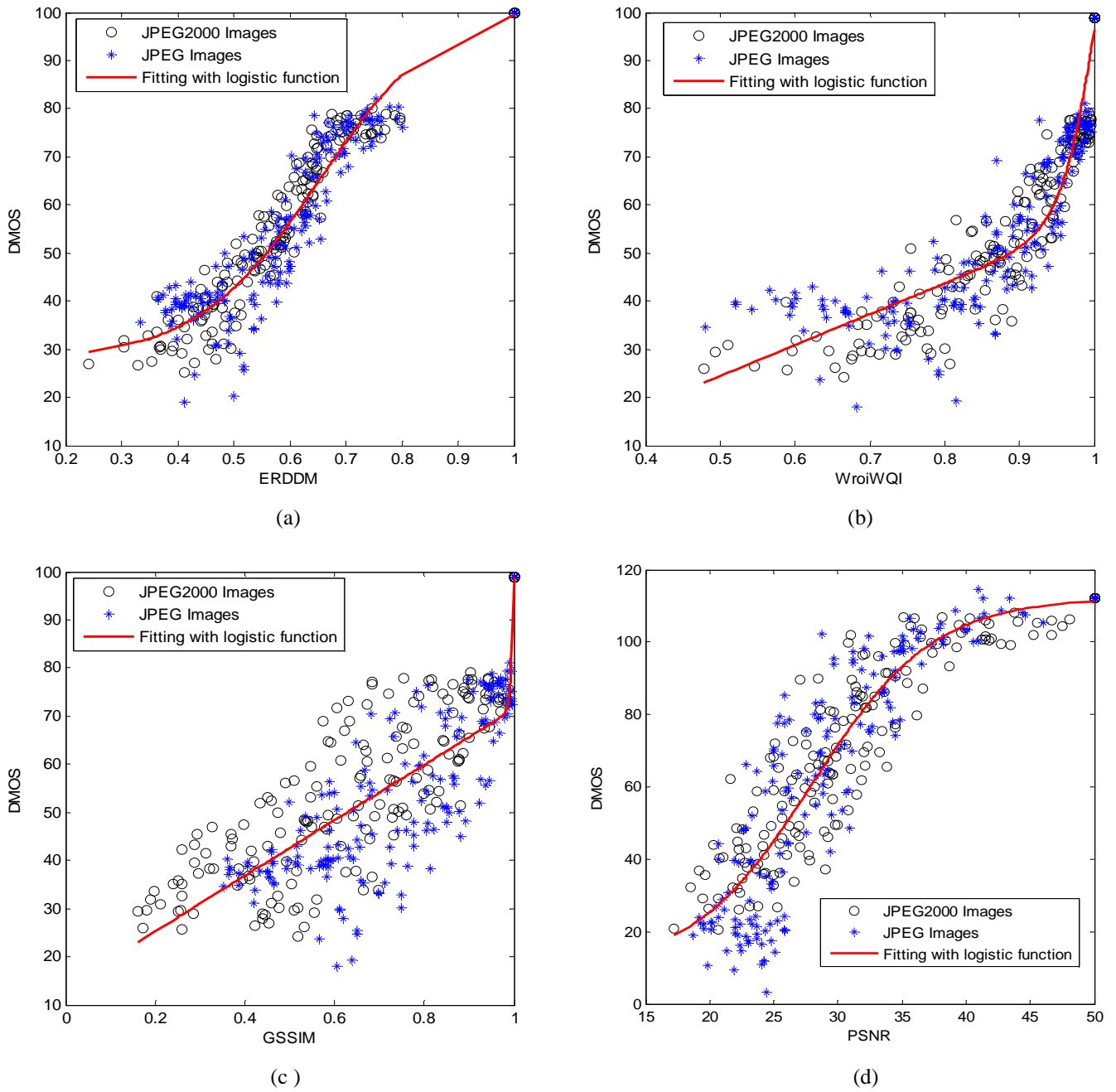
Les figures 4.11, 4.12 et les tableaux 4.6, 4.7 représentent respectivement l'erreur absolue maximale de prédiction (*MAE*) et le rapport des valeurs aberrantes (outlier ratio: *OR*). Ces mesures (*MAE* et *OR*) correspondent à la cohérence de prédiction, où les plus petites valeurs de ces métriques signifient que la performance est meilleure. On remarque d'après la figure 4.11 que la mesure *ERDDM* a la plus petite valeur de *MAE* (**4.4454**) ainsi que la meilleure valeur du *OR* d'après la figure 4.12 (**5.6823**). Ceci indique qu'*ERDDM* a la meilleure performance.



**Figure 4.13:** Graphe de *DMOS* contre les prévisions de modèles de flou gaussien, bruit blanc, erreurs de transmission dans JPEG2000 (Fastfading). (a) *ERDDM* (b) *VroiWQI* (c) *GSSIM* (d) *PSNR*.

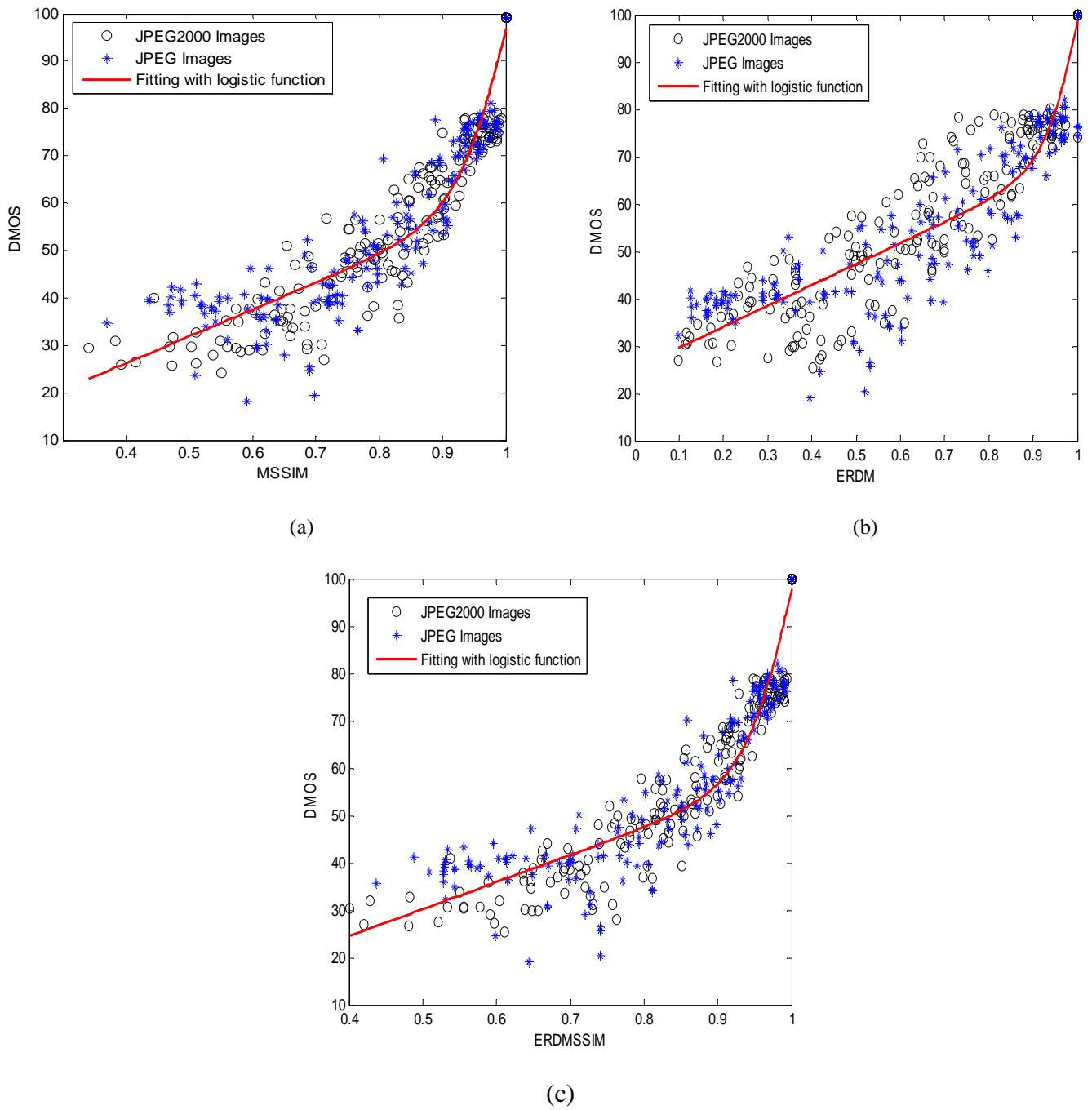


**Figure 4.14:** Graphe de *DMOS* contre les prévisions de modèles de flou gaussien, bruit blanc, Erreurs de transmission dans JPEG2000 (Fastfading). (a) *MSSIM* (b) *ERDM* (c) *ERDMSSIM*.

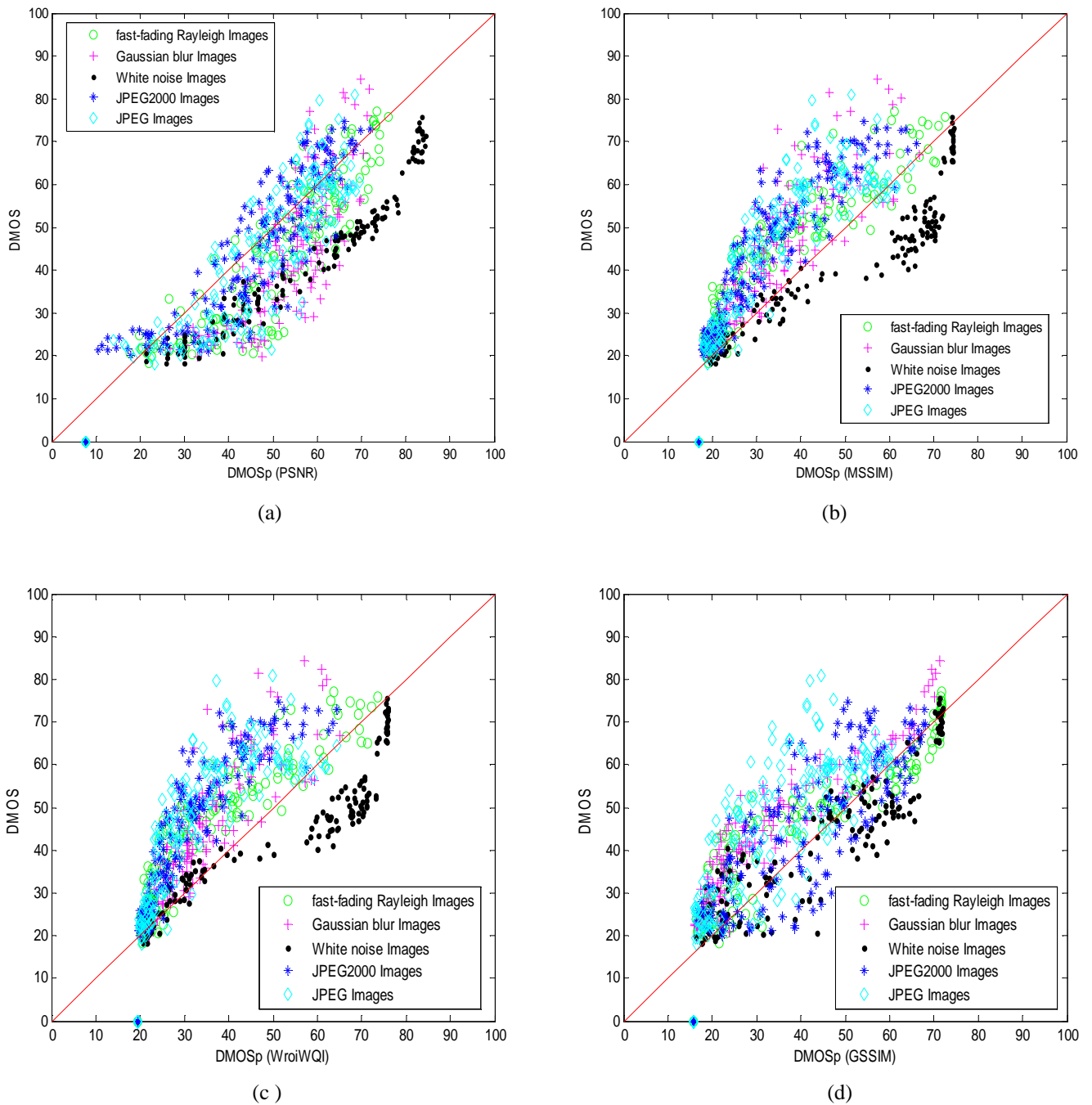


**Figure 4.15:** Graphe de *DMOS* contre les prévisions de modèles de JPEG2000, JPEG. (a) *ERDDM* (b) *VroiWQI* (c) *GSSIM* (d) *PSNR*.

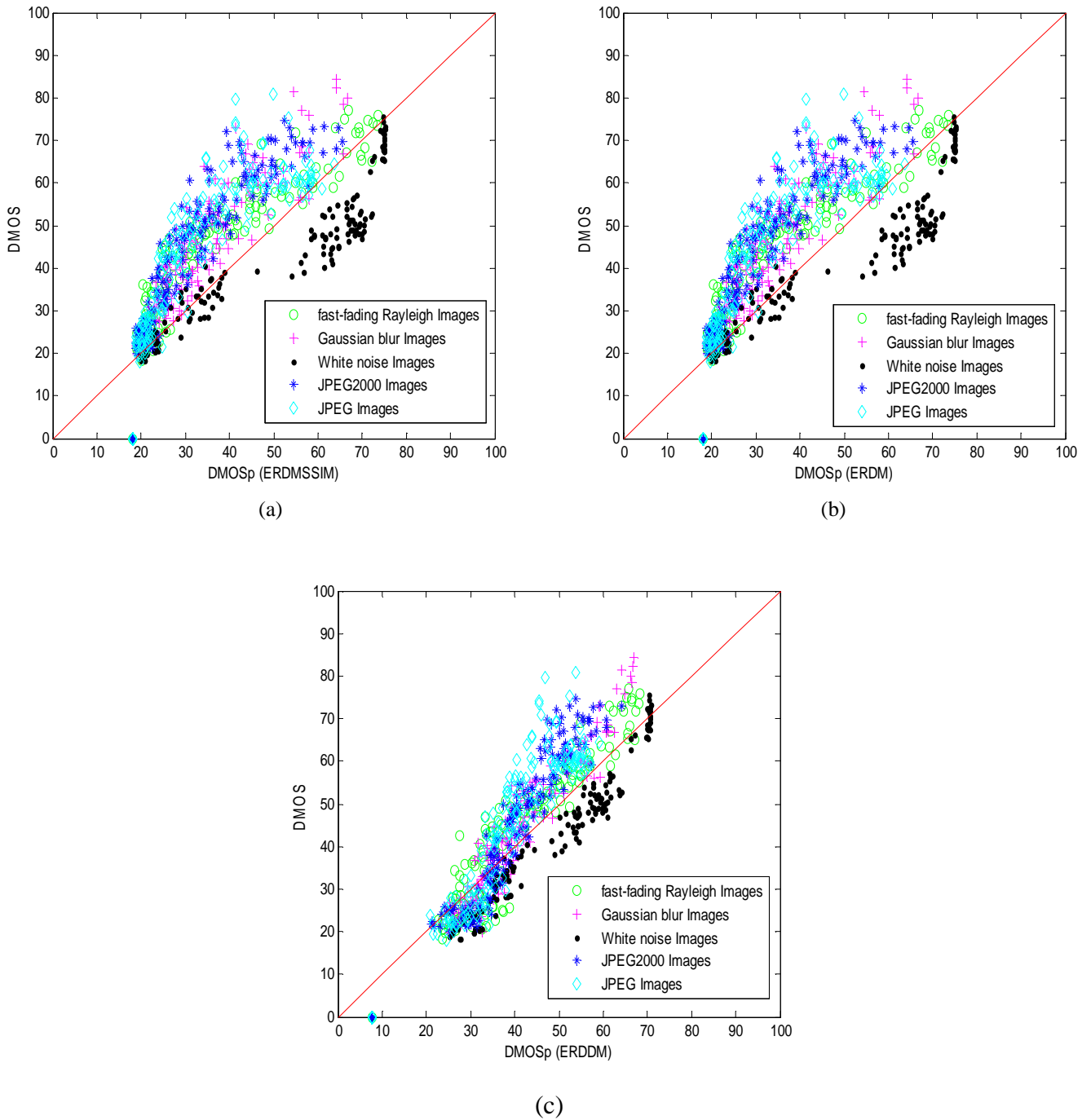




**Figure 4.16:** Graphe de  $DMOS$  contre les prévisions de modèles de JPEG2000, JPEG. (a)  $MSSIM$  (b)  $ERDM$  (c)  $ERDMSSIM$ .



**Figure 4.17:** Graphe de prévisions de qualité par les six méthodes après la régression non linéaire: (a) *PSNR* (b) *MSSIM* (c) *WroiWQI* (d) *GSSIM*.



**Figure 4.18:** Graphe de prévisions de qualité par les six méthodes après la régression non linéaire:(a) *ERDMSSIM* (b) *ERDM* (c) *ERDDM*.

### 4.5.1 Discussion

Les figures 4.13-4.18 montrent les graphes de régression de DMOS contre *PSNR*, *VroiWQI*, *GSSIM*, *MSSIM*, *ERDMSSIM*, *ERDM* et *ERDDM*. Il est clair que l'ajustement entre *ERDDM* et *DMOS* est linéaire, tandis que, comme prévu, l'ajustement entre *PSNR*, *MSSIM*, *VroiWQI*, *GSSIM* et *DMOS* est non-linéaire.

Les tableaux 4.3-4.7 montrent les résultats des cinq métriques de toutes les 982 images appartenant aux cinq déformations de la base Live.

D'après les graphes de régression, on peut voir que la méthode proposée *ERDDM* imite mieux *DMOS* que les méthodes conventionnelles.

A partir des tableaux 4.3-4.7, il en sort que les erreurs *MAE*, *RMS* et *OR* de la mesure proposée est inférieure à celles des méthodes *MSSIM*, *VroiWQI* et *GSSIM*. Des observations semblables sont faites pour la mesure de la qualité d'image par *PSNR*. En outre, l'estimation statistique du *ROCC* et de *cc* fournissent une image claire de l'efficacité de la méthode proposée par rapport à *PSNR*, *VroiWQI* et *GSSIM*. Si les valeurs du *ROCC* et du *cc* sont de près de « 1 » la qualité donnée par la métrique imite très bien *DMOS*. Puisque les valeurs du *ROCC* et du *cc* pour *PSNR* est sont inférieures à un, nous pouvons dire que *ERDDM* est meilleure que la mesure *PSNR*. Les tableaux 4.3-4.7 prouvent que la méthode *ERDDM*, même sous sa forme la plus simple, est concurrentielle avec toutes les méthodes présentées dans ce travail. *PSNR* ne corrèle pas bien avec les évaluations subjectives comme il a été mentionné par les tableaux.

Les résultats comparatifs entre *PSNR*, *MSSIM*, *VroiWQI* et *GSSIM* prouvent que les résultats d'*ERDDM* sont conformes aux points subjectifs mieux que les résultats des *GSSIM*, *MSSIM*, *VroiWQI* et *PSNR*. *GSSIM* combine trois facteurs: la perte de la corrélation, la déformation de la luminance et la déformation du contraste. La valeur calculée par cette mesure n'est pas conforme à l'évaluation subjective.

Bien que *MSSIM* ne soit pas spécifiquement conçu pour les images floues, le *cc* rapporté sur la base d'images était de 0.9355, tandis que la valeur de *cc* de la méthode *ERDDM* est 0.9798.

Les tableaux 4.3-4.7 et les figures 4.13-4.18, ont clairement montré que la méthode proposée est satisfaisante et surpasse les méthodes existantes telles que *MSSIM*, *PSNR*, *VroiWQI* et *GSSIM*.

## 4.6 Conclusion

La méthode proposée *ERDDM* a été testée sur la base d'images Live. Le comportement de la méthode a été étudié en suivant la procédure standard d'évaluation des performances utilisées par le groupe d'expert de la qualité de vidéo (*VQEG*). Cinq métriques ont été calculées (le coefficient de corrélation (*cc*), le coefficient de corrélation des rangs de Spearman (*ROCC*), le rapport des valeurs aberrantes (outlier ratio: *OR*), la racine carrée de la moyenne des erreurs (*RMS*) et l'erreur absolue maximale de prédiction (*MAE*)) afin d'évaluer la performance de la méthode proposée. Une étude comparative entre, les mesures *WroiWQI*, *GSSIM*, *MSSIM*, *ERDMSSIM*, *ERDM* et *ERDDM* a été menée. Cette comparaison nous a conduit à tirer la conclusion que la méthode *ERDDM* donne des résultats satisfaisants par rapport aux autres méthodes.

Les résultats obtenus par *ERDDM* sont compétitifs avec celles des méthodes existant dans la littérature et conformément à l'évaluation subjective.

# Conclusion générale

Dans ce travail, le problème de l'évaluation de la qualité d'images a été abordé. Des méthodes d'évaluation ont été présentées afin d'évaluer la qualité d'une image. Ces méthodes peuvent être divisées en deux classes: subjectives et objectives. Dans la pratique, l'évaluation subjective est habituellement trop inconfortable, longue et très coûteuse. Afin d'éviter ce problème, il est nécessaire d'évaluer la qualité des images compressées ; ceci est réalisé par une mesure de la dégradation en terme d'erreur moyenne quadratique (MSE) ou encore le rapport signal sur bruit (PSNR). La mesure du PSNR jusqu'à aujourd'hui est considérée comme un critère d'évaluation de la qualité le plus utilisé en traitement d'images. Cependant le PSNR est une mesure qualitative qui nécessite parfois une évaluation subjective de la dégradation. En plus, les résultats de ces méthodes ne sont pas souvent conformes à l'évaluation subjective. De ce fait, on fait appel à des méthodes qui reflètent bien l'évaluation humaine telle que MSSIM. MSSIM est basé sur l'hypothèse que le système visuel humain est fortement influencé par les structures présentées dans une scène. MSSIM mesure ainsi la dégradation des structures entre deux images à savoir dégradée et originale. Toutefois, MSSIM a été détecté pour être moins compétitif lorsqu'il est utilisé pour évaluer les images floues et bruitées.

Dans ce travail, nous avons proposé une méthode d'évaluation de la qualité d'image basée sur l'information contour, région, pixels déformés et déplacés. Cette dernière s'intègre dans les méthodes objectives d'évaluation de la qualité d'image avec référence. En outre, après avoir divisé l'image originale et déformée en des blocs de taille  $11 \times 11$  dont la région de chevauchement est en seul pixel, on détermine localement la région d'intérêt, l'information contour ainsi que le déplacement et la déformation de pixels. L'introduction de la mesure des pixels déformés améliore les résultats obtenus.

Une étude comparative entre ERDDM et *PSNR*, *WroiWQI*, *GSSIM*, *MSSIM*, *ERDMSSIM* et *ERDM* a été effectuée. Cette comparaison nous a conduit à tirer la conclusion que la méthode ERDDM donne des résultats satisfaisants par rapport aux autres méthodes.

Pour que ce travail puisse être généralisé, des améliorations sont envisageables telles que :

- L'élaboration d'une base d'images médicales comportant des images originales et dégradées.
- Extension de la méthode proposée ERDDM à une mesure d'évaluation de qualité de vidéo.
- Le problème du choix d'un espace de couleur peut être étudié par ces critères.
- Extension ERDDM à une méthode d'évaluation de la qualité d'image sans référence.



---

---

# Bibliographie

---

- [1] Z. Wang and A.C. Bovik, **Mean squared error: Love it or leave it? - A new look at signal fidelity measures**, IEEE Signal Processing Magazine, Vol: 26 No: 1, January 2009 Page(s): 98-117.
- [2] Z. Wang and A. C. Bovik, "Universal image quality index," IEEE Signal Processing Letters, Vol. 9, pp. 81-84, March 2002.
- [3] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE Image Quality Assessment Database Release 2," available at: <http://live.ece.utexas.edu/research/quality>
- [4] Z.Wang A.C. Bovik, H. R. Sheikh, and E .P. Simocelli, "Image quality assessment: From error measurement to structural similarity," IEEE Trans. Image Processing, vol. 13, no.4, pp.600- 612, Apr .2004.
- [5] S. Daly, "The Visible Difference Predictor: An Algorithm for the Assessment of Image Fidelity ". Digital Images and Human Vision, pages 179–206, 1993.
- [6] W.Osberger, " Perceptual Vision Models for Picture Quality Assessment and Compression Applications ", Queensland University of Technology, Brisbane, Australia, 1999.
- [7] Wang, Z., Lu, L., Bovik, Alan C. "Video Quality Assessment Based on Structural Distortion Measurement", Signal Processing: Image Communication, Vol. 19, No. 2, P. 121-132, 2004
- [8] Chen, G.H.[Guan-Hao], Yang, C.L.[Chun-Ling], Xie, S.L.[Sheng-Li], "Gradient-Based Structural Similarity for Image Quality Assessment", ICIP06(2929-2932).

- [9] A.J. Ahumada, C.H. Null, "Image quality: a multidimensional problem Digital Images and Human Vision" (edited by A.B. Watson), (1993) 141–148.
- [10] S.A. Klein, "Image quality and image compression: a psychophysicist's viewpoint, Digital Images and Human Vision" (edited by A.B. Watson), (1993), 73–88.
- [11] S. Santini, R. Jain, "Similarity measures", IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (9) (1999) 871–883.
- [12] M. P. Eckert and A. P. Bradley, "Perceptual quality metrics applied to still image compression," Signal Processing, vol. 70, pp. 177–200, Nov. 1998.
- [13] T. N. Pappas and R. J. Safranek, "Perceptual criteria for image quality evaluation," in Handbook of Image and Video Proc. (A. Bovik, ed.), Academic Press, 2000.
- [14] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," in The Handbook of Video Databases: Design and Applications (B. Furht and O. Marques, eds.), CRC Press, 2003.
- [15] Z. Wang, A. C. Bovik, and L. Lu. Why is image quality assessment so difficult? In Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing, Orlando, May 2002.
- [16] D. Venkata Rao, N. Sudhakar, I. Ramesh Babu, L. Pratap Reddy,"Image Quality Assessment Complemented with Visual Regions of Interest". ICCTA 2007: 681-687
- [17] Z.A. Seghir, F. Hachouf. "Image Quality Assessment based on Edge-region information and Distorted pixel for JPEG and JPEG2000", in Proc. ACIVS 2009, Bordeaux, France, Sept.28 - Oct.2, LNCS 5807, pp. 156-166.
- [18] Z.A. Seghir, F. Hachouf. "Image Quality Assessment Measure based on Region-displacement with distorted pixels for compressed Images", Third International

Conference on Electrical Engineering Design and Technologies (ICEEDT09), Oct. 31-  
Nov. 2, 2009, Sousse, Tunisia.

<http://journal.esrgroups.org/jes/iceedt/registredpapers.htm>

[19] Z.A. Seghir, F. Hachouf. “Edge-displacement measure for Image Quality Assessment”, International conference On Applied Informatics (ICAI09), November 15-17,2009, Academic centre of Bordj Bou Arreridj, Algeria.

<http://www.icaio9.org/index.php?page=ip&PHPSESSID=a7cc6c5a84fd36d8dc71c0b49b5920de>

[20] Z.A. Seghir, F. Hachouf. “Full Reference Image Quality Assessment for Compressed Image”, International Conference on System and Information Processing (ICSIP'09), May 2-4, 2009, Guelma, Algeria.

<http://www.univ-guelma.dz/icsip09/program.pdf>

[21] Z.A. Seghir, F. Hachouf, S. Chikhi, “Image Quality Assessment Measure Based on Intensity and Position“, (C V A ' 2007), 18-19-20 November 2007, faculty of Electronic engineering and computing, University of Mouloud Mammeri, Tizi-Ouzou, Algeria, pp. 223-226, 2007.

[22] Z.A. Seghir, F. Hachouf, S. Chikhi, “Full Reference Image Quality Assessment Measure based on Color and Position“, (SNIB'2008), 6-8 Mai 2008,Biskra university, Algeria, pp. 227-232, 2008.

<http://www.univ-biskra.dz/manifestations/snib2008/Sessions.htm>

[23] VQEG: Final Report from the video quality experts group on the validation of objective models of video quality assessment, FR-TV Phase II, August 2003

(<http://www.vqeg.org/>)

[24] Z.M. Parvez Sazzad, Y. Kawayoke, Y. Horita, Spatial features based no reference image quality assessment for JPEG2000 , in: Proceedings of the IEEE ICIP, Texas, US, September 2007.

[25] Z.M. Parvez Sazzad, Y. Kawayoke, Y. Horita, “No reference image quality assessment for JPEG2000 based on spatial features”, Signal Process: Image Commu. 23 (2008) 257– 268.

[26] I.v.Z. Marais, W.H. Steyn, “Robust defocus blur identification in the context of blind image quality assessment” , Signal Process: Image Commu.22 (2007) 833–844.

[27] T. Brandão, M.P. Queluz , “No-reference image quality assessment based on DCT domain statistics” , Signal Processing 88 (2008) 822–833.

[28] R. Venkatesh Babu et al, “No-reference JPEG-image quality assessment using GAP-RBF” , Signal Processing 87 (2007) 1493–1503.

[29] H. R. Sheikh, A. C. Bovik, and L. Cormack, “No-reference quality assessment using natural scene statistics: JPEG2000”, IEEE Trans. Image Process., vol. 14, no. 11, pp. 1918–1927, Nov. 2005.

[30] H. R. Sheikh, A. C. Bovik, and G. de Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics”, IEEE Trans. Image Process. 14, pp. 2117–2128, Dec. 2005.

[31] Kalpana Seshadrinathan and Alan C. Bovik. “Unifying analysis of full reference image quality assessment” . In IEEE Intl. Conf. on Image Proc., 2008.

[32] Mathieu Carnec, Patrick Le Callet, Dominique Barba, “Objective quality assessment of color images based on a generic perceptual reduced reference” , Signal Processing: Image Communication, Volume 23, Issue 4, April 2008, Pages 239-256

- [33] S. Suresh, R. Venkatesh Babu, H.J. Kim, “No-reference image quality assessment using modified extreme learning machine classifier”, *Applied Soft Computing*, Volume 9, Issue 2, March 2009, Pages 541-552
- [34] Z. Wang, H.R. Sheikh, A.C. Bovik, “No-reference perceptual quality assessment of JPEG compressed images” , in: *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, 2002, pp. I-477–I-480.
- [35] A.C. Bovik, S. Liu, “DCT-domain blind measurement of blocking artifacts in DCT-coded images” , in *Proceedings of the IEEE International Conference on Acoustic, Speech, and Signal Processing*, vol. 3, 2001, pp. 1725–1728.
- [36] T. Mei, X.-S. Hua, C.-Z. Zhu, H.-Q. Zhou, S. Li, “Home video visual quality assessment with spatio-temporal factors ”, *IEEE Transactions on Circuits and Systems for Video Technology* 17 (6) (2007) 699–706.
- [37] Y. Luo, X. Tang, “Photo and video quality evaluation: focusing on the subject”, in: *ECCV 2008, Part III, LNCS 5304*, 2008, pp. 386–399.
- [38] B.R. Venkatesh, A. Perkis, “An HVS-based no-reference perceptual quality assessment of JPEG coded images using neural networks” , in: *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, 2005, pp. I-433–I-436.
- [39] Z.A. Seghir, F. Hachouf. “Edge-region information with distorted and displaced pixels measure for image quality evaluation” in *Proc. 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010)*, Kuala Lumpur, Malaysia, May 10- 13, pp. 77-80.
- [40] Z.A. Seghir, F. Hachouf. “Edge-region information measure based on deformed and displaced pixel for Image Quality Assessment”, Elsevier, *Journal of Signal Processing*:

Image Communication (Impact Factor: 1.186), October 2011, Volume 26, Issue 8-9, pp. 534-549.

[41] Z.A. Seghir, F. Hachouf. "Image Quality Assessment Measure based on Visual Region and displacement with distorted pixels for compressed Images" *Journal of Automation and Systems Engineering*, vol 4, no. 3, pp. 196-211, September 2010.

[42] Z.A. Seghir, F. Hachouf, F. Nicolier "Distance transform measure based on edge-region information: An algorithm for image quality assessment" in *Proc. 3th Congress on Nature and Biologically Inspired Computing (IEEE, NaBIC2011)*, Salamanca, Spain, Oct 19-21, pp. 125-130.

[43] Y. Horita, Y.Kawayoke, and Z. M. Parvez Sazzad, "Image Quality Evaluation Database," available at: <http://mict.eng.utoyama.ac.jp/mictdb.html>

[44] P. Le Callet, and F. Autrusseau, , "Subjective quality assessment IRCCyN/IVC database," 2005. Available at: <http://www.irccyn.ec-nantes.fr/ivcdb/>

[45] D. M. Chandler, and S. S. Hemami, "VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2284- 2298, 2007. Available at: <http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>

[46] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics," *Advances of Modern Radioelectronics*, Vol. 10, pp. 30-45, 2009. Available at: [www.ponomarenko.info/tid2008.htm](http://www.ponomarenko.info/tid2008.htm)

[47] E. C. Larson, and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no.1, Mar. 2010. Available at: <http://vision.okstate.edu/csiq/>

- [48] Alexandre Ninassi, Patrick Le Callet, Florent Autrusseau, "Pseudo No Reference image quality metric using perceptual data hiding", in SPIE Human Vision and Electronic Imaging, vol. 6057-08, San Jose, CA, USA, January 2006
- [49] A. M. van Dijk, J. B. Martens, and A. B. Watson. "Quality assessment of coded images using numerical category scaling". Proc. SPIE, 2451: 90{101, March 1995.
- [50] <http://www.doc-etudiant.fr/Methodologie/Memoire/Rapport-Detection-des-valeurs-aberrantes-44307.html>
- [51] A. Chetouani. Vers un système d'évaluation de la qualité d'image multi-critères. Thèse de doctorat, l'université paris 13, 2010.
- [52] C. Charrier, C. Larabi, H. Saadane. Evaluation de la qualité des images. Ecole d'hiver sur l'image numérique couleur, 2005.
- [53] ITU-R Recommendation BT.500-10. Methodology for the subjective assessment of the quality of television pictures, March 2000.
- [54] J. Lubin, "The use of psychophysical data and models in the analysis of display system performance," in Digital Images and Human Vision (A. B. Watson, ed.), pp. 163-178, Cambridge, Massachusetts: The MIT Press, 1993.
- [55] J. Lubin, "A visual discrimination model for image system design and evaluation," in Visual Models for Target Detection and Recognition, E. Peli, ed., pp. 207-220, Singapore: World Scientific Publisher, 1995.
- [56] H. R. Sheikh, "Image quality assessment using natural scene statistics", PhD thesis, Faculty of the Graduate School of, The University of Texas at Austin, 2004.
- [57] Zhou Wang, Alan C. Bovik, "Modern image quality assessment", Morgan & Claypool, 2006.



- [58] C. J. van den Branden Lambrecht, "Perceptual models and architectures for video coding applications", PhD thesis, Swiss Federal Institute of Technology, Aug. 1996.
- [59] A. B. Watson, J. Hu, and J. F. III. McGowan, "DVQ: A digital video quality metric based on human vision," *Journal of Electronic Imaging*, vol. 10, no. 1, pp. 20-29, 2001.
- [60] P. C. Teo and D. J. Heeger, "Perceptual image distortion," in *Proc. IEEE Int. Conf. Image Processing*, pp. 982-986, 1994.
- [61] P. C. Teo and D. J. Heeger, "Perceptual image distortion," in *Proc. SPIE*, vol. 2179, pp. 127-141, 1994.
- [62] A. B. Watson, "DCTune: A technique for visual optimization of DCT quantization matrices for individual images," in *Society for Information Display Digest of Technical Papers*, vol. XXIV, pp. 946-949, 1993.
- [63] Alasdair McAndrew , "An Introduction to Digital Image Processing with Matlab Notes for SCM2511 Image Processing 1", School of Computer Science and Mathematics, Victoria University of Technology.
- [64] Madhuri Khambete, and Madhuri Joshi, "Blur and Ringing Artifact Measurement in Image Compression using Wavelet Transform", World Academy of Science, Engineering and Technology 26 2007
- [65] Franck Davoine , « compression d'images par fractales basée sur la triangulation de delaunay », thèse, docteur de l'institut national polytechnique de Grenoble, Décembre 1995.
- [66] Dominique Brunet, « Métriques perceptuelles pour la compression d'images étude et comparaison des algorithmes JPEG et JPEG2000 », Mémoire présenté à la Faculté des études supérieures de l'Université Laval dans le cadre du programme de maîtrise en

mathématiques pour l'obtention du grade de Maître ès sciences (M.Sc.), faculté des sciences et de génie université Laval, Québec Septembre 2007

[67] I. Bloch, Y. Gousseau, H. Maître, D. Matignon, B. Pesquet-Popescu, F. Schmitt, M. Sigelle, F. Tupin « Le traitement des images » tome 2, polycopié du cours ANIM Département TSI - Télécom-Paris version 5.0, 8 décembre 2005

[68] Beretta Patricia. « Compression d'images appliquée aux angiographies cardiaques », thèse de Doctorat. Institut National des sciences Appliquées de Lyon. 1997

[69] Gérard Puimatto. « Notions sur la compression des données dans les applications multimédias ». Note Technique. Centre National de Documentation Pédagogique, 1994.

[70] Patrick Bas, « Compression d'Images Fixes et de Séquences Vidéo », cours ENSERG/INPG, Laboratoire des Images et des Signaux de Grenoble.

<http://www.univ-sba.dz/rcam/liens/docs/courCom.pdf>

[71] Théodore Totozafiny, « compression d'images couleur pour application à la télésurveillance routière par transmission vidéo à très bas débit », thèse de Doctorat, l'université de Pau et des Pays de L'Adour, école doctorale des sciences exactes et de leurs applications, 2007.

[72] Benjamin Navaro et Cédric Rabasse, « JPEG — JPEG2000 », cours de master.

[http://clement.chatelain.free.fr/enseignements/JPEG\\_JPEG2000.pdf](http://clement.chatelain.free.fr/enseignements/JPEG_JPEG2000.pdf)

[73] Alasdair McAndrew, « An Introduction to Digital Image Processing with Matlab », Notes for SCM2511 Image Processing 1, School of Computer Science and Mathematics Victoria University of Technology.

[74] Francis Galton, « Regression Towards Mediocrity in Hereditary Stature », dans *Journal of the Anthropological Institute*, vol. 15, 1886, p. 246-263.

[75] Stephen Stigler, « Karl Pearson's theoretical errors and the advances they inspired », in *Statistical Science*, no 23, 2008, p. 261–271.

[76] V.Planchon « Traitement des valeurs aberrantes : concepts actuels et tendances générales », *Biotechnol. Agron. Soc. Environ.* 2005 p. 19-34. vol. 9 no 1.  
[www.pressesagro.be/base/text/v9n1/19.pdf](http://www.pressesagro.be/base/text/v9n1/19.pdf)

[77] Eric ALATA, "Observation, caractérisation et modélisation de processus d'attaques sur Internet", thèse de Doctorat, de l'université de Toulouse, décembre 2007

[78] F. Grubbs : *Procedures for detecting outlying observations in samples.* *Technometrics*, vol. 11, num. 1, pp. 1–21, 1969.

[79] Carletti G. (1988). « Comparaison empirique de méthodes statistiques de détection de valeurs anormales à une et à plusieurs dimensions ». Gembloux, Belgique : Faculté des Sciences agronomiques de l'état, 225 p.

[80] Munoz-Garcia J., Moreno-Rebollo JL., Pascual-Acosta A. (1990). "Outliers: a formal approach". *Int. Statist. Rev.* **58**, p. 215–226.

[81] Barnett V., Lewis T. (1994). « Outliers in statistical data ». 3<sup>rd</sup> ed. New York: John Wiley .

[82] Benjamin Bringier, « Elaboration de modèles d'évaluation de la qualité du rendu couleur », thèse de Doctorat, de l'université de Poitiers, décembre 2005.

[83] S. Lamrous, « **Compression de données** », cours de l'Université de Technologie de Belfort Montbéliard. [http://www.site-naheulbeuk.com/utbm/5\\_cours\\_compression.pdf](http://www.site-naheulbeuk.com/utbm/5_cours_compression.pdf)