

UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS
Département d'informatique et d'ingénierie

Approche par marquage pour l'évaluation de la qualité d'image dans les applications multimédias

MÉMOIRE (INF6021)
pour l'obtention du grade de
Maîtrise en informatique

PAR

BARHOUMI KHALED

Novembre 2012

Directeur de recherche : Nadia Baaziz, Ing. Ph. D.

Jury

Président du jury : Dr. Marek Zaremba

Membre du jury : Dr. Mohand Saïd Allili

Directeur de recherche : Dr. Nadia Baaziz

Table des matières

Liste des Tableaux	4
Liste des Figures	6
Liste des abréviations, sigles et acronymes	9
Remerciements	10
RÉSUMÉ	11
Chapitre 1 : Introduction	12
Chapitre 2 : Identification du problème et motivation	14
1. Les transformées DCT et DWT	14
2. Image et qualité	18
2.1. Codage et compression	24
a. JPEG et distorsion de l'image	26
2.2. Description du problème	29
2.3. Technologie du watermarking	29
Chapitre 3 : Etat de l'art	35
Chapitre 4 : Objectifs poursuivis et méthodologie	40
1. Objectifs poursuivis	40
2. Méthodologie adoptée	41
3. Méthodes du marquage multiple	42
3.1. Le marquage multiple dans le domaine transformé	43
3.2. Structuration de la marque	46
3.3. Insertion des marques	48
3.4. Détection de la marque	50
3.5. Estimation de la qualité	52
3.6. Marquage avec ajustement des facteurs de robustesse	54
Chapitre 5 : Expérimentations et résultats	57
1. Validation de la procédure d'ajustement	59
2. Génération et choix de la marque	62
3. Efficacité d'estimation de la qualité	64
4. Transcodage et l'évaluation de la qualité	68
5. Effet du bruit additif	69
6. Effet de la perte de paquets	70

7. Domaine d’insertion et les standards de compression	76
8. La performance temporelle.....	79
Chapitre 6 : MPEG et Méthodes d’évaluation de la qualité.....	82
1. Méthodologie	83
2. Insertion de la marque	84
3. Détection de la marque.....	85
4. Estimation de la qualité vidéo	85
5. Problèmes rencontrés	86
Chapitre 7 : Conclusion.....	87
Bibliographie.....	89
Annexe.....	91
Bmpeg	91
Calcul du masque JND	91
Calcul du masque perceptuel	92

Liste des Tableaux

Tableau 2. 1 : Mesure de qualité en PSNR et en JND pour les images Boat, Lena et Goldhill compressées en JPEG avec plusieurs facteurs de qualité.	20
Tableau 2. 2: Définition d'intervalles de qualité d'image en JND et attributs associés.	21
Tableau 2. 3 : Définition d'intervalles de qualité d'image en PSNR(dB) et attributs associés.	22
Tableau 2. 4 : Résultats de compression JPEG de l'image Boat.	27
Tableau 2. 5 : Résultats de compression JPEG2000 de l'image Boat.	27
Tableau 3. 1 : Résumé de quelques méthodes d'évaluation de la qualité d'image par marquage.	39
Tableau 4. 1 : Table de décision à sept intervalles PSNR pour l'évaluation automatique de la qualité d'une image test.	54
Tableau 5. 1: Valeurs des paramètres de marquage et de détection associées à chaque sous-bande DCT de l'image marquée en vue de la compression JPEG2000.	58
Tableau 5. 2 : Impact du marquage multiple Gaussien ou uniforme sur le PSNR de l'image marquée dans la DWT en vue de la compression JPEG2000.	63
Tableau 5. 3 : Résultats d'estimation de la qualité pour les images Boat, Lena, Goldhill et Peppers marquées en DWT et compressées en JPEG à différents taux.	65
Tableau 5. 4 : Résultats d'estimation de la qualité pour les images Boat, Lena, Goldhill et Peppers marquées en DWT et compressées en JPEG2000 à différents taux.	66
Tableau 5. 5 : Transcodage JPEG appliqué aux images Peppers, Goldhill, Lena et Boat préalablement marquées dans le domaine DWT.	68
Tableau 5. 6 : Transcodage JPEG appliqué aux images Peppers, Goldhill, Lena et Boat préalablement marquées dans le domaine DWT.	69
Tableau 5. 7 : Résultats de l'estimation de la qualité suite à l'ajout de bruit Gaussien pour les images Lena, Boat, Peppers et Goldhill marquées et compressées en JPEG.	70
Tableau 5. 8 : Effet de la suppression de quelques bytes du bitstream JPEG des images Lena et Boat compressées à un facteur de qualité $Q=70\%$	73
Tableau 5. 9 : Effet de la suppression (perte) de quelques bytes du bitstream JPEG2000 des images Boat et Lena à un taux de compression égal à 8.	73
Tableau 5. 10 : Effet de la modification de quelques bytes du bitstream JPEG des images Lena et Boat à un facteur de qualité de 70%	74
Tableau 5. 11 : Effet de la modification de quelques bytes du bitstream JPEG2000 des images Lena et Boat à un taux de compression égal à 8.	74
Tableau 5. 12 : Comparaison des valeurs du PSNR des images marquées lorsque le marquage multiple opère dans la DCT versus la DWT en vue de la compression JPEG.	76
Tableau 5. 13 : Comparaison des valeurs du PSNR des images marquées lorsque le marquage multiple opère dans la DCT versus la DWT en vue de la compression JPEG2000.	78
Tableau 5. 14 : Temps d'exécution (en secondes) pour le module de marquage multiple dans les deux domaines DCT et DWT. Les valeurs des facteurs de robustesse α_i sont pré-ajustées manuellement.	79
Tableau 5. 15 : Temps d'exécution (en minutes) pour le module de marquage multiple dans les deux domaines DCT et DWT. Les valeurs des facteurs de robustesse α_i sont ajustées automatiquement en vue de la compression JPEG2000.	80

Tableau 5. 16 : Temps d'exécution (en secondes) pour le module de marquage multiple dans les deux domaines DCT et DWT. Les valeurs des facteurs de robustesse α_i sont ajustées automatiquement en vue de la compression JPEG.....	81
Tableau 5. 17 : Temps d'exécution (en secondes) pour le module de détection/estimation de la qualité dans les deux domaines DCT et DWT.	81

Liste des Figures

Figure 2. 1: DCT 8x8 d'une image montrant : le coefficient DC, les 63 coefficients AC, les fréquences horizontales, verticales, et diagonales, ainsi que le balayage par ordre zig-zag (couleur bleu ciel).....	15
Figure 2. 2 : Représentation des 64 composantes fréquentielles formant la base de décomposition en DCT 8x8.	15
Figure 2. 3 : Processus de décomposition de l'image selon la transformation en ondelette : (j+1) désigne le niveau de décomposition.	17
Figure 2. 4 : Représentation d'une image par sa transformation en ondelettes. Deux niveaux de décomposition sont représentés.....	17
Figure 2. 5 : Représentation de l'image Boat par sa transformation en ondelettes à deux niveaux.	17
Figure 2. 6 : Processus d'évaluation de la qualité d'image avec référence complète (Brandão 2005).....	22
Figure 2. 7 : Processus d'évaluation de la qualité d'image sans référence (Brandão 2005).	23
Figure 2. 8 : Processus d'évaluation de la qualité d'image à référence partielle (Brandão 2005). 23	
Figure 2. 9 : Diagramme représentant les étapes principales de la compression et décompression JPEG en utilisant la transformée DCT.	25
Figure 2. 10 : Diagramme simplifié représentant les étapes principales de la compression et décompression JPEG2000 en utilisant la transformée DWT.	26
Figure 2. 11 : Compression JPEG de l'image Boat avec un taux de compression $T=24.4$ et un PSNR de 30.35 dB. Un effet de blocs est visible.	28
Figure 2. 12 : Compression JPEG2000 de l'image Boat avec un taux de compression $T=24.4$ et un PSNR de 32.03 dB.....	28
Figure 2. 13 : Représentation du processus d'insertion de la marque dans le domaine spatial.	30
Figure 2. 14 : Représentation du processus d'insertion de la marque dans le domaine transformé DCT ou DWT.....	31
Figure 2. 15 : Représentation du processus de détection de la marque par corrélation.....	31
Figure 2. 16 : Représentation du processus d'extraction de la marque et comparaison avec la marque de référence. Le TDR représente le ratio du nombre de bits corrects extraits sur le nombre de bits total de la marque.....	31
Figure 4. 1: Exemple d'images de test	42
Figure 4. 2 : Processus d'évaluation automatique de la qualité d'image.....	43
Figure 4. 3 : Décomposition de l'image hôte (de taille 512x512 pixels) en DWT à 2 niveaux et sélection de trois sous-bandes pour le marquage : les sous-bandes horizontale et verticale du niveau 1, B_1 et B_2 , et la sous-bande verticale du niveau 2, B_3	45
Figure 4. 4 : Extraction des coefficients à partir d'un bloc DCT 8x8 pour former les sous-bandes B_1 , B_2 et B_3 . Chaque sous-bande a une taille 8x8 mais ses coefficients non-nuls sont indiqués par les symboles 1, 2 et 3 pour B_1 , B_2 et B_3 respectivement. Le dernier bloc contient tous les coefficients restants (indiqués par O). La somme des 4 blocs donne le bloc DCT original.	46
Figure 4. 5 : Structuration de la marque W_1 en une matrice 8x8. La séquence pseudo-aléatoire constituante w_1 a une longueur de 8.	47

Figure 4. 6 : Structuration de la marque W2 en une matrice 8×8. La séquence pseudo-aléatoire constituante w_2 a une longueur de 16.	48
Figure 4. 7 : Structuration de la marque W3 en une matrice 8×8. La séquence pseudo-aléatoire constituante w_3 a une longueur de 8.	48
Figure 4. 8 : Localisation des coefficients DCT marqués par les marques W1, W2, et W3 dans chaque bloc DCT 8×8 de l'image hôte.	49
Figure 4. 9 : Localisation des coefficients DWT marqués dans trois sous-bandes de l'image hôte décomposée en deux niveaux DWT.	50
Figure 4. 10 : Procédure d'ajustement des facteurs de robustesse α	55
Figure 5. 1: Réponse du détecteur de marque par rapport à la variation du PSNR de l'image Boat suite à un marquage DCT et à une compression JPEG2000.	59
Figure 5. 2: Pour une valeur $\alpha=0.45$, la marque est détectable au-dessus de 35dB et non détectable au-dessous de 35dB.	60
Figure 5. 3 : Pour une valeur $\alpha = 0.102$, la marque est détectable au-dessus de 40dB et non détectable au-dessous de 40dB.	61
Figure 5. 4 : Comparaison entre les valeurs α ajustées manuellement et automatiquement pour un marquage DWT et une compression JPEG.	61
Figure 5. 5 : Comparaison entre les valeurs α ajustées manuellement et automatiquement pour un marquage DWT et une compression JPEG2000.	62
Figure 5. 6 : Exemple d'une marque W1 à distribution Gaussienne destinée au marquage d'une sous-bande DCT B_1	62
Figure 5. 7 : Exemple d'une marque W1 à distribution uniforme destinée au marquage d'une sous-bande DCT B_1	62
Figure 5. 8 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Peppers lorsque la compression visée est JPEG2000 en utilisant des marques uniformes dans la DWT...	63
Figure 5. 9 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Peppers lorsque la compression visée est JPEG2000 en utilisant des marques Gaussiennes dans la DWT.	64
Figure 5. 10 : Exemple de cascade de codeurs et décodeurs.	68
Figure 5. 11 : Processus de l'ajout de bruit Gaussien dans une image donnée.	70
Figure 5. 12 : Processus de simulation de la perte de paquets dans un bitstream JPEG ou JPEG2000.	71
Figure 5. 13 : Image Lena compressée avec JPEG à un facteur de qualité $Q=70\%$. Son PSNR est 37.06 dB.	72
Figure 5. 14 : Effet de la perte d'un byte du bitstream JPEG de l'image Lena de la Figure 39. ...	72
Figure 5. 15 : Image Lena suite à une modification de 10 bytes du bitstream JPEG et l'image de différence par rapport à l'image compressée (sans perte).	75
Figure 5. 16 : Image Lena suite à une modification de 100 bytes du bitstream JPEG2000 et l'image de différence par rapport à l'image compressée (sans perte).	75
Figure 5. 17 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Boat lorsque la compression visée est JPEG et le marquage opère dans la DCT.	77
Figure 5. 18 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Boat lorsque la compression visée est JPEG et le marquage opère dans la DWT.	77

Figure 5. 19 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Boat lorsque la compression visée est JPEG2000 et le marquage opère dans la DCT.....	78
Figure 5. 20 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Boat lorsque la compression visée est JPEG2000 et le marquage opère dans la DWT.....	78
Figure 6. 1 : Exemple d'un groupe d'image (GOP) de MPEG1 montrant les images intra (I), prédites (P) et bidirectionnelles (B).....	83
Figure 6. 2 : Processus d'estimation de la qualité d'une séquence vidéo Codée en MPEG1.....	84
Figure 6. 3 : Marquage des images intra dans une séquence vidéo destinée au codage MPEG1..	85
Figure 23 : Le codeur bbmpeg.....	91

Liste des abréviations, sigles et acronymes

BER: Bit Error Rate

DCT: Discrete Cosine Transform

DWT: Discrete Wavelet Transform

JND: Just Noticeable Difference

JPEG: Joint Photographic Experts Group

JPEG 2000: Joint Photographic Experts Group 2000

MOS: Mean Score/Opinion Observer

MPEG: Moving Picture Experts Group

PSNR: Peak Signal to Noise Ratio

QoS: Quality of Service

SS: Spread Spectrum

TDR: True Detection Rate

wPSNR: weighted PSNR

AVI : Audio Video Interleave

Remerciements

Mes remerciements vont à toute ma famille qui m'a fourni un support tout au long de mes études.

Je tiens à remercier mon directeur de recherche Mme Nadia Baaziz pour le support qu'elle m'a offert pendant la durée de ma maîtrise et pour ses directives et ses conseils.

Ce projet de recherche est développé grâce au support financier du Centre de Recherches sur les Communications (CRC, Ottawa).

RÉSUMÉ

À l'ère de la technologie de l'information et de l'internet, l'offre de nouveaux services pour la distribution de médias numériques visuels est en pleine expansion. Par conséquent, la qualité d'image perçue devient très importante pour la satisfaction des consommateurs.

En communications multimédias, l'utilisation de techniques de codage avec compression d'images et de vidéo est incontournable si on veut acheminer les données via les canaux de transmission conventionnels dont la bande passante est souvent limitée. Ce qui entraîne nécessairement une dégradation de la qualité des images décodées due à la cascade de codages appliqués ainsi qu'aux effets du canal de transmission. Ainsi la qualité perçue par l'utilisateur final peut varier de manière considérable, et des outils d'évaluation automatique de la qualité s'avèrent utiles pour mieux contrôler la qualité des médias livrés et maintenir la satisfaction du consommateur.

Le sujet de ce mémoire concerne l'évaluation automatique et sans référence de la qualité de l'image suite à l'application de standards de codage et de compressions tels que JPEG, JPEG2000 et MPEG. Plus précisément, deux méthodes opérant respectivement dans les domaines de la transformée en ondelettes et de la transformée en cosinus discrètes sont proposées. Les deux méthodes abordent l'évaluation automatique et sans référence de la qualité de l'image en exploitant de manière originale la technologie du watermarking à spectre étalé. Ainsi, des marques multiples sont destinées à mesurer la qualité de l'image; chaque marque insérée est ajustée à un degré de robustesse de façon à ce que sa détection/non-détection permette de rendre compte d'une gamme de qualité spécifique de l'image. Des tests expérimentaux ont permis de valider l'efficacité de ces méthodes et en évaluer les performances.

Chapitre 1 : Introduction

Suite aux avancements connus dans le domaine des technologies de l'information et des réseaux de distribution de données dans l'internet, les applications multimédias prolifèrent et l'offre de nouveaux services payants pour la distribution de médias numériques est en pleine expansion. Par conséquent, la création, la manipulation et la transmission de données d'images et de vidéos augmentent de manière considérable. À partir de là, la qualité perçue des médias numériques devient importante; les utilisateurs finaux (ou clients/consommateurs) expriment leur besoin d'avoir des services multimédias de qualité et les opérateurs (ou producteurs/distributeurs) tentent de fournir des produits multimédias de haute qualité pour attirer le consommateur et assurer sa satisfaction.

Dans les applications de communications multimédias (telles que la vidéo en demande), l'usage de techniques de codage avec compression d'images et vidéo est incontournable si on veut acheminer les données via les canaux de transmission conventionnels dont la bande passante est souvent limitée (câbles, lignes téléphoniques, réseaux sans fil,...). Ce qui entraîne nécessairement une dégradation de la qualité des images décodées due à la cascade de compressions appliquées (codage et transcodages) ainsi qu'aux effets du réseau de transmission (bruits, congestions, pertes de données, ...).

Ainsi, la qualité perçue par l'utilisateur final peut varier de manière considérable, ce qui va à l'encontre des attentes des utilisateurs. Ces constatations constituent de bons arguments pour le développement d'outils dédiés à l'évaluation automatique de la qualité des images et vidéo. À l'aide de tels outils, les opérateurs pourront ainsi mieux contrôler la qualité des médias délivrés et maintenir la satisfaction des consommateurs.

Le sujet de ce mémoire concerne l'évaluation automatique et sans référence de la qualité de l'image suite à l'application de standards de codage et de compressions tels que JPEG, JPEG2000 et MPEG. Plus précisément, deux méthodes opérant respectivement dans les domaines de la transformée en ondelettes et de la transformée en cosinus discrètes sont proposées. Les deux méthodes abordent l'évaluation automatique et sans référence de la qualité de l'image en exploitant de manière originale la technologie du watermarking à spectre étalé. Ainsi, des marques multiples sont destinées à mesurer la qualité de l'image; chaque marque insérée est ajustée à un degré de robustesse de façon à ce que sa détection/non-détection permette de rendre compte d'une gamme de qualité spécifique de l'image. Des tests expérimentaux ont permis de valider l'efficacité de ces méthodes et en évaluer les performances.

Dans la première section de ce document, le problème et la motivation sont définis et des concepts de base sont présentés au sujet de l'évaluation de la qualité d'image et ses

métriques, les domaines transformés et leur utilisation dans les standards de compression JPEG ainsi que le marquage numérique d'images. La deuxième section présente l'état de l'art relatif aux méthodes d'évaluation de la qualité d'image en se basant sur le marquage numérique. Les objectifs visés, la méthodologie proposée pour ce travail de recherche ainsi que les détails de développement sont exposés dans la troisième section. L'expérimentation et les résultats trouvés sont exposés à la section quatre. Cette section est suivie par la partie d'évaluation de la qualité pour le standard MPEG dans la section cinq. Finalement, la conclusion est présentée à la section six.

Chapitre 2 : Identification du problème et motivation

1. Les transformées DCT et DWT

La transformation en cosinus discrète bidimensionnelle DCT-2D ou DCT (Gonzalez and Woods 2008) est une méthode consistant à transformer l'image du domaine spatial au domaine fréquentiel en la décomposant sur une base de fonctions cosinus selon l'équation (1). La DCT est un mécanisme de base pour dé-corréler les pixels de l'image et regrouper l'énergie du signal image sur quelques coefficients *DC* et *AC* (voir la Figure 2.1). En général, les coefficients de basse fréquence ont des amplitudes non-nulles alors que plusieurs coefficients *AC* de haute fréquence ont des amplitudes faibles. Les méthodes de compression telles que JPEG prennent avantage de cette propriété pour couper l'information au niveau des hautes fréquences alors que les basses fréquences sont préservées. Rappelons que la perte d'information s'effectue par quantification des coefficients DCT ou leur mise à zéro. L'image sera reconstruite en appliquant la transformation inverse DCT^{-1} .

$$B(i, j) = \alpha_i \alpha_j \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) \cos\left(\frac{\pi(2m+1)i}{2M}\right) \cos\left(\frac{\pi(2n+1)j}{2N}\right) \quad (1)$$

Avec M et N désignant le nombre de lignes et de colonnes de l'image I ,

$$0 \leq i \leq M-1$$

$$0 \leq j \leq N-1$$

$$\alpha_i = \begin{cases} \frac{1}{\sqrt{M}}, & i = 0 \\ \frac{\sqrt{2}}{\sqrt{M}}, & 1 \leq i \leq M-1 \end{cases}$$
$$\alpha_j = \begin{cases} \frac{1}{\sqrt{N}}, & j = 0 \\ \frac{\sqrt{2}}{\sqrt{N}}, & 1 \leq j \leq N-1 \end{cases}$$

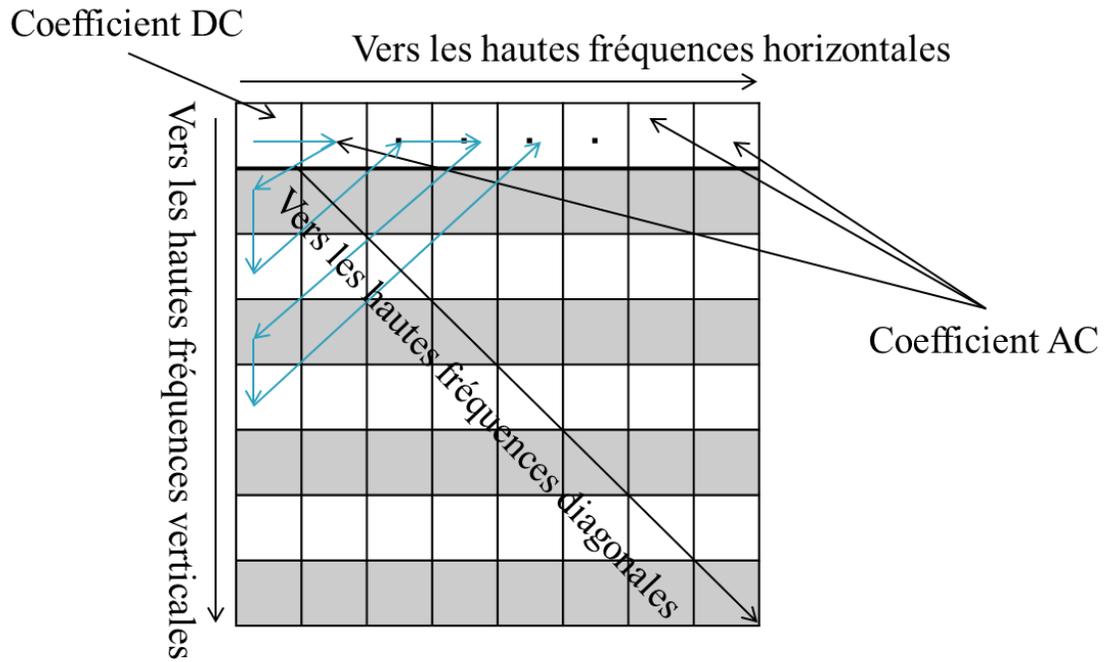


Figure 2. 1: DCT 8x8 d'une image montrant : le coefficient DC, les 63 coefficients AC, les fréquences horizontales, verticales, et diagonales, ainsi que le balayage par ordre zig-zag (couleur bleu ciel).

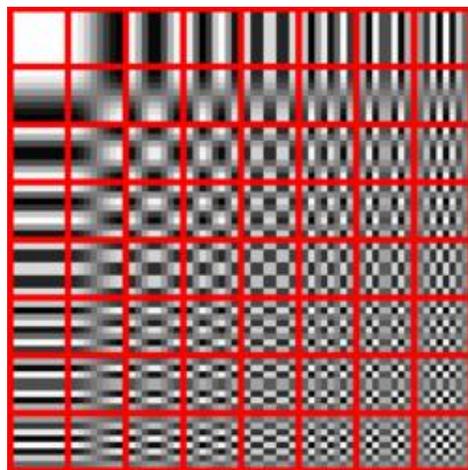


Figure 2. 2 : Représentation des 64 composantes fréquentielles formant la base de décomposition en DCT 8x8.

La transformation en ondelettes DWT (Gonzalez and Woods 2008) est une méthode consistant à transformer l'image du domaine spatial au domaine spatio-fréquentiel. La DWT se base sur des opérations de filtrage passe-haut et passe-bas et d'échantillonnage appliquées selon un algorithme arborescent et récursif (voir la Figure 2.2). Le filtrage consiste à décomposer l'image en sous-bandes et l'échantillonnage vise à réduire la résolution de chaque sous-bande au minimum requis. Pour une décomposition à un seul niveau de résolution, la DWT représente l'image sous forme de quatre sous-bandes de résolution inférieure, une représente l'image d'approximation et les trois autres montrent les détails de l'image à orientations horizontales, verticales, et diagonales tel que montré dans les Figures 2.3 et 2.4. La transformée en ondelettes inverse DWT^{-1} permet une reconstruction parfaite de l'image à partir de ses sous-bandes ondelettes.

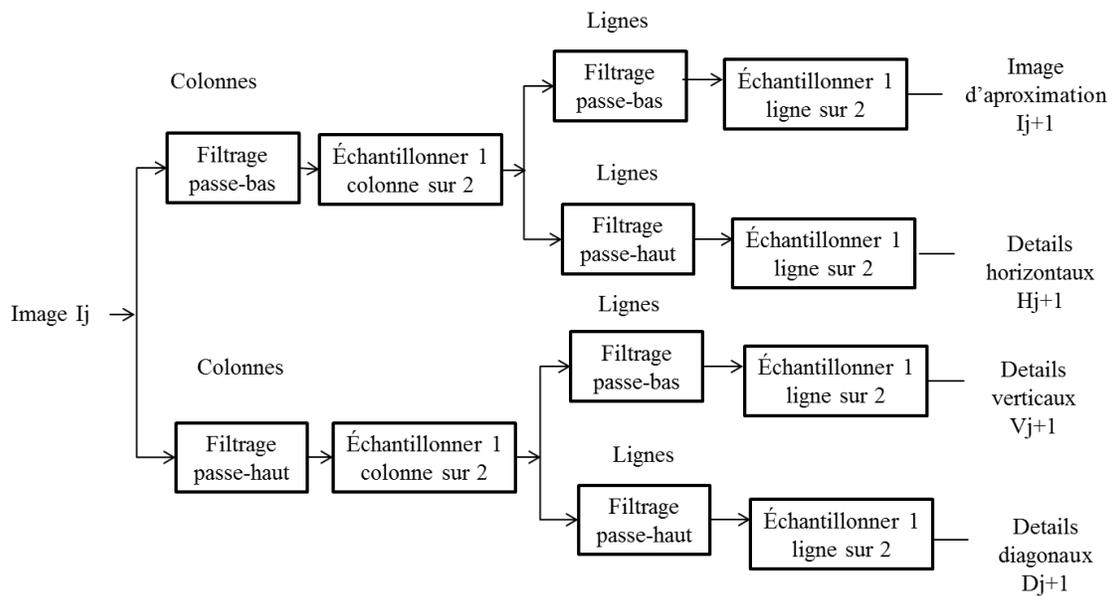


Figure 2. 3 : Processus de décomposition de l'image selon la transformation en ondelette : $(j+1)$ désigne le niveau de décomposition.

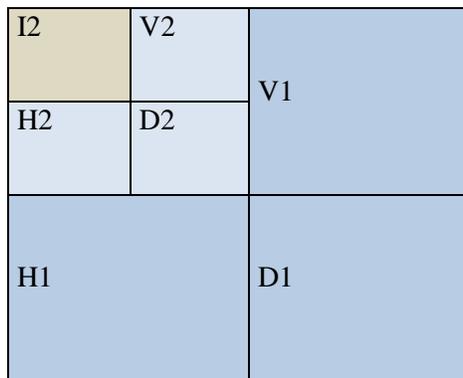


Figure 2. 4 : Représentation d'une image par sa transformation en ondelettes. Deux niveaux de décomposition sont représentés.

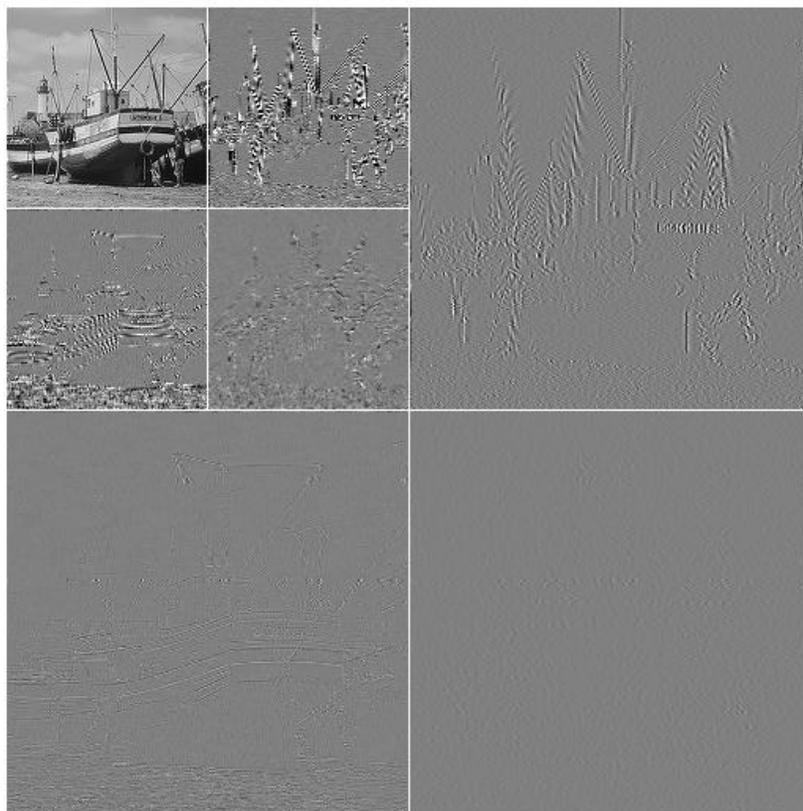


Figure 2. 5 : Représentation de l'image Boat par sa transformation en ondelettes à deux niveaux.

2. Image et qualité

L'évaluation de la qualité de l'image est très demandée à l'heure actuelle dans diverses applications de communications multimédias (Bolot and Vega-Garcia 1996; Gérard 2002). Les divers traitements de codage et de compression subis par les images, ainsi que l'impact des réseaux de communications qui sont traversés durant la transmission des données multimédias, introduisent des distorsions dans l'image et altèrent sa qualité. Les méthodes d'évaluation subjective de la qualité basées sur l'opinion moyenne d'observateurs MOS (MOS : *Mean Score/Opinion Observer*) peuvent s'avérer efficaces en termes de précision de la qualité estimée, mais le processus d'évaluation associé à cette démarche est coûteux, car il nécessite du temps, des ressources humaines et une infrastructure adéquate. Les méthodes d'évaluations objectives de la qualité utilisent des métriques pour calculer mathématiquement la qualité de l'image. Le PSNR (PSNR : *Peak Signal to Noise Ratio*) ou rapport signal à bruit crête est une métrique qui se base sur le calcul de l'erreur quadratique moyenne (EQM) entre les pixels de l'image de référence I et l'image distordue I' selon la formule suivante :

$$PSNR_{db} = 10 \log_{10} \frac{MN(2^n-1)^2}{\sum_{i=1}^M \sum_{j=1}^N (I(i,j) - I'(i,j))^2} \quad (2)$$

M et N étant le nombre de lignes et de colonnes des images I et I' . Une valeur de PSNR supérieure à 40 dB est en général un indicateur d'une très bonne qualité d'image (i.e. la dégradation subie est en général imperceptible).

Une autre métrique pour l'évaluation objective de la qualité est la distance de Watson communément appelée la métrique JND (JND : *Just Noticeable Difference*) de Watson (Watson 1993; Cox, Miller et al. 2008). Dans le domaine des études psycho-physiques, l'unité JND se définit comme étant le niveau de distorsion qui peut être perçu dans 50% des cas d'essais ou de tests expérimentaux. La métrique JND selon Watson est un modèle perceptuel qui tente d'incorporer des caractéristiques du système visuel humain dans la mesure perceptuelle de la distance entre l'image originale et l'image distordue. Trois caractéristiques du SVH sont prises en compte : la sensibilité fréquentielle, le masquage de la luminance et le masquage du contraste.

La sensibilité fréquentielle du SVH : le système visuel humain est très sensible aux changements dans les composantes fréquentielles basses et médianes de l'image. Watson définit une table de sensibilité fréquentielle t de taille 8×8 en lien avec les composantes fréquentielles des blocs DCT 8×8 de l'image (voir le Tableau 2.1). Chaque élément $t(i,j)$ correspond au coefficient DCT $B(i,j)$ et définit l'amplitude du changement qui produit

une distorsion de un JND dans l'image. Par exemple, $t(0,1)=1.01$ signifie que toute variation du coefficient $B(0,1)$ dépassant l'amplitude 1.01 va provoquer une distorsion supérieure à un JND dans l'image. En observant la table t , on peut constater que les composantes fréquentielles basses et médianes ont des petites valeurs $t(i,j)$, ce qui signifie que les changements dans ces composantes vont avoir un grand impact sur la quantité de distorsion perçue dans l'image. Il est à noter que la table de sensibilité fréquentielle t montrée ci-dessous a été calculée et présentée dans (Cox, Miller, Bloom, Fridrich & Kalker, 2008) en se basant sur l'algorithme proposé dans. Mais les paramètres choisis pour calculer cette table n'ont pas été précisés.

Tableau 2. 1: La table de sensibilité fréquentielle t définie dans (Cox, Miller, Bloom, Fridrich & Kalker, 2008).

1.40	1.01	1.16	1.66	2.40	3.43	4.79	6.56
1.01	1.45	1.32	1.52	2.00	2.71	3.67	4.93
1.16	1.45	2.24	2.59	2.98	3.64	4.60	5.88
1.66	1.52	2.59	3.77	4.55	5.30	6.28	7.60
2.40	2.00	2.98	4.55	6.15	7.46	8.71	10.17
3.43	2.71	3.64	5.30	7.46	9.62	11.58	13.51
4.79	3.67	4.60	6.28	8.71	11.58	14.50	17.29
6.56	4.93	5.88	7.60	10.17	13.51	17.29	21.15

- 1- Le masquage de la luminance : Si l'intensité moyenne du bloc image considéré est très élevée, cette région sera alors capable d'absorber de plus grandes variations car une brillance élevée procure un effet de masquage des changements subis. Le bloc DCT correspondant peut alors subir des changements plus importants que ce qui est prévu dans la table t sans être perceptibles. Pour tenir compte de ce phénomène, la table de sensibilité est alors ajustée pour donner lieu à un masque de luminance t_L défini sur toute l'image comme suit :

$$t_L(i, j, k) = t(i, j) * \left(\frac{B(0,0,k)}{B_{0,0}} \right)^{a_T} \quad (3)$$

k étant le numéro du bloc 8×8 , $B(0,0,k)$ le coefficient DC du k ème bloc, $B_{0,0}$ la moyenne de tous les coefficients DC de tous les blocs de l'image et a_T une constante valant 0.649 .

- 2- Le masquage du contraste: lorsqu'une composante fréquentielle a une énergie élevée, ceci a pour effet de réduire la visibilité d'un changement sur cette

composante. La prise en compte de ce type de masquage donne lieu à un masque de seuils comme suit :

$$s(i, j, k) = \max\{t_L(i, j, k), |B(i, j, k)|^w t_L(i, j, k)^{1-w}\} \quad (4)$$

Où w est une constante qui vaut 0.7. Notons que chaque $s(i,j,k)$ réfère à un seuil de variation que peut subir tout coefficient DCT individuel $B(i,j,k)$ avant d'atteindre une distorsion visible de un JND.

- 3- Calcul de la distance : finalement, il est important de se donner un moyen de combiner ensemble tous les effets des changements individuels et l'exprimer sous forme d'une mesure de distorsion globale. À cet effet, Watson propose la mesure de distance suivante :

$$D_{wat}(I, I') = (\sum_{i,j,k} |d[i, j, k]|^p)^{1/p} \quad (5)$$

où p est une constante dont la valeur recommandée par Watson est 4, et $d(i, j, k)$ est donné par :

$$d[i, j, k] = \frac{e[i, j, k]}{s[i, j, k]} \quad (6)$$

où $e[i, j, k]$ est la différence entre l'image originale I et l'image distordue I' .

Le Tableau 2.2 montre des expérimentations qui sont effectuées sur les images Boat, Lena et Goldhill. Ces images sont d'abord compressées en JPEG selon plusieurs facteurs de qualité, puis décompressées. La dégradation de la qualité due à cette compression-décompression est calculée en utilisant les métriques PSNR et JND. Il est intéressant d'établir un lien entre le PSNR, le JND et une description de la qualité perçue (visuellement). Ces expérimentations nous ont permis d'établir des plages JND auxquelles nous avons associé des attributs de qualité, tel que c'est montré dans le Tableau 2.3. Le Tableau 2.4 présente un exemple d'échelle PSNR-attributs telle qu'elle a été introduite dans (Baaziz, Zheng et al. 2011).

Tableau 2. 2 : Mesure de qualité en PSNR et en JND pour les images Boat, Lena et Goldhill compressées en JPEG avec plusieurs facteurs de qualité.

Image	Facteur de qualité JPEG (Q)	PSNR de l'image décodée	JND de l'image décodée
	100	58.48	3.55
	95	43.10	11.09
	90	39.30	15.99

Goldhill	85	37.54	18.32
	80	36.50	19.69
	75	35.71	20.92
	70	35.16	21.70
	65	34.66	22.46
	35	32.55	25.94
	30	32.10	26.72
	25	31.55	27.62
	20	30.86	28.78
	15	29.94	30.29
Lena	100	58.46	3.90
	95	43.78	11.00
	90	40.78	15.38
	85	39.40	17.41
	80	38.50	18.73
	75	37.80	19.71
	65	36.83	21.04
	35	34.74	23.89
	25	33.69	25.21
	20	32.94	26.12
	15	31.93	27.35
	10	30.39	29.09
Boat	100	58.49	3.74
	95	44.12	9.92
	90	40.96	14.33
	85	39.27	16.54
	80	38.15	17.91
	75	37.27	18.88
	70	36.62	19.63
	60	35.53	20.93
	35	33.41	23.52
	30	32.83	24.25
	25	32.20	25.09
	20	31.41	26.14
	15	30.35	27.57

Tableau 2. 3: Définition d'intervalles de qualité d'image en JND et attributs associés.

Attribut	Très bonne (<i>Very good</i>)	Bonne (<i>Good</i>)	Acceptable (<i>Acceptable</i>)	Mauvaise (<i>Poor</i>)
JND	≤15	15-21	21-28	>28

Tableau 2. 4 : Définition d'intervalles de qualité d'image en PSNR(dB) et attributs associés.

Attribut	Très bonne (<i>Very good</i>)	Bonne (<i>Good</i>)	Acceptable (<i>Acceptable</i>)	Mauvaise (<i>Poor</i>)
PSNR(dB)	≥ 40	40-35	30-35	< 30

Les diverses méthodes d'évaluation objective de la qualité d'image se classent dans trois catégories distinctes (Kusuma and Zepernick 2003) selon le degré d'utilisation des données de référence :

- **Évaluation de la qualité à référence complète** (voir Figure 2.5) : Ces méthodes d'évaluation de la qualité requièrent un accès total à l'image originale. Dans le cas d'une application en communications, beaucoup plus de ressources sont exigées du fait qu'il faut transmettre toutes les données de l'image originale. Pour cela, l'évaluation de qualité à référence complète est utilisée uniquement comme un moyen pour évaluer des algorithmes de traitement d'image et de vidéo (des essais dans les laboratoires). Par conséquent, on ne peut pas l'utiliser de manière pratique dans une application impliquant la transmission d'image et de vidéo.
- **Évaluation de la qualité sans référence** (voir Figure 2.6) : les algorithmes d'évaluation de la qualité accèdent seulement à l'image dégradée, pour estimer sa qualité sans avoir aucune information sur l'image originale.
- **Évaluation de la qualité à référence partielle** (voir Figure 2.7) : ces méthodes rendent secondaires la conservation de l'image originale au complet puisqu'une information partielle est suffisante. Par exemple, quelques caractéristiques de l'image dégradée sont calculées et comparées avec les caractéristiques sauvegardées correspondant à l'image originale.

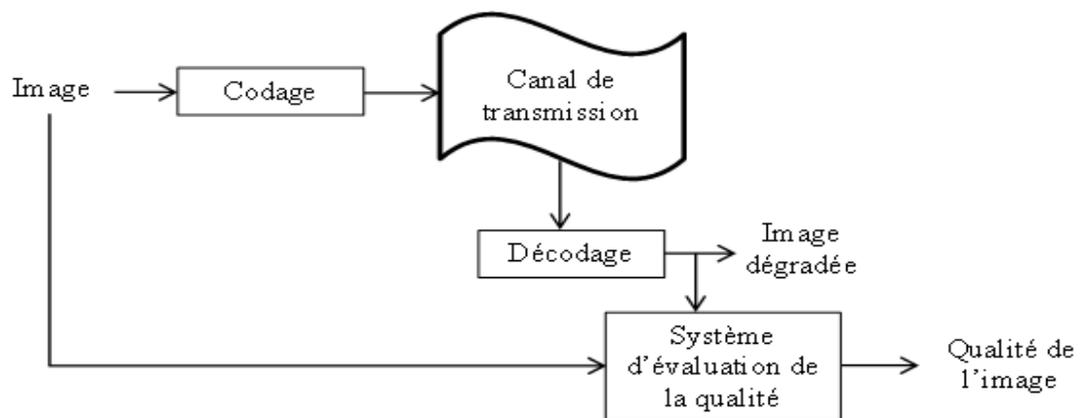


Figure 2. 6 : Processus d'évaluation de la qualité d'image avec référence complète (Brandão 2005).

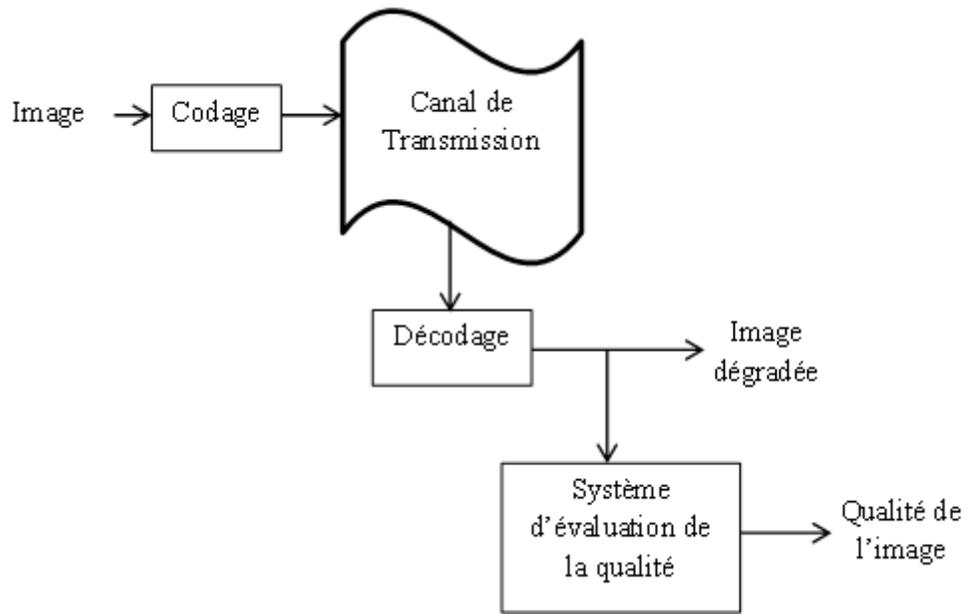


Figure 2. 7 : Processus d'évaluation de la qualité d'image sans référence (Brandão 2005).

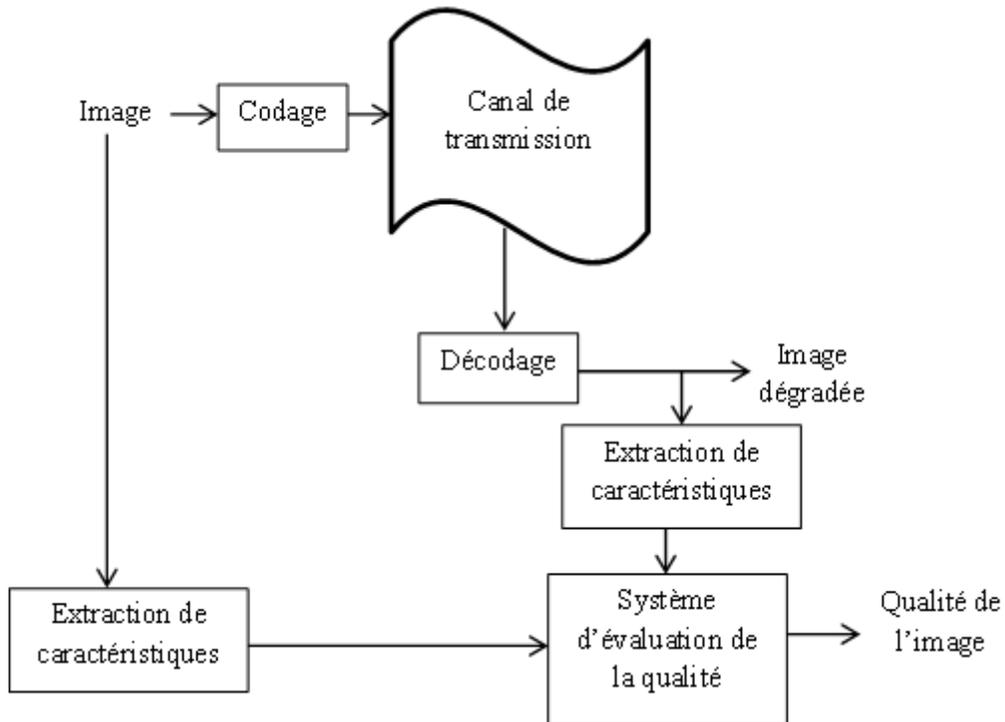


Figure 2. 8 : Processus d'évaluation de la qualité d'image à référence partielle (Brandão 2005).

2.1. Codage et compression

Les méthodes de compression et de codage d'images et vidéo visent à réduire le nombre de bits alloués à chaque pixel de l'image et à trouver une représentation binaire compacte capable d'améliorer l'efficacité du stockage ou de la transmission du média. Dans la plupart des images et des séquences d'images, les pixels voisins sont corrélés, et donc contiennent des informations redondantes. Trois types de redondances sont identifiés. D'abord, la redondance spatiale ou la corrélation entre les pixels voisins, ensuite, la redondance spectrale ou la corrélation entre les plans de différentes couleurs, et enfin, la redondance temporelle ou la corrélation entre les images adjacentes dans une séquences d'images. La compression exploite ces redondances et cherche à les diminuer autant que possible.

Les méthodes de compression se classifient en deux catégories :

- La compression sans perte : l'image reconstituée (décodée) est numériquement identique à l'image originale.
- La compression avec perte : l'image reconstituée présente une dégradation par rapport à l'image originale.

La compression d'images avec perte se base essentiellement sur une opération de quantification qui consiste à approximer les données par un ensemble fini de valeurs afin de réduire le nombre de bits requis pour leur codage. Cette opération engendrant une perte d'information est encore plus efficace lorsqu'elle est appliquée sur une image préalablement transformée dans le domaine fréquentiel où les données sont moins corrélées.

Des taux de compression élevés peuvent être atteints par les méthodes de compression avec pertes. Cela dit, cette compression est atteinte au prix de distorsions d'images qui peuvent parfois altérer de manière significative la qualité visuelle de l'image. Rappelons que le taux de compression est défini par :

$$Taux_{compression} = \frac{Taille\ de\ l'image\ originale\ (bits)}{Taille\ de\ l'image\ compressée\ (bits)} \quad (7)$$

En plus de l'appréciation visuelle, Le PSNR est communément utilisé pour mesurer la qualité de l'image décompressée. Le PSNR ou rapport signal à bruit crête est une métrique qui se base sur le calcul de l'erreur quadratique moyenne entre les pixels de l'image de référence et l'image distordue selon l'équation (2).

Une valeur de PSNR supérieure à 40 dB est en général un indicateur d'une très bonne qualité d'image (i.e. la dégradation subie est en général imperceptible).

Parmi, les standards de compression on cite JPEG et JPEG2000. Le standard de compression d'images JPEG (Gonzalez, Woods et al. 2004 ; Hwang 2009) utilise la transformation en cosinus discrète DCT comme un mécanisme de base pour dé-corréler les pixels de l'image. Dans le codeur JPEG, l'image est d'abord découpée en blocs de 8×8 pixels. Ensuite, à chacun de ces blocs est appliquée une transformation DCT qui vise à transformer ce bloc d'image du domaine spatial au domaine fréquentiel. Les coefficients transformés sont alors quantifiés puis codés dans un format binaire tel que montré dans la Figure 2.8. Lors du décodage du train de bits JPEG, l'image sera reconstruite en appliquant les opérations inverses (décodage binaire, et DCT inverse) mais la perte d'information due à la quantification ne sera jamais compensée, ce qui va se traduire par des dégradations dans l'image décodée.

JPEG2000 (Christopoulos, Skodras et al. 2000; Hwang 2009) est une autre norme de codage d'image plus flexible et plus performante. Elle est apparue après la norme JPEG en ajoutant d'autres fonctionnalités telles que : la compression sans pertes, la sélection de régions d'intérêt, l'accès et décompression aléatoire du bitstream et le traitement d'une large variété d'images. Le standard JPEG2000 utilise la transformation en ondelettes DWT lors de la compression et la DWT inverse lors de la décompression tel que montré dans la Figure 2.9. La DWT est considérée comme la plus grande rivale de la DCT parce qu'elle offre une description fréquentielle plus sélective du signal analysé du fait qu'elle décompose l'image en sous bandes fréquentielles orientées.

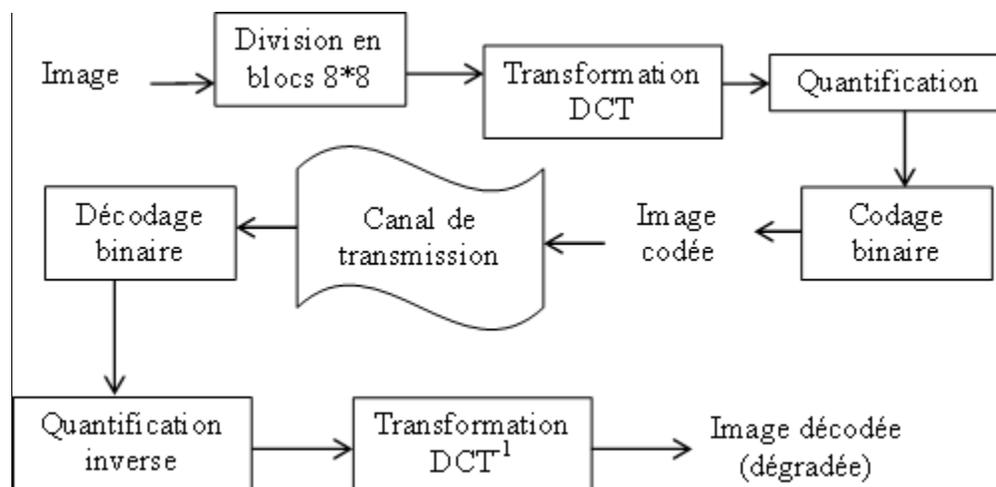


Figure 2. 9 : Diagramme représentant les étapes principales de la compression et décompression JPEG en utilisant la transformée DCT.

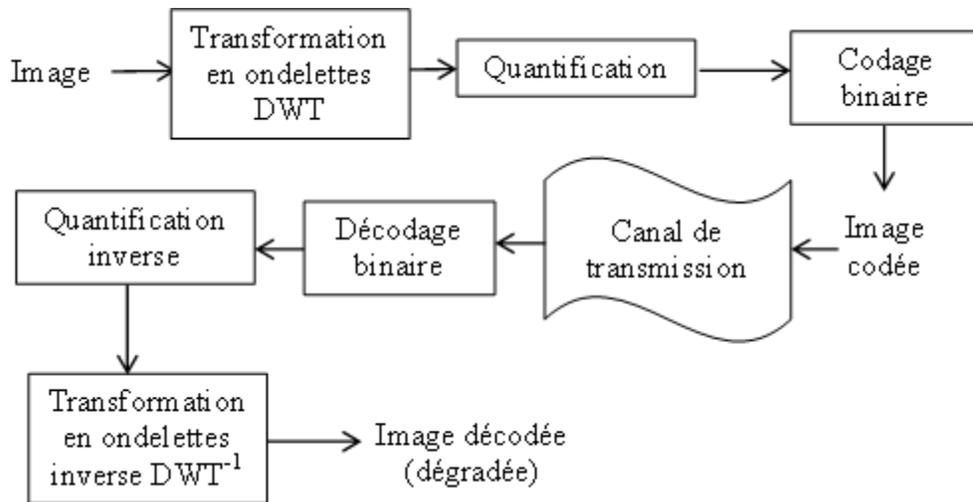


Figure 2. 10 : Diagramme simplifié représentant les étapes principales de la compression et décompression JPEG2000 en utilisant la transformée DWT.

a. JPEG et distorsion de l'image

La compression d'image avec perte utilisant les codecs JPEG ou JPEG2000 entraîne une dégradation de la qualité de l'image à cause de la quantification subie. En général, plus le taux de compression est élevé, plus l'image décodée est distordue.

Les tableaux 2.5 et 6 montrent les résultats de compression d'une image en termes de taux de compression, de PSNR et de nombre de bits par pixels en utilisant les standards JPEG et JPEG2000 (respectivement). On peut constater que le PSNR chute à fur et à mesure que le taux de compression augmente. Le standard JPEG2000 assure une meilleure qualité d'image, car pour un taux de compression de même grandeur, le PSNR de l'image JPEG2000 décodée est supérieur (d'au moins 2dB) à celui de l'image JPEG décodée.

Tableau 2. 5 : Résultats de compression JPEG de l'image Boat.

Taux de compression	PSNR de l'image décodée	Nombre de bit par pixel
1,73	58.4968	5,09
4,6	40.9639	1,93
7,9	37.2704	1,09
11,87	34.7122	0,74
16,40	32.8391	0,54
24,4	30.3561	0,35
47,6	26.0339	0,18

Tableau 2. 6 : Résultats de compression JPEG2000 de l'image Boat.

Taux de compression	PSNR de l'image décodée	Nombre de bit par pixel
2,40	54.4685	3,6
5,18	43.5542	1,7
8,90	39.2810	1
13,57	36.2888	0,7
17,80	34.2320	0,5
25,90	32.0337	0,0125
47,50	29.0462	0,004

De plus, les Figures 2.10 et 2.11 montrent que pour un même taux de compression de 24.4 atteint par JPEG et JPEG2000, l'image JPEG décodée a un PSNR plus faible et des distorsions dues à l'effet de blocs sont visibles alors que la qualité visuelle ainsi que le PSNR sont nettement meilleurs dans le cas JPEG2000.



Figure 2. 11 : Compression JPEG de l'image Boat avec un taux de compression $T=24.4$ et un PSNR de 30.35 dB. Un effet de blocs est visible.



Figure 2. 12 : Compression JPEG2000 de l'image Boat avec un taux de compression $T=24.4$ et un PSNR de 32.03 dB.

2.2. Description du problème

Dans un contexte d'applications multimédias où la surveillance automatique de la qualité d'image est hautement désirée (Bolot and Vega-Garcia 1996; Gérard 2002), il pourrait être impossible d'allouer une bande passante supplémentaire pour envoyer le contenu de l'image originale. Par conséquent, une évaluation de la qualité sans référence ou à référence partielle serait plus appropriée pour ce type d'applications.

Une des approches récentes et prometteuses pour l'évaluation de la qualité d'image sans référence s'appuie sur la technologie du marquage numérique ou watermarking; la marque insérée dans l'image originale sera soumise aux mêmes traitements subis par l'image hôte et sera donc dégradée lorsque l'image hôte est distordue par l'effet de la compression ou bien par l'effet du canal de transmission. À la réception, la marque est soit extraite ou détectée, et la mesure de sa dégradation (comparativement à la marque originale) est utilisée pour estimer la dégradation globale endurée par l'image. Il existe seulement quelques contributions (une dizaine environ) dans ce domaine.

L'ensemble des méthodes proposées partagent les faiblesses suivantes : a) manque de précision sur la qualité de l'image évaluée, et/ou b) besoin d'informations auxiliaires à transmettre pour effectuer l'évaluation. De ce fait, un effort en recherche et développement est requis pour accroître les contributions théoriques et pratiques dans ce domaine de recherche.

2.3. Technologie du watermarking

Vers les années 1990, une nouvelle technologie est apparue à savoir le marquage numérique des signaux, appelée aussi tatouage numérique ou watermarking (Nguyen and Baudry 2003; Cox, Miller et al. 2008). La plus grande partie des applications de cette technologie est reliée à des applications de sécurité des données de médias audio et visuels, tel que la protection de droit d'auteur, le contrôle de l'intégrité des données, l'authentification, etc. Le processus de marquage comporte trois étapes:

- 1- **La génération de la marque** : En général, la marque est un message transformé en une matrice W dont les valeurs sont binaires $\{\pm 1\}$ ou ternaires $\{1, 0, -1\}$.
- 2- **L'insertion de la marque** : Étant donné une image hôte I , l'insertion se fait soit dans le domaine spatial de l'image (dans les pixels) soit dans le domaine transformé de l'image (DCT, DWT, ...). De façon générale, l'insertion est une fonction qui prend en entrée l'image hôte I (ou image originale) et la marque

générée W , et délivre en sortie l'image marquée I' . Ce processus est illustré dans les Figures 2.12 et 2.13.

- 3- **La détection/extraction de la marque** : Étant donné une image test I' et une marque W , cette étape consiste à analyser l'image test I' et vérifier si la marque W est présente dans cette image. La vérification de la présence de la marque se fait par détection ou extraction :
- Détection : Mesure du degré de présence de la marque par une fonction de corrélation entre I' et W . La comparaison à un seuil de corrélation aboutit à une décision binaire (marque détectée/marque non détectée), tel qu'illustré dans la Figure 2.14.
 - Extraction : Extraire les bits de la marque W' contenue dans l'image test I' et la comparer à la marque originale W en utilisant une métrique donnée tel que le TDR (TDR : *True Detection Rate*) signifiant le nombre de bits correctement extraits sur le nombre de bits total de la marque. Une autre métrique utilisée est le BER (BER : *Bit Error Rate*) indiquant le ratio du nombre de bits extraits et erronés sur le nombre de bits total de la marque. Ce processus est illustré par la Figure 2.15.

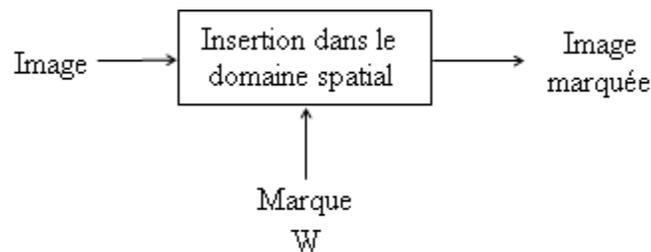


Figure 2. 13 : Représentation du processus d'insertion de la marque dans le domaine spatial.

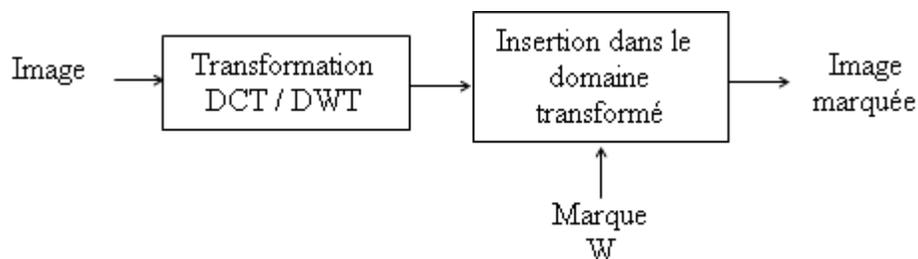


Figure 2. 14 : Représentation du processus d’insertion de la marque dans le domaine transformé DCT ou DWT.

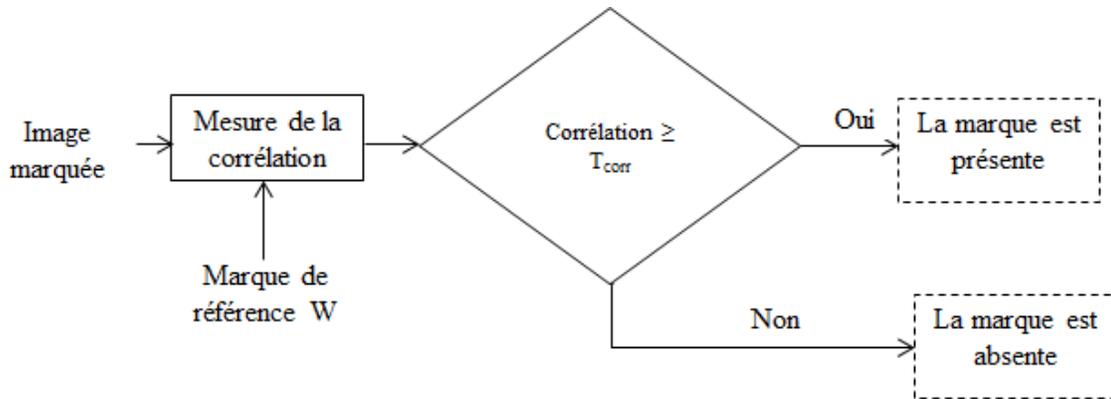


Figure 2. 15 : Représentation du processus de détection de la marque par corrélation.

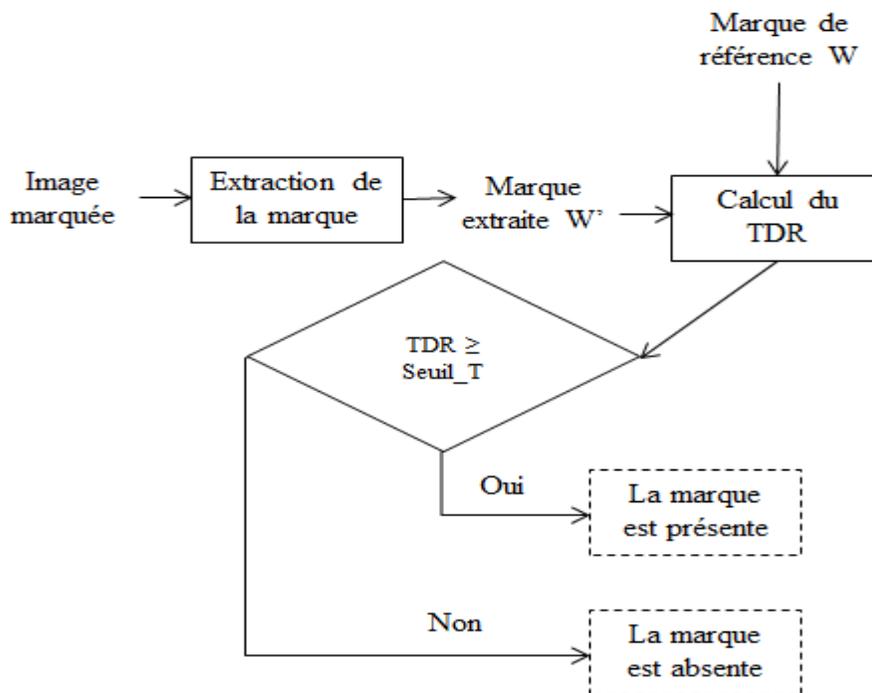


Figure 2. 16 : Représentation du processus d’extraction de la marque et comparaison avec la marque de référence. Le TDR représente le ratio du nombre de bits corrects extraits sur le nombre de bits total de la marque.

Parmi les contraintes techniques importantes du marquage, on cite :

- L’invisibilité : il faut que l’impact visuel du marquage de l’image soit très faible pour que l’image marquée reste fidèle à l’image originale.

Pour mesurer le degré de l'invisibilité, le PSNR est communément utilisé.

- La robustesse : garantit une protection de la marque, et ce en résistant aux différentes dégradations ou attaques que peut subir une image. La fragilité est le contraire de la robustesse; une marque W est fragile si sa détection échoue lors de la modification d'un seul pixel de l'image.
- La capacité : qui est le nombre de bits que l'on peut cacher dans l'image hôte. Il existe un compromis entre la capacité et l'invisibilité d'un marquage, si on insère une marque de grande taille, l'image marquée risque d'être bien dégradée et la marque perd son invisibilité.

Parmi les différentes méthodes de marquage, on rapporte le marquage à quantification et le marquage à spectre étalé SS (SS : *Spread Spectrum*). Ces deux méthodes peuvent s'appliquer aussi bien dans le domaine spatial que dans le domaine transformé de l'image. Dans ce qui suit, on considère $I(i,j)$ comme étant le coefficient hôte pouvant être un pixel d'image ou bien un coefficient transformé (DCT ou DWT).

- 1) Le marquage par quantification ([Wang, Zheng et al. 2007](#); [Cox, Miller et al. 2008](#)) consiste à modifier les coefficients hôtes en utilisant une opération de quantification avec un pas de quantification q . Dans l'exemple illustré par l'équation (8), la marque w à insérer est binaire avec des valeurs 0 (paire) ou 1 (impaire). Le pas de quantification q est un entier positif de valeur paire. Un coefficient de l'image hôte $I(i,j)$ subit l'insertion d'un bit de marque w en rendant la parité du coefficient quantifié $I'(i,j)$ identique à la parité de w .

$$I'(i,j) = \begin{cases} q * \left\lfloor \frac{I(i,j)}{q} \right\rfloor & \text{si } w \text{ pair} \\ q * \left\lfloor \frac{I(i,j)}{q} \right\rfloor + 1 & \text{si } w \text{ impair} \end{cases} \quad (8)$$

$$w = \begin{cases} 1, & \text{si } I'(i,j) \text{ est impair} \\ 0, & \text{si } I'(i,j) \text{ est pair} \end{cases} \quad (9)$$

Où $\lfloor x \rfloor$ représente la troncature à l'entier du nombre x . Notons que le pas de quantification q peut prendre la valeur $q=2$ puisque la marque est binaire. Ainsi pour l'extraction de la marque, il suffit de déterminer la parité du coefficient marqué, selon l'équation (9).

Exemple : $q=2$;

Matrice I

12	34	50	60
20	27	32	41
19	17	24	13
18	11	22	27

Marque w

1	1	0	0
1	0	1	1
0	0	0	1
1	0	1	0

Matrice I' (marquée)

13	35	50	60
21	27	33	42
19	17	24	14
19	11	23	27

- 2) L'insertion du watermark peut se faire par la méthode du spectre étalé SS (SS : *Spread Spectrum*) (Pickholtz, Schilling et al. 1982; Cox, Miller et al. 2008; Nezhadarya, Wang et al. 2009). Le Spread Spectrum est à la base une classe de techniques de modulations des signaux qui consiste à étaler l'énergie d'un signal à bande fréquentielle étroite sur une bande fréquentielle beaucoup plus large. Techniquement, cela consiste à combiner le signal à transmettre avec une séquence pseudo-aléatoire à débit très élevé (à hautes fréquences). Le signal résultant a un spectre plus étalé que celui du signal original et ressemble à un bruit. Sa démodulation n'est possible que si on connaît la séquence d'étalement. De ce fait, le spread spectrum permet d'acheminer un signal de manière confidentielle et protégée.

La technique du spectre étalé a été incorporée avec succès dans le marquage ou watermarking d'images. Plus précisément, l'étalement de spectre intervient essentiellement au niveau de la structuration de la marque à insérer. Soit un message M à un bit que l'on désire insérer dans une image hôte. Un générateur de séquences aléatoires génère une séquence de bruit pseudo-aléatoire composée de données réelles ou binaires $\{-1,1\}$ et ayant la même taille que l'image hôte. Cette séquence est alors utilisée pour moduler le message et donner lieu à une marque W dont les valeurs sont faibles et dont le spectre est étalé. Ainsi, par le biais de l'équation d'insertion ci-après, il est relativement facile d'effectuer l'insertion d'une telle marque avec invisibilité et un certain degré de robustesse contre les dégradations subies par l'image :

$$I'(i,j) = I(i,j) + \alpha * W(i,j) \quad (10)$$

Où W représente la marque à insérer et α détermine sa force ou sa robustesse. Du coté récepteur, il faut avoir W (ou la séquence pseudo-aléatoire) afin de détecter la présence/absence de la marque dans l'image.

Pour renforcer l'imperceptibilité du watermark, il est commun d'utiliser un masque de pondération, construit à partir du modèle visuel humain, qui permet de déterminer les régions pertinentes à l'insertion de la marque comme le montre l'équation suivante :

$$I'(i, j) = I(i, j) + \alpha * P(i, j) * W(i, j) \quad (11)$$

Où P représente un masque qui prend en compte les caractéristiques de la perception visuelle humaine pour localiser les régions d'insertion dans l'image.

Le masque de pondération P fournit une mesure ponctuelle de la sensibilité de l'œil humain aux perturbations locales de l'image. Son calcul implique la prise en compte de caractéristiques de masquage du HVS tel que la sensibilité à l'intensité lumineuse en niveaux de gris, le masquage fréquentiel et l'influence aux effets du bruit près des contours de l'image. Plusieurs méthodes de calcul du masque perceptuel existent dans la littérature.

Chapitre 3 : Etat de l'art

Dans la partie qui suit, nous allons dresser l'état de l'art concernant l'utilisation du marquage pour l'évaluation objective de la qualité d'image et vidéo. À notre connaissance, ce sujet a été abordé pour la première fois en 2002 par Holliman et Yeung (Holliman and Yeung 2002). Dans leur article, les auteurs énoncent les différentes exigences que doit vérifier tout algorithme d'évaluation automatique de la qualité par la technique de watermarking, à savoir : a) l'utilisation de métriques de distorsion globale qui permettent d'estimer la mesure de la distorsion due aux effets de la communication via un canal de transmission, b) la non référence à l'image originale; l'algorithme doit estimer la qualité de l'image sans avoir des informations auxiliaires sur l'image originale, et c) la contrainte sécuritaire ; l'algorithme doit prendre en compte la robustesse et l'imperceptibilité de la marque insérée.

Par la suite, un accent est mis sur les différentes approches de l'évaluation de la qualité par marquage tel que l'approche à base de quantification et celle à base de spectre étalé (Spread Spectrum) servant à marquer l'image dans le domaine spatial ou dans le domaine transformé. Les auteurs concluent que l'utilisation de marquages à base de quantification dégrade la marque un peu plus vite que lors de l'utilisation de l'approche à base de Spread Spectrum.

Depuis, on compte une dizaine de publications environ sur le sujet. Campisi et al. (Campisi, Carli et al. 2003) présentent une méthode qui adopte le marquage numérique d'une séquence vidéo pour estimer la qualité de service vidéo dans les applications de communications multimédias sans fil. Le marquage est réalisé dans le domaine transformé DCT en se basant sur la technique du spectre étalé. Le but poursuivi est d'utiliser la marque extraite au récepteur pour estimer la dégradation de la séquence vidéo due aux effets de la compression à divers débits et du canal de transmission à divers taux d'erreur (taux de pertes en bits). Plus précisément, chaque image de la séquence vidéo reçoit une marque insérée dans les fréquences médianes de ses blocs DCT, assurant ainsi une bonne invisibilité. Au récepteur, la marque de chaque image est extraite à partir de ses coefficients DCT puis comparée à la marque originale en calculant l'erreur quadratique moyenne MSE (MSE : *Mean Square Error*) indiquant ainsi le niveau d'altération de la marque. Enfin, pour estimer la qualité de la vidéo, on calcule la moyenne de tous les MSE ainsi obtenus et on la considère comme une borne inférieure du MSE de la séquence vidéo en question. Les expérimentations effectuées sur des vidéos compressés à différents débits et acheminés sur un canal de transmission à différents taux d'erreur ont montré que le MSE des watermarks augmente avec l'augmentation du taux d'erreur et la diminution du débit de transmission. De plus, le MSE moyen de chaque séquence vidéo reçue est toujours supérieur ou égal au MSE moyen des watermarks. Ces

résultats prouvent que le MSE des marques peut être utilisé comme indice de qualité de la vidéo reçue. Ainsi, l'estimation, en termes de MSE, de la distorsion subie par le watermark est utilisée pour déduire un indice de qualité du service vidéo fourni. Cependant, le système ne fournit aucune mesure précise de la qualité des images vidéo.

Carli et al. (Carli, Farais et al. 2005) proposent une amélioration de la méthode précédente en limitant l'insertion du watermark aux zones d'images qui sont perceptuellement importantes pour le système visuel humain. Ainsi le watermark inséré gagne en invisibilité. Les zones d'insertion sont sélectionnées en se basant sur trois caractéristiques perceptuelles connues pour attirer l'attention visuelle : le mouvement, le contraste et la couleur. Une carte individuelle est générée pour chacune de ces caractéristiques, puis une procédure de fusion combine les trois cartes et identifie les zones d'image pertinentes pour l'insertion de la marque. Au récepteur, l'extraction de la marque est suivie d'une comparaison avec la marque originale et un calcul du taux d'erreur ou BER (BER : *Bit Error Rate*). Une mesure subjective de la qualité vidéo est alors dérivée en suivant une approche de type MOS (MOS : *Mean Observer Score*).

Vers l'année 2006, Brandao et Queluz (Brandao and Queluz 2006) introduisent une méthode d'évaluation de la qualité des images compressée JPEG et JPEG2000 en se basant sur la technique du watermarking. La marque est un message binaire, qu'on insère dans le domaine de la transformée en cosinus discrète DCT, en utilisant une technique de quantification non uniforme adaptée au contenu fréquentiel de l'image et qui prend en compte les caractéristiques de la perception visuelle humaine. En effet, les auteurs notent que la quantification uniforme présente un inconvénient du fait que le pas de quantification doit être très petit pour atteindre l'imperceptibilité du watermark. Pour remédier à ce problème, ils proposent la quantification non uniforme combinée au modèle perceptuel de Watson. Cependant, il faut noter qu'une opération de dé-quantification est requise lors de l'extraction. L'estimation de la qualité d'une image se fait en mesurant la distorsion entre le watermark de référence et le watermark extrait en termes de MSE (MSE : *Mean Square error*) et de BER (BER : *Bit Error Rate*). Si l'image a subi une distorsion moins que le pas de quantification alors le MSE du watermark est directement converti en PSNR de l'image, sinon, le MSE du watermark est d'abord compensé par une stratégie de pondération qui inclut le BER et quelques statistiques de l'image.

En 2007, Wang et al. (Wang, Zheng et al. 2007) présentent une méthode d'évaluation de la qualité d'image basée sur le watermarking et qui vise à estimer la qualité d'images JPEG avec précision en termes de métriques objectives classiques tels que le PSNR, le PSNR pondéré wPSNR (wPSNR : *weighted PSNR*) et le JND. Le watermark est inséré dans les coefficients ondelettes de l'image en utilisant une méthode de quantification. Dans un premier temps, la courbe idéale est construite au travers d'un processus d'apprentissage. La courbe idéale reflète la correspondance entre le TDR du

watermark extrait (TDR : *True Detection Rate*) et la qualité (PSNR, JND ou wPSNR) de l'image marquée et compressée en JPEG. Dans un second temps, les paramètres d'insertion (choix de la marque, du pas de quantification et des sous-bandes DWT à marquer) sont ajustés au travers d'un processus automatique de façon à ce que le TDR du watermark extrait puisse prédire le PSNR (ou JND) de l'image compressée en utilisant la courbe idéale TDR-PSNR (ou TDR-JND). Au récepteur, en plus du watermark original utilisé pour l'image en question, le système requiert aussi la disponibilité de la courbe idéale ainsi que des paramètres auxiliaires utilisés lors de l'étape d'insertion (pas de quantification, indice de la marque utilisée). La qualité objective de l'image est estimée avec une grande précision mais un coût additionnel est nécessaire pour la transmission, pour chaque image, des données auxiliaires requises au travers d'un canal fiable.

Nezhadarya et al. (Nezhadarya, Wang et al. 2009) proposent une méthode d'évaluation de la qualité d'image qui reprend le concept de la courbe idéale présenté dans (Wang, Zheng et al. 2007) et l'applique au cas du marquage à spectre étalé (ou spread spectrum SS) opérant dans le domaine de la DCT. Le détecteur LOD (LOD : *Locally Optimum Detector*), une alternative à la corrélation linéaire, est utilisé pour détecter la marque insérée dans des images marquées et ayant subi un codage JPEG. Dans une première phase, la courbe idéale est construite au travers d'un processus d'apprentissage. La courbe idéale reflète dans ce cas la correspondance entre la valeur LOD du détecteur de marque et la qualité PSNR de l'image marquée et compressée en JPEG. Dans une deuxième phase, l'image est marquée puis compressée en JPEG en vue de sa transmission. Les paramètres d'insertion (choix de la marque, choix du facteur de robustesse de la marque) sont ajustés manuellement de façon à ce que le watermark qui sera extrait au récepteur puisse prédire le PSNR de l'image compressée JPEG en utilisant la courbe idéale LOD-PSNR. Cette méthode vise la prédiction précise du PSNR de l'image, cependant, la précision atteinte n'est pas mentionnée. De plus, cette méthode présente une dépendance par rapport à des données auxiliaires qui doivent être transmises au récepteur au travers d'un canal fiable. En effet, afin de prendre en compte plusieurs types d'images, la courbe idéale change de pente. Il est donc indispensable de transmettre la pente concernée par l'image courante comme donnée auxiliaire. De plus, l'indice de la marque utilisée ainsi que quelques paramètres requis par le détecteur LOD doivent aussi être transmis.

Les travaux de (Baaziz, Zheng et al. 2011) portent sur l'évaluation objective de la qualité d'images JPEG en termes PSNR, en proposant une approche d'évaluation sans référence exploitant le marquage multiple ; le marquage multiple consiste à insérer plusieurs marques à la fois dans une image donnée. Une sélection de sous-bandes de la transformée en ondelettes discrète DWT est utilisée pour y insérer des marques par la technique du spectre étalé et détecter leur présence par la méthode de corrélation linéaire.

Pour détecter la présence d'un ou plusieurs watermark(s) dans une image marquée le processus de détection décompose l'image marquée \hat{S} en sous-bandes DWT, mesure la corrélation linéaire R_{ld} entre chaque marque de référence W_{ld} et la sous bande marquée \hat{S}_{ld} compare la valeur de corrélation trouvée à un seuil pré-calculé T_{ld} .

Pour une image marquée et potentiellement dégradée par la compression JPEG (ou image test), le but poursuivi par ce marquage multiple n'est pas d'évaluer une valeur précise de PSNR, mais plutôt de déterminer l'intervalle PSNR auquel il appartient et, tirer par la suite, une description appropriée de la qualité de l'image.

Les résultats expérimentaux présentés attestent du succès de cette méthode d'évaluation de la qualité sans référence à l'image originale. Une des forces majeures de cette méthode est le fait qu'aucune donnée auxiliaire n'a besoin d'être transmise. De plus, les marques de références restent les mêmes pour toutes les images et peuvent donc faire partie du système de détection/évaluation. Les résultats préliminaires sont prometteurs et indiquent le potentiel de cette méthode à être intégrée dans un système de communication multimédia pour le contrôle automatique de la qualité d'images et vidéo.

Pour conclure cette partie, les principaux travaux relatifs à l'évaluation de la qualité par watermarking sont résumés dans le Tableau 3.1.

De façon générale la plupart des méthodes présentent des faiblesses soit au niveau de la précision de la qualité estimée, soit au niveau de la nécessité de transmettre des données auxiliaires pour le récepteur. En effet, dans (Campisi, Carli et al. 2003; Carli, Farais et al. 2005), on se limite à des approches subjectives manquant de précision pour l'évaluation de la qualité. Alors que dans (Wang, Zheng et al. 2005; Nezhadarya, Wang et al. 2009), des métriques sont utilisées pour estimer quantitativement la qualité de l'image, mais ceci est obtenu au prix d'une transmission de plusieurs paramètres auxiliaires en plus de l'image marquée. Enfin, la méthode proposée dans (Baaziz, Zheng et al. 2011) apporte une estimation par intervalle de PSNR (ou JND) et ne requiert aucune donnée auxiliaire à transmettre.

Tableau 3. 1 : Résumé de quelques méthodes d'évaluation de la qualité d'image par marquage.

Auteurs / Caractéristiques	(Campisi, Carli et al. 2003)	(Carli, Farais et al. 2005)	(Brandao and Queluz 2006)	(Wang, Zheng et al. 2007)	(Nezhadarya, Wang et al. 2009)	(Baaziz, Zheng et al. 2011)
Technique de marquage	SS	SS	Quantification	Quantification	SS	SS
Domaine d'insertion	DCT	DCT	DCT	DWT	DCT	DWT
Masque perceptuel	Non	Oui	Non	Non	Oui	Oui
Extraction de la marque	Oui	Oui	Oui	Oui	Non	Non
Détection de la marque	Non	Non	Non	Non	Oui	Oui
Mesure de la dégradation de la marque	MSE	BER	MSE, BER	BER	LOD-Corr	Corrélation linéaire
Métrique de qualité	Subjective	MOS	PSNR	JND, PSNR, wPSNR	PSNR	PSNR
Précision	-	-	-	0.78 dB	-	Précise par intervalles
Données auxiliaires requises	-	-	Oui	Oui (pas de quantification, indice du watermark)	Oui (indice du watermark, paramètre du détecteur LOD, pente de la courbe idéale)	Non

Chapitre 4 : Objectifs poursuivis et méthodologie

Dans ce projet de recherche, on s'est proposé d'étendre et améliorer la méthode de marquage multiple à base d'ondelettes ([Baaziz, Zheng et al. 2011](#)) pour en faire un système automatique d'évaluation de la qualité d'image et de vidéo. Pour cela, les objectifs spécifiques qu'on s'est fixé au début du projet s'énoncent comme suit :

1.Objectifs poursuivis

1. Proposer et développer une procédure qui permet le réglage automatique des valeurs des coefficients α_l afin d'atteindre le degré de robustesse requis par chacune des trois marques insérées. En effet, la méthode de référence telle que présentée dans ([Baaziz, Zheng et al. 2011](#)) effectue le réglage des coefficients de robustesse α_l manuellement. L'incorporation d'une procédure d'ajustement automatique serait un atout pour l'ensemble du système d'évaluation qui pourra alors fonctionner automatiquement. Cette procédure sera basée sur une approche itérative d'incrément/décément des facteurs α_l jusqu'à l'atteinte des contraintes r_1 et r_2 requises.
2. Étendre l'évaluation de la qualité en utilisant des métriques autres que le PSNR. On se propose d'intégrer à notre système la distance perceptuelle JND de Watson avec le tableau des intervalles de qualité de l'image tel que défini dans la sous-section 1.2.
3. Étudier l'impact du transcodage (l'influence d'une cascade de compressions/décompressions) sur l'évaluation de la qualité d'une image. Dans la pratique, l'image peut subir une cascade de codage/décodage à des taux de compression variés, car lors de son acheminement vers un utilisateur final, l'image peut passer par différentes stations (ou points intermédiaires) et chacun peut avoir des contraintes spécifiques de débits cibles. Dans notre étude, deux scénarios seront pris en compte : le premier scénario consiste à mettre en cascade deux compressions successives ayant des taux de compressions descendants (le premier taux de compression est plus élevé que le deuxième). Le deuxième scénario met en série deux compressions à taux de compression ascendants. Le but de cette étude est de déterminer la performance de notre système à évaluer correctement la qualité de l'image reçue.

4. Prendre en compte l'effet du canal de transmission : lors de la transmission des données de l'image de la source vers la destination, le canal de transmission constitue aussi une cause de dégradation de l'image. L'effet du canal de transmission peut se traduire entre autres par un ajout de bruit ou bien par une perte de paquets de données. On se propose d'étudier l'effet de ces deux types de dégradations sur la performance du système d'évaluation de la qualité. L'effet d'un canal bruité sera simulé par l'ajout d'un bruit blanc Gaussien de moyenne m et de variance σ à l'image codée. Pour la dégradation par pertes de paquets, une simulation de ce phénomène sera effectuée dans l'environnement Matlab.
5. Étudier l'influence du domaine d'insertion des marques par rapport au domaine de compression de l'image marquée. Sachant que la compression opère par quantification dans le domaine transformé (DCT pour JPEG et DWT pour JPEG2000), est-il avantageux d'insérer les marques dans le même domaine que la compression ou pas ? Intuitivement, un des avantages attendus pourrait être la rapidité à ajuster les paramètres de robustesse des marques lors de l'insertion. Un autre avantage attendu pourrait être l'augmentation de la sensibilité du détecteur de marque (mesure de corrélation) par rapport aux dégradations subies. Cette étude requiert le développement d'un nouvel algorithme de marquage multiple opérant dans le domaine DCT.
6. Étendre notre application d'évaluation automatique de la qualité aux séquences vidéo. Dans le cas du codeur MPEG, les images compressées en mode intra utilisent les mêmes techniques que le codeur JPEG, ce qui rend l'extension de notre méthode à la vidéo possible et faisable.

2. Méthodologie adoptée

L'approche proposée pour l'évaluation automatique de la qualité d'image repose sur le marquage numérique et constitue une continuation de la méthode basée sur le marquage multiple publiée dans (Baaziz, Zheng et al. 2011). Dans ce travail, on se limite à l'insertion de trois marques. L'atteinte des objectifs énoncés ci-dessus est appuyée par la proposition d'algorithmes, leur développement en programmes dans l'environnement de simulation MATLAB. Les expérimentations sont effectuées sur des images et des séquences d'images réelles à 256 niveaux de gris telles que Lena, Boat, Goldhill et Peppers (Figure 4.1), la base d'images fixes UCID (Schaefer and Stich 2003) et la base de vidéo (VIPS 2012).



Figure 4. 1: Exemple d'images de test

3. Méthodes du marquage multiple

Les méthodes proposées portent sur l'évaluation objective de la qualité d'images compressées JPEG/JPEG2000 en termes de PSNR. L'approche d'évaluation est sans référence et exploite le marquage multiple. Une transformée spatio-fréquentielle est appliquée à l'image test afin de la décomposer en plusieurs sous-bandes spatio-fréquentielles. Une sélection judicieuse de L sous-bandes est utilisée pour y insérer L marques par la technique du spectre étalé. L'insertion est effectuée sous contraintes de façon à ce que chaque marque permette le suivi d'une certaine gamme de qualité de l'image (pouvant être un intervalle de PSNR). Suite au marquage, l'image peut subir des traitements comme la compression JPEG ou JPEG2000, ce qui dégrade sa qualité ainsi que celle des marques contenues. De plus, lors de la transmission de cette image marquée, le canal de communication peut l'affecter par des opérations de transcodage, des pertes de paquets ou bien par des bruits additifs. Tout cela va contribuer à dégrader davantage la qualité de l'image.

Au niveau du récepteur, ou bien à tout point intermédiaire de la chaîne de communication, il est possible de procéder à l'estimation de la qualité de l'image en analysant les L marques contenues. L'image est alors décomposée en sous-bandes et la présence de chaque marque est détectée par une mesure de corrélation linéaire. L'ensemble des détections obtenues est par la suite converti en une mesure de la qualité de l'image test. Le schéma de marquage multiple et d'estimation de la qualité est illustré dans la Figure 4.2.

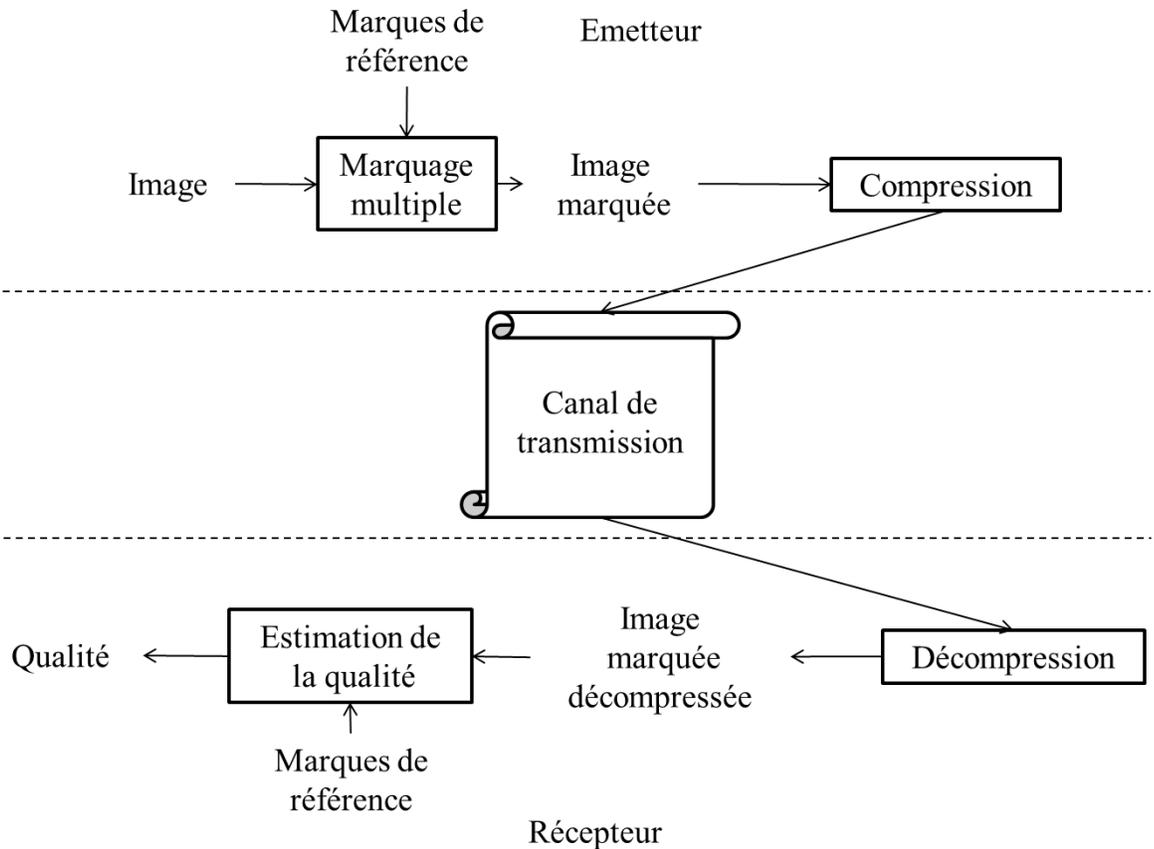


Figure 4. 2 : Processus d'évaluation automatique de la qualité d'image.

3.1. Le marquage multiple dans le domaine transformé

Dans ce mémoire, deux algorithmes d'évaluation de la qualité sont présentés ; l'un utilise la DWT alors que l'autre utilise la DCT. Le choix de ces deux transformées se justifie par le fait que JPEG utilise la DCT et que JPEG2000 utilise la DWT. Les codeurs JPEG et JPEG 2000 étant les deux dégradations principales du canal de communication que l'on étudie dans ce projet. Le marquage multiple dans chacun des domaines transformés se décrit comme suit :

a) Marquage multiple basé sur la DWT

La DWT permet de décomposer une image test en plusieurs sous-bandes selon le nombre de niveaux de résolution choisi. Plus exactement, on obtient trois sous-bandes ondelettes

à chaque niveau de résolution plus une image d'approximation tel que le montre la Figure 4.3. L'insertion de L marques distinctes se fait dans L sous-bandes ondelettes. Le choix des sous-bandes ondelettes à marquer est guidé par les constatations suivantes :

- 1- En général, la sous-bande D_1 , correspondant aux plus hautes fréquences de l'image, présente un contenu très pauvre et peu significatif. Les algorithmes de compression ont tendance à quantifier fortement ce contenu pour le moindre taux de compression visé. Par conséquent, il n'est pas recommandé de marquer cette sous-bande.
- 2- Les sous-bandes H_1 et V_1 , correspondant à des hautes fréquences de l'image, sont très sensibles aux moindres dégradations de compression subies par l'image. Les marques insérées dans ces sous-bandes seront aussi très sensibles aux moindres dégradations de l'image. De plus, la transparence des marques est facilement atteinte (à cause des hautes fréquences).
- 3- Les sous-bandes de niveau 2 contiennent les fréquences moyennes de l'image. Ces sous-bandes sont fortement affectées dans le cas de compression à taux élevé. Les marques insérées dans ces sous-bandes seront sensibles aux fortes dégradations de l'image.
- 4- Les sous-bandes de niveau 3 et plus sont de plus en plus proches des basses fréquences de l'image. Toute insertion de marque dans ces sous-bandes risque de produire des distorsions visibles dans l'image. Le marquage dans ces sous-bandes est à éviter.

Sur la base des constatations ci-dessus, la DWT d'une image nous offre donc 5 sous-bandes pour le marquage multiple. La méthode de marquage développée dans ce projet en utilise 3, plus exactement B_1 , B_2 et B_3 tel que c'est illustré dans la Figure 4.3.

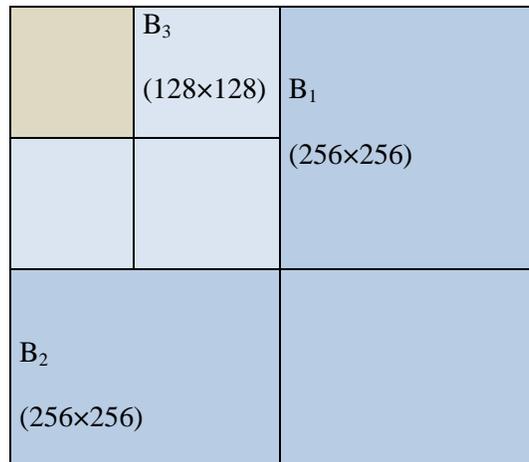


Figure 4. 3 : Décomposition de l'image hôte (de taille 512×512 pixels) en DWT à 2 niveaux et sélection de trois sous-bandes pour le marquage : les sous-bandes horizontale et verticale du niveau 1, B₁ et B₂, et la sous-bande verticale du niveau 2, B₃.

a) Marquage multiple basé sur la DCT

La DCT appliquée sur des blocs d'images de taille 8×8 permet de réaliser une transformation spatio-fréquentielle de l'image. Le choix de la taille 8×8 est populaire dans les standards de compression comme JPEG. Dans chaque bloc 8×8 de coefficients DCT, il est possible de localiser les composantes de basse, moyenne et hautes fréquences du bloc d'images correspondant, tel que le montre la Figure 4.4. D'où l'idée d'extraire et de former L sous-bandes disjointes à partir de chaque bloc DCT. Ces L sous-bandes seront dédiées à l'insertion de L marques distinctes. Le choix de la localisation de ces L sous-bandes est effectué suite à un ensemble de tests expérimentaux de telle façon à atteindre une bonne invisibilité des marques insérées (i.e. un PSNR élevé pour les images marquées). La Figure 20 illustre le choix effectué dans le cadre de ce projet pour la localisation des trois sous-bandes DCT B₁, B₂ et B₃ destinées à recevoir trois marques W_1 , W_2 et W_3 .

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0
0	0	0	0	1	1	1	0
0	0	0	0	1	1	1	0
0	0	0	0	0	0	0	0
0	0	0	0	3	0	0	0
0	0	0	0	3	0	0	0
0	0	0	0	3	0	0	0
0	0	0	0	3	0	0	0
3	3	3	3	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

0	0	0	0	0	0	2	2
0	0	0	0	0	0	2	2
0	0	0	0	0	0	2	2
0	0	0	0	0	0	2	2
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
2	2	2	2	0	0	0	0
2	2	2	2	0	0	0	0
O	O	O	O	0	O	0	0
O	O	O	O	0	O	0	0
O	O	O	O	0	O	0	0
O	O	O	O	0	O	0	0
0	0	0	0	O	0	0	O
O	O	O	O	0	0	0	O
0	0	0	0	0	0	0	O
0	0	0	0	O	O	O	O

Figure 4. 4 : Extraction des coefficients à partir d'un bloc DCT 8×8 pour former les sous-bandes B1, B2 et B3. Chaque sous-bande a une taille 8×8 mais ses coefficients non-nuls sont indiqués par les symboles 1, 2 et 3 pour B1, B2 et B3 respectivement. Le dernier bloc contient tous les coefficients restants (indiqués par O). La somme des 4 blocs donne le bloc DCT original.

3.2. Structuration de la marque

La structuration des marques à insérer consiste à bien choisir la taille de chaque marque, sa forme ainsi que le type des données et leur distribution statistique (Gaussienne ou uniforme) dans le cas du marquage dans le domaine de la DCT et dans le cas du marquage dans le domaine de la DWT.

Étant donné que la technique de marquage visée est le spectre étalé ou *spread spectrum* (voir la section 1.5), une marque est à l'origine un message binaire (soit la valeur «1») modulé par une séquence pseudo-aléatoire générée à partir d'un générateur de nombres aléatoires. La distribution statistique de cette séquence peut être uniforme ou Gaussienne. Dans le cas de la distribution uniforme, la séquence obtenue a des valeurs binaires {+1, -1} équiprobables. Dans le cas de la distribution Gaussienne de moyenne nulle et de variance 1, la séquence obtenue est à valeurs réelles. La taille de cette séquence et son organisation dépend des domaines de marquage DCT et DWT.

a) Cas de la DWT

Chaque marque à insérer W_1 , W_2 ou W_3 est une séquence pseudo-aléatoire (uniforme ou Gaussienne) organisée en matrice de taille exactement égale à la taille de la sous-bande DWT qui la reçoit. Par exemple, pour une image de taille 512×512 , les trois sous-bandes B_1 , B_2 et B_3 montrées dans la Figure 4.3 reçoivent les trois marques W_1 , W_2 et W_3 de taille 256×256 , 256×256 et 128×128 respectivement. Il est à noter que les 3 séquences pseudo-aléatoires requises sont indépendantes.

b) Cas de la DCT

Pour la DCT, les marques W_1 , W_2 et W_3 sont toutes structurées en matrices de taille 8×8 . Chaque matrice W_i contient une séquence pseudo-aléatoire de longueur S_i à la même position que la sous-bande B_i et une valeur nulle ailleurs. Les Figures 4.5, 4.6 et 4.6 montrent la structuration de chacune des trois marques. Il est à noter que les 3 séquences pseudo-aléatoires requises sont statistiquement indépendantes. L'organisation sous forme de matrices 8×8 a pour seul but de faciliter l'insertion de la marque par simple addition à un bloc 8×8 de coefficients DCT.

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	w_1	w_1	0
0	0	0	0	w_1	w_1	w_1	0
0	0	0	0	w_1	w_1	w_1	0
0	0	0	0	0	0	0	0

Figure 4. 5 : Structuration de la marque W_1 en une matrice 8×8 . La séquence pseudo-aléatoire constituante w_1 a une longueur de 8.

0	0	0	0	0	w ₂	w ₂	0
0	0	0	0	0	w ₂	w ₂	0
0	0	0	0	0	w ₂	w ₂	0
0	0	0	0	0	w ₂	w ₂	0
0	0	0	0	0	0	0	0
w ₂	w ₂	w ₂	w ₂	0	0	0	0
w ₂	w ₂	w ₂	w ₂	0	0	0	0
0	0	0	0	0	0	0	0

Figure 4. 6 : Structuration de la marque W2 en une matrice 8×8. La séquence pseudo-aléatoire constituante w2 a une longueur de 16.

0	0	0	0	w ₃	0	0	0
0	0	0	0	w ₃	0	0	0
0	0	0	0	w ₃	0	0	0
0	0	0	0	w ₃	0	0	0
w ₃	w ₃	w ₃	w ₃	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figure 4. 7 : Structuration de la marque W3 en une matrice 8×8. La séquence pseudo-aléatoire constituante w₃ a une longueur de 8.

3.3. Insertion des marques

En supposant que B_l est une sous-bande DWT ou DCT de l'image hôte qui est sélectionnée pour le marquage, l'insertion d'une marque W_l (préalablement structurée selon la technique du spectre étalé) se fait selon l'équation suivante :

$$B'_l(i, j) = B_l(i, j) + \alpha_l M_l(i, j)W_l(i, j) \quad (12)$$

Où B'_l est la sous-bande marquée, le paramètre α_l est le facteur de robustesse de la marque et M_l est un masque perceptuel. Le masque M_l prend en compte les caractéristiques du système visuel humain pour guider l'insertion de la marque dans les zones pertinentes de l'image afin de garantir une meilleure invisibilité des marques

insérées. Les détails sur le calcul des masques perceptuels dans le cas du marquage DCT ou DWT sont donnés en annexe de ce document.

Le paramètre α_l sert à ajuster la robustesse de chaque marque. Il est à noter que l'augmentation de la valeur de ce paramètre affecte l'invisibilité de la marque.

Dans le cas de la DCT, l'insertion d'une seule marque W_l consiste à appliquer l'équation (12) à la sous-bande B_l de chaque bloc DCT 8×8 de l'image hôte. Dans le cas de la DWT, l'insertion d'une seule marque W_l consiste à appliquer l'équation (12) à la sous-bande B_l de la DWT de l'image hôte.

Le marquage multiple consiste à insérer plusieurs marques, soit ($l = 2, 3, 4, \dots L$). Les Figures 4.8 et 4.9 illustrent le cas d'insertion de 3 marques.

Une fois les marques insérées, les sous-bandes ainsi marquées sont replacées dans la représentation spatio-fréquentielle de l'image (DCT ou DWT), et la transformée inverse est utilisée pour reconstruire l'image marquée.

0	0	0	0	3	0	2	2
0	0	0	0	3	0	2	2
0	0	0	0	3	0	2	2
0	0	0	0	3	0	2	2
3	3	3	3	0	1	1	0
0	0	0	0	1	1	1	0
2	2	2	2	1	1	1	0
2	2	2	2	0	0	0	0

Figure 4. 8 : Localisation des coefficients DCT marqués par les marques W_1 , W_2 , et W_3 dans chaque bloc DCT 8×8 de l'image hôte.

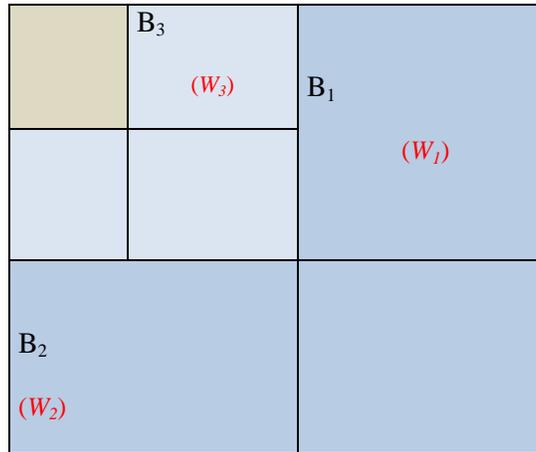


Figure 4. 9 : Localisation des coefficients DWT marqués dans trois sous-bandes de l'image hôte décomposée en deux niveaux DWT.

3.4. Détection de la marque

Pour détecter la présence d'une ou plusieurs marques dans une image marquée (ou image test), le processus de détection décompose l'image marquée en sous-bandes DWT pour le cas du marquage DWT (ou en blocs DCT 8×8 pour le cas du marquage DCT respectivement), mesure la corrélation linéaire R_l entre chaque marque de référence W_l et la sous-bande DWT marquée B'_l selon l'équation (13) ou bien selon l'équation (14) pour le cas de marquage DCT, puis compare chaque valeur de corrélation trouvée à un seuil pré-calculé T_l .

$$R_l = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N B'_l(i, j, 1) W_l(i, j) \quad \text{pour la DWT} \quad (13)$$

ou bien

$$R_l = \frac{1}{KS_l} \sum_{k=1}^K \sum_{i=1}^8 \sum_{j=1}^8 B'_l(i, j, k) W_l(i, j) \quad \text{pour la DCT} \quad (14)$$

M et N étant les dimensions des sous-bandes DWT. Dans l'équation (14), l'indice k indique le k ième bloc DCT de l'image test ayant au total K blocs DCT et S_l est le nombre de coefficients marqués dans chaque sous-bande DCT. S_l correspond aussi à la longueur de la séquence pseudo-aléatoire de la marque W_l .

Le module de détection de marque a été amélioré en lui apportant plus de précision. Nous sommes passés d'une détection binaire (0 ou 1) telle que présentée dans ([Baaziz, Zheng et al. 2011](#)) à une détection ternaire plus précise (0, 1, 2). La détection et le processus de décision se formulent comme suit :

$$\text{Si } (R_l \geq T_l + \Delta T)$$

$D_l=1$; La marque est présente dans l'image.

$$\text{Sinon Si } (T_l - \Delta T < R_l \leq T_l + \Delta T) \text{ alors}$$

$D_l=2$; La marque est probablement présente dans l'image.

$$\text{Sinon Si } (R_l < T_l - \Delta T) \text{ alors}$$

$D_l=0$; La marque est absente de l'image.

ΔT est une petite variation du seuil T_l qu'on a fixée expérimentalement à 5% ou 7% de T_l .

Il est important de mentionner que le détecteur n'a besoin que de l'image marquée et les marques de référence W_l . Les seuils T_l sont calculés pour chaque image test au niveau du détecteur (ou récepteur) à l'aide des équations suivantes :

$$T_l = 3.97 \sqrt{2\sigma_l^2} \quad (17)$$

avec

$$\sigma_l^2 = \frac{1}{(MN)^2} \sum_{i=1}^M \sum_{j=1}^N B_l'(i, j, 1)^2 \quad \text{pour la DWT} \quad (18)$$

ou bien

$$\sigma_l^2 = \frac{1}{(KS_l)^2} \sum_{k=1}^K \sum_{i=1}^8 \sum_{j=1}^8 B_l'(i, j, k)^2 \quad \text{pour la DCT} \quad (19)$$

M et N étant les dimensions des sous-bandes DWT. Dans l'équation (19), l'indice k indique le k ième bloc DCT de l'image test ayant au total K blocs DCT et S_l est le nombre de coefficients marqués dans chaque sous-bande DCT.

En effet, la valeur du seuil de décision T_l repose sur l'énergie de la sous-bande marquée, et est déterminée par un calcul a posteriori issu du critère de *Neyman-Pearson* (Barni, Bartolini et al. 2001) effectué pour chaque sous-bande marquée. Ce calcul suppose une probabilité d'erreur de détection de 10^{-8} (qui veut dire une erreur de détection sur 100000000 détections). T_l se formule essentiellement comme le montrent les équations (17) et (18) pour la DWT et les équations (17), (19) pour la DCT.

3.5. Estimation de la qualité

Pour une image marquée et potentiellement dégradée par la compression JPEG/JPEG2000 (ou image test), le but poursuivi par ce marquage multiple est d'évaluer la qualité de l'image en termes de PSNR. Plus précisément, il s'agit de déterminer l'intervalle de PSNR correspondant à sa qualité et dériver par la suite, une description appropriée de la qualité de l'image. Pour cela, la méthode d'estimation de la qualité requiert la définition préalable d'intervalles de qualité. Dans le Tableau 2.4 sont illustrés les quatre intervalles de PSNR choisis ainsi que les attributs de qualité associés tels que présentés dans (Baaziz, Zheng et al. 2011). Il est convenu qu'au-dessus de 40 dB, le PSNR est le témoin d'une très bonne qualité d'image (Very Good). En dessous de 30 dB, la qualité de l'image correspondante est mauvaise (Poor). Enfin entre 30 dB et 35 dB, la qualité est acceptable (acceptable) alors qu'entre 35 dB et 40 dB, la qualité de l'image est plutôt bonne (Good). Cependant, il est précisé que le choix de ces intervalles est arbitraire et peut être différent d'un domaine d'application à un autre.

Dans le cadre de ce travail, nous avons choisi des gammes de qualité un peu plus précises en définissant sept intervalles de PSNR. Dans le Tableau 4.1 sont illustrés les sept intervalles de PSNR choisis ainsi que les attributs de qualité associés. Il est convenu qu'au dessus et autour de 40 dB, le PSNR est le témoin d'une très bonne qualité. En dessous de 30 dB, le PSNR est le témoin d'une mauvaise qualité. Par la suite, entre 35 dB et 40 dB, et autour de 35 dB, le PSNR est le témoin d'une bonne qualité. Enfin, entre 30 dB et 35 dB et autour de 30 dB, la qualité est acceptable.

Pour définir ces sept intervalles, trois seuils de PSNR sont requis : $T1_PSNR=40$ dB, $T2_PSNR = 35$ dB et $T3_PSNR= 30$ dB. À chaque seuil de PSNR est associé une marque W_l destinée pour une sous-bande DWT/DCT particulière B_l . Le marquage est alors utilisé pour insérer les trois marques en accordant à chacune un degré de robustesse suffisant pour survivre à une certaine quantité de dégradation de l'image et disparaître dès que la dégradation subie dépasse un certain niveau mesuré en termes de PSNR de l'image hôte. Plus précisément, l'insertion des trois marques W_l selon l'équation (12) est

contrôlée par l'ajustement fin des trois facteurs de robustesse α_1 , α_2 et α_3 de façon à répondre aux deux contraintes suivantes :

- **Contrainte r_1** : lorsque l'image test est une image marquée mais n'ayant pas subi de dégradation additionnelle, l'ensemble du marquage effectué doit être imperceptible et chaque marque doit être détectable à tout moment ($D_1 = 1$, $D_2 = 1$ et $D_3 = 1$).
- **Contrainte r_2** : Lorsque l'image test est une image marquée et ayant subi une dégradation (due à une compression JPEG/JPEG2000), si le PSNR de l'image test chute au-dessous d'un des seuils prédéfinis, ceci doit nécessairement s'accompagner d'une non-détection de la marque associée. Autrement dit, la détection de la marque W_1 doit échouer (i.e., $D_1=0$) dès lors que $PSNR < T1_PSNR$. De la même façon la détection de la marque W_2 (W_3) doit échouer dès lors que $PSNR < T2_PSNR$ ($T3_PSNR$) respectivement.

Du côté du récepteur, évaluer et caractériser la qualité d'une image test (supposée marquée) passe par ces trois étapes :

- 1- Tout d'abord, les trois seuils de détection T_1 , T_2 et T_3 sont calculés selon les équations (17) et (18) pour le cas de la DWT et selon les équations (17) et (19) pour le cas de la DCT.
- 2- Puis, trois processus de détection par corrélation sont exécutés afin de détecter la présence/absence de chacune des trois marques W_1 , W_2 et W_3 . Les résultats de la détection sont mis dans les variables de décision ternaires D_1 , D_2 et D_3 respectivement.
- 3- Enfin, l'évaluation de la qualité de l'image test est donnée par la lecture de la table de décision ci-dessous (voir Tableau 4.1). Par exemple, si le triplet $(D_1, D_2, D_3) = (0, 1, 1)$ alors l'image est de *bonne* qualité et son PSNR est entre 35 dB et 40 dB, et si le triplet $(D_1, D_2, D_3) = (0, 2, 1)$ alors l'image est de *bonne* qualité et son PSNR est autour de 35 dB.

Tableau 4. 1 : Table de décision à sept intervalles PSNR pour l'évaluation automatique de la qualité d'une image test.

D_1	1	2	0	0	0	0	0
D_2	-	-	1	2	0	0	0
D_3	-	-	-	-	1	2	0
<i>Qualité</i>	<i>Très bonne</i>		<i>Bonne</i>		<i>Acceptable</i>		<i>Mauvaise</i>
<i>PSNR (dB)</i>	>40	≈40	35-40	≈35	30-35	≈30	<30
Légende : '-' : 0 ou 1 ou 2							
'≈40' : autour de 40 dB							

3.6. Marquage avec ajustement des facteurs de robustesse

Le facteur de robustesse α_l joue un rôle très important pour l'évaluation de la qualité. C'est grâce à l'ajustement progressif des facteurs α_l qu'on arrive à régler la robustesse de chaque marque insérée de façon à répondre aux contraintes de marquage r_1 et r_2 . La procédure d'ajustement est itérative et comporte deux niveaux ; le niveau 1 s'occupe à répondre à la contrainte r_1 et le niveau 2 s'occupe de l'atteinte de la contrainte r_2 . Pour vérifier l'atteinte de la contrainte r_1 , il suffit de s'assurer que les marques insérées sont toutes détectables et que le PSNR de l'image marquée dépasse largement les 42 dB, ce qui est un bon indicateur de l'invisibilité de la marque.

Une fois la contrainte r_1 atteinte, on fait subir à l'image marquée une série de compressions JPEG/JPEG2000 à taux variables et on prélève les variables de détection D_l ainsi que les PSNR des images marquées et compressées. Si la contrainte r_2 est atteinte alors l'image marquée est prête, sinon, alors il faut ajuster les facteurs α_l , refaire le marquage, et les deux niveaux de vérification. Ce processus est illustré dans la Figure 4.10.

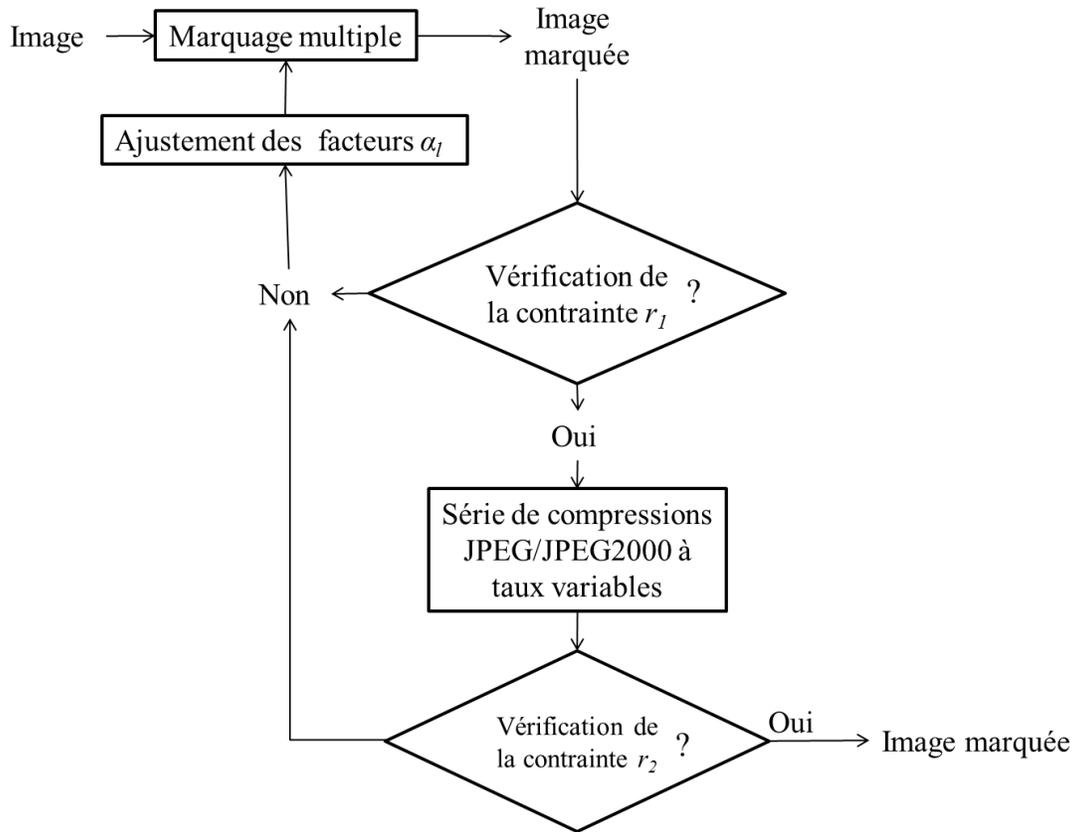


Figure 4. 10 : Procédure d'ajustement des facteurs de robustesse α_l .

3.6.1. Ajustement manuel

Le réglage des facteurs de robustesse α_l peut se faire de façon manuelle et répétitive jusqu'à atteindre les contraintes r_1 et r_2 . Ce type d'ajustement requiert un temps énorme pour atteindre les résultats souhaités. Car, à chaque fois que les deux contraintes ne sont pas atteintes alors il faut refaire tout le marquage avec de nouvelles valeurs des α_l . Il appartient à l'utilisateur de savoir dans quel sens il faut varier progressivement les valeurs.

3.6.2. Ajustement automatique

Le réglage des facteurs de robustesse α_l peut se faire de façon automatique par la mise au point d'une procédure itérative d'ajustement. La procédure d'ajustement automatique suit le diagramme de la Figure 4.10. La première partie de cette procédure ajuste

itérativement les facteurs de robustesse jusqu'à atteindre la contrainte r_1 . La deuxième partie continue à ajuster les facteurs de robustesse jusqu'à l'atteinte de la contraintes r_2 tout en ayant la contrainte r_1 vérifiée. Le développement de la procédure d'ajustement automatique des facteurs de robustesse prend en compte les points suivants :

- Au début de la procédure, les facteurs α_l sont initialisés à des valeurs arbitraires.
- Les facteurs α_l sont ajustés séparément l'un de l'autre et séquentiellement jusqu'à atteindre la contrainte désirée.
- Une itération peut ajuster un facteur α_l en l'incrémentant ou en le décrémentant d'un pas fixe $\Delta\alpha$ préalablement fixé.
- Il est préférable de commencer par l'ajustement du facteur assurant la robustesse la plus élevée, soit α_3 , car ce facteur est de plus grande amplitude et souvent lent à ajuster.
- L'ajustement du facteur α_2 peut perturber l'ajustement précédemment effectué pour α_3 . De même l'ajustement du facteur α_1 peut perturber l'ajustement effectué pour α_2 . C'est pour cela qu'un deuxième cycle d'ajustement avec un pas de variation $\Delta\alpha$ plus fin est effectué afin de remédier à ce problème de perturbation.

Chapitre 5 : Expérimentations et résultats

Dans cette partie, nous allons montrer les résultats trouvés suite à différentes expérimentations effectuées avec les deux méthodes introduites pour l'évaluation automatique de la qualité d'images sans référence. Le plan de test comprend les parties suivantes:

- Validation du marquage multiple et de l'estimation par détection des marques.
- Validation de la procédure d'ajustement automatique des facteurs α_l .
- Étude de l'impact du choix des marques Gaussiennes et uniformes.
- Évaluation de l'efficacité d'estimation de la qualité d'image.
- Étude de l'effet du transcodage, du bruit additif et de la perte de paquets.
- Analyse de l'impact du domaine d'insertion par rapport au domaine de compression (JPEG/JPEG2000).
- Comparaison des deux méthodes introduites en termes de performance temporelle du marquage et de l'estimation de la qualité.
- Application au codage MPEG de séquences vidéo.

L'ensemble des tests a utilisé des images à niveaux de gris (comme Lena, Goldhill, Boat, et Peppers). Les distorsions principales du canal de communication sont représentées par les standards de compressions JPEG et JPEG2000 disponibles dans MATLAB. Quelques distorsions additionnelles sont aussi considérées comme le transcodage, le bruit additif et la perte de paquets. Le marquage multiple repose soit la DCT ou bien la DWT disponibles dans MATLAB. La DWT utilise les filtres d'ondelettes biorthogonales bior3-7.

Le Tableau 5.1 montre entre autres les paramètres d'un exemple typique de marquage multiple de l'image Boat suivi d'une détection des trois marques insérées. L'insertion de trois marques Gaussiennes est effectué dans les blocs DCT 8×8 de l'image hôte. Le triplet des facteurs de robustesse $\alpha = (0.055, 0.183, 0.165)$ a été préalablement ajusté (manuellement) en vue d'une compression JPEG2000. Le degré d'invisibilité des marques insérées est mesuré par le PSNR (dB) de l'image marquée (en comparaison avec l'image originale (hôte)). Le détecteur utilise les marques originales W_1, W_2 et W_3 , mesure pour chaque marque, une corrélation R_l et la compare à un seuil pré-calculé T_l selon les équations 13-18 et la section 3.2.4. Les résultats de la détection sont donnés dans les triplets (D_1, D_2, D_3) .

Remarquons que le Tableau 5.1 illustre bien l'atteinte de la contrainte r_l , puisque chaque image marquée jouit d'un PSNR élevé (plus que 42 dB) qui témoigne de

l'invisibilité des marques insérées tout en ayant des marques toutes détectables avec $(D_1, D_2, D_3) = (1, 1, 1)$.

Lorsque l'image marquée subit une compression JPEG ou JPEG2000, son PSNR chute et ceci affecte les marques contenues. Au niveau du détecteur de marque, les valeurs de corrélation se trouvent ainsi diminuées. Ce qui peut affecter les résultats de la détection. La Figure 5.1 montre les courbes de variation de la corrélation et du seuil de détection en fonction du PSNR de l'image marquée et compressée sous JPEG2000.

On peut voir des intersections dans des points bien choisis (30dB, 35dB, et 40dB), si on prend l'exemple de 30dB la conclusion est la suivante ; au-dessus de 30dB, la marque est détectable, et au-dessous de 30dB, la marque est absente. En fait, ces valeurs sont les seuils de détection prédéfinis, l'atteinte de ces valeurs est due aux valeurs α choisies. Lors de notre travail on a essayé de faire un ajustement manuel afin de trouver des valeurs α efficaces pour pouvoir répondre aux contraintes r_1 et r_2 .

Tableau 5. 1: Valeurs des paramètres de marquage et de détection associées à chaque sous-bande DCT de l'image marquée en vue de la compression JPEG2000.

Transformée	Paramètres	Lena	Peppers	Goldhill	Boat
DCT	Robustesse				
	α_1	0.035	0.065	0.055	0.055
	α_2	0.38	0.29	0.17	0.183
	α_3	0.20	0.86	0.155	0.165
	PSNR (dB)				
	Image marquée	46.90	46.39	46.40	49.91
	Corrélation				
	R_1	0.8132	0.5979	0.7489	1.1505
	R_2	1.3818	1.0609	0.4784	0.5375
	R_3	2.7017	5.5008	2.9928	2.6290
	Seuil				
	T_1	0.2766	0.3532	0.3520	0.3059
T_2	0.1000	0.1237	0.1526	0.1385	
T_3	0.5621	0.6808	0.6934	0.7969	
Détection					
(D_1, D_2, D_3)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	

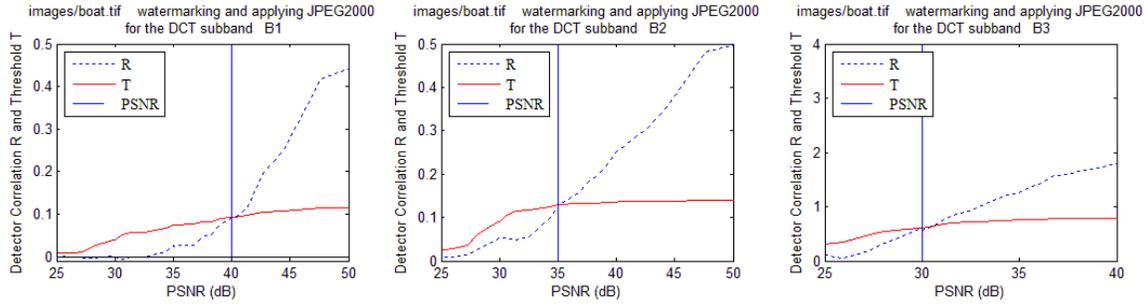


Figure 5. 1: Réponse du détecteur de marque par rapport à la variation du PSNR de l'image Boat suite à un marquage DCT et à une compression JPEG2000.

1.Validation de la procédure d'ajustement

Nous rappelons que le facteur de robustesse α joue un rôle très important pour notre système d'évaluation de la qualité des images. Si par exemple, on augmente ce facteur, on remarque que la marque dans l'image résiste mieux aux effets des dégradations subies par l'image hôte, et si on diminue la valeur de α on remarque que la marque devient plus fragile et a tendance à disparaître plus vite suite aux dégradations subies par l'image hôte. Dans les deux Figures 5.2 et 5.3, l'image Lena est marquée deux fois dans le domaine DWT. La première fois, la sous-bande B_1 reçoit une marque avec une robustesse $\alpha = 0.450$. La deuxième fois, la sous-bande B_1 reçoit une marque avec une robustesse plus faible $\alpha = 0.102$. Dans chacun des cas, l'image marquée subit diverses compressions JPEG. Pour chaque compression, la dégradation résultante est mesurée en calculant le PSNR entre l'image originale et l'image marquée compressée. Au même temps, la présence de la marque est détectée par la fonction de corrélation linéaire R comparée au seuil de détection T (voir la section 3.3.4). Les courbes des Figures 5.2 et 5.3 représentent d'une part, la variation de la fonction de corrélation R et du seuil de détection T en fonction du PSNR de l'image marquée et compressée en JPEG. Nous pouvons voir que la fonction de corrélation est croissante avec le PSNR. Le seuil de détection varie très peu avec le PSNR. Pour $\alpha = 0.450$, la marque est présente ($D=1$) dans l'image tant que la dégradation de compression subie est au-dessus de 35 dB. Alors que pour le cas $\alpha = 0.102$, la marque est présente dans l'image uniquement si la dégradation subie est au-dessus des 40 dB. Ainsi, on montre un résultat important : en ajustant la valeur du facteur de robustesse α , il est possible de fixer le seuil PSNR qui fait disparaître la marque et la rend non détectable ($D=0$) pour un type de dégradation donnée tel que la compression JPEG.

Dans le cas du marquage multiple à 3 marques et de la dégradation JPEG, l'ajustement des 3 facteurs de robustesse α_l consiste à trouver les bonnes valeurs α_l permettant la disparition de la marque W_1 dès que le PSNR est en dessous de 40 dB, la marque W_2 dès que le PSNR est en dessous de 35 dB et la marque W_3 dès que le PSNR est en dessous de 30 dB. Le PSNR étant calculé entre l'image originale et l'image marquée et compressée en JPEG. Pour une image hôte donnée, l'ajustement α_l d'une image est différent selon que la dégradation visée est JPEG ou JPEG2000. Il est aussi différent selon que le marquage se fait dans le domaine DCT ou DWT.

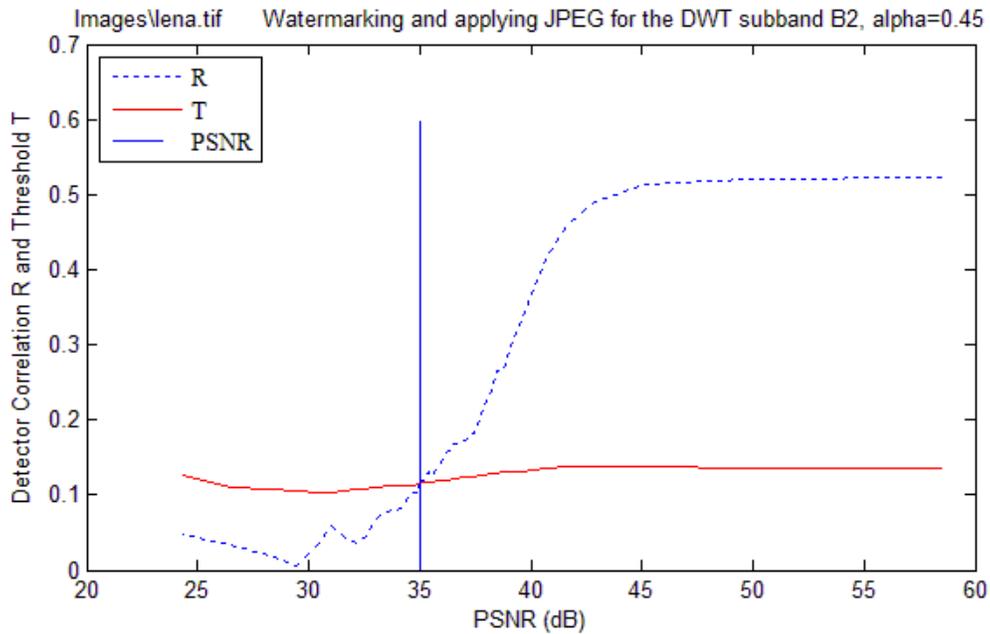


Figure 5. 2: Pour une valeur $\alpha = 0.45$, la marque est détectable au-dessus de 35dB et non détectable au-dessous de 35dB.

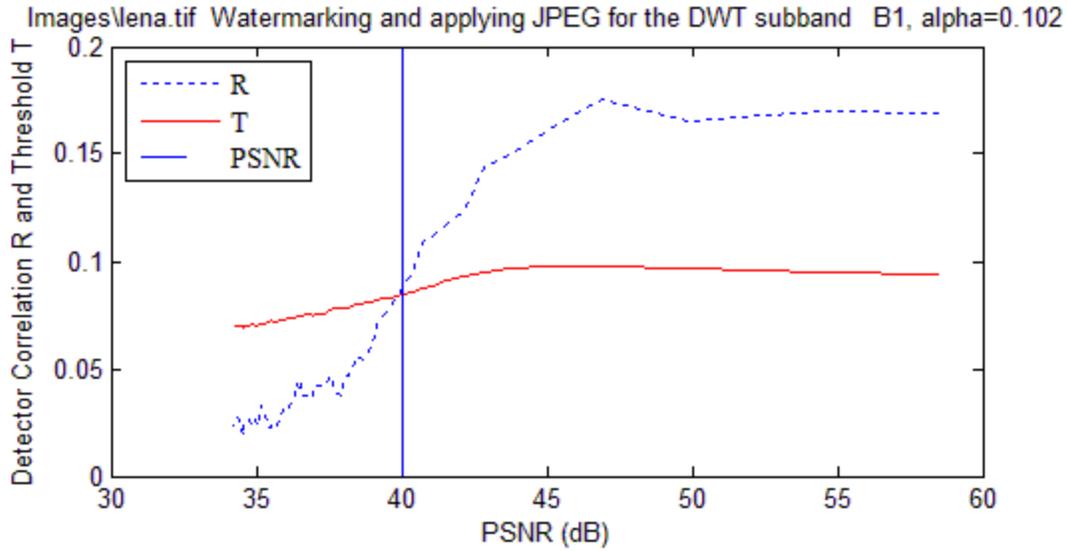


Figure 5. 3 : Pour une valeur $\alpha = 0.102$, la marque est détectable au-dessus de 40dB et non détectable au-dessous de 40dB.

Dans le cas du marquage multiple, l'ajustement des facteurs de robustesse α_i peut se faire manuellement ou de manière automatique. Le critère d'arrêt est l'atteinte des contraintes r_1 et r_2 tel que cela est expliqué dans la sous-section 3.3.6. La validation de la procédure d'ajustement automatique des α_i est illustrée par les Tableaux 5.2 et 5.3. Dans chacun des cas, on montre que les valeurs des facteurs α_i fixées automatiquement sont sensiblement les mêmes que celles trouvées manuellement. Ces résultats montrent que la procédure d'ajustement automatique ajuste très bien les valeurs α_i pour répondre aux contraintes r_1 et r_2 .

Tableau 5. 2 : Comparaison entre les valeurs α_i ajustées manuellement et automatiquement pour un marquage DWT et une compression JPEG.

$\alpha_1, \alpha_2, \alpha_3$	Manuel			Automatique		
Lena	0.056	0.176	1.029	0.051	0.172	1.020
Peppers	0.080	0.340	0.934	0.081	0.340	0.940
Goldhill	0.061	0.112	0.683	0.060	0.114	0.780
Boat	0.058	0.168	0.910	0.059	0.170	0.950

Tableau 5. 3: Comparaison entre les valeurs α ajustées manuellement et automatiquement pour un marquage DWT et une compression JPEG2000.

$\alpha_1, \alpha_2, \alpha_3$ Image	Manuel			Automatique		
Lena	0.059	0.140	1.200	0.059	0.146	1.180
Peppers	0.092	0.300	1.060	0.090	0.278	0.998
Goldhill	0.058	0.122	0.96	0.059	0.120	0.950
Boat	0.05	0.131	1.26	0.50	0.135	1.270

2. Génération et choix de la marque

Chaque marque utilisée comporte une séquence pseudo-aléatoire. La distribution statistique des séquences pseudo-aléatoires peut être uniforme à valeurs $\{+1, -1\}$ ou bien Gaussienne à moyenne nulle et à variance unitaire tel que cela est expliqué à la sous-section 3.3.2. Le générateur de séquences pseudo-aléatoires est disponible dans MATLAB. Les Figures 5.6 et 5.7 montrent des exemples de ces marques.

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	-3,3418	-1,9879	0
0	0	0	0	-2,2937	-0,6470	-1,9366	0
0	0	0	0	-3,6781	0,9554	0,4100	0
0	0	0	0	0	0	0	0

Figure 5. 4 : Exemple d'une marque W1 à distribution Gaussienne destinée au marquage d'une sous-bande DCT B_1 .

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	1	-1	0
0	0	0	0	1	-1	1	0
0	0	0	0	-1	1	-1	0
0	0	0	0	0	0	0	0

Figure 5. 5 : Exemple d'une marque W1 à distribution uniforme destinée au marquage d'une sous-bande DCT B_1 .

Nous avons effectué le marquage de plusieurs images en utilisant des marques Gaussiennes, puis une autre série de marquages utilisant des marques uniformes. D'après les résultats obtenus, on remarque qu'en utilisant des marques Gaussiennes, le PSNR des images marquées est meilleur qu'avec des marques uniformes. La différence peut dépasser les 4 dB tel que cela est montré dans le Tableau 5.4. Cette constatation est valide aussi bien pour un marquage DWT qu'un marquage DCT.

Au niveau des réponses du détecteur par rapport au PSNR, les fonctions de corrélation se comportent de manière similaire avec des marques Gaussiennes qu'avec des marques uniformes (voir les courbes ci-dessous).

Notons que pour chacun des cas de marquage multiple DCT et DWT avec $L=3$, trois marques de références W_1 , W_2 et W_3 sont fixées une bonne fois pour toutes les images à marquer. Ces marques sont aussi disponibles dans le détecteur de marques sur une base permanente. Ainsi, le système ne requiert pas de transmettre les marques de référence au récepteur.

Tableau 5. 4 : Impact du marquage multiple Gaussien ou uniforme sur le PSNR de l'image marquée dans la DWT en vue de la compression JPEG2000.

Marques uniformes		Marques Gaussiennes	
Image	PSNR (dB)	Image	PSNR (dB)
Boat	45.20	Boat	49.72
Goldhill	46.58	Goldhill	51.28
Lena	44.33	Lena	47.15
Mandrill	49.01	Mandrill	48.49
Peppers	49.97	Peppers	48.43

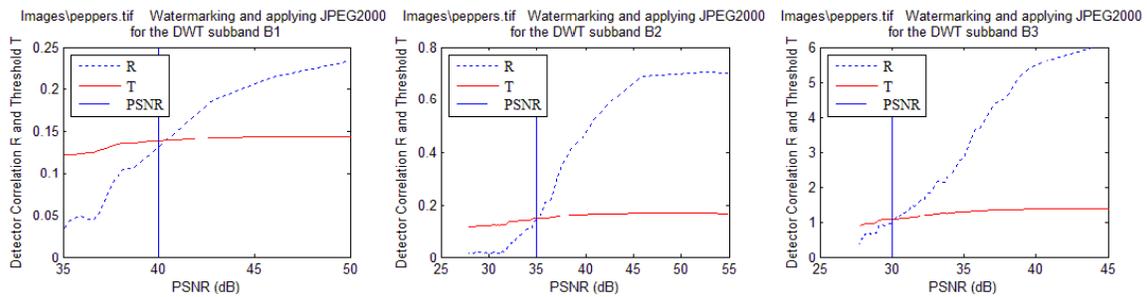


Figure 5. 6 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Peppers lorsque la compression visée est JPEG2000 en utilisant des marques uniformes dans la DWT.

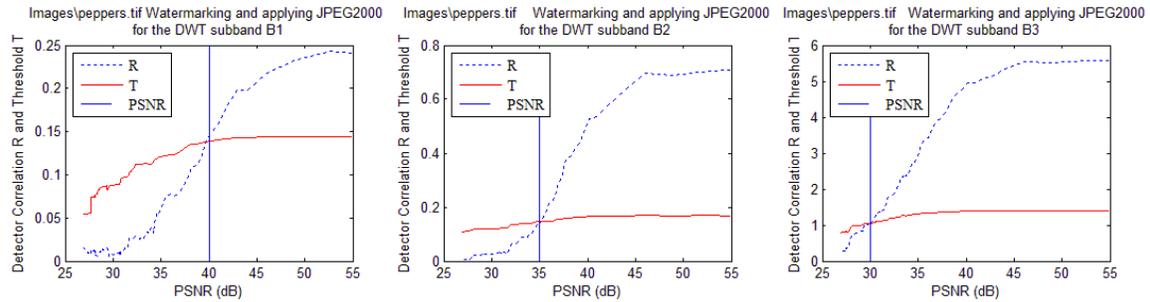


Figure 5. 7 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Peppers lorsque la compression visée est JPEG2000 en utilisant des marques Gaussiennes dans la DWT.

3. Efficacité d'estimation de la qualité

La validation des méthodes d'estimation de la qualité proposées est donnée par une statistique faite sur différentes images à des différents taux de compression pour les deux standards de compression JPEG et JPEG2000. Les résultats sont faits sur des images tests à niveau de gris de taille 512×512 ; Lena, Goldhill, Boat, Peppers et de la base de données des images UCID (Schaefer and Stich 2003). Les tableaux suivants montrent les résultats trouvés pour un marquage multiple dans le domaine DCT et dans le domaine DWT. Dans ces tableaux on trouve le véritable PSNR $PSNR_{vrai}$ calculé entre l'image originale et l'image marquée compressée, le $PSNR$ estimé par nos méthodes d'estimation de la qualité, la décision D_{vrai} qui aurait due être prise et la décision estimée D au niveau du détecteur de marques.

À partir de la base d'images UCID (Schaefer and Stich 2003), 40 images en niveaux de gris ont été sélectionnées et marquées selon les deux méthodes (DWT et DCT). Chaque image a subi 13 tests de compression JPEG/JPEG2000 et d'estimation de la qualité.

En nous basant sur les résultats obtenus (dont ceux apparaissant dans les tableaux ci-dessous), on conclue que les deux méthodes d'estimation de la qualité estiment correctement la qualité des images soumises que ce soit pour la distorsion JPEG ou JPEG2000. Le taux d'estimation correcte est autour de 94%. De plus, la plupart des cas d'estimations erronées sont reliés d'une façon ou d'une autre aux cas de décision $D_i=2$. En effet, souvent la décision $D_i=2$ est erronée au profit d'une décision $D_i=1$ ou $D_i=0$. Dans ces cas, c'est souvent un manque de précision sur l'estimation qu'une véritable erreur d'estimation. Par exemple, pour une image à un PSNR de 34,97dB on estime qu'il est dans l'intervalle 30-35dB au lieu de décider qu'il est autour de 35 dB (≈ 35 dB).

Tableau 5. 5 : Résultats d'estimation de la qualité pour les images Boat, Lena, Goldhill et Peppers marquées en DWT et compressées en JPEG à différents taux.

Image	Facteur de qualité Q	$PSNR_{vrai}$	$PSNR_{estimé}$	Estimation correcte	D_{vrai}	D
Boat	98	49.83	> 40	1	(1 1 1)	(1 1 1)
	92	41.87	> 40	1	(1 1 1)	(1 1 1)
	89	40.49	\approx 40	1	(2 1 1)	(2 1 1)
	80	38.11	35 -- 40	1	(0 1 1)	(0 1 1)
	70	36.58	35 – 40	1	(0 1 1)	(0 1 1)
	50	34.68	\approx 35	1	(0 2 1)	(0 2 1)
	43	34.11	30 – 35	1	(0 0 1)	(0 0 1)
	42	34.01	30 – 35	1	(0 0 1)	(0 0 1)
	40	33.82	30 – 35	1	(0 0 1)	(0 0 1)
	28	32.59	30 – 35	1	(0 0 1)	(0 0 1)
	20	31.38	30 – 35	1	(0 0 1)	(0 0 1)
	17	30.78	30 – 35	1	(0 0 1)	(0 0 1)
	11	28.84	<30	1	(0 0 0)	(0 0 0)
	7	28.41	<30	1	(0 0 0)	(0 0 0)
6	26.80	<30	1	(0 0 0)	(0 0 0)	
Taux d'estimation correcte				15/15		
Lena	98	49.81	> 40	1	(1 1 1)	(1 1 1)
	92	41.57	> 40	1	(1 1 1)	(1 1 1)
	90	40.74	> 40	1	(1 1 1)	(1 1 1)
	89	40.39	> 40	0/1	(2 1 1)	(1 1 1)
	85	39.36	35 -- 40	1	(0 1 1)	(0 1 1)
	70	37.25	35 -- 40	1	(0 1 1)	(0 1 1)
	50	35.37	35 -- 40	1	(0 1 1)	(0 1 1)
	43	35.28	\approx 35	1	(0 2 1)	(0 2 1)
	42	35.28	\approx 35	1	(0 2 1)	(0 2 1)
	40	35.06	\approx 35	1	(0 2 1)	(0 2 1)
	28	34.03	30 -- 35	1	(0 0 1)	(0 0 1)
	20	32.90	30 -- 35	1	(0 0 1)	(0 0 1)
	15	31.89	30 -- 35	1	(0 0 1)	(0 0 1)
	10	30.37	\approx 30	1	(0 0 2)	(0 0 2)
9	29.91	\approx 30	1	(0 0 2)	(0 0 2)	
Taux d'estimation correcte				15/15		
Taux d'erreur				1/15		
Goldhill	98	49.79	> 40	1	(1 1 1)	(1 1 1)
	92	40.36	\approx 40	1	(2 1 1)	(2 1 1)
	86	37.85	35 – 40	1	(0 1 1)	(0 1 1)
	80	36.48	35 – 40	1	(0 1 1)	(0 1 1)
	70	35.15	\approx 35	1	(0 2 1)	(0 2 1)
	60	34.23	30 – 35	1	(0 0 1)	(0 0 1)
	50	33.56	30 – 35	1	(0 0 1)	(0 0 1)
	40	32.88	30 – 35	1	(0 0 1)	(0 0 1)
	30	32.08	30 – 35	1	(0 0 1)	(0 0 1)
	25	31.54	30 – 35	1	(0 0 1)	(0 0 1)
	20	30.85	30 – 35	1	(0 0 1)	(0 0 1)
	17	30.34	30 – 35	0/1	(0 0 2)	(0 0 1)
	14	29.71	<30	0/1	(0 0 2)	(0 0 0)

	9	28.28	<30	1	(0 0 0)	(0 0 0)
	8	27.93	<30	1	(0 0 0)	(0 0 0)
Taux d'estimation correcte				13/15		
Taux d'erreur				2/15		
Peppers	98	49.77	> 40	1	(1 1 1)	(1 1 1)
	94	41.73	> 40	1	(1 1 1)	(1 1 1)
	91	39.52	≈ 40	1	(2 1 1)	(2 1 1)
	89	38.70	35 – 40	1	(0 1 1)	(0 1 1)
	80	36.92	35 – 40	1	(0 1 1)	(0 1 1)
	70	35.98	35 – 40	1	(0 1 1)	(0 1 1)
	60	35.28	35 – 40	0/1	(0 2 1)	(0 1 1)
	45	34.49	30 – 35	1	(0 0 1)	(0 0 1)
	30	33.52	30 – 35	1	(0 0 1)	(0 0 1)
	25	33.03	30 – 35	1	(0 0 1)	(0 0 1)
	20	32.39	30 – 35	1	(0 0 1)	(0 0 1)
	17	31.91	30 – 35	1	(0 0 1)	(0 0 1)
	14	31.27	30 – 35	1	(0 0 1)	(0 0 1)
	9	29.70	≈ 30	1	(0 0 2)	(0 0 2)
	8	29.26	<30	1	(0 0 0)	(0 0 0)
Taux d'estimation correcte				14/15		
Taux d'erreur				1/15		
Taux global d'estimation correcte				56/60=94%		

Tableau 5. 6 : Résultats d'estimation de la qualité pour les images Boat, Lena, Goldhill et Peppers marquées en DWT et compressées en JPEG2000 à différents taux.

Image	Taux de compression	$PSNR_{vrai}$	$PSNR_{estimé}$	Estimation correcte	D_{vrai}	D
Boat	2	53.65	> 40	1	(1 1 1)	(1 1 1)
	7	40.00	≈40	1	(2 1 1)	(2 1 1)
	7.5	39.53	≈40	1	(2 1 1)	(2 1 1)
	8	39.10	≈40	1	(0 1 1)	(0 1 1)
	10	37.50	35 -- 40	1	(0 1 1)	(0 1 1)
	12	36.13	35 -- 40	1	(0 1 1)	(0 1 1)
	13	35.57	35 -- 40	1	(0 1 1)	(0 1 1)
	14	35.15	≈ 35	1	(0 2 1)	(0 2 1)
	15	34.7	30 -- 35	0/1	(0 2 1)	(0 1 1)
	30	31.13	30 -- 35	1	(0 0 1)	(0 0 1)
	35	30.37	≈ 30	1	(0 0 2)	(0 0 2)
	38	30.02	≈ 30	1	(0 0 2)	(0 0 2)
	40	29.79	<30	0/1	(0 0 2)	(0 0 0)
	45	29.33	<30	1	(0 0 0)	(0 0 0)
	50	28.84	<30	1	(0 0 0)	(0 0 0)
Taux d'estimation correcte				13/15		
Taux d'erreur				2/15		
	2	54.13	> 40	1	(1 1 1)	(1 1 1)
	4	44.63	> 40	1	(1 1 1)	(1 1 1)
	6	41.76	> 40	1	(1 1 1)	(1 1 1)
	7	40.88	> 40	1	(1 1 1)	(1 1 1)

Lena	9	39.54	≈ 40	1	(2 1 1)	(2 1 1)
	12	38.25	35 -- 40	1	(0 1 1)	(0 1 1)
	13	37.87	35 -- 40	1	(0 1 1)	(0 1 1)
	14	37.56	35 -- 40	1	(0 1 1)	(0 1 1)
	15	37.30	35 -- 40	1	(0 1 1)	(0 1 1)
	25.5	34.80	≈ 35	1	(0 2 1)	(0 2 1)
	27	34.66	≈ 35	1	(0 2 1)	(0 2 1)
	40	32.78	30 -- 35	1	(0 1 1)	(0 1 1)
	72	30.42	≈ 30	1	(0 0 2)	(0 0 2)
	75	30.13	≈ 30	1	(0 0 2)	(0 0 2)
	95	27.27	<30	1	(0 0 0)	(0 0 0)
Taux d'estimation correcte				14/15		
Taux d'erreur				1/15		
Goldhill	2	51.96	> 40	1	(1 1 1)	(1 1 1)
	5	39.55	≈ 40	1	(2 1 1)	(2 1 1)
	6	38.23	35 -- 40	1	(0 1 1)	(0 1 1)
	7	37.31	35 -- 40	1	(0 1 1)	(0 1 1)
	11	34.72	30 -- 35	0/1	(0 2 1)	(0 0 1)
	12	34.27	30 -- 35	1	(0 0 1)	(0 0 1)
	13	33.92	30 -- 35	1	(0 0 1)	(0 0 1)
	14	33.64	30 -- 35	1	(0 0 1)	(0 0 1)
	26	31.15	30 -- 35	1	(0 0 1)	(0 0 1)
	28	30.82	30 -- 35	1	(0 0 1)	(0 0 1)
	30	30.62	30 -- 35	1	(0 0 1)	(0 0 1)
	40	29.75	≈ 30	1	(0 0 2)	(0 0 2)
	45	30.02	≈ 30	1	(0 0 2)	(0 0 2)
	72	28.11	<30	1	(0 0 0)	(0 0 0)
	80	27.86	<30	1	(0 0 0)	(0 0 0)
95	27.43	<30	1	(0 0 0)	(0 0 0)	
Taux d'estimation correcte				15/16		
Taux d'erreur				1/16		
Peppers	2	51.96	> 40	1	(1 1 1)	(1 1 1)
	5	39.55	≈ 40	1	(2 1 1)	(2 1 1)
	7	37.31	35 -- 40	1	(0 1 1)	(0 1 1)
	9	35.79	35 -- 40	1	(0 1 1)	(0 1 1)
	11	34.27	30 -- 35	1	(0 0 1)	(0 0 1)
	14	33.64	30 -- 35	1	(0 0 1)	(0 0 1)
	14	33.64	30 -- 35	1	(0 0 1)	(0 0 1)
	26	31.15	30 -- 35	1	(0 0 1)	(0 0 1)
	28	30.82	30 -- 35	1	(0 0 1)	(0 0 1)
	30	30.62	30 -- 35	1	(0 0 1)	(0 0 1)
	40	29.75	≈ 30	1	(0 0 2)	(0 0 2)
	70	28.11	<30	1	(0 0 0)	(0 0 0)
	80	27.86	<30	1	(0 0 0)	(0 0 0)
	90	27.58	<30	1	(0 0 0)	(0 0 0)
Taux d'estimation correcte				14/14		
Taux global d'estimation correcte				57/60=95%		

4. Transcodage et l'évaluation de la qualité

Dans la pratique, l'image peut subir une cascade de codage/décodage à des taux de compression variés, car lors de son acheminement vers un utilisateur final, l'image peut passer par différentes stations (ou points intermédiaires) et chacun peut avoir des contraintes spécifiques de débits cibles. Dans notre étude, deux scénarios seront pris en compte : le premier scénario consiste à mettre en cascade deux compressions successives JPEG/JPEG20000 ayant des taux de compression descendants (le premier taux de compression T_{c1} est plus élevé que le deuxième taux de compression T_{c2}). Le deuxième scénario met en série deux compressions JPEG/JPEG2000 à taux de compression ascendants. Le but de cette étude est de déterminer la performance de notre système à évaluer correctement la qualité de l'image reçue.

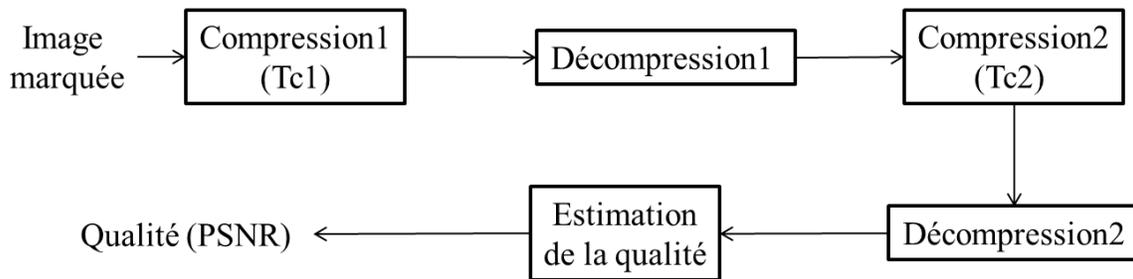


Figure 5. 8 : Exemple de cascade de codeurs et décodeurs.

- Cas : $T_{c1} < T_{c2}$

Tableau 5. 7 : Transcodage JPEG appliqué aux images Peppers, Goldhill, Lena et Boat préalablement marquées dans le domaine DWT.

Image	Tc1	Tc2	PSNR calculé après Tc1	PSNR calculé après Tc2	Détection	PSNR _{estimé}
Peppers	5	25	40.85	33.97	$D=(0,0,1)$	[30-35]
Goldhill	5	25	39.55	31.15	$D=(0,0,1)$	[30-35]
Lena	5	25	43.08	34.86	$D=(0,0,1)$	[30-35]
Boat	5	25	42.77	31.85	$D=(0,0,1)$	[30-35]

- Cas : $Tc1 > Tc2$

Tableau 5. 8 : Transcodage JPEG appliqué aux images Peppers, Goldhill, Lena et Boat préalablement marquées dans le domaine DWT.

Image	Tc1	Tc2	PSNR calculé après Tc1	PSNR calculé après Tc2	Détection	PSNR _{estimé}
Peppers	25	5	34.00	33.99	$D=(0,0,1)$	[30-35]
Goldhill	25	5	31.15	31.15	$D=(0,0,1)$	[30-35]
Lena	25	5	34.86	34.85	$D=(0,0,1)$	[30-35]
Boat	25	5	31.85	31.85	$D=(0,0,1)$	[30-35]

Suite aux résultats obtenus incluant ceux figurant dans les Tableaux 5.7 et 5.8, nous remarquons que lorsque l'image marquée subit une succession de compressions à différents taux de compression, le PSNR de l'image décodée reflète principalement l'effet de la compression ayant le plus grand taux de compression.

Le système d'évaluation de la qualité réussit très bien à estimer le PSNR de l'image finale malgré la cascade de compressions subie que ce soit pour le cas de compression JPEG ou de compression JPEG2000.

5. Effet du bruit additif

Le bruit additif est une distorsion de l'image qui peut venir de l'effet du canal de communication. Pour simuler l'effet du bruit additif, tel que montré dans la Figure 5.11, l'image marquée et compressée en JPEG/JPEG2000 est d'abord décompressée puis un bruit Gaussien de moyenne m et de variance ν lui est ajouté à l'aide de la fonction Matlab '*imnoise ()*'. L'image ainsi bruitée est présentée au module d'estimation de la qualité. Par ailleurs, le calcul du PSNR entre l'image marquée et l'image marquée compressée et bruitée permet de vérifier si l'estimation effectuée est correcte. Le Tableau 5.9 présente quelques résultats de cette simulation. Nous remarquons que l'estimation de la qualité est correcte si le bruit additif est très faible avec une moyenne nulle et une variance en dessous de 10^{-4} . Lorsque la variance ν est plus élevée, la qualité estimée est alors erronée.

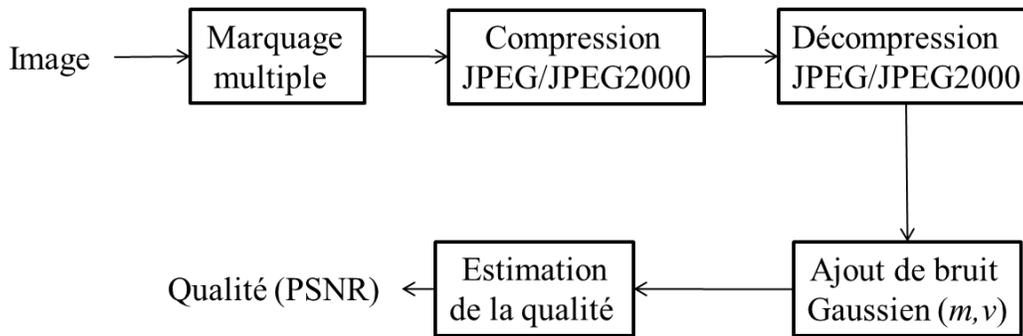


Figure 5. 9 : Processus de l’ajout de bruit Gaussien dans une image donnée.

Tableau 5. 9 : Résultats de l’estimation de la qualité suite à l’ajout de bruit Gaussien pour les images Lena, Boat, Peppers et Goldhill marquées et compressées en JPEG.

Image	Facteur de qualité	PSNR _{vrai} (compressée et bruitée)	Décision prévue	Moyenne du bruit	Variance du bruit	Décision	PSNR estimé
Lena	75	36.32	(0, 1, 1)	0	0.0001	(0, 1, 1)	35—40
Boat	60	33.43	(0, 0, 1)	0	0.0001	(0, 0, 1)	30—35
Peppers	75	31.19	(0, 0, 1)	0.001	0.001	(0, 0, 0)	<30
Goldhill	80	37.93	(0, 1, 1)	0	0,00001	(0, 1, 1)	35—40

6. Effet de la perte de paquets

La perte de paquets est un problème qui peut arriver lors de la transmission de données. Le bitstream JPEG ou JPEG2000 à décoder se trouve ainsi erroné ou amputé de quelques bytes. Le décodeur dispose normalement d’une stratégie de compensation de cette perte, mais comme elle n’est pas standard, elle peut donc différer d’un décodeur à un autre. Dans la plupart des cas, la perte de données se traduit par une distorsion apparaissant dans l’image décodée. Nous nous intéressons à l’étude de l’effet de la perte de paquets du bitstream compressé sur l’évaluation de la qualité de l’image. Pour simuler la perte de paquets deux techniques ont été utilisées :

- a) Suppression de plusieurs bytes quelque part au milieu du bitstream : le bitstream se trouve ainsi amputé de quelques données et sa taille est plus petite. Dans le cas du décodeur JPEG de MATLAB, l’image décodée affiche tout un décalage des pixels à partir de l’endroit où les bytes ont été perdus.

b) Modification de plusieurs bytes quelques part au milieu du bitstream : sachant que lors d'une transmission par paquets, les entêtes de paquets contiennent des numéros indiquant leur succession dans le temps. Il est donc facile de repérer un paquet perdu (ou retardé) et de le remplacer par un paquet de substitution. Ce dernier peut contenir des données nulles. Pour simuler cette situation, nous avons choisi de modifier quelques bytes au milieu du bitstream et de constater l'effet sur l'image décodée puis sur l'évaluation de la qualité de cette image. Notons que dans ce cas, le bitstream à décoder garde exactement la même taille.

Les résultats présentés ci-dessous sont faits sur des images préalablement marquées dans le domaine DCT, compressées en JPEG ou JPEG2000 puis altérées par la perte de paquets (suppression ou modification). La décompression JPEG ou JPEG2000 permet de constater l'effet de l'altération subie et d'appliquer l'estimation de la qualité telle que c'est illustré dans la Figure 5.12.

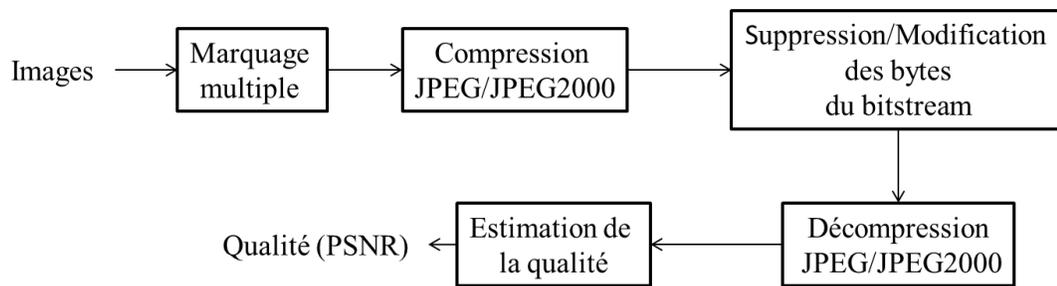


Figure 5. 10 : Processus de simulation de la perte de paquets dans un bitstream JPEG ou JPEG2000.

Dans le cas de la perte de paquets par suppression de bytes du bitstream JPEG, l'image décodée comporte des décalages. La Figure 5.13 présente l'image Lena suite à une compression JPEG avec un facteur de qualité de 70%. La Figure 5.14 présente la même image Lena mais suite à une compression JPEG de 70% de qualité suivie d'une suppression de 1 byte. Cette perte a résulté en un décalage du reste de l'image à décoder, ce qui fausse le calcul du PSNR sans pour autant dégrader la qualité perçue. Toutefois, l'estimation de la qualité est correcte tant que le nombre de bytes perdus est inférieur à 10 bytes comme le montre le Tableau 5.10.

Dans le cas de la compression JPEG2000, dès qu'un byte est supprimé du bitstream, l'image décodée est dégradée et son PSNR est fortement diminué. L'estimateur de qualité semble bien détecter cette dégradation comme le montrent les résultats du Tableau 5.11.



Figure 5. 11 : Image Lena compressée avec JPEG à un facteur de qualité $Q=70\%$. Son PSNR est 37.06 dB.



Figure 5. 12 : Effet de la perte d'un byte du bitstream JPEG de l'image Lena de la Figure 39.

Tableau 5. 10 : Effet de la suppression de quelques bytes du bitstream JPEG des images Lena et Boat compressées à un facteur de qualité Q=70%.

Image	Nombre de bytes perdus	PSNR avant la perte de bytes	PSNR après la perte de bytes	Décision de l'estimateur de qualité	Qualité estimée en dB
Lena	1	37.06	37.01	(0, 1, 1)	[35-40]
	10	37.06	35.92	(0, 1, 1)	[35-40]
	50	37.06	15.59	(0, 0, 1)	[30-35]
	100	37.06	13.78	(0, 0, 1)	[30-35]
Boat	1	36.56	36.46	(0, 1, 1)	[35-40]
	10	36.56	36.32	(0, 1, 1)	[35-40]
	50	36.56	16.24	(0, 0, 1)	[35-40]
	100	36.56	17.78	(0, 0, 1)	[35-40]

Tableau 5. 11 : Effet de la suppression (perte) de quelques bytes du bitstream JPEG2000 des images Boat et Lena à un taux de compression égal à 8.

Image	Nombre de bytes perdus	PSNR avant la perte de bytes	PSNR après la perte de bytes	Décision de l'estimateur de qualité	Qualité estimée en dB
Lena	1	39.74	27.79	(0, 0, 1)	[30-35]
	10	39.74	11.23	(0, 0, 0)	<30
	50	39.74	10.06	(0, 0, 0)	<30
	100	39.74	9.57	(0, 0, 0)	<30
Boat	1	38.91	21.95	(0, 0, 1)	[30-35]
	10	38.91	18.74	(0, 0, 0)	<30
	50	38.91	11.59	(0, 0, 0)	<30
	100	38.91	10.05	(0, 0, 0)	<30

Dans le cas de la perte de paquets par modification de bytes du bitstream JPEG, l'image décodée comporte une similarité à l'image marquée compressé s'il s'agit d'une modification inférieure à 10 bytes. Par contre, si le nombre de bytes modifié est supérieur à 10 l'image présente un décalage. La Figure 5.15 présente deux images; une image de Lena compressée JPEG avec une modification d'un byte et l'image de différence entre l'image compressée et l'image compressée et altérée.

Si la compression/décompression est JPEG2000, on constate que la modification de bytes engendre des distorsions de l'image dont la localisation est en lien avec les sous-bandes ondelettes. Le PSNR chute et l'estimateur de qualité arrive à estimer le PSNR tant que le nombre de bytes modifiés ne dépasse pas les 100 bytes comme le montre le Tableau 5.12.

La Figure 5.16 présente deux images; à gauche, l'image Lena marquée, compressée en JPEG2000 et suivie d'une modification de 100 bytes, et à droite, sa différence avec l'image Lena marquée et compressée. L'image résultante après la modification des bytes présente quelques distorsions légères sur l'image.

Tableau 5. 12 : Effet de la modification de quelques bytes du bitstream JPEG des images Lena et Boat à un facteur de qualité de 70%.

Image	Nombre de bytes modifiés	PSNR avant la modification des bytes	PSNR après la modification des bytes	Décision de l'estimateur de qualité	Qualité estimée en dB
Lena	1	37.09	36.88	(0, 1, 1)	[35-40]
	10	37.09	35.69	(0, 1, 1)	[35-40]
	100	37.09	14.63	(0, 0, 1)	[30-35]
Boat	1	36.52	35.88	(0, 1, 1)	[35-40]
	10	36.52	35.42	(0, 1, 1)	[35-40]
	100	36.52	15.32	(0, 0, 1)	[35-40]

Tableau 5. 13 : Effet de la modification de quelques bytes du bitstream JPEG2000 des images Lena et Boat à un taux de compression égal à 8.

Image	Nombre de bytes modifiés	PSNR avant la modification des bytes	PSNR après la modification des bytes	Décision de l'estimateur de qualité	Qualité estimée en dB
Lena	10	38.71	38.69	(0, 1, 1)	[35-40]
	100	38.71	29.64	(0, 1, 1)	[35-40]
	1000	38.71	21.32	(0, 0, 1)	[30-35]
Boat	10	37.40	36.92	(0, 1, 1)	[35-40]
	100	37.40	36.69	(0, 1, 1)	[35-40]
	1000	37.40	23.79	(0, 0, 1)	[35-40]



Figure 5. 13 : Image Lena suite à une modification de 10 bytes du bitstream JPEG et l'image de différence par rapport à l'image compressée (sans perte).



Figure 5. 14 : Image Lena suite à une modification de 100 bytes du bitstream JPEG2000 et l'image de différence par rapport à l'image compressée (sans perte).

D'après les résultats présentés ci-dessus nous concluons qu'avec le standard JPEG2000, l'estimateur de qualité fonctionne bien si le nombre de bytes à modifier est inférieur à 100 bytes. Par contre, l'estimation est erronée lorsque la modification dépasse les 100 bytes. Dans le cas de la compression JPEG, cette limite se fixe à 10 bytes. Cela peut être dû en partie à l'encodeur/décodeur de Matlab qui ne prend pas en charge la compensation d'erreur.

Cette conclusion reste la même lorsque le marquage multiple se fait dans le domaine de la DWT au lieu de la DCT.

7. Domaine d'insertion et les standards de compression

Parmi les objectifs fixés dans le cadre de ce projet était d'étudier l'influence du domaine d'insertion des marques par rapport au domaine de compression de l'image marquée. Sachant que la compression opère par quantification dans le domaine transformé (DCT pour JPEG et DWT pour JPEG2000), est-il avantageux d'insérer les marques dans le même domaine que la compression ou pas ?

a) Cas de la compression JPEG

Suite à un ensemble de tests sur plusieurs images, nous avons remarqué que lorsque la distorsion attendue est une compression JPEG, le marquage multiple dans le domaine de la DWT est meilleur. En effet, le PSNR des images préalablement marquées (indiquant l'invisibilité des marques) est nettement meilleur. Le gain obtenu par rapport à un marquage dans la DCT dépasse les 2 dBs comme le montre le Tableau 5.14. De plus, la réponse du détecteur de marques dans la DWT a une fonction de corrélation qui s'écarte très bien du seuil de détection, alors que dans la DCT, la fonction de corrélation a tendance à se confondre avec le seuil dès que le PSNR est inférieur au PSNR cible. Cette constatation est bien illustrée dans les Figures 5.17 et 5.18.

Tableau 5. 14 : Comparaison des valeurs du PSNR des images marquées lorsque le marquage multiple opère dans la DCT versus la DWT en vue de la compression JPEG.

Image	DCT	DWT
Lena	43.92	48.74
Boat	48.44	47.19
Goldhill	45.75	49.76
Peppers	43.31	49.25

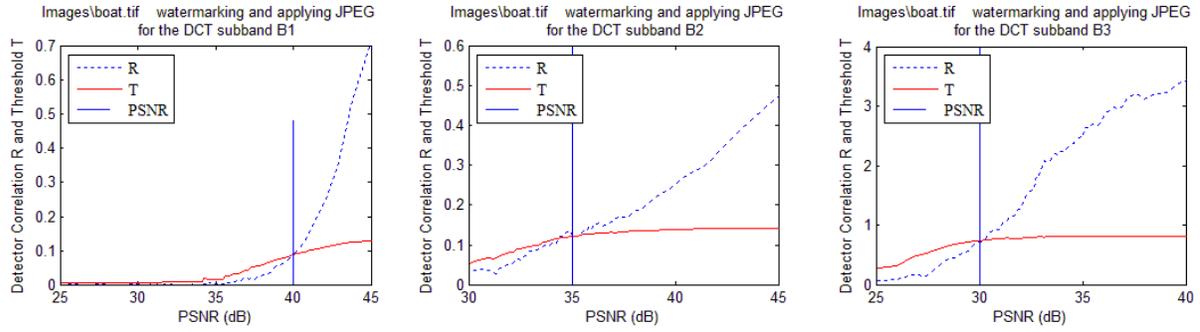


Figure 5. 15 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Boat lorsque la compression visée est JPEG et le marquage opère dans la DCT.

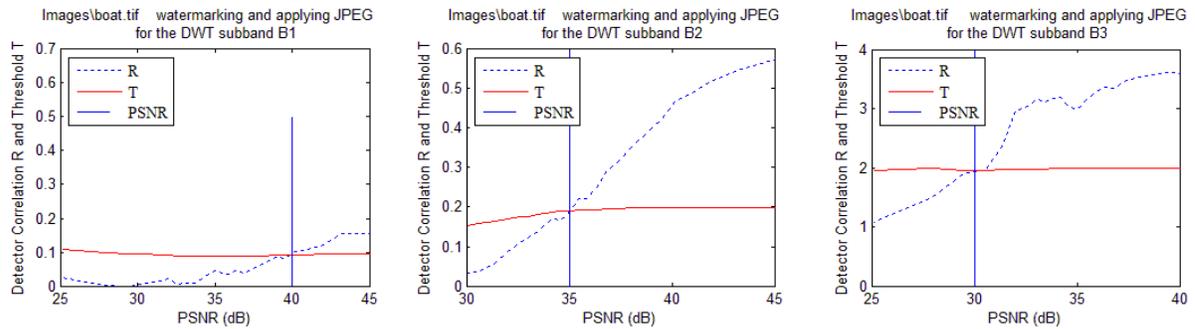


Figure 5. 16 : Réponse du détecteur par corrélation en fonction du PSNR de l'image Boat lorsque la compression visée est JPEG et le marquage opère dans la DWT.

b) Cas de la compression JPEG2000

Suite à un ensemble de tests sur plusieurs images, nous avons remarqué que lorsque la distorsion attendue est une compression JPEG2000, le marquage multiple dans le domaine de la DCT est meilleur. En effet, le PSNR des images préalablement marquées (indiquant l'invisibilité des marques) est nettement meilleur. Le gain obtenu par rapport à un marquage dans la DWT dépasse un dB dans la plupart des cas, comme le montre le Tableau 5.15. Au niveau de la réponse du détecteur de marques, les fonctions de corrélation se comportent de manière favorable aussi bien dans le cas de la DCT que dans le cas de la DWT. Cette constatation est bien illustrée dans les Figures 5.19 et 5.20.

Tableau 5. 15 : Comparaison des valeurs du PSNR des images marquées lorsque le marquage multiple opère dans la DCT versus la DWT en vue de la compression JPEG2000.

Image	DCT	DWT
Lena	46.89	46.25
Boat	49.91	45.69
Goldhill	50.97	47.52
Peppers	46.39	45.37

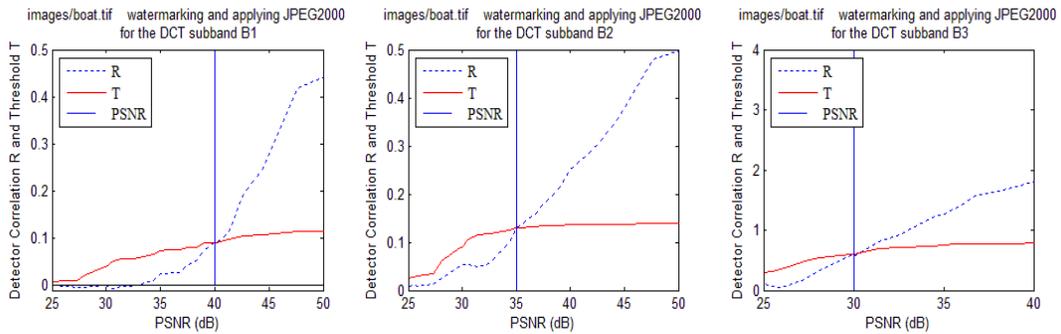


Figure 5. 17 : Réponse du détecteur par corrélation en fonction du PSNR de l’image Boat lorsque la compression visée est JPEG2000 et le marquage opère dans la DCT.

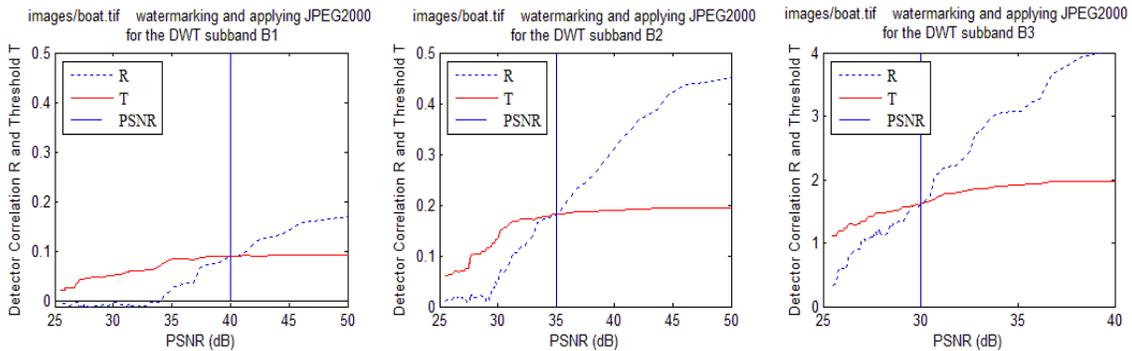


Figure 5. 18 : Réponse du détecteur par corrélation en fonction du PSNR de l’image Boat lorsque la compression visée est JPEG2000 et le marquage opère dans la DWT.

Suite à cette étude, nous pouvons conclure qu’il est préférable que le domaine d’insertion des marques soit différent du domaine de la compression. Donc, si JPEG est prévu comme distorsion du canal de communication, il serait mieux d’adopter un marquage multiple dans le domaine DWT. Inversement, le marquage dans le domaine DCT serait plus adapté à la compression JPEG2000. Ceci étant dit, il s’agit juste d’une

préférence. En effet, les deux méthodes de marquage multiples introduites dans ce projet restent correctes et efficaces quelques soit le type de compression (JPEG ou JPEG2000) adopté.

8. La performance temporelle

La performance en temps des deux méthodes proposées est donnée dans les trois tableaux ci-dessous. L'exécution des deux méthodes de marquage multiple est réalisée sur une machine dont le système d'exploitation est le Microsoft Windows Professionnel 7, 2.5GHz dual-core Intel Core i5, Turbo Boost up to 3.1GHz, 4GB 1600MHz pour la mémoire. Les images test sont des images en niveaux de gris de taille 512×512 pixels.

Les résultats relatifs au temps d'exécution du module de marquage multiple avec des valeurs α_i pré-ajustées manuellement, sont données dans le Tableau 5.16. Le temps d'exécution comporte la lecture et l'écriture des images. Ces résultats sont obtenus suite à une exécution répétitive sur toutes les images (le marquage est répété 10 fois pour chaque image) puis la moyenne du temps d'exécution est calculée.

Tableau 5. 16 : Temps d'exécution (en secondes) pour le module de marquage multiple dans les deux domaines DCT et DWT. Les valeurs des facteurs de robustesse α_i sont pré-ajustées manuellement.

Image	DWT	DCT
	Temps d'exécution moyen (en secondes)	
Lena	2.9, 3.8, 3.2, 3.2, 3.1	4.2, 3.3, 3.5, 3.4, 3.7
	3.4, 3.6, 2.8, 3.2, 2.9	3.3, 3.4, 3.5, 3.2, 3.6
Peppers	3.1, 3.1, 3.7, 3.5, 3.4	3.7, 3.3, 3.6, 3.9, 3.2
	2.9, 3.4, 3.2, 3.5, 3.3	3.3, 3.6, 3.5, 3.4, 3.1
Goldhill	3.2, 3.1, 2.7, 3.1, 3.2	3.2, 3.4, 3.1, 2.9, 3.1
	3.1, 2.7, 3.1, 3.2, 3.1	3.1, 3.6, 3.5, 2.8, 3.4
Boat	4.5, 3.2, 2.8, 3.1, 2.9	4.1, 3.7, 3.1, 3.9, 3.8
	3.2, 3.1, 2.9, 3.3, 3.7	3.6, 3.5, 3.2, 3.5, 3.3
Temps moyen/image	3.2	3.5

Dans le cas où les valeurs α_l sont ajustées automatiquement, les résultats relatifs au temps d'exécution du module de marquage multiple sont donnés dans le Tableau 5.17 et 26. Le temps d'exécution inclut la lecture des images et la compression/décompression des images. Le standard de compression utilisé est le JPEG2000. Remarquons que chaque itération de la procédure d'ajustement est accompagnée d'une compression/décompression JPEG2000 de l'image, ce qui ralentit énormément le marquage multiple.

Tableau 5. 17 : Temps d'exécution (en minutes) pour le module de marquage multiple dans les deux domaines DCT et DWT. Les valeurs des facteurs de robustesse α_l sont ajustées automatiquement en vue de la compression JPEG2000.

Image	DCT	DWT
	Temps d'exécution moyen (en minutes)	
Lena	1.6, 2.1, 2.2, 1.5, 1.6	1.6, 1.4, 1.1, 1.2, 1.6
	2.1, 2.2, 1.6, 1.5, 2.1	1.5, 1.1, 1.2, 1.3, 1.9
Peppers	1.6, 2.3, 2.1, 1.7, 1.9	1.4, 1.2, 1.1, 1.5, 1.4
	2.3, 2.4, 1.8, 1.4, 2.1	1.2, 1.7, 1.9, 1.1, 1.6
Goldhill	1.6, 1.4, 2.3, 2.5, 1.7	1.7, 1.1, 1.3, 1.5, 1.3
	1.4, 1.7, 2.1, 2.5, 1.8	1.6, 1.3, 1.4, 1.1, 1.9
Boat	1.3, 1.2, 2.1, 1.5, 2.1	1.4, 1.2, 1.5, 1.4, 1.2
	1.7, 1.3, 1.6, 1.9, 2.	1.2, 1.6, 1.1, 1.3, 1.9
Temps moyen/image	1.9	1.4

Tableau 5. 18 : Temps d'exécution (en secondes) pour le module de marquage multiple dans les deux domaines DCT et DWT. Les valeurs des facteurs de robustesse α_i sont ajustées automatiquement en vue de la compression JPEG.

Image	DCT	DWT
	Temps d'exécution moyen (en minutes)	
Lena	2.1, 2.2, 1.6, 1.5, 2.1	1.3, 1.2, 1.1, 1.1, 1.2
Peppers	2.3, 2.4, 1.8, 1.4, 2.1	1.2, 1.0, 1.1, 2.1, 1.3
Goldhill	1.6, 1.4, 2.3, 2.5, 1.7	1.1, 0.9, 1.2, 1.2, 0.9
Boat	1.3, 1.2, 2.1, 1.5, 2.1	1.4, 1.3, 1.1, 1.3, 0.9
Temps moyen/image	1.9	1.3

Concernant le module de détection des marques et estimation de la qualité de l'image, les temps d'exécution obtenus sont présentés dans le Tableau 5.18. Il est à noter que le temps de lecture des fichiers d'images et de décodage JPEG/JPEG2000 n'a pas été pris en compte.

Tableau 5. 19 : Temps d'exécution (en secondes) pour le module de détection/estimation de la qualité dans les deux domaines DCT et DWT.

Images	DCT	DWT
	Temps d'exécution moyen (en secondes)	
Lena	1.01, 1.12, 0.92, 1.06, 1.12	0.15, 0.11, 0.12, 0.14, 0.13
Boat	1.13, 1.12, 1.09, 1.11, 1.21	0.12, 0.11, 0.13, 0.10, 0.09
Goldhill	1.01, 1.05, 1.09, 1.10, 1.12	0.14, 0.15, 0.11, 0.12, 0.11
Peppers	1.09, 1.11, 0.95, 1.02, 1.12	0.12, 0.11, 0.08, 0.13, 0.11
Temps moyen/image	1.02	0.12

Les résultats montrent que la méthode d'évaluation de la qualité des images, en utilisant le marquage multiple dans le domaine DWT est plus rapide que celle utilisant la DCT. Par conséquent, il est prometteur pour les séquences vidéo en temps réel. En effet, à 0.12 secondes de temps d'exécution, nous pouvons envisager, par exemple, une estimation en temps réel de la qualité d'une séquence vidéo de 30 images par seconde à raison de deux images marquées à chaque seconde.

Chapitre 6 : MPEG et Méthodes d'évaluation de la qualité

Une séquence vidéo présente une redondance dans l'image (spatiale, psychovisuelle et de codage) et une redondance temporelle ou inter-images ; dans de nombreuses séquences vidéos, de nombreuses images restent inchangées ou bien changent très peu puisque les changements sont dus essentiellement aux objets en mouvement dans la scène. La compression vidéo vise à réduire les redondances. Les standards de compression vidéo, comme MPEG1 ou MPEG2 utilisent principalement des techniques comme la transformée DCT, la prédiction intra-image, l'estimation de mouvement, la prédiction inter-image et/ou compensation de mouvement, la quantification et le codage entropique. Dans cette partie on est intéressé par le standard de compression MPEG1 qui est apparu dans les années 1990-91. Le débit typique est 1.5 Mbits/sec pour des images de type SIF (360×240 pixels) et des séquences de 30 images par seconde.

Dans MPEG1, une séquence d'images est en général organisée en groupes d'images ou GOP. Chaque GOP peut contenir de 12 à 15 images codées selon trois modes possibles (voir la Figure 6.1):

- Images Intra (**I**)
- Images Prédites (**P**)
- Images Bidirectionnelles (**B**)

Une image intra (I) est codée par un algorithme de compression très similaire à JPEG sans aucune référence aux autres images. Le taux de compression est moins élevé que ceux des autres modes P et B. Le mode Intra permet l'accès aléatoire à la séquence compressée et facilite les modes de lecture accélérée (*Fast Forward, Fast Backward*). Dans un GOP à 15 images, il est fréquent d'avoir une seule image Intra, ce qui équivaut à deux images intra par seconde de vidéo.

L'image prédite (P) est une image prédite par compensation de mouvement à partir de l'image I ou P précédente. Son décodage dépend du décodage de l'image I ou P précédente. L'image bidirectionnelle (B) est une image prédite par compensation de mouvement à partir des images passée et future I ou P, requiert un buffer où on garde 3 images : courante, passée et future I ou P et impose un ordre de codage/décodage différent de l'ordre d'affichage. Ainsi, cette image limite la propagation de l'erreur lors du décodage, puisqu'aucun autre codage ne dépend d'une image B. La Figure 6.1 montre un exemple d'un GOP à 12 images. Notons que l'ordre de codage et de stockage pour ce GOP est: I P B B B P B B B I B B B, alors que l'ordre d'affichage pour ce groupe d'images est: I B B B P B B B P B B I.

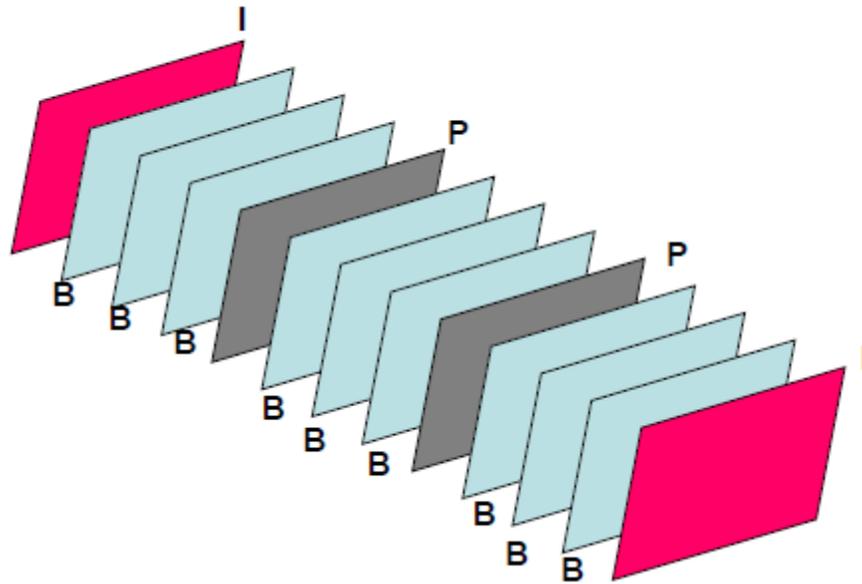


Figure 6. 1 : Exemple d'un groupe d'image (GOP) de MPEG1 montrant les images intra (I), prédites (P) et bidirectionnelles (B).

Dans ce contexte, nous allons proposer un schéma permettant d'évaluer la qualité des séquences vidéo en utilisant les méthodes de marquage multiples proposées. Pour ce faire, nous avons besoin de séquences d'images (soit une séquence d'images SIF de taille 360×240). L'exécution de la simulation requiert l'environnement Matlab ainsi qu'un codeur MPEG1. Il est à noter que MATLAB version 2012 ne propose pas encore de codeur MPEG1. Seul le décodage MPEG1 est possible à travers la fonction `mmreader()`. Cette fonction support plusieurs formats vidéo dont le AVI '.avi', le MPEG1 '.mpg', et le Windows Media Video '.wmv', '.asf', '.asx', ... Cette fonction permet aussi de donner une information générale sur la séquence vidéo tel que la durée, le nom, le type, etc.

1. Méthodologie

L'approche proposée pour l'évaluation automatique de la qualité d'images dans une séquence vidéo repose sur les méthodes de marquage multiples présentées dans ce mémoire. L'atteinte de cet objectif est appuyée par le développement d'un programme dans l'environnement de simulation MATLAB (la version de MATLAB utilisée est R2011), et la disponibilité d'un standard de compression MPEG1 comme le `bbmpeg` (voir l'annexe de ce document). Les expérimentations sont effectuées sur des séquences d'images réelles à 256 niveaux de gris et la base de vidéo ([VIPS 2012](#)) est utilisée à cette fin. Le problème que nous avons rencontré dans cette partie est que le codeur MPEG1

n'opère pas dans l'environnement MATLAB. Seul le décodage MPEG1 est possible dans MATLAB. Le processus de marquage d'une séquence vidéo et d'estimation de sa qualité comporte les étapes suivantes (voir aussi la Figure 6.2) :

- Extraction des images intra de la séquence d'images.
- Marquer les images intra.
- Réinsertion des images intra dans la séquence d'images.
- Codage MPEG1 en utilisant *bbmpeg*.
- Décodage MPEG1 en utilisant MATLAB.
- Conversion vers une série d'images.
- Extraction des images marquées.
- Estimation de la qualité de la séquence vidéo.

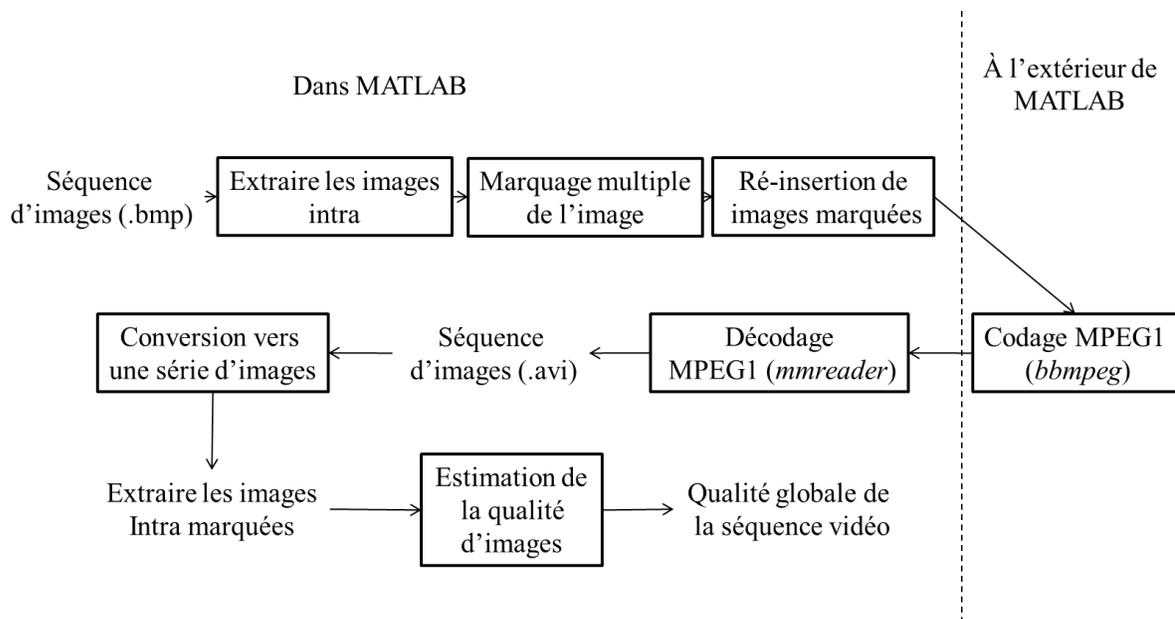


Figure 6. 2 : Processus d'estimation de la qualité d'une séquence vidéo Codée en MPEG1.

2. Insertion de la marque

Le processus de marquage multiple est limité aux images prévues pour un mode de codage MPEG1-intra comme le montre la Figure 6.3. La raison est que le codage MPEG1-intra partage le même algorithme que JPEG. Le marquage multiple est donc effectué sur chaque image intra de la séquence comme si cette dernière était une image fixe destinée à la compression JPEG. En d'autres mots, l'ajustement des facteurs de robustesse α_l des trois marques insérées se fait en vue de la compression JPEG. Une fois

ces images marquées, elles sont intégrées dans la séquence d'images préalablement au codage MPEG1.

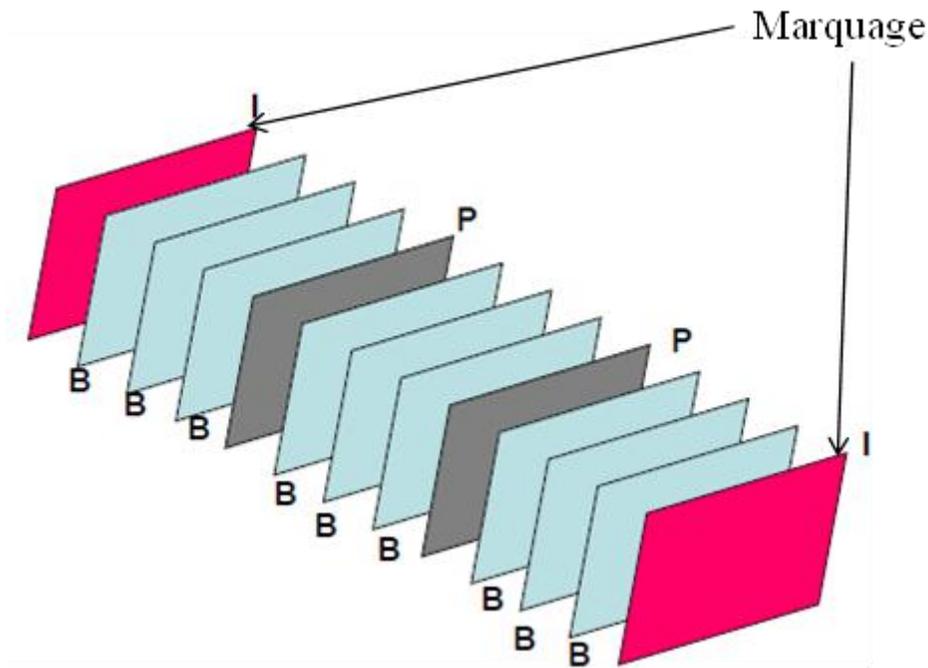


Figure 6. 3 : Marquage des images intra dans une séquence vidéo destinée au codage MPEG1.

3. Détection de la marque

Une fois la séquence d'images est codées en MPEG1, les images en mode intra sont décodées par MPEG1 et soumises à l'estimation de la qualité. Les trois marques contenues dans chaque image intra sont donc détectées et le triplet de détection est converti en une décision de qualité en termes de PSNR selon le Tableau 4.1.

4. Estimation de la qualité vidéo

Ayant estimé la qualité des images intra de la séquence (soit une à deux images par seconde), il est alors possible de se faire une idée globale de la qualité de la séquence vidéo en affichant sa qualité toutes les 0.5 secondes. Puisque les images I sont en général moins compressées que les images P et B, alors la qualité des images I dans un GOP constituent une borne supérieure pour la qualité de tout le GOP. Ce schéma est encore

plus intéressant lorsque le GOP contient une seule image intra. L'estimation de la qualité de la totalité du vidéo est donnée par la qualité des images intra la plus fréquente. Par exemple si on a 10 images marquées dans une séquence vidéo, et que 8 images ont une qualité entre 30—35 dB et les deux autres ont une qualité entre 35—40 dB, donc la qualité de la séquence vidéo est celle de la qualité d'image la plus fréquente, dans ce cas 30—35 dB.

5. Problèmes rencontrés

Suite aux tests effectués avec MPEG1, le principal problème rencontré est lié au fait que le codage et le décodage MPEG1 ne se réalisent pas dans le même environnement (le codage se fait avec *bbmpeg* et le décodage se fait avec MATLAB). Par conséquent, lors du marquage multiple de l'image, l'ajustement des facteurs de robustesse α_i se fait en considérant une distorsion de type JPEG au lieu d'une distorsion de type MPEG1-intra. Ceci a malheureusement conduit à des estimations de qualité qui sont erronées dans la plupart des cas. Il faudra donc disposer d'un codeur MPEG1-intra afin de procéder à une simulation correcte et complète.

Chapitre 7 : Conclusion

Dans ce mémoire, nous avons présenté notre travail qui concerne le développement de méthodes d'évaluation automatique sans référence de la qualité d'images compressées en JPEG ou en JPEG2000. Deux méthodes basées sur le marquage multiple sont proposées et étudiées : une nouvelle méthode opérant dans le domaine de la transformée DCT et une version améliorée de la méthode d'évaluation présentée dans ([Baaziz, Zheng et al. 2011](#)) et opérant dans le domaine de la DWT.

Afin de valider notre travail nous avons suivi les objectifs énoncés dans la section 3.1. Concernant la méthode d'évaluation de la qualité basée sur le marquage multiple dans la DWT, une des améliorations réalisées est au niveau du module de détection de marques et décision. En effet, la décision binaire (0, 1) est passée à une décision ternaire (0, 1, 2), ce qui a permis d'augmenter la précision de l'estimation en passant de 4 intervalles de qualité PSNR à sept. Une autre amélioration a consisté en l'introduction d'une procédure automatique pour l'ajustement des facteurs de robustesse lors du marquage multiple.

La nouvelle méthode d'évaluation automatique de la qualité des images JPEG ou JPEG2000 proposée dans ce mémoire est basée sur le marquage multiple dans le domaine de la DCT. Le masque JND est utilisé pour guider l'insertion invisible des marques. Des sous-bandes DCT ont été adéquatement choisies pour l'insertion des marques multiples. Pour les deux méthodes, l'ensemble des expérimentations effectuées nous ont amenés aux résultats suivants :

- Les marques à distribution Gaussienne sont préférables aux marques à distribution uniforme en termes de PSNR de l'image marquée.
- L'estimation de la qualité est correcte lorsque l'image a subi une cascade de codeurs/décodeurs, par contre, la distorsion par bruit Gaussien additif fausse l'estimation lorsque les paramètres du bruit (moyenne et variance) sont un peu grands.
- Le module d'estimation de la qualité réagit bien contre la perte de paquets par modification de bytes du bitstream compressé lorsque le nombre de bytes modifié est limité à une dizaine de bytes.
- Il est recommandé de choisir un domaine de marquage qui est différent du domaine de la compression. Par exemple, si on sait que la compression est JPEG, il est préférable de faire le marquage dans la DWT et non pas dans la DCT.
- Les deux méthodes proposées peuvent être utilisées pour les images JPEG et JPEG2000 et les séquences vidéo codées en Motion-JPEG ou bien en Motion-JPEG2000.

- Une extension possible est le développement de ces méthodes avec une performance temps réel.

Les résultats expérimentaux présentés attestent du succès de ces méthodes d'évaluation de la qualité sans référence à l'image originale. Une des forces majeures de ces méthodes est le fait qu'aucune donnée auxiliaire n'a besoin d'être transmise. De plus, les marques de référence restent les mêmes pour toutes les images et peuvent donc faire partie du module de détection/évaluation. Les résultats sont prometteurs et indiquent le potentiel de ces méthodes à être intégrées dans un système de communication multimédia pour le contrôle automatique de la qualité d'images et vidéo.

Bibliographie

Baaziz, N., D. Zheng, et al. (2011). Image quality assessment based on multiple watermarking approach. IEEE 13th International Workshop on Multimedia Signal Processing (MMSP), Pages: 1-5.

Barni, M., F. Bartolini, et al. (2001). "Improved wavelet-based watermarking through pixel-wise masking." IEEE Transactions on Image Processing **10**(5): 783-791.

Bolot, J. C. and A. Vega-Garcia (1996). Control mechanisms for packet audio in the Internet. Fifteenth Annual Joint Conference of the IEEE Computer Societies. Networking the Next Generation (INFOCOM), Pages: 232-239.

Brandão, T. (2005). "Image communication quality assessment based on watermarking, an overview." Retrieved 2011, from http://www.img.lx.it.pt/~mpq/TFC_05_06.pdf.

Brandao, T. and P. Queluz (2006). Towards Objective Metrics for Blind Assessment of Images Quality. 2006 IEEE International Conference on Image Processing, Pages: 2933 - 2936.

Campisi, P., M. Carli, et al. (2003). "Blind quality assessment system for multimedia communications using tracing watermarking." IEEE Transactions on Signal Processing, Pages: 996-1002.

Carli, M., M. C. Q. Farais, et al. (2005). Quality assessment using data hiding on perceptually important areas. IEEE International Conference on Image Processing (ICIP), Pages: III - 1200-3.

Christopoulos, C., A. Skodras, et al. (2000). "The JPEG2000 still image coding system: an overview." IEEE Transactions on Consumer Electronics, Pages: 1103-1127.

Cox, I. J., M. L. Miller, et al. (2008). Digital Watermarking and Steganography, Morgan Kaufmann.

Gérard, I. (2002). "Une contribution à la conception et à la mise en œuvre d'applications sous contraintes de QoS temporelle, réparties, adaptables." Retrieved 2011, from <http://perso.telecom-paristech.fr/~demeure/PUBLIS/habil.pdf>.

Gonzalez, R. C. and R. E. Woods (2008). Digital Image Processing, Prentice Hall.

Gonzalez, R. C., R. E. Woods, et al. (2004). Digital Image Processing using MATLAB, Pearson/Prentice Hall.

Holliman, M. J. and M. M. Yeung (2002). Watermarking for automatic quality monitoring SPIE Security and Watermarking of Multimedia, Pages: 458-469.

Hwang, J.-N. (2009). Multimedia Networking from Theory to Practice Cambridge University Press.

Kusuma, T. M. and H. J. Zepernick (2003). A reduced-reference perceptual quality metric for in-service image quality assessment. Joint First Workshop on Mobile Future and Symposium on Trends in Communications (SympoTIC), Pages: 71-74.

Nezhadarya, E., Z. J. Wang, et al. (2009). Image quality monitoring using spread spectrum watermarking. 16th IEEE International Conference on Image Processing (ICIP), Pages: 2233-2236.

Nguyen, P. and S. Baudry (2003). "Le tatouage de données audiovisuelles." Retrieved 2011, from www.cairn.info/revue-les-cahiers-du-numerique-2003-3-page-135.htm.

Pickholtz, R., D. Schilling, et al. (1982). "Theory of Spread-Spectrum Communications--A Tutorial." IEEE Transactions on Communications, Pages: 855-884.

Schaefer, G. and M. Stich (2003). "UCID - Uncompressed Colour Image Database." Retrieved 2012, from <http://www-staff.lboro.ac.uk/~cogs/datasets/UCID/ucid.html>.

VIPS (2012). "Database: Images & Video Clips (2)." Retrieved 2012, from http://see.xidian.edu.cn/vips/database_Video.html#image.

Wang, S., D. Zheng, et al. (2005). A digital watermarking and perceptual model based video quality measurement. IEEE Conference on Instrumentation and Measurement Technology (IMTC), Pages: 1729 - 1734.

Wang, S., D. Zheng, et al. (2007). "An image quality evaluation method based on digital watermarking." IEEE Transactions on Circuits and Systems for Video Technology, **Pages:** 98-105.

Watson, A. B. (1993). Visually optimal DCT quantization matrices for individual images. Data Compression Conference (DCC), Pages: 178-187.

Annexe

Bbmpeg

Il s'agit d'un logiciel codeur et décodeur MPEG1 qui permet le codage d'une séquence vidéo en format MPEG1. Les images intra sont à chaque 15 images (Les GOP sont de longueur 15).

Bbmpeg est un programme qui s'exécute dans le système Microsoft Windows et qui permet de convertir des fichiers multimédia en fichiers MPEG (.mpg). Parmi les formats qui peuvent être convertis on cite le format AVI (Audio Video Interleave).

C'est un programme gratuit disponible à l'adresse Web : http://www.01net.com/telecharger/windows/Multimedia/encodeurs_et_decodeurs/fiches/10507.html

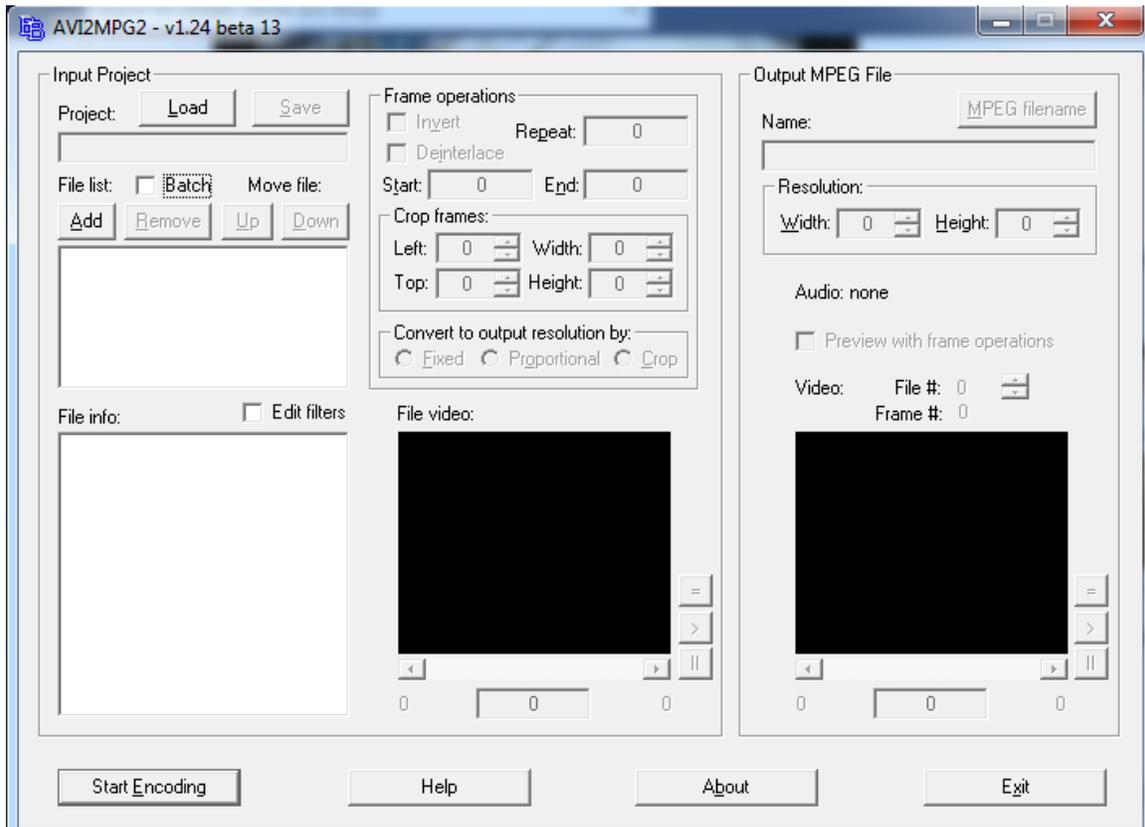


Figure 1 : Le codeur bbmpeg.

Calcul du masque JND

Le masque JND est un masque perceptuel utilisé pour guider l’insertion d’une marque dans les blocs DCT 8×8 d’une image. Le calcul du masque JND est fortement inspiré d’une partie de la distance de Watson telle que c’est développé dans la section 1.2 de ce document. Les étapes de calcul sont présentées dans ce qui suit :

Tableau 1 : La table de sensibilité fréquentielle t définie dans (Cox, Miller, Bloom, Fridrich & Kalker, 2008).

1.40	1.01	1.16	1.66	2.40	3.43	4.79	6.56
1.01	1.45	1.32	1.52	2.00	2.71	3.67	4.93
1.16	1.45	2.24	2.59	2.98	3.64	4.60	5.88
1.66	1.52	2.59	3.77	4.55	5.30	6.28	7.60
2.40	2.00	2.98	4.55	6.15	7.46	8.71	10.17
3.43	2.71	3.64	5.30	7.46	9.62	11.58	13.51
4.79	3.67	4.60	6.28	8.71	11.58	14.50	17.29
6.56	4.93	5.88	7.60	10.17	13.51	17.29	21.15

1- Calcul du masque de luminance

$$t_L(i, j, k) = t(i, j) * \left(\frac{B(0,0,k)}{B_{0,0}} \right)^{a_T} \quad (3)$$

2- Calcul du masque de contraste

$$s(i, j, k) = \max\{t_L(i, j, k), |B(i, j, k)|^w t_L(i, j, k)^{1-w}\} \quad (4)$$

Où w est une constante qui vaut 0.7, $t_L(i, j, k)$ est le masque de luminance, $B(i, j, k)$ est le K^{eme} bloc.

Calcul du masque perceptuel

Le masque de pondération fournit une mesure ponctuelle de la sensibilité de l’œil humain aux perturbations locales de l’image. Son calcul implique la prise en compte de caractéristiques de masquage du HVS tel que la sensibilité à l’intensité lumineuse en niveaux de gris, le masquage fréquentiel et l’influence aux effets du bruit près des contours de l’image. Le calcul se fait en utilisant une transformée en ondelettes de l’image selon les équations suivantes:

$$M_{ld}(i, j) = [P_d(2^{l-1}i, 2^{l-1}j)], \quad d=1,2,3; \quad l=1..L$$

$$P_d(i, j) = \frac{1}{2} B(i, j) T(i, j)^{0.2} E(d, i, j)^{0.2}$$

Où

$$B(i, j) = (1/256) C_L(i, j) + 1$$

$$T(i, j) = \text{Var}\{C_L(i + y, j + x)\}_{x=0,1;y=0,1}$$

$$E(d, i, j) = \frac{1}{4} \Sigma_l \Sigma_x \Sigma_y [C_{ld}(i + y, j + x)^2]_{x=0,1;y=0,1;l=1\dots L}$$

Les C_{ld} sont les sous-bandes d'une transformée en ondelettes de l'image où d indique la direction (horizontale, verticale et diagonale) et l indique le niveau de résolution. C_L est l'image d'approximation (bande de basses fréquences) au niveau L . Le facteur B est proportionnel aux valeurs de luminance de l'image d'approximation. Le facteur T capture l'activité de la texture au sein d'un voisinage dans la bande des basses fréquences alors que le facteur E mesure la présence de contours orientés dans les sous-bandes directionnelles pour $d=1, 2$ ou 3 . Les facteurs B , T et E sont alors fusionnés (par multiplication) dans un masque perceptuel P directionnel. On obtient alors un masque perceptuel M_{ld} pour chaque sous-bande C_{ld} de la décomposition en ondelettes de l'image en utilisant les sous-échantillonnages requis. Ce masque a été initialement proposé dans ([Barni, Bartolini et al. 2001](#)). Une amélioration a été proposée dans ([Baaziz, Zheng et al. 2011](#)) en utilisant entre autres la transformée en ondelettes stationnaire SWT.