



**HAL**  
open science

# Masque psychovisuel à base d'ondelettes pour le Watermarking

Meina Amar

► **To cite this version:**

Meina Amar. Masque psychovisuel à base d'ondelettes pour le Watermarking. Autre. Université d'Orléans; Université Ibn Zohr (Agadir, Maroc). Faculté des sciences, 2018. Français. NNT : 2018ORLE2022 . tel-02096070

**HAL Id: tel-02096070**

**<https://theses.hal.science/tel-02096070>**

Submitted on 11 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



جامعة ابن زهر  
+08/86144 290 3800  
UNIVERSITÉ IBN ZOHR



## Université d'Orléans

École doctorale Mathématiques,  
Informatique, Physique Théorique et  
Ingénierie des Systèmes (MIPTIS)  
Laboratoire PRISME  
Pôle IRAuS, Axe Image et Vision

## Université Ibn Zohr

Faculté des Sciences d'Agadir  
Centre des Etudes Doctorales IBN ZOHR  
Formation doctorale Mathématiques,  
Informatique et Applications  
Laboratoire IRF-SIC

THÈSE EN COTUTELLE INTERNATIONALE présentée par

# AMAR Meina

Soutenue le 21 Avril 2018

Pour obtenir le grade de  
**Docteur de l'université d'Orléans**  
**Docteur de l'Université Ibn Zohr**

Discipline : **Informatique**  
Spécialité : **Traitement d'images numériques**

## Masque psychovisuel à base d'ondelettes pour le Watermarking

### THÈSE dirigée par :

DOUZI Hassan  
HARBA Rachid

Professeur, Université Ibn Zohr (Maroc)  
Professeur, Université d'Orléans (France)

### RAPPORTEURS :

SADGAL Mohamed  
JACQUET Gérard

Professeur, Université C. Ayyad Marrakech (Maroc)  
Professeur, Université de St Etienne (France)

### Président du JURY :

MAMMASS Driss

Professeur, École Supérieure de Technologie  
d'Agadir (Maroc)



# Remerciement

Je tiens tout d'abord à exprimer ma profonde gratitude à mes directeurs de thèse, Monsieur Hassan Douzi professeur d'enseignement supérieur à l'université Ibn Zohr Agadir, et Monsieur Rachid Harba professeur des universités à l'université d'Orléans, pour m'avoir proposé ce sujet riche d'informations. J'aimerais les remercier pour leur disponibilité envers moi ainsi que pour leurs conseils judicieux qui m'ont permis de surmonter les difficultés et de mener à bien ce travail.

Je tiens à exprimer mes profonds remerciements au professeur Driss Mammass professeur d'enseignement supérieur à l'université Ibn Zohr Agadir, qui m'a fait l'honneur de présider le jury. Je tiens à remercier profondément Monsieur Mohamed Sadgal professeur d'enseignement supérieur à l'Université Cadi Ayyad de Marrakech pour l'attention qu'il a manifestée à l'égard de cette recherche en s'engageant à être rapporteur. Je souhaiterais aussi adresser mes remerciements à Monsieur Gérard Jacquet professeur des universités à l'université Jean Monnet de Saint-Étienne, pour avoir rapporté ma thèse ainsi que pour le temps qu'il a consacré pour la lecture de ce manuscrit.

Je tiens à remercier sincèrement Monsieur Frédéric Ros HDR, chercheur associé à l'université d'Orléans, qui m'a apporté une aide précieuse ainsi que le temps qu'il m'a accordé pour m'écouter et discuter avec moi.

Je voudrais remercier également Monsieur Mohamed El Hajji professeur assistant au centre régional des métiers de l'éducation et de la formation et le docteur Rabia Riad, pour l'aide qu'ils m'ont apportée.

Ces remerciements seraient incomplets si je n'en adressais pas à tous les collègues du laboratoire IRF-SIC du laboratoire PRISME pour leur soutien moral ainsi que le cadre agréable dans lequel ils m'ont permis de préparer ce travail dans une ambiance chaleureuse et conviviale.

J'exprime ici ma reconnaissance à ma famille qui me soutient depuis toujours pour que je puisse continuer mes études. Je tiens enfin à remercier tout particulièrement quelqu'un que j'ai rencontré bien avant de débiter ce travail de thèse, pour qui mon admiration n'a jamais cessé depuis malgré la distance, merci Sidi.

# Table des matières

Remerciement.....	i
Table des matières .....	ii
Table des figures.....	v
Liste des tableaux .....	ix
Glossaire .....	x
Introduction générale.....	1
Chapitre 1 .....	5
1. Etat de l’art sur le tatouage numérique.....	5
1.1. Introduction.....	5
1.2. Aspect général du tatouage numérique .....	5
1.2.1 Principe du tatouage numérique .....	5
1.2.2 Phase d’insertion .....	6
1.2.3 Phase d’extraction .....	8
1.2.3 Contraintes du tatouage numérique.....	9
1.3 Domaine d’insertion .....	11
1.3.1 Domaine spatial.....	12
1.3.2 Domaine fréquentiel.....	12
1.4 Les attaques.....	24
1.5 Outils d’évaluation des méthodes de tatouage numérique.....	27
1.5.1 Evaluation de la qualité visuelle.....	28
1.5.2 Evaluation de la robustesse du tatouage.....	28
1.6 Types d’applications .....	28
1.7 Les récentes applications dans le domaine de tatouage .....	30
1.8 Conclusion .....	33
Chapitre 2 .....	34
2. Le système visuel humain – Description et modélisation pour le tatouage numérique.....	34
2.1. Introduction.....	34

2.2. Le système visuel Humain (SVH) .....	35
2.2.1 Structure du SVH .....	35
2.2.2 Perception de la couleur .....	37
2.2.3 Perception de la luminance .....	38
2.2.4 Le contraste .....	41
2.3 Les masques psychovisuels .....	43
2.3.1 Le masquage.....	43
2.3.2 Evaluation de la qualité visuelle.....	48
2.4 Conclusion .....	57
Chapitre 3 .....	58
3. Contribution dans le tatouage psycho-visuel : JND par discrimination de texture à base d'ondelettes de Faber-Schauder.....	58
3.1. Introduction.....	58
3.2 La discrimination de texture fondée sur une représentation d'ondelettes FSDWT à échelles mixées .....	60
3.3. Le masquage psychovisuels.....	63
3.3.1 Masque de luminance de Chou <i>et</i> Li.....	63
3.3.2 Masque de contour de Chou <i>et</i> Li .....	65
3.3.3 Le masque de texture de Qi <i>et al.</i> .....	67
3.3.4. Masque psychovisuel basé sur la transformé de Faber-shauder à échelles mixées .....	68
3.4 Schéma d'insertion du tatouage .....	71
3.5 Test sur des images de synthèses .....	72
3.5.1 Base expérimentale d'image .....	73
3.5.2 Evaluation de l'imperceptibilité .....	76
3.6 Evaluation et comparaison sur images réelles .....	78
3.6.1 Évaluation des performances .....	81
3.4 Conclusion .....	96
Chapitre 4 .....	98
4. Nouveau masque de luminance.....	98

4.1 Introduction.....	98
4.2 Test subjectif.....	98
4.3 Modélisation de la courbe expérimental .....	100
Conclusion .....	102
Conclusion générale .....	103
Perspectives .....	105
Bibliographie .....	106
Bibliographie de l’auteur .....	115

## Table des figures

Figure 1-1 : Schéma général du tatouage numérique .....	6
Figure 1-2: Schéma additif .....	7
Figure 1-3: Schéma substitutif.....	8
Figure 1-4: Tatouage robuste.....	10
Figure 1-5: Tatouage fragile .....	10
Figure 1-6: Compromis entre la robustesse, la visibilité et la capacité d'insertion .....	11
Figure 1-7: Image Barbara et son spectre de Fourier .....	13
Figure 1-8: Exemple d'insertion dans les fréquences moyennes de DCT.....	15
Figure 1-9: Schéma du tatouage dans le domaine d'invariance Fourier-Mellin.....	17
Figure 1-10: Décomposition et reconstruction par la transformée en ondelettes (un seul niveau) [2].....	19
Figure 1-11: Décomposition en ondelettes sur trois de niveaux de résolution.....	19
Figure 1-12: la Transformation par lifting scheme (signal 1D).....	20
Figure 1-13: La représentation à échelles séparée d'une décomposition successive par la transformée en ondelettes discrète (jusqu'à trois niveaux). .....	22
Figure 1-14: Décomposition à échelles séparée par la transformée en ondelettes discrète de l'image Lena.....	22
Figure 1-15: La représentation à échelles mixées de FSDWT de l'image Lena .....	23
Figure 1-16: La classification des attaques que peut subir un document tatoué selon [2]. .	25
Figure 2-1 : Coupe légende de l'œil [51] .....	35
Figure 2-2: La rétine [52] .....	36
Figure 2-3 : Forme des bâtonnets et cônes .....	37
Figure 2-4: Estimations des spectres d'absorption des trois cônes [54] .....	38
Figure 2-5: Comportement des photorécepteurs face à l'intensité lumineuse.....	39
Figure 2-6: Expérience de Weber-Fechner.....	40
Figure 2-7: Le tracé de la loi de variation du seuil de perception .....	40
Figure 2-7: mire sinusoïdale de Cambell et Robson .....	42
Figure 2-8: Illustration du CSF en fonction de la fréquence spatiale et la luminance du fond [63] .....	43
Figure 2-9: L'image Lena (à gauche) et une copie de celle-ci avec un bruit uniforme ajouté avec une variance uniforme.....	44



Figure 2-10: Méthode à simple stimulus.....	52
Figure 2-11: Méthode à double stimulus.....	52
Figure 3-1 : (a) Représentation à échelle séparée de l'image Barbara, (b) : La représentation à échelles mixées de FSDWT de l'image Barbara .....	61
Figure 3-2: (a) exemple d'images synthétiques composées de régions : uniformes, de contours, texturés. La couleur bleue correspond à des régions uniformes avec un nombre de coefficients dominants variant de 0 à 2. La couleur verte correspond aux régions de contours avec un nombre de coefficients dominants variant de 3 à 4. Enfin, la couleur rouge correspond aux régions texturées avec un nombre de coefficients dominants variant de 5 à 9.....	62
Figure 3-3: Image correspondant au coefficient dominant pour Lena, Mandrill et Barbara. Où la couleur bleu correspond aux zones uniformes, la couleur rouge pour les zones texturées et la couleur verte pour les contours.....	63
Figure 3-4: Seuils de visibilité dus à la luminance d'arrière-plan [66].....	64
Figure 3-5: Implémentation du masquage de la luminance sur l'image Barbara, les parties blanches du masquage qui représentent les zones claires et obscures correspondent aux régions avec une forte force d'insertion du tatouage.....	65
Figure 3-6: (a) Stimuli pour déterminer les seuils de visibilité dus au changement de luminance, (b) seuils de visibilité de différence à différentes luminances de l'arrière-plan [66] .....	66
Figure 3-7: Implémentation du masquage de contours l'image Barbara.....	67
Figure 3-8 : Implémentation du masquage de texture sur l'image Barbara .....	68
Figure 3-9: Masquage final qui combine la transformée en ondelettes à échelle mixte, pour la discrimination des zones uniformes de contour ou de texture, avec les JND de Chou et de Qi.....	69
Figure 3-10: Algorithme du JNDEM.....	70
Figure 3-11: Exemple d'application de notre JNDEM.....	70
Figure 3-12: Exemple du processus d'insertion de tatouage par Étalement de spectre pour un message de 4 bit.....	71
Figure 3-13: Comparaison entre deux marques produites par étalement de spectre : une marque de tatouage spatiale $w_i$ et une marque de tatouage couplé avec le masque psychovisuel JNDEM. On voit bien que la deuxième marque s'adapte aux zones d'insertion .....	72
Figure 3-14: Exemples d'images de la base expérimentale ALOT.....	73

Figure 3-15: Exemples d'images de la base expérimentale MBT.....	73
Figure 3-16: Discriminateur de texture de quelques images des deux bases ALOT et MBT. .....	74
Figure 3-17: Image originale et images tatouée sans utilisation de masque et image tatouée avec masque JNDEM .....	75
Figure 3-18: Courbe résultat du WPSNR et SSIM de 112 images de la base MBT .....	77
Figure 3-19: Courbe résultat du WPSNR et SSIM de 112 images de la base ALOT .....	77
Figure 3-20: Illustrations des différents masques de référence JND et que notre JNDEM sur l'image Barbara : (a) JND de Chou & Li [66], (b) JND de Qi & al [69], (c) JND de Barni & al [77], (d) JND de Watson [70], (e) notre JNDEM .....	79
Figure 3-23: Courbe résultat du WPSNR de 50 images tatouées .....	82
Figure 3-24: Courbe résultat de SSMI de 50 images tatouées .....	82
Figure 3-25: Résultat du test subjectif. (a) JNDEM comparé à Chou & al, (b) JNDEM comparé à Qi & al, (c) JNDEM comparé à Barni & al, (d) JNDEM comparé à Watson....	83
Figure 3-26: Schéma d'un processus de test de tatouage utilisant des modèles perceptuel	86
Figure 3-27: Valeurs de détection du BER pour les images tatouées après compression JPEG : (a) 80%, (b) 10%. .....	86
Figure 3-28: Exemple sur l'image Barbara de l'attaque Jpeg avec un QF=10% .....	87
Figure 3-29: Valeurs de détection du BER pour les images tatouées après attaque AWGN : (a) variance 0.1, (b) de variance 0.5. ....	88
Figure 3-30: Exemple sur l'image Barbara de l'attaque AWGN avec une variance $\sigma = 0.188$	88
Figure 3-31: Valeurs de détection du BER pour les images tatouées après rotation : (a) de 30°, (b) de 45° .....	89
Figure 3-32: Exemple sur l'image Barbara de l'attaque rotation avec un angle de rotation de 45° .....	89
Figure 3-33: Valeurs de détection du BER pour les images tatouées après cropping : (a) de 3%, (b) de 5%. .....	90
Figure 3-34: Exemple sur l'image Barbara de l'attaque cropping en gardant 3% de l'image .....	91
Figure 3-35: Valeurs du BER pour les images tatouées après filtre moyenneur : (a) 3x3, (b) 7x7. ....	92
Figure 3-36: Exemple sur l'image Barbara de l'attaque filtre moyenneur d'une fenêtre 3x3. .....	92

---

Figure 3-37: Valeurs du BER pour les images tatouées après filtre médian : (a) 3x3, (b) 7x7. .....	93
Figure 3-38: Exemple sur l'image Barbara de l'attaque filtre médian d'une fenêtre 3x3... 93	93
Figure 3-39: Valeurs de détection BER pour les images tatouées après : (a) égalisation d'histogramme et (b) ajustement de contraste. ....	94
Figure 3-40: Exemple sur l'image Barbara de l'attaque égalisation d'histogramme. ....	94
Figure 3-41: Exemple sur l'image Barbara de l'attaque ajustement de contraste.....	95
Figure 4-1: (a) L'expérience dans [66], (b) notre expérience.....	98
Figure 4-2 : Résultat des tests subjectifs la courbe en bleu la reproduction de l'expérience dans [66], la courbe en rouge notre expérience. ....	99
Figure 4-3: La courbe noire représente le polynôme de degré n avec n variant de 5 à 8..	100
Figure 4-4: (a) Seuils de visibilité expérimentale dus à la luminance de fond. (b) Seuils de visibilité dus à la luminance de l'article [66] .....	101
Figure 4-5: Exemple de comparaison entre le masque de luminance de l'article [66] et le masque de luminance que nous avons développé. ....	102

## Liste des tableaux

Tableau 2-1 : Échelles de qualité et de dégradation de l'UIT-R.....	52
Tableau 3-1: Valeurs du MSSIM pour images tatouées de la base ALOT .....	75
Tableau 3-2: Valeurs du MSSIM pour images tatouées de la base MBT .....	76
.Tableau 3-3: Valeurs de WPSNR pour 50 images tatouées .....	84
Tableau 3-4: Valeurs de SSIM pour 50 images tatouées .....	84
Tableau 3-5: Valeurs MOS pour 50 images tatouées .....	84
Tableau 3-6: Valeurs BER de 50 images après l'attaque Jpeg.....	87
Tableau 3-7: Valeurs BER de 50 images après l'attaque AWGN.....	88
Tableau 3-8: Valeurs BER de 50 images après l'attaque rotation.....	90
Tableau 3-9: Valeurs BER de 50 images après l'attaque cropping.....	91
Tableau 3-10: Valeurs BER de 50 images après l'attaque filtre moyenneur .....	92
Tableau 3-11: Valeurs BER de 50 images après l'attaque filtre médian.....	93
Tableau 3-12: Valeurs BER de 50 images après les attaques EH et AC.....	95

---

# Glossaire

- AWGN** : Additive white Gaussian noise (bruit blanc gaussien additif)
- BER** : bit error rate
- CVH** : Caractéristiques visuelles humaines (CVH)
- CWT** : Complexe wavelet transformation (CWT)
- DCT** : Discrete cosine transform (Transformée en cosinus discrète)
- DFT** : Discrete Fourier transform (Transformation de Fourier discrète)
- DSCQS** : Double Stimuli Continuous Quality Scale
- DWT** : Discrete wavelet transforms (Transformée en ondelettes discrète)
- EQM** : Erreur quadratique moyenne
- FFT** : Fast Fourier Transform (Transformation de Fourier rapide)
- FMT** : Fourier Mellin transform (Transformation de Fourier-Mellin)
- FSDWT** : Faber-Schauder discret wavelet transformation
- JPEG** : Joint Photographic Experts Group
- JND** : Just Noticeable Distortion
- JNDEM** : JND à base d'Echelle Mixée
- KLT** : Karhunen-Loève transformation (KLT)
- MOS** : Mean Opinion Score
- NVF** : Noise Visibility Function
- PSNR** : Peak Signal to Noise Ratio
- QIM** : Quantization index modulation
- RVB** : Rouge Vert Bleu
- SSCQS** : Single Stimulus Continuous Quality Scale
- SSIM** : Structural Similarity
- SVD** : Singumlar value decomposition
- SVH** : Système visuel humain
- UQI** : Universal Quality Index
- WPSNR** : Weighted Peak Signal-to-Noise Ratio

## **Introduction générale**

L'accès de plus en plus facile au contenu multimédia via les moyens modernes de communication ainsi que la disponibilité de nombreux outils d'édition puissants offrent de grandes opportunités pour diffuser et commercialiser des signaux, des images et des vidéos. Il a également rendu plus facile que jamais la violation du droit d'auteur. Sans sécurité, les documents numériques peuvent être facilement dupliqués et distribués sans le consentement du propriétaire. On peut en effet les stocker, les copier, les modifier et enfin les diffuser illégalement sans qu'ils perdent de leur qualité. Il est donc primordial de trouver des solutions pour empêcher ou tout du moins minimiser le piratage des œuvres multimédia. Ces dispositifs doivent s'appliquer à tous types de supports numériques.

Depuis la seconde moitié des années 1990, la dissimulation des données numériques « digital data hiding » a reçu une attention particulière de la part de la communauté des technologies de l'information et de la communication car elle permet d'apporter des solutions de sécurité de diffusion des œuvres. Le tatouage numérique appelé en anglais digital watermarking est apparu comme nouvel axe de recherche qui fait partie de ce domaine scientifique très récent et ne cesse de prendre de l'importance au sein de la communauté scientifique. Le tatouage numérique apparaît comme un moyen de sécurité supplémentaire, à assurer un accès autorisé, à faciliter l'authentification du contenu ou empêcher la reproduction illégale.

L'idée de base du tatouage numérique consiste à cacher dans un document multimédia une information de copyright, ou d'autres messages de vérification, invisible ou inaudible suivant la nature du document permettant d'assurer un service de sécurité (e.g. droits d'auteur, intégrité, traçabilité, non répudiation, etc.). Ce processus s'apparente à un rajout de bruit, appelé aussi marque. Le tatouage est théoriquement indépendant du format de fichier. Il est aussi censé être détecté ou extrait même si le document a subi des modifications (ou attaques) comme le décryptage, le déchiffrement, la compression, la conversion numérique-analogique ou les changements de format de fichier.

Le tatouage numérique d'images a trois objectifs essentiels :

1. L'imperceptibilité: la marque de tatouage insérée doit être imperceptible à l'œil. Contrairement au marquage visible, elle ne nuit pas à l'esthétique d'une image.

2. La Robustesse : la marque de tatouage doit résister aux attaques lors de la manipulation et l'usage courant du document ou lors d'attaques malveillantes.
3. La capacité : Le tatouage contient suffisamment d'espace pour englober l'information nécessaire.

Bien que de nombreux algorithmes de tatouage numérique aient été proposés ces dernières années, un certain nombre de problèmes subsistent. Le problème principal est de trouver un compromis entre la robustesse et l'imperceptibilité. En effet, pour améliorer la robustesse, il faut généralement augmenter la force d'insertion du tatouage et donc avoir une perte au niveau de l'imperceptibilité et vice versa. Concevoir des algorithmes pour un tel problème d'optimisation est toujours un défi pour la communauté du tatouage.

Une solution prometteuse à ce problème difficile est de prendre en compte les propriétés du Système Visuel Humain (SVH) pour développer des masques psychovisuels. L'idée de base est de définir, localiser et exploiter des régions d'intérêt, où le l'œil humain est moins sensible à l'insertion du tatouage. En effet, il est permis d'augmenter la force d'insertion du tatouage dans de telles régions dans le but d'augmenter la robustesse et ce, sans nuire à la qualité visuelle de l'image. Il faut donc localiser ces régions en tenant compte des limites du SVH.

Barni a résumé des observations et des expériences subjectives pour les images à niveau de gris. Dans son travail il présente les imperfections et limitations du SVH et explique comment les exploiter pour renforcer la force d'insertion de la marque. Barni a défini quatre règles fondamentales :

- Règle 1 : les perturbations (ou la marque) sont beaucoup moins visibles sur les régions fortement texturées que sur les zones uniformes.
- Règle 2 : les contours sont plus sensibles à l'ajout de bruit que les régions très texturées mais moins sensibles que les zones uniformes.
- Règle 3 : les perturbations sont moins visibles sur les régions sombres et très claires.
- Règle 4 : l'œil est plus sensible aux perturbations sur les hautes fréquences verticales ou horizontales que celles à  $\pm 45^\circ$ .

Un des moyens pour mesurer cette sensibilité de l'œil humain est le calcul du seuil de détection visuel. Ce seuil représente la plus petite distorsion que l'œil humain peut déceler, il est appelé JND (Just-Noticeable-Difference) et calculé en exploitant certaines limitations

du SVH, telles que la sensibilité au contraste, l'adaptation de luminance et le masquage de contraste. Le modèle JND ainsi calculé est utilisé pour contrôler l'intensité du tatouage, tout en assurant la performance de tatouage optimale. De nombreuses méthodes de tatouage psychovisuel ont été proposées. Elles sont toutes basées sur une association de règles loin d'être bien justifiée soit elles n'intègrent pas l'intégralité des règles de Barni.

## Contribution de la thèse

Dans cette thèse nous présentons la conception et le développement d'un nouveau schéma de tatouage psychovisuel basé sur un modèle JND. Ce modèle est développé en utilisant la transformée d'ondelette à échelles mixées (JNDEM) qui associe les avantages d'une transformée en ondelette à une localisation spatiale précise. Cela permet de discriminer trois types de régions dans les images candidates pour le tatouage : les régions uniformes, les contours et les régions texturées. On peut alors moduler la force d'insertion du tatouage en fonction du type de la région. Dans chaque région nous allons sélectionner les modèles JND de la littérature les plus performants tout en tenant compte des 4 règles de Barni. Cette méthode n'est pas fondée sur des règles empiriques mais sur une approche logique du problème.

Une nouvelle méthode de tatouage d'image perceptive est ainsi proposée. Elle combine les avantages du domaine d'ondelettes avec l'utilisation d'une représentation à échelles mixées pour la séparation des régions et l'insertion modulée du tatouage dans le domaine spatial avec des méthodes performantes existantes.

Le modèle JNDEM proposé sera testé en tatouage numérique sur une base de 50 images naturelles et comparé à quatre autres modèles perceptifs de la littérature. Nous allons évaluer la qualité visuelle des images tatouées via des métriques objectives en plus d'un test subjectif. Nous allons aussi évaluer la robustesse de la méthode de tatouage proposé en appliquant diverses attaques numériques.

## Organisation de la thèse

Cette thèse est organisée en quatre chapitres.

Dans le premier chapitre nous présenterons le contexte général, les aspects principaux et les terminologies liées au tatouage invisible des images numériques. Ensuite nous détaillerons les méthodes et les domaines d'insertion, ainsi que les différents types d'attaques possibles.



On distinguera les schémas de tatouage additif (le plus souvent par étalement de spectre), le schéma substitutif ou certaines composantes de l'image sont remplacées par la signature. Ces deux classes de schémas peuvent être appliqués dans le domaine spatial de l'image ou bien dans le domaine transformé comme la Transformée de Fourier Discrète ou encore les transformations en ondelettes. Enfin nous dresserons un panorama des problématiques associées aux différents types d'attaques qui existent dans la littérature.

Le deuxième chapitre décrit les différentes composantes du SVH, ainsi que les composants fondamentaux qui sont essentiels au développement de nouveaux modèles psychovisuels. Ensuite une section a été consacrée à certains modèles JND importants du SVH qui sont déjà caractérisés et exploités avec succès dans le domaine du traitement d'image. Enfin on a introduit l'évaluation subjective de la qualité des images, en présentant les différents protocoles ainsi que les principaux facteurs influençant l'appréciation humaine.

Le troisième chapitre, le cœur de cette thèse, est consacré à la description de notre modèle JND basée sur un discriminateur de texture à base d'ondelette de Faber-Schauer à échelles mixées. Ce discriminateur permet de séparer les différentes zones de l'image à savoir, uniforme, texture et contour. Ceci permet d'appliquer des modèles JND spécifiques, pour chaque type de région, considérés comme des références dans la communauté du tatouage numérique. Nous avons aussi pris en compte la règle 4 de Barni concernant la sensibilité angulaire du SVH à l'activité locale. Nous avons testé à la fois la qualité visuelle et la robustesse aux attaques de la méthode de tatouages proposés et nous l'avons comparée à quatre autres méthodes de pointe avec différents domaines d'insertion.

Le chapitre 4 propose une perspective liée à un nouveau test psychovisuel adapté au tatouage et non pas déduit d'expériences liées à la compression d'images.

Enfin dans la conclusion, nous récapitulerons les principales contributions de ce travail de thèse avant d'exposer les perspectives envisagées.

Cette thèse s'est effectuée dans le cadre d'une cotutelle entre l'université d'Orléans et l'université d'Agadir. Elle a bénéficié d'un support du projet PPR2-CNRST: Développement de prototypes pour des systèmes d'authentification basés sur la biométrie du visage ainsi que support de la Région Centre Val de Loire pour son soutien partiel à la mobilité internationale.

# **Chapitre 1**

## **1. Etat de l'art sur le tatouage numérique**

### **1.1. Introduction**

La révolution numérique entraîne une circulation accrue des documents multimédia (images, vidéos, textes, son, etc.). L'ampleur de ce phénomène est telle que des questions fondamentales se posent désormais quant à la protection et au contrôle des données échangées. En effet, de par leur nature numérique les documents multimédia peuvent être dupliqués, modifiés, transformés et diffusés très facilement. Dans ce contexte, le tatouage numérique (Digital Watermarking en anglais) apparaît comme une alternative pouvant s'avérer efficace et complémentaire pour aider à établir une sécurité supplémentaire, à assurer un accès autorisé, à faciliter l'authentification du contenu ou empêcher la reproduction illégale.

Dans ce chapitre nous décrirons, dans un premier temps, le contexte général, les aspects principaux et les terminologies liés au tatouage invisible des images numériques. Ensuite nous détaillerons les méthodes et les domaines d'insertion, ainsi que les différents types d'attaques possible. On distinguera les schémas de tatouage additif (le plus souvent par étalement de spectre), le schéma substitutif ou certaines composantes de l'image sont remplacées par la signature. Ces deux classes de schémas peuvent être appliqués dans le domaine spatial de l'image ou bien dans le domaine transformé comme la Transformée de Fourier Discrète ou encore les transformations en ondelettes. Enfin nous dresserons un panorama des problématiques associées aux différents types d'attaques qui existent dans la littérature. La marque insérée doit résister aux divers traitements que subit l'image ainsi qu'aux traitements malveillants, dont le but est d'empêcher la détection de la marque.

### **1.2. Aspect général du tatouage numérique**

#### **1.2.1 Principe du tatouage numérique**

Le principe du tatouage numérique consiste à insérer une marque invisible appelée aussi signature, message, ou tatouage, dans une image, une vidéo, un signal audio ou d'autres documents numériques. Le tatouage numérique est composé de trois phases principales. La

phase d'insertion, la transmission (ou attaque) et la phase de détection (ou extraction). Dans la phase d'insertion une marque est rajoutée au document numérique, cette marque peut être une séquence aléatoire, un logo binaire ou bien un message en fonction de l'application visée. Donc, et cette remarque sera importante par la suite, un tatouage peut être considéré comme un ajout d'un bruit dans l'image. Plus ce bruit est fort, plus la marque sera visible : on parle alors de la force du marquage. Dans les débuts du tatouage, cette force était fixée pour arriver à un PSNR d'environ 40dB. Ensuite il faut détecter (ou extraire) la présence de la marque en la comparant avec la marque originale. Au cours de la transmission, le document peut subir des modifications qui peuvent l'altérer. Pour faire un bon tatouage numérique il faut tenir compte des contraintes (la robustesse, l'imperceptibilité et la capacité d'insertion) et du contexte (schéma d'insertion, champ d'insertion, type d'attaque) (figure 1-1).

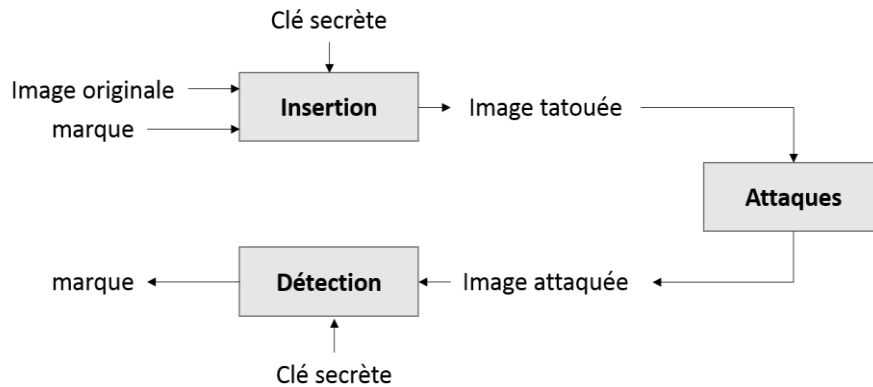


Figure 1-1 : Schéma général du tatouage numérique

Dans la suite nous présentons les propriétés principales du tatouage numérique.

### 1.2.2 Phase d'insertion

L'ajout de la marque de tatouage se décompose en deux étapes :

- le codage de la marque  $W$  : elle est générée à partir d'un message  $m$  qui constitue les informations qu'on souhaite véhiculés et d'une clé  $k$  de cryptage.
- la deuxième étape est l'insertion de la marque. Il existe en général deux manières pour insérer une marque  $W$  dans une image  $I_0$  : un schéma d'insertion additif ou un schéma d'insertion substitutif.

Nous pouvons modéliser cette phase d'insertion par la fonction suivante :

$$I_w = \varepsilon(I_o, w, k) \quad (1-1)$$

Où  $I_w$  est l'image tatouée ;  $I_o$  l'image originale.

$W$  la marque ;  $k$  la clé de cryptage.

### 1.2.2.1 Schéma additif

Les méthodes additives sont les plus nombreuses et les plus utilisées. Elles consistent principalement à ajouter la marque aux composantes du document en utilisant l'une des équations suivantes [1].

$$\begin{cases} I_w = I_o + W \\ I_w = I_o \times (1 + W) \\ I_w = I_o \times e^W \end{cases} \quad (1-2)$$

La figure 1-2 montre le schéma complet d'une méthode additive. La première étape est la génération d'une marque  $W$  qui est composée d'une séquence pseudo-aléatoire modulant souvent un message  $M$ . La seconde étape est l'addition de la marque dans les valeurs de l'image, suivant l'une des équations de (Eq1-2). Cette incrustation peut se faire directement sur l'image (dans le domaine spatial) ou dans un domaine de transformé  $T$  approprié.  $T$  possède des propriétés d'invariance ou de dispersion facilitant l'insertion et rendant  $W$  robuste et invisible.

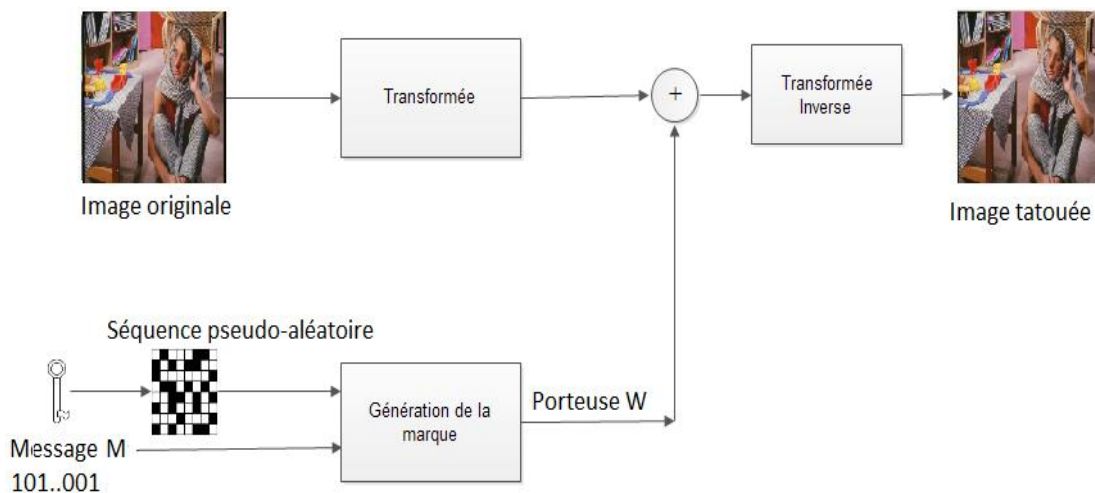


Figure 1-2: Schéma additif

Remarque : à l'œil humain, l'image tatouée est identique à l'image originale, alors qu'en réalité elles sont différentes.

### 1.2.2.2 Schéma substitutif

Contrairement aux méthodes additives, le tatouage substitutif se propose de substituer un élément de l'image originale par un autre tatoué (voir figure 1-3). La marque à insérer est obtenue en appliquant une contrainte (une mesure de similarité, une propriété géométrique, un critère d'ordre ...) afin de le faire correspondre au message que l'on souhaite transmettre. On distingue dans ces méthodes le tatouage par quantification appelé aussi tatouage substitutif avec dictionnaire et le tatouage par contrainte qui consiste à imposer un ensemble de contraintes aux données marquées [2].

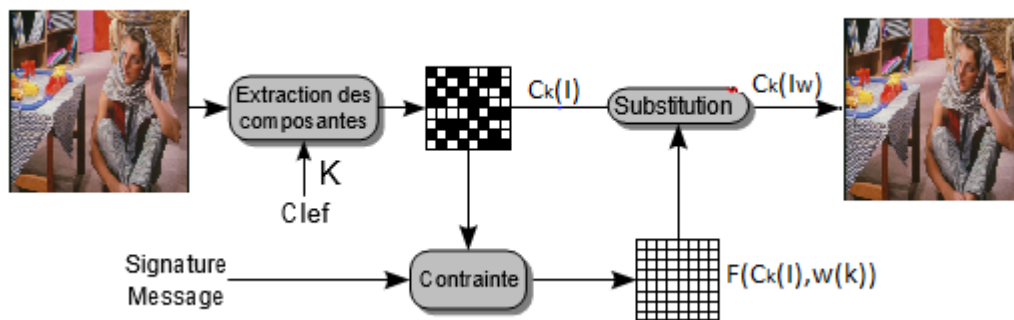


Figure 1-3: Schéma substitutif

### 1.2.3 Phase d'extraction

Au niveau de la détection nous aurons soit la marque extraite  $\hat{W}$  ou bien un message indiquant si une marque est détectée : 1 si il y a une marque ou 0 s'il y en a pas.

Il existe trois types de détection :

- **Détection non aveugle :** Dans ce type de détection nous disposons de l'image originale, l'image tatouée ainsi que la marque insérée. A la sortie du détecteur nous pouvons avoir soit la marque détectée soit une mesure de sa présence
  - Si le détecteur extrait la marque, il est dit de Type I, on a alors :

$$(I_0; I_w; K) \rightarrow \hat{W}$$

- Si le détecteur est une mesure de présence de la marque (sa sortie sera 1 si la marque est détectée 0 sinon), la détection est alors une application de type II avec :

$$(W; I_0; I_w; K) \rightarrow (0 \text{ ou } 1)$$

Ce type de détection est bien évidemment incompatible avec des applications visant à vérifier l'intégrité ou l'authenticité de l'image, ou à assurer la vérification en temps réel du copyright (problème lié au temps d'accès à la base de données contenant les informations originales)

- **Détection semi-aveugle** : Dans ce type de détection nous ne disposons pas de l'image originale et nous aurons un message indiquant oui ou non la présence de la marque.

$$(W; I_w; K) \rightarrow (0 \text{ ou } 1)$$

- **Détection aveugle** : Une détection aveugle (appelée parfois publique) extrait la marque insérée sans avoir besoin de l'image originale et de la marque d'insertion.

$$(I_w; K) \rightarrow \hat{W}$$

### 1.2.3 Contraintes du tatouage numérique

Le plus important dans un tatouage numérique c'est de garantir un bon compromis entre l'invisibilité, la robustesse de la marque et la capacité d'insertion.

#### 1.2.3.1 Imperceptibilité

Appelée aussi la distorsion d'insertion [3]. Il s'agit de faire en sorte que l'impact visuel résultant à l'insertion de la marque dans un document soit le plus faible possible afin que l'image tatouée soit visuellement équivalente (fidèle) à l'image originale pour un observateur humain. De plus, la marque insérée ne devrait pas affecter la qualité du document. En ce qui concerne les images tatouées, une évaluation de la qualité visuelle nécessite une analyse du SVH. Plusieurs mesures objectives et subjectives ont été proposées dans la littérature pour mesurer la qualité visuelle d'un document tatoué. Une description détaillée de ses mesures sera traité dans le chapitre 2.

#### 1.2.3.2 Robustesse

La robustesse représente la capacité du tatouage à résister aux dégradations du document tatoué. La marque doit être difficile à supprimer. Si un utilisateur malveillant tente de supprimer la marque (en termes de dégradations visibles inacceptables et/ou d'utilisation

commerciale rendue impossible) ou bien si l'image subit des dégradations non intentionnelles (on trouve dans cette catégorie la compression JPEG, les conversions de format en général et les changements de résolution entre autres) ceci doit causer une forte dégradation au niveau de la qualité de l'image. Il existe trois types de tatouages selon le critère de robustesse

- **Tatouage robuste** : Un tatouage est dit robuste s'il tente de préserver la marque contre tout type d'attaque, malgré de fortes modifications, le message doit pouvoir être lu correctement. Toutefois si la marque est supprimée ou endommagé ceci conduit également à la distorsion de l'image (voir figure 1-4).

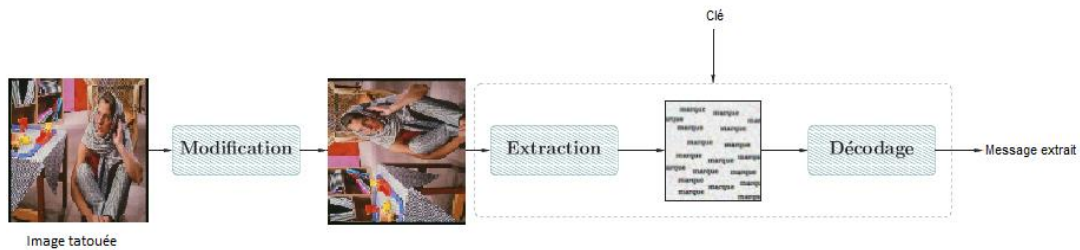


Figure 1-4: Tatouage robuste

- **Tatouage fragile** : C'est un tatouage qui permet de détecter si l'image a été modifiée. La moindre modification du document se répercute fortement sur la marque extraite et on peut en déduire que le document n'est pas authentique (figure 1-5).

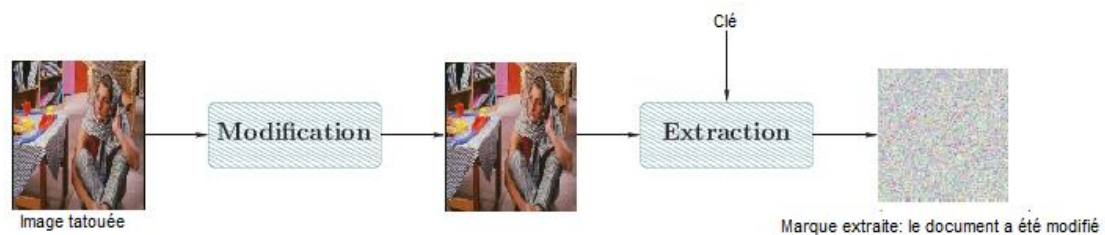


Figure 1-5: Tatouage fragile

- **Tatouage semi-fragile** : C'est une combinaison entre un tatouage robuste et fragile. Il permet de détecter les modifications tout en restant robuste face aux attaques.

### 1.2.3.3 Capacité d'insertion

Elle désigne la taille maximale de la marque qui peut être caché dans une image ou un document numérique sans risque de causer des dégradations statistiquement détectables. Cette quantité varie selon l'application visée. Plus la capacité est grande plus la déformation est grande. En général, la taille de la marque est généralement fixée et elle est la même à la phase d'insertion et d'extraction.

On peut remarquer que ces trois critères (imperceptibilité, robustesse et capacité) sont contradictoires (voir figure 1-6). Si l'on augmente par exemple la force de la marque on va avoir un tatouage plus robuste, mais cela aura en contrepartie pour effet de rendre ce dernier également plus visible. De la même manière, si l'on augmente la taille du message à insérer cela se fera au détriment de la robustesse. Il est donc nécessaire de trouver le meilleur compromis possible entre ces trois paramètres en fonction de l'application visée.

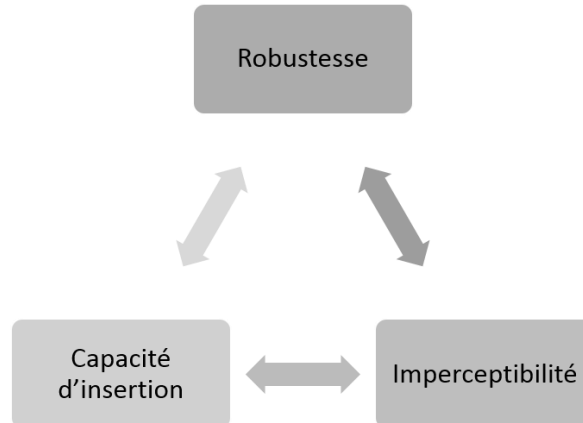


Figure 1-6: Compromis entre la robustesse, la visibilité et la capacité d'insertion

## 1.3 Domaine d'insertion

L'une des étapes les plus importantes de la mise en œuvre d'un algorithme de tatouage est probablement le choix d'un espace ou domaine de tatouage. Pour les deux types de schémas d'insertion abordés dans la section précédente, il est possible de travailler dans les trois domaines de représentation : spatiale, fréquentielle et multirésolution. Le choix de l'une ou



de l'autre technique se fera directement en fonction des objectifs des applications et de leurs performances.

### **1.3.1 Domaine spatial**

Le tatouage numérique dans le domaine spatial a été utilisé en premier. Dans ce domaine on exploite la valeur numérique des pixels pour insérer une marque. En général on insère la marque dans les LSB de l'image ceci permet à ce que la marque soit insérée sans affecter l'image [4]. L'outil spatial utilise le spectre bleu comme emplacement pour stoker la marque. Le choix du spectre bleu est due au fait que l'œil humaine est peu sensible aux modifications dans ce domaine [3]. Bender & al [5] ont proposé un algorithme appelé Patchwork son principe repose sur la différence de luminance observé entre deux ensemble de pixel déterminées à l'aide d'une clé. Pitas [6] présente un procédé de tatouage d'image dans le domaine spatial qui modifie les LSB en augmentant le contraste entre deux sous-ensembles aléatoires des pixels. Wei Lu et Hongtao Lu [7] ont fourni un nouveau schéma de tatouage pour les images numériques à l'aide du sous-échantillonnage et de la factorisation matricielle non négative. À l'origine, un sous-échantillonnage est utilisé pour créer une séquence de sous-image. Ensuite, la factorisation matricielle non négative est appliquée pour décomposer la séquence en fonction de la similarité des colonnes de la séquence de sous-image. Une séquence pseudo-aléatoire gaussienne est insérée dans les coefficients de décompositions factorisées.

L'avantage du domaine spatial est que les opérations d'insertion et de détection sont peu coûteuses en termes de calcul [8]. Le principal inconvénient avec le domaine spatial reste sa fragilité en face du filtrage et aussi il résiste très mal aux attaques géométriques ceci a poussé les chercheurs à se diriger vers d'autres domaines d'insertion pour le tatouage robuste.

### **1.3.2 Domaine fréquentiel**

Par rapport aux techniques de domaine spatial, les techniques de domaine fréquentiel sont plus utilisées de nos jours. L'objectif de ces techniques est d'insérer le tatouage dans les coefficients spectraux de l'image. Les transformées les plus couramment utilisées sont la transformée de Fourier discrète (DFT), la transformée de cosinus discrète (DCT) et la transformée d'ondelettes discrètes (DWT). Ces algorithmes ont pour avantages en général d'être robustes aux compressions telles que JPEG pour les images. La DFT est aussi utilisée

pour sa résistance face à certaines transformations géométriques. Nous allons détailler le cas de la DFT et DCT.

### 1.3.2.1 Transformée de Fourier Discrète

La Transformée de Fourier Discrète ou bien Discrete Fourier Transform (DFT) en anglais décompose l'image numérique  $I$  de taille  $N_1 \times N_2$  en plusieurs ondes sinusoïdales par la relation suivante :

$$F(u, v) = \sum_{m=0}^{N_1-1} \sum_{n=0}^{N_2-1} I(m, n) e^{-j\left(\frac{2\pi}{N_1}\right)um} e^{-j\left(\frac{2\pi}{N_2}\right)vn} \quad (1-3)$$

La transformée de Fourier  $F$  est complexe. En général, on calcule le module  $M$  et la phase  $P$  données par les équations suivantes :

$$\text{Où } M(u, v) = |F(u, v)| = \sqrt{(\text{Re}(u, v))^2 + (\text{Im}(u, v))^2}$$

$$P(u, v) = \arg(F(u, v)) = \tan^{-1}\left(\frac{\text{Im}(u, v)}{\text{Re}(u, v)}\right)$$

où  $Im$  et  $Re$  représentent respectivement la partie imaginaire et la partie réelle de la transformée de Fourier de l'image  $I$ . En général, on représente uniquement le module dont la répartition fréquentielle est illustrée sur la figure 1-7. Parmi les avantages de la transformée de Fourier on trouve l'invariance par translation. De plus, une rotation dans le domaine spatial se traduit par une rotation identique du module [9].



Figure 1-7: Image Barbara et son spectre de Fourier

Ros & al proposent dans [10] une méthode qui consiste à insérer une marque générée d'une manière pseudo aléatoire à moyenne nulle. Dans la phase d'insertion la même marque est insérée deux fois dans deux sous bandes des fréquences moyennes. La phase de détection consiste à calculer la corrélation entre la marque générée et les coefficients de Fourier de l'image tatouée. Une méthode de tatouage présentée dans [11] est basée sur l'utilisation d'anneaux particuliers. Cette technique rend la méthode insensible aux transformations de type translation, rotation et changement d'échelle. L'algorithme d'insertion consiste à appliquer la transformée de Fourier à une image en niveaux de gris de dimensions  $N \times N$ . Puis la marque est insérée dans le module de la DFT. Le tatouage se réalise dans les moyennes fréquences de l'espace de Fourier. Licks et Jordan [12] ont proposé une autre méthode qui consiste à insérer une marque circulaire dans le module de la transformée de Fourier. Si l'image a subi une rotation, le spectre et la marque subit la même rotation. En conséquence, la marque aura subi un décalage cyclique. Lors de la phase de la détection la corrélation croisée est calculée entre la marque originale et les coefficients tatoués décalés. Dans [13] Poljicak *et al* ont proposé une amélioration de cette méthode et dans [14] les auteurs ont amélioré la méthode en ajoutant un prétraitement de l'image avant l'insertion de la marque. Ce prétraitement consiste à diminuer la variance du vecteur qui supporte la marque. La réduction de la variance a été réalisée soit par une modification directe des coefficients dans lesquels la marque est insérée ou par sélection des coefficients.

### 1.3.2.2 Transformée de Cosinus Discrète

Les méthodes basées sur la DCT figurent en bonne place parmi les méthodes développées pour le Watermarking [15]. La compression JPEG fait partie des attaques qui détruisent le tatouage. Comme son principe est basé sur la DCT, les auteurs de ces méthodes espèrent ainsi en travaillent dans le domaine de la DCT mieux contrecarrer les attaques liées à la compression JPEG. De plus il est possible de bénéficier des études psycho-visuelles déjà menées en codage de sources pour gérer les problèmes de visibilité comme par exemple les travaux de Watson [16] et Cox & al [3].

La transformée en cosinus discrète d'une image  $I$  de taille  $N_1 \times N_2$ , notée par  $F(p, q)$ , est donnée par :

$$F(p, q) = \frac{2\Lambda(p)\Lambda(q)}{\sqrt{N_1 N_2}} \sum_{m=0}^{N_1-1} \sum_{n=0}^{N_2-1} I(m, n) \cos\left(\frac{\pi(2m+1)p}{2N_1}\right) \cos\left(\frac{\pi(2n+1)q}{2N_2}\right) \quad (1-4)$$

$$\text{Avec } \Lambda(E) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } E = 0 \\ 1 & \text{sinon} \end{cases}$$

Cette transformée est souvent calculée sur des blocs de l'image de taille  $8 \times 8$ , soit 64 coefficients. Ces coefficients sont répartis sur trois zones : basses, moyennes et hautes fréquences. Dans la figure 1-8 nous avons un exemple d'insertion du tatouage dans les moyennes fréquences du domaine DCT.

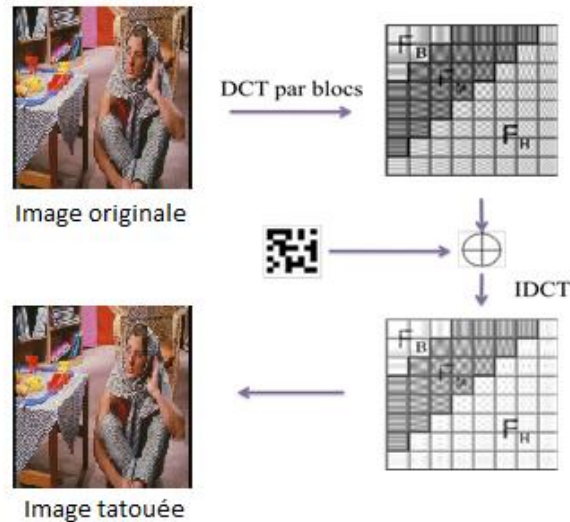


Figure 1-8: Exemple d'insertion dans les fréquences moyennes de DCT

Kosh et Zhao proposent l'idée de coder le watermark dans le domaine fréquentiel avec la DCT [17]. Suhail et Obaidat [18] ont présenté un algorithme de tatouage basé sur la DCT ; l'image originale est divisée en blocs de taille  $8 \times 8$  pixels puis la DCT est appliquée à chaque bloc. Le message, constitué d'une séquence aléatoire, est inséré dans les fréquences moyennes de la DCT de chaque bloc. Dans [19] les auteurs proposent d'améliorer la méthode print/scan (PS) de Cheng *et al.* Ils résolvent le problème d'échec d'extraction de données provoqué par un déplacement de pixels et proposent d'insérer le tatouage dans la bande du milieu de coefficients DCT basée sur le fait que la moyenne des coefficients DCT est égale à zéro avant et après le PS. En outre, une méthode de mapping est conçue pour distribuer le changement de moyenne aux pixels de l'image originale en fonction de la complexité locale. Ils utilisent aussi le modèle JND mentionné dans la méthode de Cheng *et al.*, afin de garantir la qualité visuelle. Dans [20] un algorithme de tatouage aveugle basé sur une modification du coefficient DCT basée sur un bloc est proposé ; la différence entre les deux coefficients DCT des blocs successifs à la même position est calculée pour l'insertion de la marque. En

fonction de la nature de bit de la marque (1 ou 0) à insérer la différence est modifiée. Ceci se fait en modifiant l'un des deux coefficients DCT afin d'apporter la différence dans une zone prédéterminée. Cette zone de différence est ensuite utilisée pour l'extraction du bit de la marque.

### **1.3.2.3 Autres transformée**

Outre les FFT et DCT précédemment cité, d'autres transformations inversibles classiques en traitement d'images ont été aussi envisagées, sans apporter en pratique une amélioration significative des performances ou des invariances géométriques. Nous citons par exemple, l'utilisation de la décomposition en valeurs singulières (SVD) [21, 22] et la transformation de Karhunen-Loève (KLT) [23]. Ces approches de domaine transformées ont été proposées pour concevoir des schémas de tatouages robustes à des classes spécifiques de déformations.

Dans [24] une méthode est proposée qui utilise une procédure d'adaptation pour calculer le facteur d'échelle ou la force de mise à l'échelle en utilisant la fonction sigmoïde dans le domaine de transformation de Hadamard. La valeur du facteur de mise à l'échelle est régie par un paramètre de contrôle. Le paramètre de commande peut être ajusté pour rendre le système de tatouage soit visible ou invisible.

Une nouvelle méthode de tatouage non-aveugle des images est proposée dans [25] ; basée sur la décomposition en valeur singulières (SVD) des images dans le domaine de la transformation complexe d'ondelettes (CWT) car la CWT offre une capacité supérieure à celle du domaine des ondelettes réelles. La modification des sous-bandes appropriées conduit à un système de tatouage qui préserve favorablement la qualité.

Les transformations géométriques sont actuellement les attaques les plus efficaces pour empêcher la détection de la signature dans l'image marquée. Ce constat a conduit les tatoueurs d'images à chercher un espace transformée invariant aux transformées géométriques. Comme nous l'avons vu dans la section précédente, le module de la transformée de Fourier possède une invariance aux translations (qui sont en réalité des translations spatiales circulaires). Maintenant, concernant la rotation et le changement d'échelle, le spectre d'amplitude est affecté par ces transformations. En effet, l'extension des axes d'une image, avec un facteur dans le domaine spatial provoque une mise à l'échelle inverse dans le domaine fréquentiel. En outre, la rotation de l'image d'un angle dans le domaine spatial se traduit par une rotation du même angle du spectre d'amplitude de la

transformé de Fourier. Pour faire face aux problèmes de rotation et de changement d'échelle, la transformation de Fourier-Mellin (FMT) [26,27] s'avère être une autre alternative offrant des propriétés intéressantes ; invariance par translation, rotation et mise à l'échelle. Cette transformation est obtenue par une transformée de Fourier suivie par une transformation en coordonnées logarithmique-polaires puis une deuxième transformée de Fourier. Pour que la transformation soit inversible, les phases calculées après chaque transformée de Fourier sont conservées et réutilisées lors du retour au domaine spatial. Le module de la transformée de Fourier est invariant par translation. Sa conversion en coordonnées log-polaires permet de convertir le changement d'échelle et la rotation en translation horizontale et verticale. La deuxième transformation de Fourier permet d'obtenir un domaine invariant à la translation, rotation et le changement d'échelle. Zhao dans [27] propose un algorithme de tatouage d'image qui utilise une transformée de radon et une transformation analytique de Fourier-Mellin pour construire des moments invariables. Ces moments invariables sont utilisés pour concevoir et détecter le tatouage. L'insertion et la détection de la marque se font de façon classique dans le domaine transformé de l'image. La figure 1-9 présente le schéma de tatouage contenant les étapes de transformations de l'image du domaine spatial jusqu'au domaine invariant.

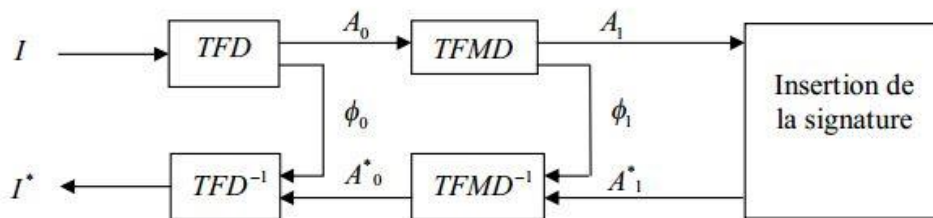


Figure 1-9: Schéma du tatouage dans le domaine d'invariance Fourier-Mellin.

#### 1.3.2.4 Transformation en ondelettes

La transformée de Fourier est un outil largement utilisé pour de nombreux domaines scientifiques, elle permet de connaître le comportement fréquentiel d'un signal mais perd toutes les informations relatives au temps. L'idée d'utiliser la transformée de Fourier à court terme, développée par Gabor en 1946. Cette transformée consiste à considérer le signal autour d'un temps  $t$ . Ce signal est analysé ensuite par une fenêtre glissante  $g(u-t)$  centrée sur cet instant  $t$  en appliquant sa transformée de Fourier. Mais la transformée de Gabor permet seulement de traiter les signaux avec des résolutions homogènes, à cause du choix

fixe de la fenêtre glissante. Ceci a donné naissance à une nouvelle analyse temps-fréquence qui peut supporter toutes les résolutions c'est l'analyse en ondelettes.

#### 1.3.2.4.1 Transformé d'ondelette discrète

La transformée en ondelettes est née de la convergence des travaux théoriques très anciens, notamment les travaux de Haar (1910), de Littlewood et Paley (1930), de Zygmund (1930), de Gabor (1940), puis vers 1960 de Calderon, et des idées récentes pour le traitement numérique de certains signaux par Morlet (le premier à avoir proposé le nom d'ondelettes, 1982), ou pour le développement d'outils mathématiques utilisés en physique théorique par Grossmann (1983) [9].

Par définition, les ondelettes sont des fonctions gérées à partir d'une ondelette mère  $\psi$ , par dilatation et translations. Ainsi, la décomposition en ondelettes fait intervenir deux paramètres qui sont le facteur d'échelle  $a$  et le facteur de translation  $b$ . Le paramètre d'échelle  $a$  permet d'obtenir des ondelettes, à partir d'une ondelette mère, comprimées (support réduit) ou dilatées (support étendu). Les ondelettes comprimées sont utilisées pour déterminer les composantes hautes fréquences tandis que les ondelettes dilatées permettent de déterminer les composantes basses fréquences. Le paramètre  $b$ , quant à lui, permet d'analyser par translations successives le signal jusqu'à ce que celui-ci soit entièrement parcouru [28]. La transformée en ondelettes continue, est définie par:

$$F(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (1-5)$$

Où  $\psi^*$  est le conjugué complexe de l'ondelette mère.

L'analyse multirésolution dyadique permet de donner une transformation d'ondelettes discrète pour l'analyse d'un signal en différentes bandes de fréquences dyadiques. Le principe de l'algorithme est un filtrage en cascade pyramidale qui à chaque étape prélève les détails et analyse l'approximation du signal restante et on réitère l'opération en grossissant son échelle d'un facteur de deux à chaque étape, jusqu'à obtenir une description complète du signal. L'un des éléments fondamentaux de l'analyse multirésolution est l'introduction d'une matrice de dilatation  $D$  qui définit le "processus" de lissage lors d'un changement de résolution. Nous présentons dans la figure 1-10 et la figure 1-11 l'algorithme de décomposition/synthèse rapide (DWT) d'un signal  $x(n)$  tel qu'il a été proposé par S. Mallat [29]. Le calcul de l'approximation passe-bas et des coefficients d'ondelettes à l'échelle  $l$  se

résume à la convolution (filtrage) des coefficients de l'approximation passe-bas à l'échelle  $l - 1$  suivie d'une opération de décimation suivant  $D$  :  $h_0(n)$  représente le filtre passe-bas,  $h_1(n)$  est le filtre passe-haut,  $2 \downarrow$  représente l'opération de décimation d'un facteur 2 et  $2 \uparrow$  représente l'interpolation qui consiste à intercaler un zéro entre deux échantillons. En suivant le même raisonnement que pour la décomposition, on obtient que la reconstruction de l'approximation passe-bas à l'échelle  $l$  se résume à la convolution des coefficients de l'approximation passe-bas et des coefficients d'ondelettes à l'échelle  $l + 1$  précédée d'une opération d'interpolation suivant  $D$ . Comme pour l'analyse, ce processus peut se réitérer permettant ainsi de reconstruire la séquence initiale à partir de tous les coefficients d'ondelettes et la dernière approximation passe-bas.

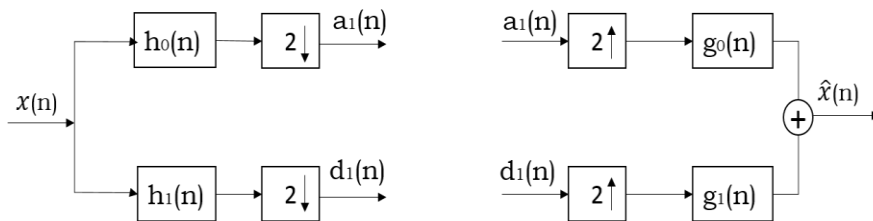


Figure 1-10: Décomposition et reconstruction par la transformée en ondelettes (un seul niveau) [2]

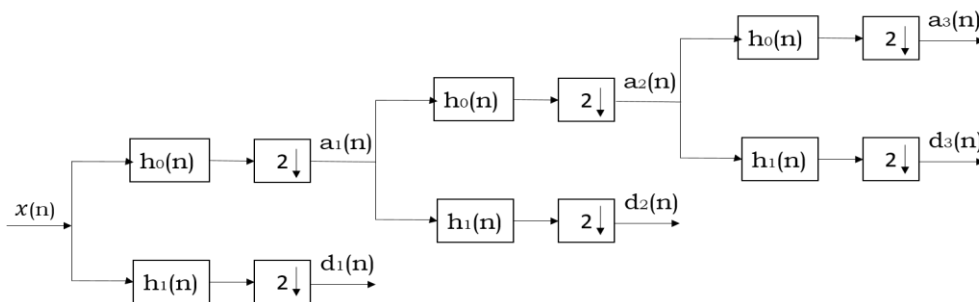


Figure 1-11: Décomposition en ondelettes sur trois de niveaux de résolution

### 1.3.2.4.2 Algorithme Lifting-Scheme 2D rapide de la transformée d'ondelettes de Faber-Schauder

Le Lifting-Scheme a été proposé en 1995 par Sweldens [30,31], dite ondelettes de deuxième génération. Elles permettent d'effectuer la transformée en ondelettes sans le banc de filtres, ce dernier est remplacé par trois étapes (figure 1-12).

Le signal (l'image) est considéré comme une séquence  $x^0 = (x_{i,j}^0)_{i,j \in \mathbb{Z}}$ . L'algorithme de la transformée en ondelettes de Faber-Schauder (FSDWT) de  $x^0$  est exprimé par une



décomposition non-orthogonales mais cette transformée est plus simple à exprimer par un schéma lifting [32].

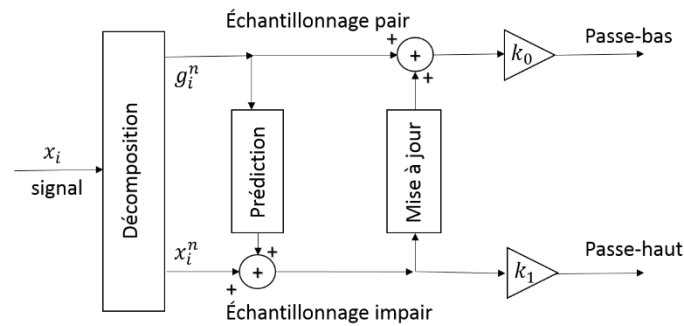


Figure 1-12: la Transformation par lifting scheme (signal 1D)

-la première étape consiste à transformer le signal d'entrée  $x_i$  en séquence, un échantillonnage pair (approximation) et trois impairs (détails) (Lazy wavelet transform)

$$\begin{cases} g^{1,0} = x_{2i+1,2j}^0 \\ g^{2,0} = x_{2i,2j+1}^0 \\ g^{3,0} = x_{2i+1,2j+1}^0 \\ x^{1,0} = x_{2i,2j}^0 \end{cases}$$

-la prédiction : les échantillons impairs  $g_i^n$  sont prédits à partir des échantillons pairs voisins  $x_i^n$ ,

$$\begin{cases} g_{ij}^{1,1} = g_{ij}^{1,0} - \frac{1}{2}(x_{ij}^{1,0} + x_{i+1,j}^{1,0}) \\ g_{ij}^{2,1} = g_{ij}^{2,0} - \frac{1}{2}(x_{ij}^{1,0} + x_{i,j+1}^{1,0}) \\ g_{ij}^{3,1} = g_{ij}^{3,0} - \frac{1}{4}(x_{ij}^{1,0} + x_{i+1,j}^{1,0} + x_{i,j+1}^{1,0} + x_{i+1,j+1}^{1,0}) \end{cases}$$

-la mise à jour :  $x_i^n$  est réalisée avec les séquences prédites précédemment par une pondération,

Cette technique a pour avantages de :

- réduire la complexité de l'implantation associée à un coût de stockage mémoire plus faible ;
- un schéma algorithmique plus "intuitif" (la reconstruction parfaite du signal).

L'image qui en résulte comporte une image d'approximation et trois images de détails. Le niveau suivant est réalisé en appliquant de nouveau la transformation sur l'image d'approximation.

La transformé d'ondelette de Faber-Schauder est donné par l'algorithme suivant :

$$FSWT: \begin{cases} x_{ij}^0 = x_{ij} & \text{pour } i, j \in Z \\ & \text{pour } 1 \leq k \leq N \text{ et } i, j \in Z \\ & x_{ij}^0 = x_{2i,2j}^{k-1} \\ & g_{ij}^k = (g_{ij}^{k1}, g_{ij}^{k2}, g_{ij}^{k3}) \\ & g_{ij}^{k1} = x_{2i+1,2j}^{k-1} - \frac{1}{2}(x_{2i,2j}^{k-1} + x_{2i+2,2j}^{k-1}) \\ & g_{ij}^{k2} = x_{2i,2j+1}^{k-1} - \frac{1}{2}(x_{2i,2j}^{k-1} + x_{2i+2,2j+2}^{k-1}) \\ & g_{ij}^{k3} = x_{2i+1,2j+1}^{k-1} - \frac{1}{4}(x_{2i,2j}^{k-1} + x_{2i,2j+2}^{k-1} + x_{2i+2,2j}^{k-1} + x_{2i+2,2j+2}^{k-1}) \end{cases} \quad (1-6)$$

La transformé inverse de cette transformation est donnée par l'algorithme

$$IFSWT: \begin{cases} & \text{pour } 0 \leq k \leq N - 1, \quad i, j \in Z \\ & x_{2i,2j}^k = x_{ij}^k \\ & x_{2i+1,2j}^k = g_{ij}^{k+1,1} + \frac{1}{2}(x_{ij}^{k+1} + x_{i+1,j}^{k+1}) \\ & x_{2i,2j+1}^k = g_{ij}^{k+1,2} + \frac{1}{2}(x_{ij}^{k+1} + x_{i,j+1}^{k+1}) \\ & x_{2i+1,2j+1}^k = g_{ij}^{k+1,3} + \frac{1}{4}(x_{ij}^{k+1} + x_{i+1,j}^{k+1} + x_{i,j+1}^{k+1} + x_{i+1,j+1}^{k+1}) \end{cases} \quad (1-7)$$

Où  $(x_{ij}^0)$  représente l'image à transformer,  $(x_{ij}^k)$  et  $(g_{ij}^k)$  représentent l'approximation et les détails des coefficients d'ondelettes à l'échelle k.

#### 1.3.2.4.3 Représentation des coefficients d'ondelettes

La transformée en ondelettes discrète décompose une image en quatre sous-bandes, à savoir une sous-bande d'approximation LL et trois sous-bandes de détails : LH, HH et HL, correspondant, respectivement aux détails verticaux, diagonaux et horizontaux. La lettre H correspond au filtrage passe-haut et la lettre L à celui du passe-bas appliqués de façon séparable sur les lignes et les colonnes. La décomposition de la sous bande d'approximation LL permet d'obtenir une représentation sous forme pyramidale. La plupart des auteurs visualisent le résultat de la transformée en ondelettes par cette séquence d'image pyramidal désignée par le nom de représentation à échelles séparées. Sur la figure 1-13 la représentation à échelles séparées de la décomposition successive par la transformée en ondelettes discrète d'une image quelconque jusqu'à trois niveaux de résolution avec les sous-bandes

correspondantes. Ensuite sur, la figure 1-14 une présentation de la décomposition par la transformée en ondelettes discrète de l'image Lena en 2 niveaux de résolution.

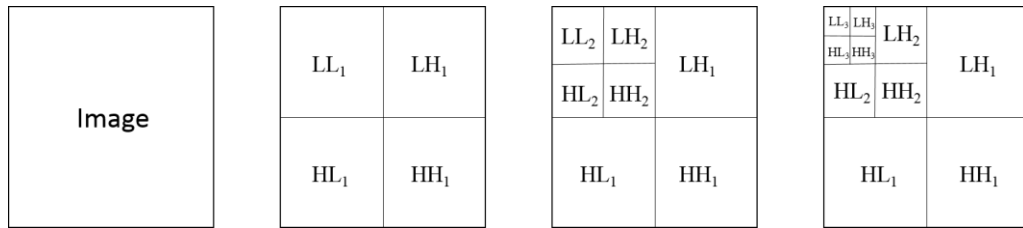


Figure 1-13: La représentation à échelles séparée d'une décomposition successive par la transformée en ondelettes discrète (jusqu'à trois niveaux).

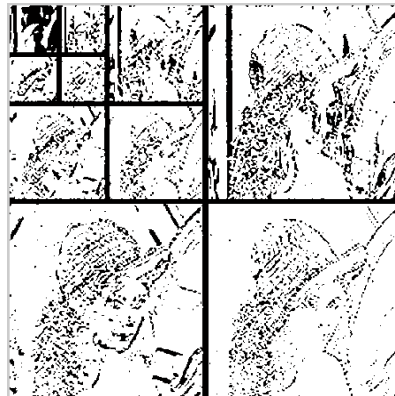


Figure 1-14: Décomposition à échelles séparée par la transformée en ondelettes discrète de l'image Lena

Il existe une autre visualisation des coefficients d'ondelettes, c'est la "représentations à échelles mixée" qui consiste à affecter la valeur de chaque coefficient d'ondelette au pixel où la fonction de base associée atteint son maximum [32]. Les transformations par les schémas lifting, comme celle de Faber-Shauder, sont en fait des transformations linéaires. Donc, elles redistribuent différemment l'information contenue dans les valeurs des pixels des images originales. Ainsi, il est plus naturel de visualiser les coefficients issus de cette redistribution sur une seule image de la même manière que l'image originale. Pour réaliser ceci, on utilise le même principe naturel qui est utilisé pour visualiser les images exprimées dans la base canonique. Dans la représentation à échelle mixte, chaque coefficient d'ondelettes est positionné à l'endroit où la fonction d'ondelettes associée est localisée. Une seule image transformée résulte (au lieu d'images pyramidales) ayant la même taille que l'image originale comme illustré à la figure 1-15.

Un regard attentif sur l'image des coefficients d'ondelettes à échelles mixées montre que la concentration des coefficients dominants est plus élevée dans les régions texturées et des contours. Dans les régions uniformes, les coefficients dominants sont dispersés. Par

conséquent, dans la représentation de l'échelle mixée, la densité des coefficients dominants peut être utilisée comme discriminateur entre les zones uniformes, texturées et de contours [32,33].



Figure 1-15: La représentation à échelles mixées de FSDWT de l'image Lena

#### 1.3.2.4.4 Quelques algorithmes de tatouages utilisant la transformé en ondelette

De nombreuses approches de tatouage d'image numérique basée sur le domaine DWT comme domaines d'insertion ont été proposées.

Dans le travail de [34], qui utilise des images médicales, on a proposé une technique de tatouage aveugle. Le tatouage est inséré en tant qu'information médicale secrète dans l'image originale couleur en utilisant la DWT. Le tatouage est inséré dans les sous-bandes de niveau inférieur (LL) du canal bleu de l'image originale. Le processus d'insertion est: la décomposition de l'image originale en trois canaux de couleur, l'insertion du tatouage dans la sous-bande LL du canal de couleur bleu de l'image originale sur la base de trois transformées en ondelettes discrètes (DWT). Ensuite, ils ont compressé les données avec des techniques Zigzag pour convertir deux dimensions en une image de dimension 1. Enfin, une DWT inverse est appliqué pour obtenir l'image tatouée perceptuelle. Dans le processus d'extraction, l'image tatouée est décomposée sur la base de DWT, la sous-bande LL est divisée en trois canaux de couleur. Le canal bleu est choisi. Ensuite, la technique Zigzag inverse est appliquée pour obtenir une image de canal bleu à deux dimensions. Finalement, le DWT inverse est appliqué pour obtenir l'image tatouée. Le résultat expérimental montre que la méthode est imperceptible et donne une haute qualité sans attaques. Mais il montre aussi que la méthode a une faible robustesse face aux attaques de cropping, de filtrage et de bruit.

Dans [35], Ahmad *et al* ont proposé des techniques de tatouage non-aveugles dans le domaine fréquentiel à trois niveaux DWT. La marque de tatouage est insérée dans la sous-bande basse fréquence en utilisant une technique de mélange. Dans la phase d'insertion, trois niveaux de DWT sont appliqués à l'image originale et à l'image tatouée. Ensuite, ils intègrent le tatouage dans l'image originale en utilisant l'interpolation linéaire. Les résultats expérimentaux montrent que l'approche fournit une imperceptibilité acceptable et elle est robuste contre les attaques fréquentes, mais elle a une faible robustesse contre les attaques de bruit.

[36] a proposé une technique de tatouage d'image numérique dans le domaine fréquentiel basée sur la DWT. L'image originale et l'image tatouée sont décomposées en deux niveaux DWT. Ensuite, ils ont calculé la matrice de caractéristiques de l'image et la fonction de hachage appliquée pour chaque sous-bande (LL, LH, HL, HH). Ensuite, en fonction de la fonction XOR, ils ont intégré le tatouage dans l'image originale. Finalement, la DWT inverse est appliquée. Les données de coefficient produites sont quantifiées pour obtenir le hachage d'image. L'approche proposée a atteint un PSNR élevé (bonne qualité) avec une faible complexité de calcul et un temps d'exécution réduit.

## 1.4 Les attaques

Un des critères fondamentaux à prendre en compte lors de la conception d'un algorithme de tatouage numérique est la robustesse de la marque aux diverses attaques. Les attaques dans le tatouage numérique sont définies comme l'ensemble des opérations qui peuvent rendre la marque indétectable/indécodable, qu'elles soient bienveillantes ou malveillantes. Il existe deux grandes catégories d'attaques :

**Attaques bienveillantes :** Elles regroupent les manipulations effectuées par un utilisateur qui n'a pas initialement pour objectif d'empêcher la détection ou l'extraction de la marque. L'objectif de ces traitements est de modifier ou masquer certaines caractéristiques de l'image mais pas de supprimer la marque. Il est impossible de faire une étude exhaustive de ces traitements car ils sont trop nombreux, mais nous nous contenterons de présenter les plus couramment rencontrés.

**Attaques malveillantes :** Elles regroupent les opérations qui ont pour objectifs de supprimer ou d'empêcher la détection ou l'extraction correcte de la marque.

Dans la littérature, plusieurs classifications des attaques de tatouages ont été proposées. Par exemple, celles présentées par S. Voloshynovskiy & al [37] et Cox & al [3].

En se basant sur [37] et [3], nous proposons dans la figure 1-10 une classification en trois catégories [2] :

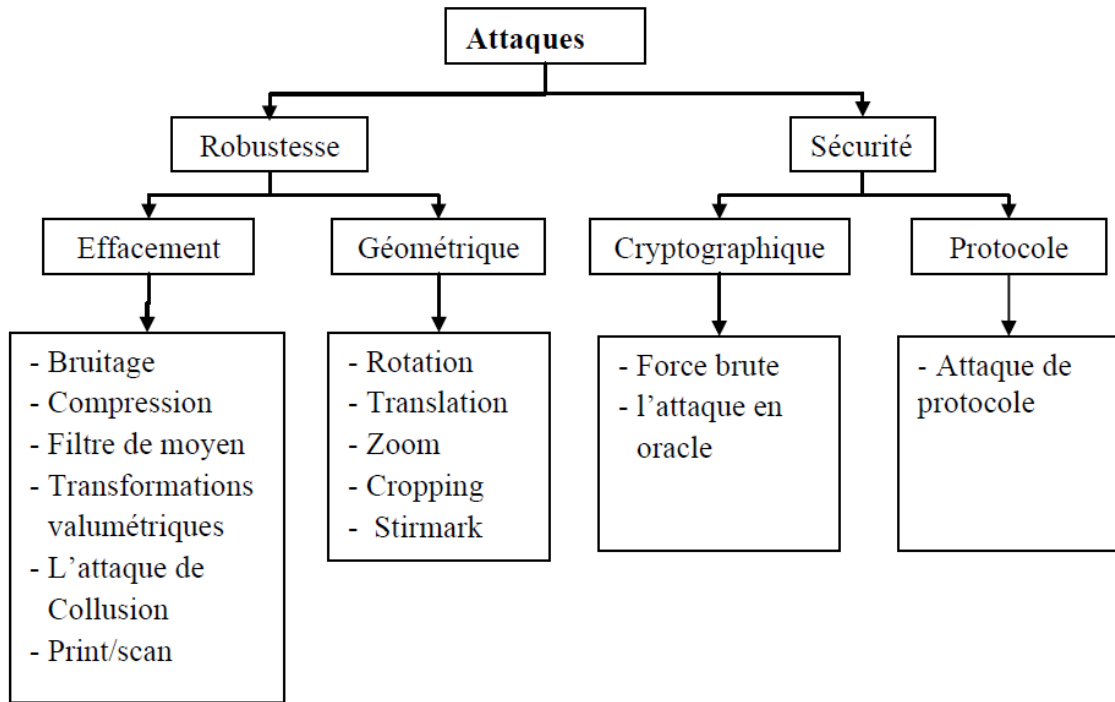


Figure 1-16: La classification des attaques que peut subir un document tatoué selon [2].

### Attaque d'effacement

- Ajout de bruit (gaussien, aléatoire uniforme) : l'ajout involontaire ou délibéré d'un bruit dans l'image peut, lorsqu'il est suffisamment important, avoir pour effet de masquer la marque.
- Filtres : les filtres sont un des outils de base du traitement d'image. Ils sont principalement utilisés pour améliorer l'aspect d'une image, en rendant, par exemple, celle-ci plus « douce » (filtres passe-bas, médian, « anti-aliasing », etc.), ou en faisant ressortir des détails (rehaussement des contours). Les filtres ont généralement pour effet d'atténuer le tatouage dans l'image.
- Compression : tout système de tatouage d'image doit pouvoir résister jusqu'à un certain niveau de compression. La compression a pour objectif de réduire autant que possible la quantité de données d'un document. Ce processus consiste à éliminer les composantes perceptuellement moins significatives en

préservant les composantes importantes du document. En effet, la majorité des données qui circulent sur les réseaux sont sous forme compressée (avec ou sans perte). Jpeg et JPEG2000 sont les deux algorithmes de compression d'image les plus utilisés actuellement.

- Quantification des couleurs: cette opération est couramment appliquée lorsque l'on convertit une image dans un format avec une palette de couleur réduite (Gif par exemple). Bien souvent, elle est suivie d'un « dithering » (tramage) de manière à simuler des couleurs supplémentaires, qui a pour effet de rajouter du bruit à l'image.
- Correction gamma et égalisation d'histogramme : ce type d'opération est fréquemment utilisé pour améliorer le contraste d'une image en changeant la valeur de la luminance de chaque pixel de l'image par une fonction linéaire ou non-linéaire, afin d'améliorer l'aspect visuel de l'image.
- Passage en niveaux de gris: la conversion d'une image couleur en niveaux de gris peut poser des problèmes entre autres aux systèmes de tatouage d'image qui insèrent la marque dans une composante de l'image autre que la luminance (comme le bleu par exemple).
- Print/scan: dans le cadre d'images fixes, ce type de conversion correspond principalement à une impression suivie d'une numérisation à l'aide d'un scanner [38]. Les distorsions subies par l'image sont multiples et dépendent de la qualité des équipements utilisés.

### **Transformations géométriques globales et locales**

Ce genre de transformations n'a pas pour effet d'éliminer la marque mais plutôt rend la marque indétectable même si elle reste encore dans l'image. Il existe plusieurs transformations géométriques, on peut citer par exemple les transformations géométriques affines (translation, rotation et changement d'échelle), le recadrage ou 'cropping' en anglais, qui consiste à supprimer une partie de l'image et par conséquent une partie de la marque. Il existe aussi des transformations géométriques locales comme l'attaque StirMark. Cette attaque consiste à une succession de distorsions géométriques aléatoires appliquées localement à plusieurs endroits dans l'image [39].

- Les transformations affines (locales ou globales): les principales sont les translations, les rotations et les changements d'échelle.

- Les symétries axiales (horizontale et verticale): elles ne sont pas forcément décelables si l'image présente naturellement un axe de symétrie ou aucune information textuelle, mais peuvent suffire à piéger l'algorithme d'extraction, si celui-ci ne les prend pas en compte.
- Recadrage et extraction: ces opérations visent à ne préserver que la partie « intéressante » de l'image. Elles sont généralement problématiques dans la mesure où elles introduisent une désynchronisation (similaire à celle produite par une translation) et où la taille de l'image recadrée risque d'être insuffisante pour contenir une marque robuste.
- Suppression de lignes et de colonnes : ces manipulations sont généralement invisibles, mais suffisent à créer un décalage significatif pouvant rendre difficile l'extraction du tatouage

### **Attaques de nature Cryptographique**

Moins courantes, ces attaques suivent le modèle des attaques classiques en cryptographie. Certaines de ces attaques, telles que les « Brute Force Attacks », ont pour objectif de découvrir la clé secrète utilisée pour insérer la marque, en essayant de manière exhaustive toutes les clés possibles. Bien évidemment, ce genre d'attaque est très coûteux en temps de calcul, et n'est réellement efficace que sur des algorithmes utilisant des clés de petite taille

### **Attaques de protocole**

Les attaques de protocoles se distinguent des autres familles d'attaques, dans la mesure où leur but n'est pas de détruire ou d'empêcher la détection de la marque par des manipulations de l'image. Ces attaques s'en prennent directement aux protocoles de l'application elle-même. Une des premières attaques de ce type a été proposée par Craver *et al.* [40].

## **1.5 Outils d'évaluation des méthodes de tatouage numérique**

Un système de tatouage doit satisfaire différentes exigences en matière de transparence, de robustesse et de capacité afin de garantir un compromis entre ses trois contraintes. En effet, aucun cahier des charges ne donne réellement des valeurs fixées pour la capacité, la qualité visuelle de l'image tatouée ou les différentes attaques auxquelles l'algorithme du tatouage doit être robuste. Par conséquent, dans cette section ainsi que dans le reste de cette thèse,



nous abordons le problème de l'évaluation de la performance de deux critères majeurs : l'imperceptibilité du tatouage et la robustesse.

### **1.5.1 Evaluation de la qualité visuelle**

Comme déjà mentionné la qualité d'une image tatouée est l'une des contraintes principales pour valider une méthode de tatouage numérique. Plusieurs mesures de qualité ont été proposées dans la littérature. Dans le chapitre suivant une description détaillée de certaines métriques objectives peuvent être trouvée ainsi que les métriques subjectives.

### **1.5.2 Evaluation de la robustesse du tatouage**

Bien que les algorithmes de tatouage aient émergé et évolué, l'évaluation des performances se réalise de plusieurs façons et qui s'avère être une tâche difficile. Ceci est dû à l'absence d'une méthode d'évaluation commune et aussi dans la nécessité d'avoir une normalisation des attaques contre lesquelles les algorithmes doivent être testés. Dans le cadre de mettre en ordre le processus d'évaluation, plusieurs outils ont été proposés dans la littérature. Le but de ces logiciels est double. Tout d'abord, la mise en œuvre des manipulations spécifiques dont le but est de détruire la marque contenue dans une image ou d'empêcher sa détection. Deuxièmement, permettre d'avoir des plates-formes communes et efficaces pour la comparaison de différents algorithmes de tatouage. Parmi ces logiciels nous pouvons citer certains utilitaires opérant sur de nombreuses attaques, tels que StirMark benchmark développé par Kutter et Peticolas [37], CheckMark benchmark proposé par Voloshynovskiy *et al.* [35], OptiMark benchmark réalisé par Nikolaidis *et al.* [39]. StirMark est le plus utilisé et le plus éprouvé de tous.

## **1.6 Types d'applications**

Le Watermarking offre trois avantages, il est imperceptible, ne peut pas être déplacé ou converti et une fois inséré dans une image il subit les mêmes transformations que l'image. Ces trois aspects sont les principales raisons qui permettent l'intégration du tatouage dans diverses applications. Jusqu'à ce jour il n'existe pas une méthode de tatouage qui soit à la fois imperceptible, robuste et qui puisse insérer une grande quantité d'information. Il est nécessaire de prendre en compte les besoins de l'application visée lors de la conception d'un algorithme de tatouage numérique. Plusieurs applications industrielles de tatouage numérique sont proposées dans la littérature. Cox *et al.* présentent dans leur livre [3] une

description détaillée des différentes applications du tatouage numérique. Dans la suite, nous citons quelques champs d'application du tatouage numérique.

- **Le contrôle de diffusion (Broadcast monitoring)** : Les annonceurs paient un prix élevé pour présenter leurs produits lors d'une parenthèse publicitaire pendant des manifestations sportives importantes ou des films. Par conséquent, ils veulent s'assurer que leurs annonces ont été effectivement diffusées. Cette application permet aux propriétaires ou aux distributeurs du contenu de suivre la diffusion des émissions sur la télévision ou sur Internet de leur contenu.
- **Identifier le propriétaire (Owner identification)** : La protection des droits d'auteur a été une des premières applications étudiée en tatouage numérique des images. Ce service reste cependant toujours d'actualité et concerne encore une grande majorité des publications. Cette application permet d'insérer des informations relatives à l'auteur par exemple comme marque pour garantir les droits d'auteurs en cas de litige par exemple.
- **Preuve de propriété (Proof of ownership)** : Cette application utilise le Watermarking pour donner des preuves pour des documents où il y a un conflit car elle fournit une protection du document même lorsqu'il est diffusé de sorte qu'elle identifie le propriétaire du document.
- **Suivi des transactions (Transaction tracking)** : Soit une image diffusée sous forme chiffrée, interdisant qu'une personne puisse avoir accès à son contenu lors de la transmission de celle-ci. Au niveau du destinataire, il peut remettre illégalement en circulation cette image, il sera alors possible de remonter à la source du délit grâce à l'identifiant du destinataire caché dans l'image. Donc cette application utilisant le Watermarking permet d'identifier des gens qui ont eu un contenu légalement mais qui l'ont utilisé d'une manière illégale.
- **Identification du contenu (Content authentication)** : Si on prend comme exemple une image, il serait facile de supprimer une partie de l'image. Ceci peut changer la sémantique de l'image originale. Ainsi, pour éviter ce traitement illicite, une marque est insérée dans le document pour vérifier plus tard si leur contenu n'a pas été modifié. Si la marque n'est pas détectée d'une manière correcte, on pourra conclure que le document a été modifié.
- **Contrôle de copie (Copy control)** : C'est une application très importante dont le but est de contrôler ou d'empêcher la copie illégale d'un document protégé. Utiliser le

Watermarking pour prévenir les utilisateurs que le contenu d'un document ne peut pas être enregistré. L'objectif est de détecter la présence d'un copyright (une marque) pour contrôler ou rendre la copie de l'œuvre extrêmement difficile

- **Contrôle des appareils (Device control)**: Cette application utilise le Watermarking pour fabriquer des dispositifs, tels que les jouets, qui réagissent au contenu affiché.

## 1.7 Les récentes applications dans le domaine de tatouage

On a vu dans la section 1.6 que le tatouage numérique peut être utilisé dans plusieurs types d'applications. Au cours de ces dernières années le tatouage numérique a été largement développé. Des combinaisons avec d'autres technologies ont vu le jour, comme avec le Deep-learning ou des applications d'authentification des objets en 3D et Vidéo entre autre.

Parmi les applications du tatouage qui ont été beaucoup développées nous trouvons l'imagerie médicale. Le tatouage numérique peut être utilisé dans des images médicales à plusieurs fins : protéger les informations du patient contre les personnes non autorisées ; authentification de l'image du patient.

Dans [42] Dong *et al.* ont proposé un schéma d'hierarchies multiples robustes pour les images médicales basées sur la DWT et la DCT, qui peut résoudre efficacement les problèmes de la façon de déterminer la région d'intérêt (ROI) (c'est la partie de l'image médicale contenant des informations cruciales et importantes pour le diagnostic, cette région est généralement définie par un médecin) et d'améliorer la capacité de dissimulation. L'algorithme combine le vecteur de fonctionnalité visuelle des images, la technologie de cryptage avec l'authentification de tiers et évite le processus ennuyeux pour la sélection du ROI. Cet algorithme peut améliorer la sécurité, la confidentialité et l'intégrité de l'image médicale selon les besoins de l'application. Par la suite Dhaou *et al.* ont proposé dans [43], une nouvelle approche invisible et robuste de tatouage d'images médicales qui insère les données cryptées du patient au niveau de la région de non intérêt RONI (qui peut être généralement l'arrière-plan noir de l'image, ou bien une partie gris en plus de l'arrière-plan noir.) en utilisant un tatouage robuste. En plus, la clé utilisée lors du cryptage sera insérée au sein de la région d'intérêt ROI en utilisant un algorithme fragile. Ceci permettra de maximiser la robustesse du tatouage, de protéger les données du patient et aussi de contrôler l'intégrité des informations importantes contenues dans l'image à tatouer.

En ce qui concerne les applications de sécurité pour l'authentification plusieurs auteurs ont proposé des méthodes de tatouage novatrices. Kamran *et al.* [44] proposent d'insérer un bit de message correct uniquement. En raison de cette proposition, le schéma de codage et de permutation BCH (Bose–Chaudhuri–Hocquenghem) n'est pas nécessaire, ce qui augmente la capacité d'insertion de la marque. Kamran *et al.* [44] ont également proposé une approche novatrice qui utilise le concept d'échantillonnage en aval pour l'amélioration de la performance. L'échantillonnage en aval fournit deux versions sous-échantillonnées de l'image originale, c'est-à-dire la référence et la dissimulation de données. Ensuite, les blocs sont générés en utilisant ces deux versions sous-échantillonnées. L'insertion se fait dans les blocs à l'aide d'une modification d'histogramme. De plus, pour rendre la technique aveugle les auteurs ont intégré la carte de localisation (LM) dans l'image tatouée. Dans [45] Ralaivao *et al.* ont proposé une étude d'une nouvelle technique de transfert sécurisé d'image numérique se basant sur la combinaison de la technologie de chiffrement et de tatouage d'image. Dans ce cas, celui qui intercepte une image cryptée sans en avoir le droit ne pourra rien faire de cette image parce qu'il n'a pas la clé de chiffrement. Aussi, grâce à la technique de tatouage, la clé de chiffrement est cachée dans l'image cryptée elle-même et le transfert de la clé de chiffrement vers le destinataire ne sera plus un problème. Le schéma de cryptotatouage proposé dans cet article doit sa simplicité d'une part à l'algorithme de cryptage adopté pour l'image ne faisant intervenir que l'opérateur logique XOR tout en assurant un niveau de sécurité élevé et d'autre part, l'utilisation du SVD dans le schéma du tatouage. Plus récemment Haddada *et al.* dans [46] proposent une nouvelle approche de renforcement du tatouage numérique qui assure un bon compromis entre le niveau de sécurité des données biométriques d'une personne et la complexité de calcul du système proposé, tout en maintenant un espace de stockage réduit et une bonne qualité visuelle de l'image hôte tatouée. Ils valident l'approche proposée sur deux modalités biométriques, l'empreinte digitale et le visage, qui sont parmi les plus utilisées et les plus matures dans les systèmes biométriques de tatouage. Le niveau de sécurité amélioré est assuré par un système combiné de tatouage. En fait, le même algorithme de tatouage est appliqué deux fois successivement. Tout d'abord, le visage d'un individu est marqué par les caractéristiques locales, les minuties, de son empreinte digitale. Deuxièmement, l'image du visage précédemment tatouée est insérée dans l'image originale de l'empreinte digitale en tant qu'informations d'identité supplémentaires.

Comme la sécurité des objets 3-D est de plus en plus portée à l'attention du public grâce à l'expansion de nouvelles technologies multimédia telles que l'impression en 3-D. Dans le développement des systèmes de crypte-sécurité des objets 3D, on peut identifier deux directions principales représentées par la cryptographie et le tatouage numérique. Un bon système de sécurité doit être conforme au format, doit conserver le débit binaire d'origine et, dans la mesure du possible, il doit être réversible. La méthodologie du tatouage a l'avantage de s'assurer que le message caché intégré peut être vérifié à n'importe quel stade de traitement tel que la transmission, le stockage et la visualisation des supports d'intégration. Dans [47] Itier *et al.* ont examiné les travaux antérieurs en matière de sécurité 3-D et ils ont aussi testé la crypto-sécurité d'une méthode de tatouage 3-D qui intègre l'information par minimisation de la distorsion de la surface de maille.

Les réseaux de neurones profonds ont récemment réalisé des progrès significatifs. Le partage de modèles qualifiés de ces réseaux de neurones profonds est très important dans la progression rapide de la recherche ou du développement de systèmes de réseaux neuronaux profonds. Dans le même temps, il est nécessaire de protéger les droits des modèles partagés. À cette fin, les dernières technologies qui s'intéressent au Deep-learning ou à l'intelligence artificielle commencent à se combiner avec le tatouage numérique. C'est-ce-que nous retrouvons dans [48] où Heydari *et al.* proposent un nouvel algorithme non supervisé pour la classification des messages multimédias et les trancher pour se déplacer entre les nœuds à l'aide de l'apprentissage par machine. Le procédé proposé comprend plusieurs étapes qui incluent : 1- les informations du nœud du capteur dans les données d'image en utilisant des algorithmes pour ajouter de manière sécurisée le tatouage résistant aux attaques. 2- les données agrégées sont des informations agrégées. 3- Dans les principaux capteurs, les fonctionnalités de haut niveau et les algorithmes d'apprentissage en profondeur fournis par les modèles formés, puis catégoriser les images sous la forme de pièces continues sont envoyés sur le réseau. Outre la segmentation des images utilisant des algorithmes d'apprentissage approfondi, le volume de paquets multimédia au niveau du réseau a été réduit et la vitesse de transfert de données augmente également. Dans [49] Uchida *et al.* proposent d'utiliser une technologie de tatouage numérique pour protéger la propriété intellectuelle ou détecter la violation de la propriété intellectuelle des modèles qualifiés. Premièrement, ils formulent un nouveau problème : intégrer des tatouages dans des réseaux neuronaux profonds. Ensuite ils définissent également les exigences, les situations d'insertion et les types d'attaque pour le tatouage vers des réseaux de neurones profonds. Deuxièmement, ils

proposent un cadre général pour insérer la marque dans les paramètres du modèle à l'aide d'un régulateur de paramètres. Leur approche ne nuit pas à la performance des réseaux dans lesquels un tatouage est inséré.

Les progrès que connaît le tatouage numérique en le combinant avec d'autres technologies surtout celles de Deep-learning n'est qu'à son début. Les exemples vus au-dessus ne sont que quelques un parmi tant d'autres.

## **1.8 Conclusion**

Ce chapitre a présenté un état de l'art sur le tatouage des images numériques. Nous nous sommes intéressés aux notions liées aux techniques du tatouage numérique. Nous avons aussi abordé les différentes phases de la conception d'une méthode de tatouage. Un schéma de tatouage permet d'insérer une marque d'identification dans une image. Ce schéma se compose de deux phases : la phase d'insertion qui consiste à insérer (ou à cacher) une marque d'identification et la phase de détection qui consiste à détecter ou à décoder la marque insérée. Nous avons présenté aussi une classification des techniques du tatouage selon différents critères : les champs d'application, les types d'algorithmes et les domaines d'insertion. Nous avons vu que différents domaines peuvent être exploités, souvent pour répondre aux exigences de la multitude d'attaques auxquelles les algorithmes peuvent être soumis et aussi le type d'application envisagée. Effectivement, la plupart des attaques, qu'elles soient intentionnelles ou bienveillantes, peuvent être assez dévastatrices. Ainsi les techniques du tatouage peuvent utiliser le domaine spatial où l'insertion modifie directement les valeurs des pixels, mais la plupart des techniques de tatouage passent souvent par une transformée. Les plus populaires sont la transformée de Fourier, la DCT ou la transformée en ondelettes.

Le chapitre suivant est spécialement dédié à l'étude du système visuel humain qui permet de proposer des méthodes de tatouage psychovisuel qui sont les plus performantes actuellement.

## **Chapitre 2**

# **2. Le système visuel humain – Description et modélisation pour le tatouage numérique**

### **2.1. Introduction**

Nous avons vu dans la section 1.2.3 l'importance de l'invisibilité ainsi que de la robustesse du tatouage. Pour satisfaire ces deux contraintes antagonistes, les schémas de tatouage numérique s'orientent de plus en plus vers l'exploitation des modèles du Système Visuel Humain (SVH). La vision est un sens essentiel de l'être humain ; environ 80 à 90% des neurones dans le cerveau humain sont impliqués dans la perception visuelle [50]. En tenant compte des propriétés et des limites du SVH, les images peuvent être plus efficacement analysées, traitées et transmises. Pour atteindre ces objectifs, il faut comprendre la structure du SVH ainsi que ses diverses propriétés afin de pouvoir modéliser avec précision le SVH et utiliser ces modèles perceptuels pour le traitement de l'image. Les modèles visuels utilisés dans le domaine du tatouage d'images sont identiques à ceux développés dans le cadre du codage d'images. Ces modèles exploitent essentiellement la sensibilité au contraste du SVH ainsi que ses capacités de masquage pour assurer aussi bien l'imperceptibilité que la robustesse du tatouage. Toutefois, une différence doit être faite entre l'utilisation de modèles visuels à des fins de compression et leur utilisation dans des applications de tatouage. Effectivement, le domaine de la compression exploite les modèles visuels dans le but de supprimer les composantes les moins significatives de l'image permettant de l'alléger sans toutefois modifier sa qualité. Pour les schémas de tatouage d'images, l'objectif est de modifier judicieusement les composantes peu visibles, qui ne seront pas remarquées par un observateur.

L'objectif de ce chapitre est de donner un aperçu de la structure du SVH et se limite à certaines questions générales pertinentes pour le traitement d'image et le tatouage numérique en particulier.

## 2.2. Le système visuel Humain (SVH)

Le SVH est un système complexe et pas encore totalement compris. Néanmoins, il peut être considéré comme un système de transformation d'informations en données exploitables par le cerveau. En première approximation, le SVH prend en considération les informations de texture, de luminosité et de contours comme éléments de base.

### 2.2.1 Structure du SVH

Le SVH peut être subdivisé en quatre composants majeurs correspondant à la voie empruntée par l'information visuelle :

- l'œil qui capte l'information lumineuse,
- la rétine qui convertie cette information en signaux électriques décodables par le cerveau,
- l'information est transmise par la suite, au cortex visuel à travers les nerfs optiques et les corps genouillés,
- enfin le cortex visuel est chargé du décodage et du traitement des informations.

#### 2.2.1.1 L'œil

Du point de vue optique, l'œil humain fonctionne comme une caméra. La lumière entre dans l'œil par un système optique composé de cornée et de lentille (correspondant à l'objectif d'une caméra) et, par conséquent, suit les principes physiques de la réfraction et de la diffraction. L'angle d'ouverture est contrôlé par la pupille et l'iris tandis que la rétine équivaut à la couche photosensible (film). La figure 2-1 montre une coupe latérale de l'œil humain.

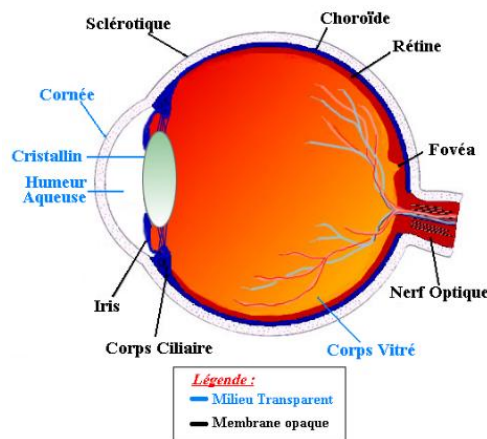


Figure 2-1 : Coupe légende de l'œil [51]



### 2.2.1.2 La rétine

La rétine est la première structure intervenant dans la phase de traitement de l'information visuelle et peut être considérée comme une extension du cortex. Son rôle principal consiste à convertir les signaux lumineux en signaux électriques qui seront ensuite transmis au cortex visuel par l'intermédiaire d'un système de nerfs optiques. La rétine est constituée de deux grandes parties : les neurones rétiniens et les photorécepteurs. La lumière doit traverser plusieurs couches de neurones de la rétine avant d'atteindre les photorécepteurs qui sont dans le côté opposé. La figure 2-2 illustre une organisation simplifiée de la rétine.

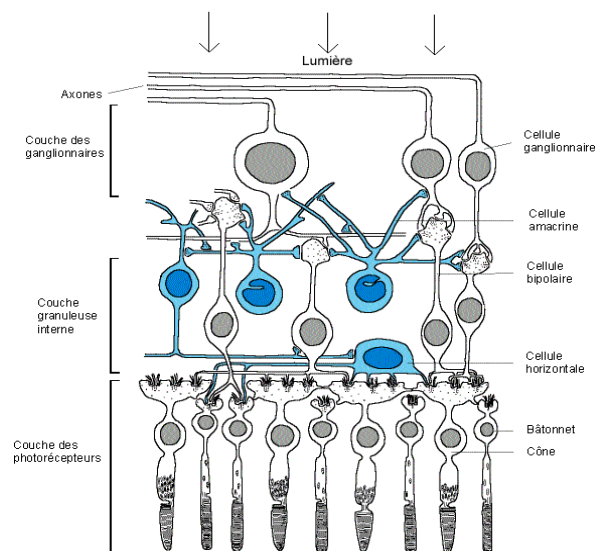


Figure 2-2: La rétine [52]

#### 2.2.1.2.1 Les photorécepteurs

Ce sont les cellules qui se situent sur la couche la plus profonde, par rapport à l'arrivée de la lumière, aussi appelées cellules sensorielles. La rétine est en effet "inversée", car la lumière doit traverser la rétine avant de pouvoir atteindre les photorécepteurs, sensibles à la lumière. Cette couche comporte environ 130 millions de cellules photosensibles différentes, portant des noms reflétant leur forme (voir figure 2-3)

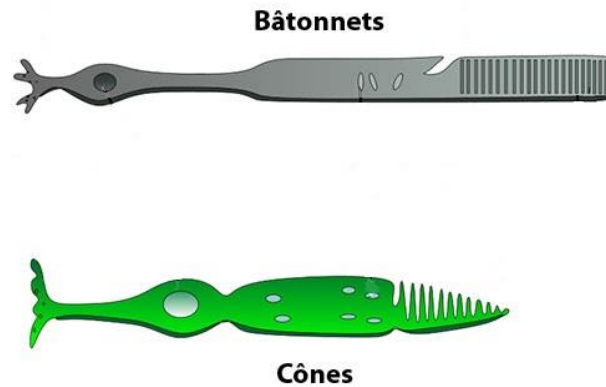


Figure 2-3 : Forme des bâtonnets et cônes

- **Les bâtonnets**

Ils constituent environ 95% de ces cellules, soit au nombre de 120 millions [53], sont responsables de la vision nocturne, et ne sont sensibles qu'à la différence entre obscurité et lumière. Par contre, ils ont la plus grande sensibilité, et sont par cela adaptés à de faibles quantités de lumière.

- **Les cônes**

Les cônes, cellules sensorielles plus grandes, forment les 5% restants des photorécepteurs [53]. Les cônes sont responsables de la vision diurne, et font la différence entre les couleurs. Les cônes sont présents en majorité au niveau de la "tache jaune", dans la région centrale de la rétine. Au milieu de celle-ci se trouve la fovéa, une légère dépression où sont présents uniquement des cônes, très serrés, et où les autres couches sont rejetées à la périphérie, pour laisser pénétrer la lumière plus facilement. La vision est à ce niveau plus précise, plus détaillée et plus sensible aux mouvements que sur le reste de la rétine. C'est d'ici que provient la plupart de l'information visuelle arrivant au cerveau.

## **2.2.2 Perception de la couleur**

Il existe trois différents types de cônes qui peuvent être classés en fonction de la sensibilité spectrale de leurs photochimiques. Ces trois types sont désignés par L, M et S-cônes, ce qui correspond à leur sensibilité aux longues, moyennes et courtes longueurs d'onde, respectivement. Par conséquent, les cônes sont parfois également appelées cônes rouge,

verte et bleu. Les estimations des spectres d'absorption de ces trois types de cônes sont représentées sur la figure 2-4.

Les cônes donnent une vision trichromatique (RVB)

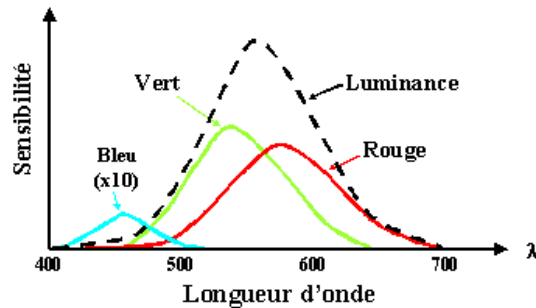


Figure 2-4: Estimations des spectres d'absorption des trois cônes [54]

Toutes les autres couleurs peuvent être produites sous forme d'une combinaison linéaire de ces trois couleurs. Au cours des dernières expériences psychophysiques de la vision des couleurs, il est suggéré que les sorties des trois cônes sont transformés dans un canal achromatique (luminance) et deux canaux chromatiques (couleurs). Ce canal Achromatique est traité plus ou moins séparément des autres canaux. En outre, le SVH est beaucoup plus sensible aux variations du canal achromatique que dans les canaux chromatiques.

C'est pourquoi la majorité des modèles à base du SVH proposés dans la littérature ne traitent que des signaux de luminance. Pour la même raison, nous choisissons de travailler uniquement sur le composant de la luminance dans le reste de la thèse.

### 2.2.3 Perception de la luminance

La compréhension du processus de la perception visuelle est un élément à considérer lors de la modélisation du SVH. La perception d'une zone de l'image engendre trois types de sensations. Les sensations de teinte et de saturation sont liées à la chromaticité de la zone observée alors que la sensation de luminosité reflète la luminance perçue. Alors que les cônes nécessitent une certaine luminosité, les bâtonnets réagissent déjà à partir de très basses luminances. L'adaptation de luminance (également appelée masquage de luminance) peut être expliquée comme la variation de la sensibilité du SVH à une large gamme de luminance du fond. Plusieurs processus contribuent à la perception d'un large dynamique de la luminance [ $3 \cdot 10^{-2} \text{cd/m}^2$  à  $3 \cdot 10^5 \text{cd/m}^2$ ]. La figure 2-5 représente comportement des

photorécepteurs face à l'intensité lumineuse. A partir de la figure ci-dessous nous pouvons constater qu'il existe trois types de domaine :

- le domaine scotopique : la vision de nuit avec une luminance  $10^{-6} < L < 10^{-3}$  cd/m<sup>2</sup>,
- le domaine mésopique: vision crépusculaire avec une luminance  $10^{-3} < L < 10$  cd/m<sup>2</sup>,
- le domaine photopique: la vision de jour avec une luminance  $10 < L < 10^{10}$  cd/m<sup>2</sup>.

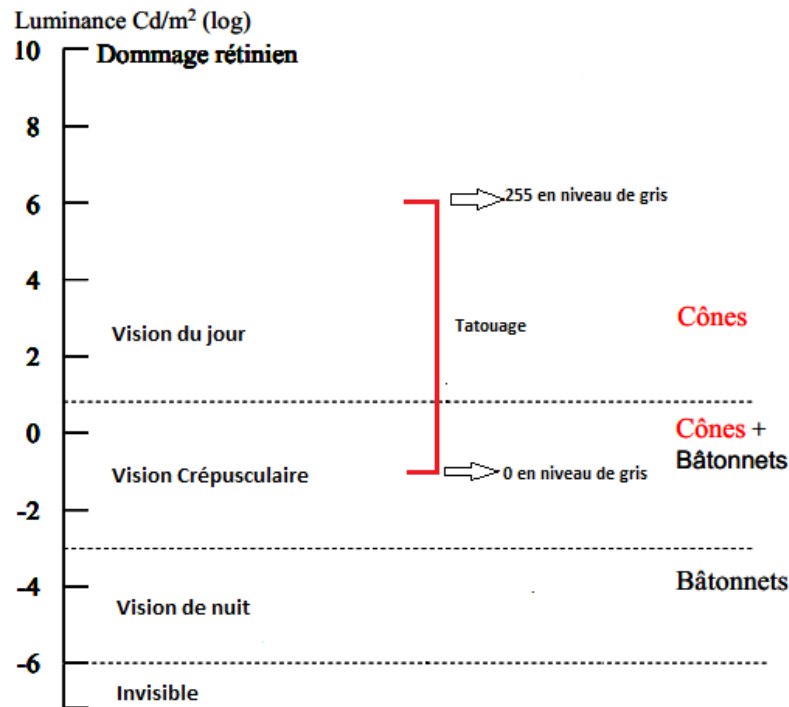


Figure 2-5: Comportement des photorécepteurs face à l'intensité lumineuse

Une des premières modélisations quantitatives de la perception de la luminance a été proposée par Weber [55]. Dans cette expérience classique utilisée pour déterminer la capacité du système visuel humain pour la discrimination de la luminosité consiste à avoir un sujet qui regarde une surface plane avec un petit cercle au centre (figure 2-6). La luminosité du cercle est augmentée jusqu'à que le sujet puisse le distinguer. Cette expérience est conduite pour des luminances entre -1 et 6 en échelle logarithmique qui correspond à la plupart des situations réelles. On considère ainsi qu'une image numérique sera dans cette gamme de luminance, avec le niveau de gris le plus faible autour de -1 en le niveau 255 autour de 6.

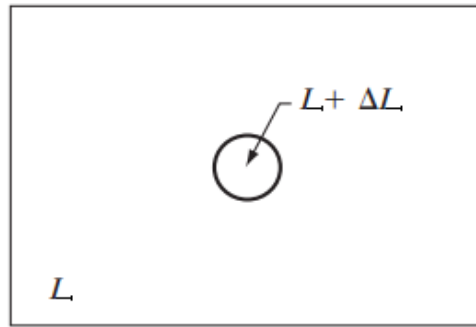


Figure 2-6: Expérience de Weber-Fechner

La courbe tracée sur la figure 2-7 représente le tracé log-log de la variation de l'incrément de luminance juste perceptible  $\Delta L$  (ou JND pour Just Noticeable Difference) en fonction de la luminance du fond  $L$ . Pour de faible luminosité, le JND est élevé et il diminue quand l'intensité de la lumière augmente.

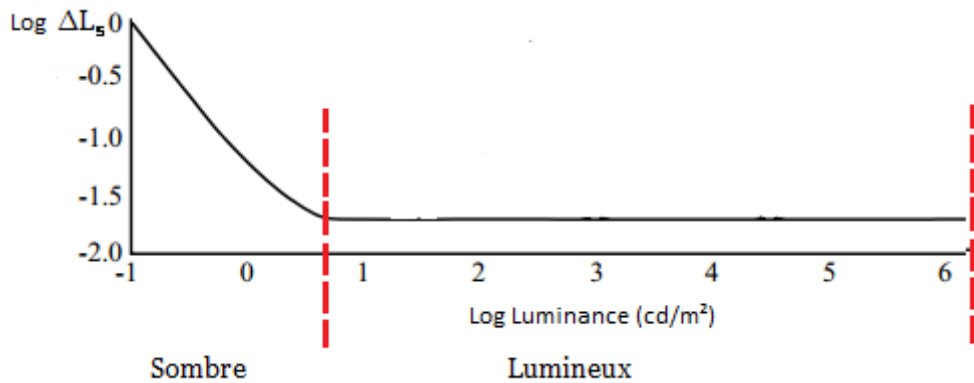


Figure 2-7: Le tracé de la loi de variation du seuil de perception

D'après les résultats obtenus par l'expérience faite par Weber nous constatons que l'œil humaine est moins sensible aux variations qui arrivent dans les zones sombres. Par contre il est très sensible aux modifications dans une zone à lumineuses. Ceci nous conduit à conclure qu'au niveau du tatouage numérique qu'une insertion dans les zones sombres sera moins visible qu'une insertion dans les zones lumineuses.

## 2.2.4 Le contraste

C'est l'appréciation subjective de la différence d'apparence entre deux parties du champ visuel vues simultanément ou successivement. Il peut s'agir d'un contraste de couleur, d'un contraste de luminance, d'un contraste simultané ou successif.

L'un des contrastes les plus anciens et les plus utilisés est celui de Michelson [56]. Dans son expérience, il introduit un rapport qui quantifie le contraste avec  $L_{min}$  et  $L_{max}$  qui désignent respectivement les valeurs minimum et maximum de la luminance, ce contraste est donné par l'équation suivante :

$$C_M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \quad (2-1)$$

Un autre concept de contraste basé sur la luminance est celui de Weber-Fechner [54] donné par la loi de variation du seuil de perception ci-dessous.

$$C = \frac{\Delta L}{L} \quad (2-2)$$

Le contraste de Weber-Fechner est resté très longtemps la seule référence dans le domaine jusqu'à ce que Moon et Spencer étendent la notion de contraste de luminance au cas d'objet noyé dans un fond non uniforme. L'idée principale de Moon et Spencer est d'appliquer le principe de Holladay [57] selon lequel tout fond non uniforme peut être remplacé par un autre de luminance uniforme et produisant le même effet au niveau de la perception :

$$C(x, y) = \frac{|L(x, y) - Lu_S(x, y)|}{Lu_S(x, y)} \quad (2-3)$$

où  $L(x, y)$  est la luminance du pixel,  $Lu_S(x, y)$  est la luminance de l'environnement immédiat, calculée comme la luminance moyenne de huit voisins du pixel  $(x, y)$  dans une fenêtre de 3x3 :

$$Lu_S(x, y) = \frac{1}{8} \sum_{k, l = -1; k, l \neq 0}^L Lu(x + k, y + l)$$

Depuis les expériences Moon-Spencer, un certain nombre d'études déterminent une expression analytique pour le contraste de Moon-Spencer qui pourrait être utilisé pour des images réelles [58, 59, 60].

En se basant sur la sensibilité fréquentielle du SVH, Peli définit un contraste par bande de fréquences spatiales [61]. L'image est décomposée en plusieurs canaux au moyen d'un banc de filtres passe bande de type Gabor, par exemple :

$$H_k(r) = \frac{1}{2}(1 + \cos(\pi \cdot \log_2 r - \pi \cdot k))$$

où  $r$  est la fréquence radiale.

Pour chaque sous-bande  $k$ , le contraste est donné par :

$$C_k(x, y) = \frac{h_k * I(x, y)}{\sum_{i=0}^{i=k-1} [h_i * I(x, y)]} \quad (2-4)$$

Par la suite pour étudier le comportement du SVH face à la sensibilité au contraste des expériences ont été faites qui utilisent le plus souvent le contraste de Michelson.

#### **2.2.4.1 Fonction de sensibilité au contraste (CSF)**

L'une des premières expériences sur la fonction de sensibilité au contraste (CSF ou Contrast Sensitivity Function) a été proposée par Campbell et Robson [62]. Une mire sinusoïdale est utilisée comme illustré ci-dessous.

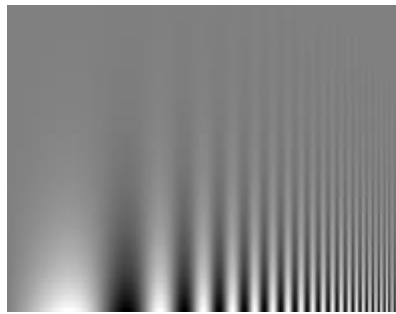


Figure 2-8: mire sinusoïdale de Campbell et Robson

La sensibilité au contraste permet de répondre à la question suivante : à partir de quels niveaux de fréquence et d'intensité un motif ne peut plus être distingué d'un champ uniforme. Dans cette mire, la luminance des pixels est modulée sinusoïdalement le long de la dimension horizontale. La fréquence de modulation augmente exponentiellement de la gauche vers la droite. Le contraste varie également de 100% à environ 0,5% (de bas en haut).

La luminance des pics et des creux reste constante le long d'un trajet horizontal au travers de l'image. Par conséquent, si la détection du contraste est dictée uniquement par le contraste de l'image, les barres lumineuses et sombres alternées devraient avoir la même hauteur partout dans l'image. Cependant, les barres semblent plus hautes au milieu de l'image que sur les côtés. Cette enveloppe de visibilité en forme de U inversé est la fonction de sensibilité au contraste (CSF). Par conséquent, l'enveloppe en forme de U inversé reflète la propriété du système visuel.

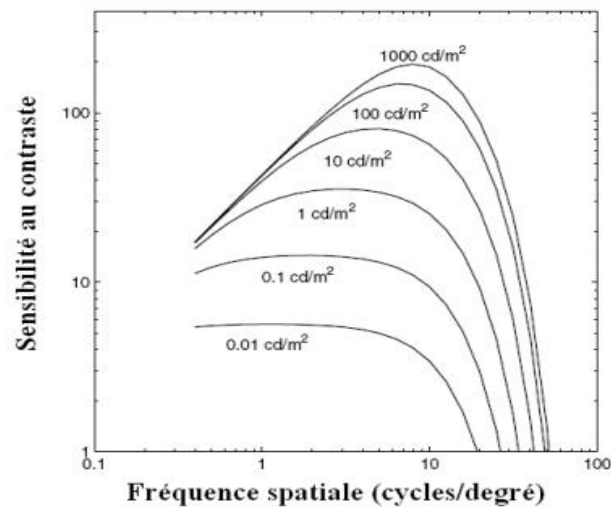


Figure 2-9: Illustration du CSF en fonction de la fréquence spatiale et la luminance du fond [63]

On constate que la CSF peut être considérée comme un filtre passe-bande dans lequel la sensibilité est la plus élevée aux moyennes fréquences autour de 2 à 4 cycles / degré et descend très rapidement aussi bien pour les basses fréquences et les hautes fréquences (figure 2-8). En liant cette conclusion avec le tatouage numérique, nous constatons que nous pouvons insérer plus dans les moyennes fréquences qui en traitement d'image représentent les zones de texture et insérer moins dans les basses et hautes fréquences qui représentent les zones uniformes et les zones de contours respectivement.

## **2.3 Les masques psychovisuels**

### **2.3.1 Le masquage**

Les signaux visuels interagissent entre eux et souvent de manière non linéaires. L'effet non linéaire le plus important est l'effet de masquage. Le phénomène de masquage fait référence aux changements de visibilité d'un stimulus (ou signal masqué) en fonction de la présence



d'un autre (signal masquant) [64]. La visibilité du stimulus peut être soit moins visible, et dans ce cas on parle d'effet de masquage, soit plus visible, et on parle alors d'effet de facilitation. Ce phénomène est accentué quand le signal masqué et le signal masquant sont d'orientations et de fréquences voisines. La notion de masquage est donc liée à l'orientation et à la fréquence des signaux que l'on observe. L'étude du SVH ci-dessus permet de retenir certains éléments exploitables dans le domaine du tatouage des images. Certains aspects, tels que la notion de contraste, la perception de la luminance ou encore l'effet de masquage, peuvent apporter des éléments utiles au domaine du tatouage.

Pour illustrer ceci, considérons pour les exemples les deux images de la figure 2-9 : à gauche, l'image populaire Lena est représentée, à droite une copie de celle-ci avec ajout d'un bruit de variance uniforme. Comme on peut le voir, le bruit ajouté est complètement imperceptible sur les plumes du chapeau de Lena, il au contraire très facilement perçu sur les zones plates de l'image (par exemple le fond ou les épaules de Lena) et il est légèrement perceptible autour des contours d'objets. En outre, le bruit est moins visible sur des régions très sombres et très lumineuses.



Figure 2-10: L'image Lena (à gauche) et une copie de celle-ci avec un bruit uniforme ajouté avec une variance uniforme

Barni [65] résume ces observations qui intègrent l'ensemble des connaissances du SVH dans les quatre règles suivantes pour une image en niveau de gris :

- Règle 1: les perturbations sont moins visibles sur les régions texturées que sur des zones uniformes.
- Règle 2: les contours sont plus sensibles à l'addition du bruit que les régions très texturées mais moins sensibles que les zones uniformes.

- Règle 3: les perturbations sont moins visibles dans les régions sombres et très lumineuses.
- Règle 4: l'œil est plus sensible aux perturbations sur les hautes fréquences verticales ou horizontales que celles à  $\pm 45^\circ$ .

Le tatouage perceptuel utilise ces propriétés pour masquer la marque efficacement. Dans la littérature, on distingue différents types de masquage : le masquage de luminance (adaptation de la luminance) et le masquage de contraste (contours et textures). Le masque final sera appelé JND (Just noticeable difference).

L'utilisation du modèle JND pour ajuster la force d'insertion d'une manière optimale et adaptative par rapport aux zones de l'image est devenue une direction intéressante dans la communauté du tatouage car il s'agit d'une clé pour résoudre le compromis entre la robustesse et la transparence. Dans la sous-section suivante, nous décrivons quelques modèles psychovisuels importants qui sont conçus dans différents domaines, c'est-à-dire le domaine spatial, DCT, et ondelettes.

### **2.3.1.1 Masques psychovisuels dans le domaine spatial**

Un JND typique dans le domaine spatial considère deux facteurs majeurs pour chaque pixel d'une image : le masquage de luminance et le masquage de contraste.

Chou et Li [66] ont proposé un modèle JND dans le domaine spatial développé pour la compression des images. Ce modèle prend en compte l'adaptation de la luminance et le masquage de la texture. Le modèle d'adaptation à la luminance est inspiré de l'expérience de Moon-Spencer mais adapté aux images naturelles. Le masquage de la texture est déterminé en fonction des activités spatiales locales de la région (par exemple le calcul des gradients autour du pixel). La détection de la texture est effectuée à l'aide de filtres passe-haut simples dans quatre directions. Le JND final est alors considéré comme l'effet dominant entre le masquage de la texture et le masquage de la luminance :

$$JND_f(x, y) = \max(JND_l(x, y), JND_t(x, y)) \quad (2-5)$$

où  $JND_l(x, y)$  and  $JND_t(x, y)$  sont le masque de la luminance et le masque de la texture respectivement.

Dans [67,68], Yang *et al.* ont apporté quelques modifications sur le modèle de Chou et Li et l'ont utilisée dans le tatouages des séquences vidéos. Tout d'abord, ils ont modifié le

masquage de la texture en intégrant un facteur pour distinguer entre les régions de contours et de texture, car le SVH est très sensible aux distorsions autour des contours mais pas dans les textures. Deuxièmement, le masque final composé du masquage de la luminance et du masquage de la texture est maintenant décrit comme une combinaison additive de facteurs individuels :

$$JND_f(x, y) = JND_l(x, y) + JND_t(x, y) - C_{l,t}(x, y) \min(JND_l(x, y), JND_t(x, y)) \quad (2-6)$$

où  $C_{l,t}$  ( $0 < C_{l,t} < 1$ ) est un facteur empirique qui explique l'effet de chevauchement entre ces deux types de masquage

Qi *et al.* dans [69], ont proposé un modèle JND pour le tatouage des images. Ils ont utilisé le JND proposé par Chou et Li comme étant leur masque de luminance. Pour le masque de la texture ils ont proposé de faire un calcul de moyenne des pixels et pour le calcul du masque de contour un calcul du Laplacien a été fait.

Cependant tous ces modèles considèrent uniquement les trois premières règles résumées par Barni sans tenir compte de la sensibilité à l'orientation décrite par la quatrième règle. Le JND développé par Chou et Li à l'avantage d'être basé sur des expériences psychovisuels mais son inconvénient réside dans le fait qu'il ne différencie pas entre les zones de contours et de textures. L'inconvénient du JND proposé par Yang & al est que leur JND final est basé sur un choix non expliqué. Et pour Qi & al l'inconvénient majeurs de leur modèle JND et de considérer tout le masque final de Chou & Li comme un masque de Luminance.

### **2.3.1.2 Masques psychovisuels dans le domaine DCT**

L'un des modèles JND les plus connus basé sur DCT est celui développé par Watson *et al.* [70]. Ce modèle est devenu une référence citée par de nombreux chercheurs. Dans ce modèle, le seuil JND pour chaque sous-bande DCT est calculé en incorporant les mécanismes les plus importants du SVH tels que la fonction de sensibilité au contraste (CSF), l'adaptation de luminance et le masquage de contraste :

Le CSF est intégré par la procédure de détermination du seuil de visibilité pour chaque fréquence de base. Cette procédure s'effectue au moyen de tests subjectifs décrits dans [71, 72]. Dans ces expériences, chaque fréquence est d'abord transformée en fréquence spatiale (des cycles / unité de degré). Ensuite, les seuils de détection (seuil de contraste) pour les fréquences de base unique ont été mesurés par une méthode TAFC (Two Alternative Forced

Choice). Pour illustrer l'effet de l'adaptation de la luminance, Watson *et al.* ont utilisé une méthode inspirée des expériences réalisées par Ahumada *et al.* [73]. Cependant, pour des raisons de simplicité, ils se rapprochent de la formule proposée par Ahumada *et al.* par une fonction de puissance plus simple :

$$t_{i,j,k} = t_{i,j} \left( \frac{c_{0,0,k}}{c_{0,0}} \right)^{a_L}$$

Où :

- $a_L$  est un facteur empirique pour contrôler le degré de masquage,
- $c_{0,0,k}$  est le coefficient DC du DCT pour le bloc  $k$ ,
- $c_{0,0}$  est le coefficient continu de l'image entière correspondant à la luminance moyenne de l'affichage (1024 pour une image de 8 bits).
- $t_{i,j,k}$  est le seuil de détection (qui représente le CSF et l'adaptation de la luminance) de la fréquence  $(i, j)$  dans le bloc  $k$
- $t_{i,j}$  est le seuil de visibilité qui relève uniquement du CSF.

Le calcul de  $t_{i,j}$  doit être effectué à la luminance moyenne de l'affichage.

$$JND_f(i, j, k) = \max \left[ t_{i,j,k}, t_{i,j,k} \cdot \frac{c_{i,j,k}}{t_{i,j,k}} \right]^{a_c} \quad (2-7)$$

où  $c_{i,j,k}$  est le coefficient DCT à la position  $(i, j)$  du bloc  $k$ ,  $a_c$  est un facteur empirique qui contrôle le degré de masquage.

Dans [74], Zhang *et al.* ont donné quelques ajustements sur le modèle Watson en modifiant l'adaptation de luminance et les étapes de masquage du contraste. Tout d'abord, ils ont adapté le modèle d'adaptation à la luminance de Chou et Li [66]. Cela semble donner plus de précision car le SVH est plus sensible dans des régions à luminosité moyenne que dans des régions très sombres ou très lumineuses. Deuxièmement, ils ont proposé de supprimer le masquage de contraste pour les blocs de contours puisque le SVH est supposé être sensible au changement sur les contours.

Même si le modèle JND proposé par Watson reste très utilisé mais son utilisation dans le tatouage numérique n'est pas avantageuse. C'est dû au fait qu'il est basé sur la DCT Cette approche basée sur un bloc entraîne la perte de corrélation entre les blocs [3]. Malgré les bons résultats rapportés par leur méthode, la complexité du modèle de Zhang & al est

préoccupante. La classification des blocs utilisés est assez complexe dans laquelle de nombreux paramètres sont choisis sans aucune explication.

### **2.3.1.3 Masques psychovisuels dans le domaine DWT**

Dans [75], Z. Liu a proposé un modèle JND pour la transformée en ondelettes discrètes de Daubechies 9/7. Les fonctions de correction de luminance et de masquage de contraste utilisées dans ce modèle sont exactement les mêmes que celle du modèle de Watson (avec les mêmes paramètres). Pour la modélisation CSF, ils utilisent les seuils de détection d'ondelettes mesurés dans les expériences de Watson *et al.* [76].

Barni et Bartolini [77] ont proposé un tatouage perceptuel dans le domaine DWT (Discrete Wavelet Transform). La faiblesse de la méthode provient de la pondération dite perceptuelle de 3 termes qui est loin d'être bien justifié. Aussi comme l'insertion est faite uniquement dans le niveau de résolution le plus élevé, les informations sur le tatouage peuvent être facilement effacées par un attaquant potentiel.

Nguyen *et al.* [78] a proposé un modèle JND dans le domaine DWT en utilisant une représentation pyramidale. Il intègre trois fonctions : sensibilité au contraste, adaptation à la luminance et masquage de contraste, mais ne prend pas aussi en compte la règle 4 de Barni sur l'orientation.

## **2.3.2 Evaluation de la qualité visuelle**

Cette section traite le problème liée à la mesure de l'impact perceptuel d'une image tatouée. En fait, la transparence d'un tatouage peut être évaluée soit par la notion de la fidélité, soit par celle de la qualité. Ainsi avant d'aborder l'état de l'art sur les méthodes d'évaluation de la qualité visuelle d'image, rappelons d'abord les notions de fidélité et qualité afin de lever toute ambiguïté.

- **Fidélité:** se réfère à la notion de conservation de l'information, c'est une mesure de similitude (ou de différence) entre le contenu original et tatoué. Une fidélité élevée signifie que le contenu tatoué est très similaire à sa version originale ou différemment.
- **Qualité:** La qualité d'image est quant à elle une notion très liée à la perception visuelle humaine. C'est une mesure absolue de l'acceptabilité, effectuée sur le contenu tatoué sans avoir référence au contenu original. Il y a aussi une dimension

liée à notre faculté d'apprentissage. En effet, notre capacité à évaluer la qualité d'une image sans avoir recours à l'originale tient de ce que notre système possède la faculté d'emmagasiner et d'apprendre une quantité énorme d'information. Évaluer la qualité d'une image c'est lui associer un ou plusieurs qualificatifs permettant de situer sa position relative dans un référentiel défini par notre sens et selon l'application envisagée.

L'évaluation de la qualité des images tatouées est habituellement effectuée en utilisant deux méthodes : évaluation subjective et objective.

L'évaluation subjective fait appel à des observateurs humains pour évaluer (ou comparer) la qualité d'une image selon un protocole bien défini. Elle est considérée comme le moyen le plus fiable et donc la référence pour comparer les différentes métriques proposées.

L'évaluation objective fait référence aux méthodes basées sur l'analyse et la mesure quantitative du niveau de dégradation au moyen de métriques directement liées au signal physique. La tendance actuelle est d'exploiter au mieux les connaissances des mécanismes de la perception visuelle de façon à tendre vers les résultats de l'appréciation subjective. En s'appuyant sur ces définitions, nous présentons dans ce qui suit les méthodes d'évaluations de qualité subjectives et objectives.

### **2.3.2.1 Evaluation subjective**

L'évaluation subjective implique des techniques qui utilisent les êtres humains pour faire le jugement. Ce dernier est amené à juger la qualité des images qui lui sont présentées selon une grille d'évaluation à plusieurs niveaux d'appréciation. À l'issue de ces tests, une note subjective appelée MOS (Mean Opinion Score) est obtenue. Cependant, le jugement humain peut varier considérablement d'un individu à l'autre en fonction de la sensibilité personnelle. Par conséquent, les tests subjectifs exigent un grand nombre de sujets et d'essais. De plus, de telles expériences sont de nature statistique et devraient être réalisées conformément aux procédures psychophysiques. Donc lors de la réalisation de ces tests, il est important de porter une attention particulière à certains facteurs pouvant altérer les jugements des participants. Certaines procédures courantes habituellement utilisées pour un test subjectif sont décrites ci-dessous.

### **2.3.2.1.1 Les facteurs influant**

L'UIT-T P.910 [79] et l'UIT-R BT.500-11 [80] sont des normes proposées pour une évaluation subjective de la qualité subjective de l'image et de la vidéo, mais peuvent également être inspirées et utilisées dans le tatouage numérique. Ces normes spécifient certaines questions à considérer dans un test subjectif tel que les procédures, les conditions, les matériaux, les méthodes d'analyse des résultats, etc. Pour s'assurer de la fiabilité de l'évaluation subjective, il convient de tenir compte de certains facteurs pouvant influencer grandement le jugement des observateurs [81]. Nous listons ici les plus pertinents.

**Distance d'observation :** La visibilité d'un stimulus dépend fortement de la distance d'observation. Il est recommandé de fixer cette distance à 6 fois la hauteur de l'image projetée et de la garder durant tout le temps du test.

**Conditions de visualisation :** L'éclairage lié à la luminosité de la salle ainsi que l'environnement jouent aussi un rôle très important dans la perception des stimuli. Un éclairage élevé risque d'aveugler l'observateur et ainsi affaiblir sa perception.

**Ecran :** Il est nécessaire de faire le calibrage de l'écran afin de reproduire les couleurs désirées. Sans cette calibration, les couleurs affichées peuvent être différentes du stimulus d'origine. Il faut aussi tenir compte des artefacts liés à la technologie de l'écran utilisé durant le test [82].

**Le choix des images :** Un échantillon d'images de contenus visuels assez variés et le plus représentatif possible des scènes selon l'application visée doit être utilisé.

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

**Les observateurs :** Le choix des observateurs doit être établi en fonction de l'application visée par les tests et le plus représentatif possible de la population moyenne concernée. Des

experts et des non-experts du domaine de l'application peuvent être admis. Il est généralement admis que les tests sont statistiquement valides si le nombre des participants est supérieur ou égal à 15.

En conclusion, pour minimiser les erreurs de jugements des observateurs, tous ces paramètres doivent être pris en compte et notés attentivement.

### **2.3.2.1.2 Les protocoles**

Différentes méthodes d'évaluation subjective ont été définies par l'ITU [81]. On retrouve essentiellement trois types de protocoles : les tests à simple stimulus permettant de juger la qualité visuelle des images sans avoir recours aux informations relatives à l'image originale. Les tests à double stimuli où il est demandé à un observateur de juger la similarité selon la qualité visuelle entre deux images et les tests comparatifs dont l'objectif principal est de comparer deux ou plusieurs images. Ces méthodes sont décrites ci-après.

- **Simple stimulus**

La méthode à simple stimulus, appelé "Single Stimulus Continuous Quality Scale (SSCQS)" permet de mesurer la qualité d'une image. Dans ce type de méthode, une seule image ou séquence d'images est présentée à l'observateur qui donne une note à l'ensemble de la présentation. Il est à noter que le mode de présentation, illustré par la figure 2-10, se voit ainsi limité. Les images sont présentées une par une avec un délai entre deux présentations. Ce temps permet à l'observateur de noter la qualité visuelle de l'image. Dans ce cas, les observateurs attribuent à une image ou une séquence d'images une catégorie choisie parmi un ensemble de catégories définies d'un point de vue sémantique. Les catégories peuvent traduire la présence ou l'absence d'une dégradation, par exemple, pour établir le seuil de dégradation. Les échelles par catégories permettant d'évaluer la qualité de l'image et la dégradation de l'image, qui ont été utilisées dans la plupart des cas sont celles proposées par l'UIT-R (Tableau 2-1).



Echelle à cinq notes			
Qualité		Dégradation	
5	Excellent	5	Imperceptible
4	Bon	4	Perceptible mais non gênant
3	Assez bon	3	Légèrement gênant
2	Médiocre	2	Gênant
1	Mauvais	1	Très gênant

Tableau 2-1 : Échelles de qualité et de dégradation de l'UIT-R

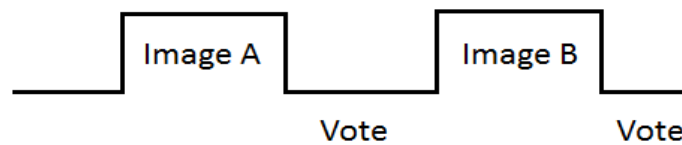


Figure 2-11: Méthode à simple stimulus.

- **Double stimulus**

Cette méthode, appelée « Double Stimuli Continuous Quality Scale » (DSCQS), a pour but de mesurer la qualité ou fidélité d'une image par rapport à sa version originale. L'image originale est présentée à l'observateur, suivi d'un écran gris et puis de sa version dégradée (ou pas). Un second écran gris est affiché, il permet à l'observateur de faire son vote. Un exemple de présentation est illustré par la figure 2-11. Il est à noter que le temps de présentation des deux images doit être identique. Une autre procédure consiste à présenter à l'observateur une paire d'images, chacune provenant de la même source, l'une dégradée et l'autre provenant directement de la source. L'observateur est prié d'évaluer la qualité des deux images au cours de séances qui durent au plus 30 min. A la fin des séances expérimentales, on calcule les moyennes pour chaque condition expérimentale et chaque image d'essai présentée.

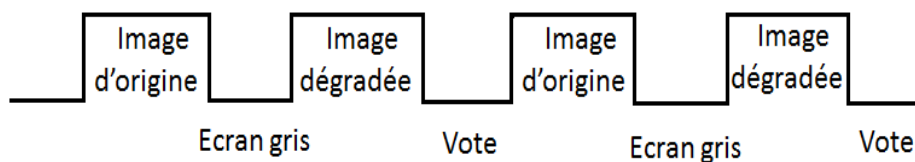


Figure 2-12: Méthode à double stimulus.

- **Comparatif**

C'est le protocole expérimental que nous avons utilisé dans les tests subjectifs pour l'évaluation de la qualité visuelle de notre modèle de tatouage.

L'objectif des méthodes comparatives est de comparer la qualité de deux ou plusieurs images. Avec cette méthode nous souhaitons quantifier la perception des différences entre les images sans avoir recours aux informations liées aux stimuli. De ce fait, l'image jugée avoir la meilleure qualité visuelle par l'observateur peut être de mauvaise qualité. Une fois les tests terminés, les résultats obtenus sont analysés et combinés de manière à calculer une note unique pour chaque image décrivant sa qualité moyenne. Cette note appelée « Mean Opinion Score » (MOS) est donnée par la formule suivante :

$$Mos(i) = \frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} Note_i(j) \quad (2-8)$$

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

Un intervalle de confiance est souvent associé à chaque note MOS, permettant ainsi de réduire l'impact d'éventuelles erreurs. Il est généralement fixé à 95% et est donné par la relation suivante [80]:

$$[MOS(i) - e_j, MOS(i) + e_j]$$

où

$$e_j = 1,96. \sigma_j$$

et

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

### 2.3.2.2 Evaluation objective

Bien que l'évaluation subjective soit la solution appropriée et la mesure de référence pour l'évaluation de l'imperceptibilité du tatouage, elle est généralement inconmode, coûteuse et longue à effectuer, et donc pas toujours facile à utiliser. De ce fait plusieurs travaux ont été menés dans ce sens, donnant ainsi naissance à une abondance de méthodes d'évaluation de

la qualité d'image objective. Certaines sont basées sur des approches mathématiques telles que l'erreur quadratique moyenne. D'autres intègrent quelques caractéristiques et descripteurs de l'image (le contraste, le contenu fréquentiel, les structures géométriques) ou tentent de reproduire les processus d'appréciation de la qualité d'image via des modèles des mécanismes basé sur le SVH. Ainsi, l'évaluation objective est utilisée comme méthode alternative. L'objectif de cette méthode est d'évaluer la qualité / fidélité des contenus visuels au moyen d'outils automatiques (métriques objectives) sans utilisation d'observateurs humains. Un aperçu de certaines de ces mesures objectives est donné par la suite

Les mesures de qualité objective proposées dans la littérature peuvent être classées en 3 catégories selon que l'on dispose ou non de l'image originale. On a alors les métriques avec référence (Full Reference (FR)), sans référence (No Reference (NR)) et avec référence réduite (Reduced Reference (RR)).

**Mesures avec référence :** sont utilisées pour mesurer la qualité entre une image originale et sa version dégradée. Le score final de qualité est basé sur des informations extraites à la fois de l'image originale et de la version déformée. Généralement, aucune information liée au type de dégradation n'est exploitée. Elles sont donc supposées être universelles. La plupart des métriques sont développées avec référence. Bien que cette approche donne généralement de bonnes performances, la nécessité de l'image originale est parfois une contrainte limitant son attrait.

**Mesures sans référence :** comme cela peut être facilement compris, elles sont utilisées pour mesurer la qualité d'une image de test sans référence. Orientés vers une dégradation spécifique, les mesures sans référence sont les plus captivantes car elles estiment la qualité de l'image en utilisant uniquement l'image dégradée. Cependant, le manque complet d'informations de référence rend la conception de ces paramètres une tâche très difficile. Habituellement, on suppose que le type de dégradation contenu dans l'image est connu *a priori*.

**Mesures avec référence réduite :** elles ont été développées afin de surmonter les contraintes liées aux mesures avec et sans référence, en donnant une solution intermédiaire. Ce type de mesure permet de quantifier la qualité d'une image à partir d'un certain nombre de fonctions

de l'image originale et sans aucune information *a priori* sur le type de dégradation contenu dans l'image. Bien que ces méthodes soient relativement accessibles pour certaines applications, elles restent les moins populaires.

Nous choisissons de présenter six métriques de qualité d'image Full Reference (FR) populaires, à savoir EQM, PSNR, SSIM [83], UQI [84], WPSNR [85,86], SVD [87]. Comme mentionné dans la première section, il n'y a pas de métrique objectif spécifiquement conçu pour le tatouage d'image.

Nous proposons de classer ces métriques en trois sous-ensembles : Les métriques basées sur l'erreur quadratique moyenne (EQM), celles basées sur l'analyse des structures locales de l'image et celles inspirées du SVH.

### 2.3.2.2.1 Approche EQM

L'erreur quadratique moyenne est l'une des métriques les plus connues et utilisées à ce jour. Elle est obtenue en calculant la moyenne quadratique du signal erreur ou distorsion. Elle est donnée par :

$$EQM = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \|I_o(m,n) - I_w(m,n)\|^2 \quad (2-9)$$

où  $I_o$  et  $I_w$  sont respectivement l'image originale et l'image tatouée de dimension  $M \times N$ .

- **PSNR** (Peak Signal To Noise Ratio) : est une métrique qui calcule le MSE (Mean Square Error) entre les images originales et déformées sa valeur est souvent donnée en décibels (dB). Malgré ses mauvaises performances, le PSNR est largement utilisée non seulement dans la communauté de qualité d'image, mais aussi dans de nombreux autres domaines.

$$PSNR = 10 \times \log_{10} \left( \frac{d^2}{EQM} \right) \quad (2-10)$$

où  $d$  est la dynamique de l'image (la valeur maximum possible pour un pixel). Dans le cas standard d'une image dont les pixels sont codés sur 8 bits,  $d = 255$ .

### 2.3.2.2.2 Approche structurelle

Pour résoudre les problèmes liés à la faiblesse des mesures de type EQM, basées sur le calcul du pixel, des auteurs se sont orientés vers les métriques de qualité basées sur une analyse des

structures locales du signal image. Parmi les métriques développées que nous trouvons dans la littérature, nous avons celles proposées par Wang *et al.* appelées UQI et SSIM [83,84]

- **UQI** (Universal Quality Index) est un indice universel d'indice de qualité objectif, conçu pour modéliser toute distorsion en tant que combinaison de trois facteurs : perte de corrélation, distorsion de luminance et distorsion de contraste. UQI est une métrique mathématiquement définie sans modèle HVS incorporé et donc faible complexité.
- **SSIM** (Structural SIMilarity) est une métrique objective entièrement basée sur la métrique UQI. Cela suppose que la perception visuelle humaine soit très adaptée pour extraire l'information structurelle d'une scène. L'indice SSIM est basé sur une combinaison de comparaison de luminance, de contraste et de structure. Les comparaisons sont effectuées sur les fenêtres locales, la qualité globale de l'image est calculée en moyenne sur ces fenêtres locales. SSIM est devenu une mesure très connue pour l'évaluation de la qualité de l'image perceptive et a été étendue dans diverses directions. Dans [88], Wang *et al.* a proposé une version multi-échelle de SSIM (MSSIM) où les images sont filtrées à basse fréquence et échantillonnées à la baisse par un facteur de deux et le contraste et la structure sont calculés pour chaque niveau de sous échantillonnage

$$SSIM = \frac{1}{W} \sum_{i=1}^W l(i) \cdot c(i) \cdot s(i), \quad (2-11)$$

$$\text{où} \quad l(i) = \frac{\overline{2x(i) \cdot y(i)} + C_1}{\overline{x(i)^2 + y(i)^2} + C_1}, \quad c(i) = \frac{2\sigma_x(i) \cdot \sigma_y(i) + C_2}{\sigma_x^2(i) + \sigma_y^2(i) + C_2}, \quad s(i) = \frac{2\sigma_{xy}(i) + C_3}{\sigma_x(i) + \sigma_y(i) + C_3}$$

$$C_1 = (L \cdot K_1)^2, C_2 = (L \cdot K_2)^2, C_3 = \frac{C_2}{2},$$

où  $L$  est la dynamique des niveaux de gris.  $K_1$  et  $K_2$  sont deux constantes fixées par les auteurs à 0.01 et 0.03, respectivement.  $\sigma_x$  et  $\sigma_y$  peuvent être considéré comme des estimations du contraste de  $x$  et  $y$ .

### 2.3.2.2.3 Approche basé sur le SVH

Nous avons ici quelques métriques inspirées du SVH. Ces méthodes ont permis de se rapprocher du comportement du SVH.

- **SVD** (Singular Value Decomposition) est une métrique de qualité d'image mathématique pure utilisant la SVD. La différence entre les valeurs singulières entre

les blocs est calculée. L'indice de distorsion finale est obtenu par une simple sommation sur tous les blocs.

- **WPSNR** (weighted PSNR) : est une métrique de qualité d'image conçue pour l'évaluation de la transparence du tatouage dans le cadre du projet de référence Checkmark. Cette métrique prend en compte une propriété implicite que le SVH est moins sensible dans les régions à activité élevée (texture) que dans les régions uniformes. À cette fin, les auteurs ont modifié le PSNR par un facteur appelé NVF (Noise Visibility Function) qui caractérise l'énergie d'image locale. L'expression du WPSNR donné par [89] est la suivante :

$$WPSNR = 10 \log_{10} \frac{\max(I_o)^2}{\|(I_w - I_o) \cdot NVF\|^2} \quad (2-12)$$

avec

$$NVF = \frac{1}{1 + \theta \sigma_{I_o}^2(i, j)}$$

où  $\sigma_x^2(i, j)$  désigne la variance locale de l'image dans une fenêtre centrée sur le pixel avec des coordonnées  $(i, j)$  et  $\theta$  est un paramètre de réglage qui joue le rôle du réglage du contraste.

## 2.4 Conclusion

L'objectif de ce chapitre n'était pas de fournir une description détaillée du SVH, mais plutôt de décrire de manière synthétique le SVH afin de fonder les développements méthodologiques et expérimentaux dans le reste de la thèse. À cet égard, les principales caractéristiques du SVH ont été étudiées : dans une première partie, nous avons introduit les différentes composantes du SVH, ainsi que les constituants fondamentaux qui sont essentiels au développement de modèles psychovisuels. Ensuite une section a été consacrée à certains modèles JND importants du SVH qui sont déjà caractérisés et exploités avec succès dans le domaine du traitement d'image et en particulier pour le tatouage. Enfin nous avons introduit l'évaluation subjective de la qualité des images, en présentant les différents protocoles ainsi que les principaux facteurs influençant l'appréciation humaine. Comme nous l'avons constaté, l'étude du système visuel humain permet de définir en tout point de l'image les forces maximales applicables lors du tatouage. Ceci permet donc d'insérer la marque adaptée au seuil de visibilité, garantissant ainsi le meilleur compromis entre l'invisibilité et la robustesse.

## **Chapitre 3**

### **3. Contribution dans le tatouage psycho-visuel : JND par discrimination de texture à base d'ondelettes de Faber-Schauder**

#### **3.1. Introduction**

Le tatouage d'image psycho-visuel consiste à utiliser la connaissance du Système Visuel Humain (SVH) pour choisir la force du tatouage en fonction des propriétés locales de l'image. Barni a synthétisé ces travaux et a défini quatre règles fondamentales pour exploiter les limites du SVH pour des applications de compression ou de tatouage :

- Règle 1 : les perturbations sont beaucoup moins visibles sur les régions fortement texturées que sur les zones uniformes.
- Règle 2 : les contours sont plus sensibles à l'ajout de bruit que les régions très texturées mais moins sensible que les zones uniformes.
- Règle 3 : les perturbations sont moins visibles sur les régions sombres et très claires.
- Règle 4 : l'œil est plus sensible aux perturbations sur les hautes fréquences verticales ou horizontales que celles à  $\pm 45^\circ$ .

Nous avons vu dans le chapitre précédent plusieurs méthodes de masques psycho-visuels, notamment celles de Chou & Li et celle de Barni & al, pour la compression et le tatouage numérique. Mais toutes ces méthodes présentent des lacunes qui se manifestent notamment par un choix des règles de calcul de la force non suffisamment justifiés car éloignées des règles de base, par de nombreux paramètres ad-hoc, et aussi par le fait qu'elles ne tiennent pas toujours compte de la sensibilité à l'orientation.

A la lecture des 4 règles de Barni, nous constatons que la classification d'une zone de l'image en texture, zone uniforme ou contour joue un rôle essentiel. Nous proposons une approche logique qui consiste en premier lieu à classer chaque pixel en fonction en trois types de régions dans les images candidates pour le tatouage : les régions uniformes, les contours et

les régions texturées. Nous proposons de faire ce classement en utilisant la transformée d'ondelette à échelle mixée, et plus particulièrement les ondelettes de Faber-schauer (FSDWT). Ceci nous va permettre de proposer un nouveau JND à base d'échelle mixée (JNDEM) qui module la force d'insertion du tatouage en fonction du type de la région. Une nouvelle méthode de tatouage d'image perceptive est ainsi proposée. Elle combine les avantages du domaine d'ondelettes avec l'utilisation d'une représentation à échelles mixées pour la séparation des régions et l'insertion modulée du tatouage dans le domaine spatial avec des méthodes existantes qui ont fait leur preuve. La transformation en ondelettes Faber-Schauer est une transformation d'ondelettes obliques qui se distingue par sa simplicité avec des opérations arithmétiques bien adaptées au traitement multiéchelle d'images. La représentation à échelles mixées des images transformées permet de représenter toutes les échelles dans une seule image en plaçant chaque coefficient d'ondelettes à l'endroit où sa fonction de base d'ondelettes est localisée. Les transformations d'ondelettes ont pour effet de décorréler l'information contenue dans l'image traitée. En effet l'information se concentre sur les coefficients d'ondelettes avec des valeurs élevées qui correspondent à des variations de niveaux de gris élevées dans l'image. Ces coefficients, où l'information est concentrée, sont appelés coefficients dominants. Une forte densité de ces coefficients dominants indique des régions à activité élevée, c'est-à-dire des textures. A l'opposé, une faible densité de coefficients dominants correspond à des régions uniformes. Les contours correspondent à une situation intermédiaire entre les textures et les régions uniformes. Les zones uniformes, de contours et de textures sont différenciées en fonction de la densité des coefficients d'ondelettes dominants.

Nous utilisons ces propriétés pour remédier aux problèmes rencontrés avec d'autres méthodes psychovisuelles en choisissant le masque JND en fonction du type de la région traité. Notre méthode offre donc la possibilité de choisir le masque optimal adapté à l'emplacement du pixel en considérant le masquage de luminance de contours ou de textures.

Nous avons d'abord testé notre JND sur deux bases de texture. Ensuite le modèle JND proposée a été testée en tatouage numérique sur une base de 50 images naturelles et comparé à d'autres modèles perceptifs de la littérature. Nous avons évalué la qualité visuelle des images tatouées via des métriques objectives en plus d'un test subjectif. Nous avons aussi évalué la robustesse de la méthode de tatouage proposé en appliquant divers attaques numériques.



## **3.2 La discrimination de texture fondée sur une représentation d'ondelettes FSDWT à échelles mixées**

S. Mallat [29] a montré dans ses travaux sur les ondelettes que si on considère une ondelette mère du type dérivée de gaussienne, la transformation en ondelettes peut être considérée comme un détecteur multiéchelles de contour de type Canny ou Marr-Hildrith. Ceci peut être étendu, en général à l'ensemble des transformations d'ondelettes : pour un seuil donné les coefficients dominants (qui dépassent ce seuil) correspondent à une activité élevée dans l'image pour les différentes échelles. Ceci permet de distinguer les régions uniformes, de contour ou celles qui sont texturées en fonction de la densité de ces coefficients dominants.

Nous avons choisi la transformée en ondelettes de Faber-Schauer (FSDWT) [33], qui constitue un bon compromis pour la discrimination régionale (régions uniformes, de contour et de texture), comme expliqué ci-après :

Pour les régions uniformes la transformation FSDWT est un compromis entre la transformation de Haar qui ne considère comme uniforme que les régions stationnaires et les transformations de Daubechies d'ordre supérieure à 2 qui incluent des situations beaucoup plus compliquées qui nécessitent beaucoup plus de calcul. En effet la méthode FSDWT permet d'inclure aussi les régions avec une variation linéaire avec un algorithme rapide et simple comparable à celui de Haar. En ce qui concerne les contours et les textures, Haar considère uniquement les contours francs, tandis que FSDWT inclut également des contours de type rampe.

En résumé, nous privilégions les ondelettes FSDWT, car la transformation de Haar est trop rigide pour faire face à des situations du monde réel, et les transformations d'ondelettes de types Daubechies sont trop sophistiquées lors d'un choix entre les régions uniformes, les contours ou les régions texturées.

Dans la représentation à échelles mixées, chaque coefficient d'ondelettes est positionné à l'endroit où la fonction d'ondelettes associée est localisée. On obtient une seule image de la transformée d'ondelettes (au lieu des images pyramidales) de la même taille que l'image originale (voir figure 3-1). Cette image à échelles mixées permet de distinguer les zones texturées et les zones de contour de l'image originale [33, 90]. Un examen fin de l'image transformée permet de voir que la concentration des coefficients dominants est plus élevée

dans les régions texturées que dans celles des contours. Dans les régions uniformes, très peu de coefficients dominants sont présents. Par conséquent, la densité des coefficients dominants peut être utilisée comme un discriminateur viable entre les zones uniformes, texturées et de contours de l'image originale [33,90].

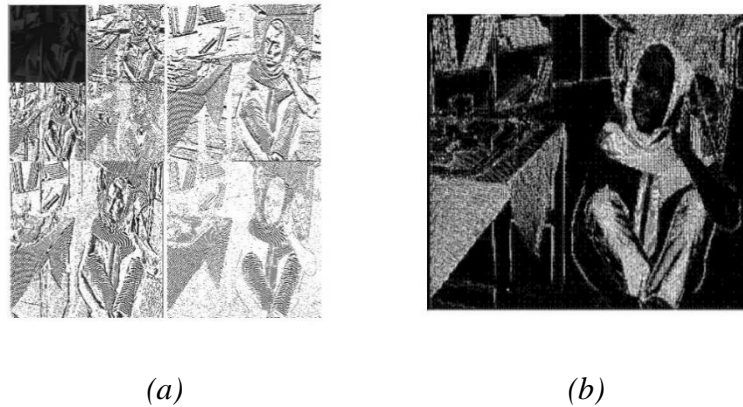


Figure 3-1 : (a) Représentation à échelle séparée de l'image Barbara, (b) : La représentation à échelles mixées de FSDWT de l'image Barbara

Nous procédons en deux étapes. La première consiste à calculer la densité des coefficients dominants : nous considérons qu'un coefficient est significatif lorsque sa valeur absolue dépasse un seuil  $Sc$  choisi empiriquement. En effet Mallat [29] et [2] suggèrent, en général pour les images naturelles, un seuil de 10% pour les coefficients d'ondelettes qui caractérisent l'image ; La seconde étape consiste alors à séparer les différentes zones de l'image à savoir, uniformes, texturés et de contours, en se basant sur la densité des coefficients dominants. Là aussi un seuil est choisi en séparant l'intervalle des coefficients dominants en trois parties.

Concrètement, nous commençons par transformer l'image en utilisant le schéma lifting de Faber-Schauer à échelles mixées (voir Chapitre 1). Nous sélectionnons ensuite les 10% coefficients avec les valeurs absolues les plus élevés ce qui nous donne une image des coefficients dominants. Enfin, pour chaque pixel de l'image des coefficients dominants, nous déterminons le nombre de coefficients dominants observés dans une fenêtre  $3 \times 3$  centrée sur le pixel. Pour la plupart des images naturelles, lorsque la fenêtre glissante est positionnée au milieu d'une région uniforme (région parfaitement plane ou linéaire), 0 ou 1 coefficient dominant est présent dans la fenêtre. Lorsque la fenêtre est centrée sur un contour de forte intensité, le nombre de coefficient dominant est élevé avec un maximum de 6. Le cas des régions texturées est plus complexe à analyser facilement mais on peut considérer qu'un maximum de 9 coefficients dominant est possible.

Dans notre méthode, les différentes régions de l'image sont caractérisées par le nombre des coefficients dominants dans une fenêtre  $3 \times 3$  glissante de la manière suivante :

- de 0 à 2 coefficients dominants, la région est considérée comme uniforme,
- 3 et 4 coefficients dominants correspondent à une région de contour,
- 5 à 9 coefficients dominants, la région est texturée.

Ces nombres de coefficients dominants ont été choisis d'après les résultats expérimentaux réalisés sur un grand nombre d'images naturelles. De plus, le nombre de coefficients dominants dans la fenêtre correspond, en divisant par 9, à une notion subjective de densité des coefficients dominants et peut être attaché à la notion d'activité de région telle que définie par Chou et Li [66].

La figure 3-2 montre, avec des images synthétiques transformées par ondelettes, les différentes régions (uniforme, contour et texturées) après discrimination par la densité des coefficients dominants. Sur la figure 3-3 la même analyse est présentée pour l'image Lena, Mandrill et Barbara.

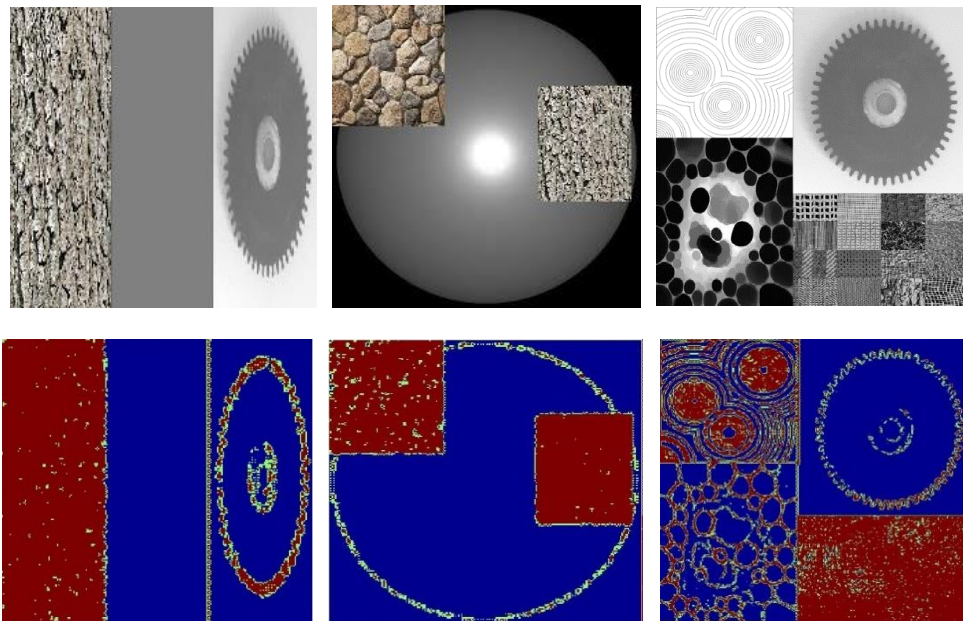


Figure 3-2: (a) exemple d'images synthétiques composées de régions : uniformes, de contours, texturées. La couleur bleue correspond à des régions uniformes avec un nombre de coefficients dominants variant de 0 à 2. La couleur verte correspond aux régions de contours avec un nombre de coefficients dominants variant de 3 à 4. Enfin, la couleur rouge correspond aux régions texturées avec un nombre de coefficients dominants variant de 5 à 9.

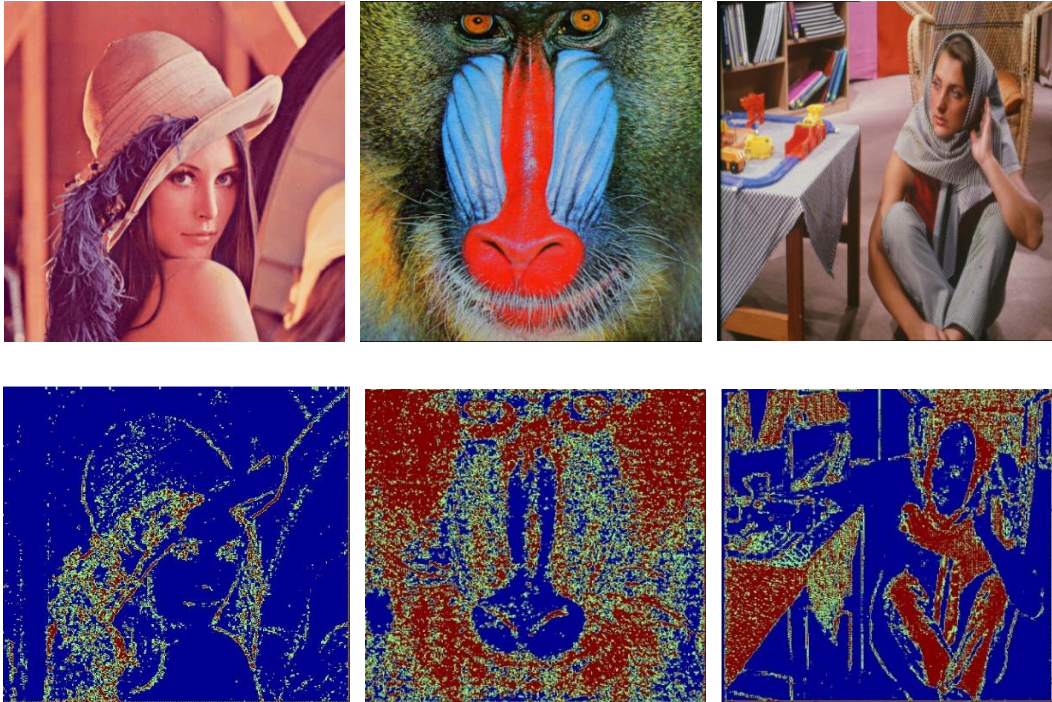


Figure 3-3: Image correspondant au coefficient dominant pour Lena, Mandrill et Barbara. Où la couleur bleu correspond aux zones uniformes, la couleur rouge pour les zones texturées et la couleur verte pour les contours.

### 3.3. Le masquage psychovisuels

Dans les sections suivantes nous allons présenter les différents masquages utilisés pour développer notre modèle JND à savoir les masques de luminance et de contour de Chou et Li [66] ainsi que le masque de texture de Qi *et al.* [69].

#### 3.3.1 Masque de luminance de Chou *et Li*

Chou et Li [66] ont proposé un masquage de luminance dans lequel le seuil de visibilité dû à la luminance de l'arrière-plan constitue le masque de luminance  $JND_L(x, y)$ . Pour développer ce masque ils ont effectué un test subjectif, dans lequel un carré de  $32 \times 32$  se trouve au centre d'un arrière-plan de luminance uniforme (background luminance  $b_g$ ) est présenté à des observateurs. Ils ont modifié la luminance du carré jusqu'à ce qu'un observateur puisse le détecter. Le résultat du test subjectif du seuil de visibilité en fonction de la luminance de l'arrière-plan est illustré sur la figure 3-4.

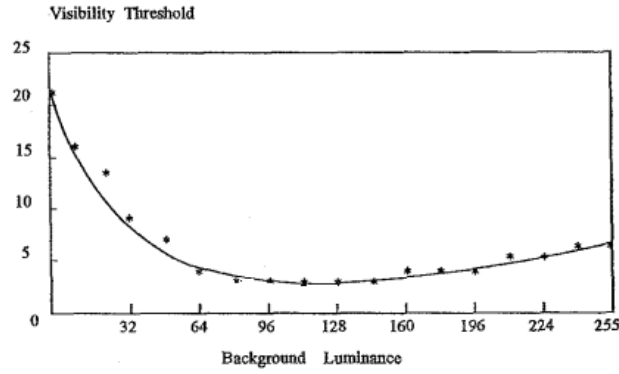


Figure 3-4: Seuils de visibilité dus à la luminance d'arrière-plan [66]

Pour les images à niveaux de gris de 8 bits, le masquage de la luminance pour le seuil correspondant à une faible luminance de l'arrière-plan (inférieur à 127) est modélisé par une fonction parabolique alors que l'autre partie (supérieur à 127) est approximée par une fonction linéaire :

$$JND_L(x, y) = \begin{cases} 17 \left( 1 - \left( \frac{b_g(x, y)}{127} \right)^{1/2} \right) + 3 & \text{for } b_g(x, y) \leq 127 \\ \frac{3}{128} (b_g(x, y) - 127) + 3 & \text{for } b_g(x, y) \geq 127 \end{cases} \quad (3-1)$$

$$b_g(x, y) = \frac{1}{32} \sum_{i=1}^5 \sum_{j=1}^5 I_o(x - 3 + i, y - 3 + j) B(x, y),$$

$$B(i, j) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 2 & 0 & 2 & 1 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Dans l'équation (3.1),  $0 \leq x < H$ ,  $0 \leq y < W$ ,  $H$  et  $W$  désignent respectivement la hauteur et la largeur de l'image  $I_o$ .  $b_g(x, y)$  est la luminance moyenne de l'arrière-plan qui est calculée en utilisant le filtre moyenneur  $B$ .

L'avantage du masque de luminance de Chou & Li est qu'il utilise la luminance locale et donne directement la valeur du pixel redondante du masquage. Le masquage montre également que lorsque la valeur du pixel d'origine est inférieure ou supérieure à un certain niveau, le seuil de masquage augmentera. Cela correspond à la règle, déjà vue dans le chapitre 2, relative à la diminution de la sensibilité des yeux lors d'une saturation en luminosité ou dans l'obscurité. Lorsque le niveau de gris est 0, c'est à dire totalement sombre,

le seuil de masquage est de 20. Cela signifie que si on change la valeur de pixel d'origine de 0 à 20, la vision humaine ne fera pas la différence. Lorsque le niveau de gris atteint 255 ce qui indique qu'il est lumineux, le seuil de masquage est 6. Ceci suggère, en particulier, que l'insertion d'un tatouage sera plus tolérée dans les zones obscures que les zones éclairées. Cependant, dans la gamme moyenne de niveaux de gris, le seuil du masquage est très faible avec un seuil de masquage d'environ 3 ce qui limite la force d'insertion d'un tatouage.

La figure 3-5 montre un résultat de ce masquage de la luminance.



Figure 3-5: Implémentation du masquage de la luminance sur l'image Barbara, les parties blanches du masquage qui représentent les zones claires et obscures correspondent aux régions avec une forte force d'insertion du tatouage.

### **3.3.2 Masque de contour de Chou et Li**

Le masquage psychovisuel des contours est déduit de tests perceptuels [66]. La procédure et le résultat du test sont présentés sur la figure 3-6. Les résultats expérimentaux relient le seuil de différence de luminance à la hauteur du contour de luminance. Ces résultats indiquent notamment que le seuil de visibilité dépend non seulement de la hauteur du contour mais aussi de la luminance de l'arrière-plan.

A partir de ces résultats expérimentaux, un modèle de masquage spatiale est calculé à chaque pixel, où le poids diminue à mesure que la distance entre le pixel voisin et le pixel central augmente.

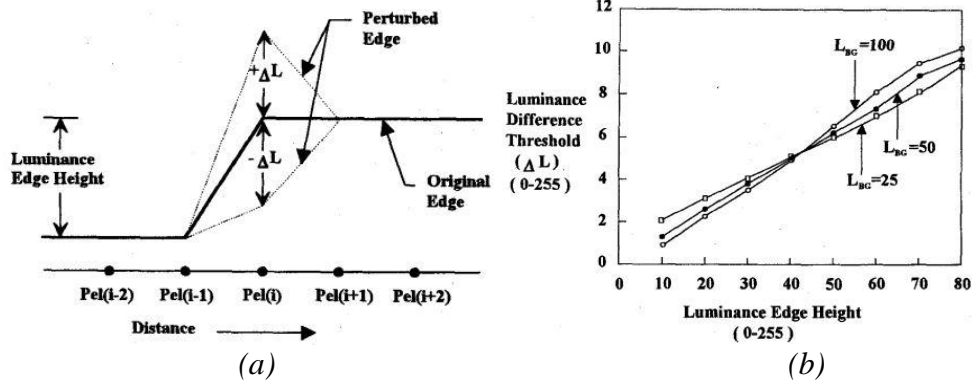


Figure 3-6: (a) Stimuli pour déterminer les seuils de visibilité dus au changement de luminosité, (b) seuils de visibilité de différence à différentes luminances de l'arrière-plan [66]

Chou et Li ont proposé de modéliser ce comportement par la fonction linéaire,  $JND_E(x, y)$  suivante :

$$JND_E(b_g(x, y), m_g(x, y)) = m_g(x, y)\alpha b_g(x, y) + \beta b_g(x, y) \quad (3-2)$$

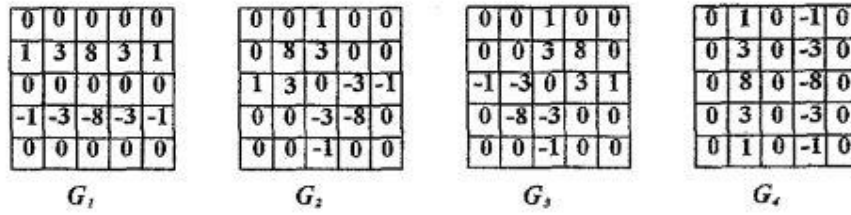
$$\alpha(b_g(x, y)) = b_g(x, y) \cdot 0.0001 + 0.115,$$

$$\beta(b_g(x, y)) = 1/2 - b_g(x, y) \cdot 0.01,$$

$$m_g(x, y) = \max_{k=1,2,3,4} \{grad_k(x, y)\},$$

$$grad_k(x, y) = \frac{1}{16} \sum_{i=1}^5 \sum_{j=1}^5 I(x-3+i, y-3+j) \cdot G_k(x, y),$$

où  $b_g$  et  $m_g$  sont la luminosité de l'arrière-plan moyenne et la différence de luminosité moyenne pondérée maximale entourant le pixel en position  $(x, y)$ , respectivement. Comme on le voit sur la figure 3-6(b), la pente approximative de la ligne qui relie le seuil de visibilité à la différence de luminosité tend à augmenter légèrement à mesure que la luminosité de l'arrière-plan augmente. Les paramètres  $\alpha(x, y)$  et  $\beta(x, y)$  sont les fonctions dépendantes de l'arrière-plan qui spécifient la pente de la ligne et l'intersection avec l'axe du seuil de visibilité.  $G_k(x, y)$  pour  $0 < k < 5$  est un opérateur pour calculer les variations moyennes de luminosité dans quatre directions :



La figure 3-7 montre le résultat d'implémentation du masquage de contour.



Figure 3-7: Implémentation du masquage de contours l'image Barbara.

### 3.3.3 Le masque de texture de Qi *et al.*

Le masquage empirique de texture que Qi *et al* [69] utilise la valeur absolue de la distance entre chaque pixel et la valeur moyenne locale des pixels dans une fenêtre glissante comme c'est indiqué dans l'équation. 3.3 :

$$JND_T(I_o) = |I_o(i, j) - \bar{I}_o(i, j)| \quad (3-3)$$

$$|\bar{I}_o(i, j)| = \frac{1}{(2L + 1)^2} \sum_{k=-L}^L \sum_{l=-L}^L I_o(i + k, j + l)$$

où  $JND_T$  est le masque de texture,  $I_o(i, j)$  est le pixel à la position  $(i, j)$ ,  $(2L + 1)^2$  représente le nombre de pixels dans la fenêtre glissante.  $L$  est choisi égale à 1 c'est à dire que la taille de la fenêtre coulissante est de 3 x 3.



Nous présentons le résultat de l'implémentation du masquage de la texture, sur la figure 3-8. Les zones blanches de l'image du masque, qui correspondent aux régions texturées, ont une marge élevée pour l'insertion d'un tatouage à cause de l'insensibilité de l'œil humaine aux variations dans ces régions comme on l'a détaillé dans le chapitre 2.

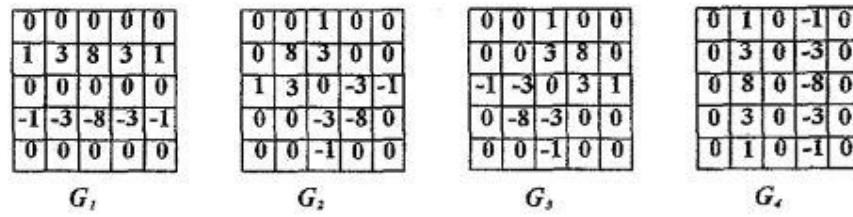


Figure 3-8 : Implémentation du masquage de texture sur l'image Barbara

### 3.3.4. Masque psychovisuel basé sur la transformé de Faber-shauder à échelles mixées

En résumé, nous proposons une nouvelle méthode de tatouage d'image perceptive qui combine les avantages du domaine d'ondelettes et du domaine spatial. Une représentation en ondelettes à échelles mixées est appliquée. Les zones uniformes, contours et texturées sont différenciées en fonction de la densité des coefficients d'ondelettes dominants [91]. Le masque de luminance développé par Chou et Li est appliqué dans le cas de zones uniformes. Le masquage développé par Qi *et al.* est choisi dans les régions texturées, finalement le masque de contours développé par Chou et Li est choisi pour les zones de contours.

Comme on a vu dans le chapitre 2, l'étude du comportement du SVH suggère en effet que l'œil est moins sensible aux stimuli à orientation diagonale (comme ceux décrits par les coefficients d'ondelettes dans les sous-bandes diagonales de la DWT). Barni suggère d'introduire un facteur d'orientation pour tenir compte de cette insensibilité dans le calcul du JND ; Ainsi pour une orientation des stimuli d'approximativement de  $\pm 45^\circ$ , le masque de texture et celui des contours sont multipliés par  $\sqrt{2}$  [65]. Pour notre JNDEM l'orientation de chaque pixel est déterminée en utilisant les filtres G1 à G4 suivant :



La figure 3-9 présente le masque perceptuel proposé.

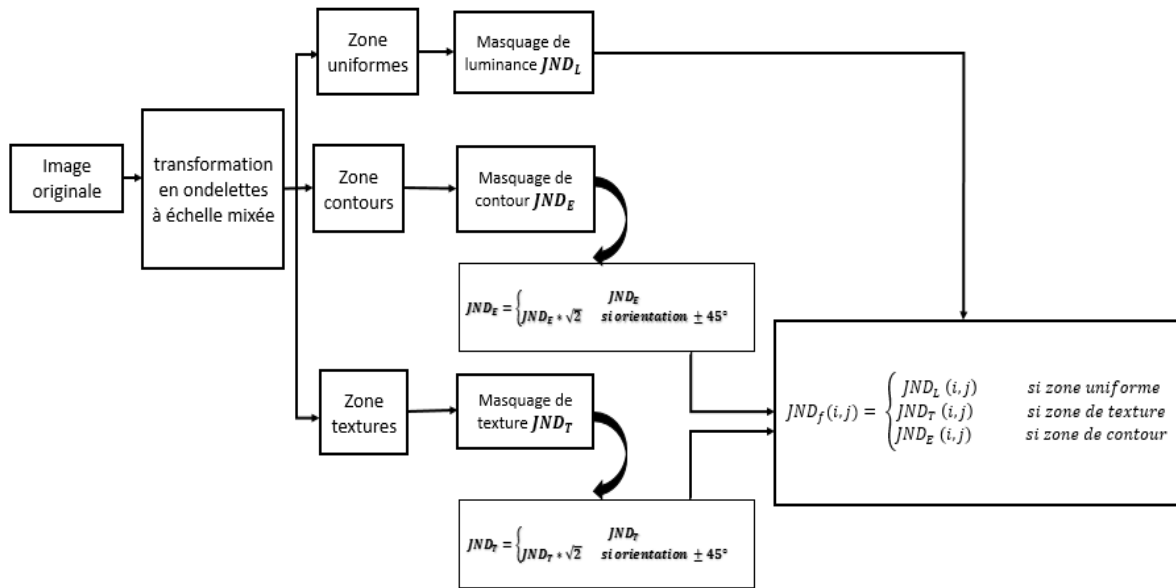


Figure 3-9: Masquage final qui combine la transformée en ondelettes à échelle mixte, pour la discrimination des zones uniformes de contour ou de texture, avec les JND de Chou et de Qi.

L'algorithme de la méthode peut être présenté comme suit (figure 3-10) :

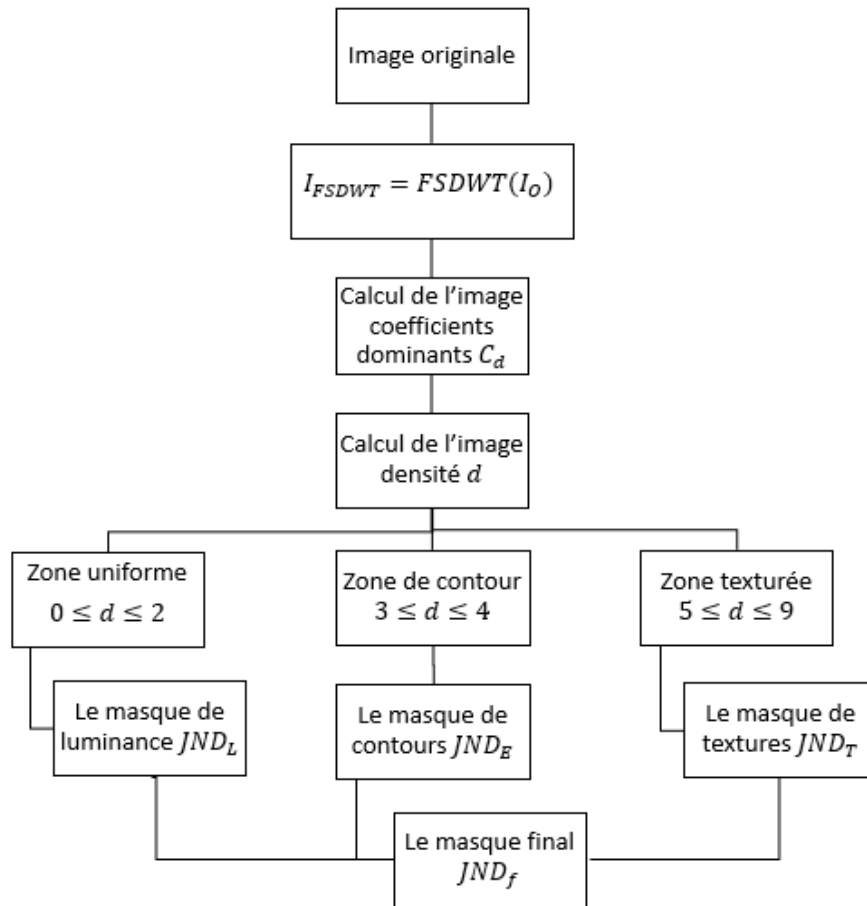


Figure 3-10: Algorithme du JNDEM

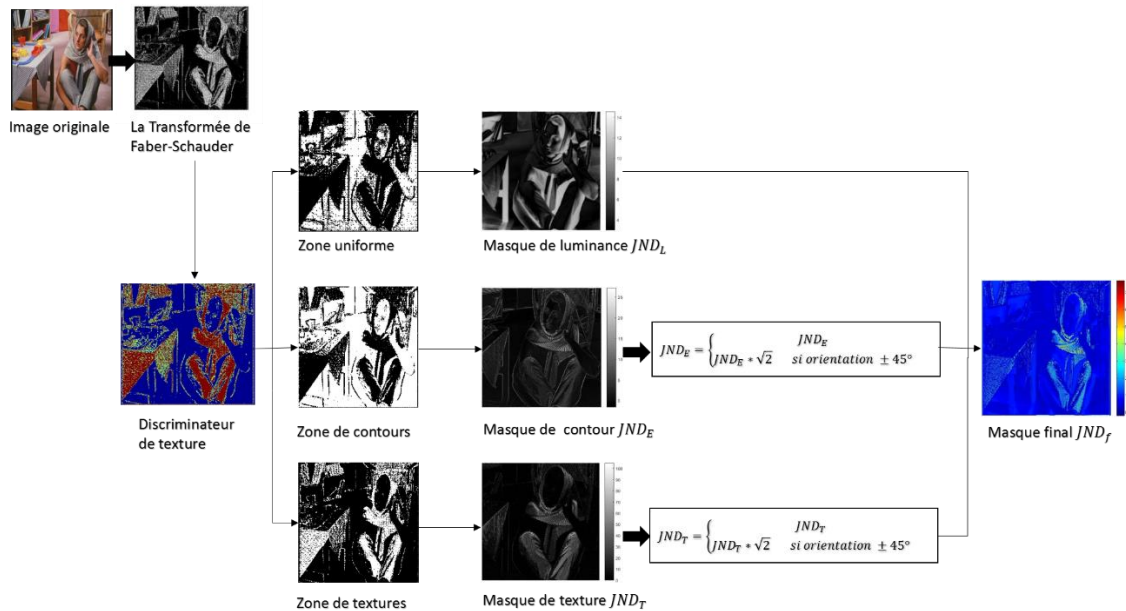


Figure 3-11: Exemple d'application de notre JNDEM

### 3.4 Schéma d'insertion du tatouage

Dans notre méthode, le tatouage a été inséré dans le domaine spatial avec une technique de spectre étalé comme décrit dans [3]. Pour insérer une séquence de  $m$  bits, une image pseudo aléatoire est divisée en  $m$  blocs mutuellement exclusifs. Chaque bloc est modulé en fonction du bit qui lui est affecté. Dans la figure 3-15 nous avons un exemple d'insertion pour une séquence de 4 bits.

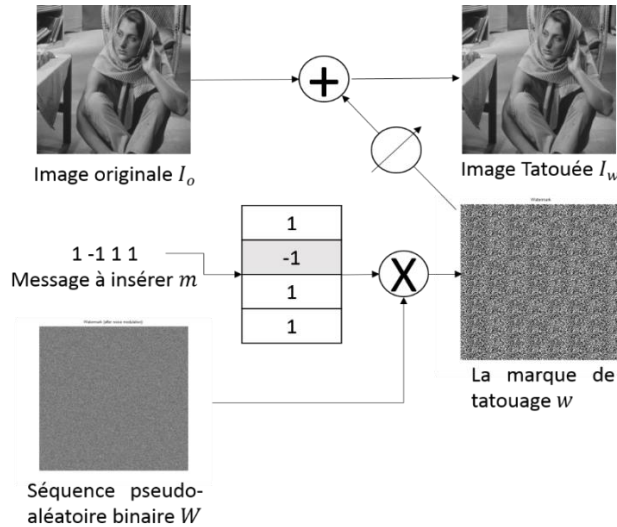


Figure 3-12: Exemple du processus d'insertion de tatouage par Etalement de spectre pour un message de 4 bit

Dans nos essais de comparaison, nous avons utilisé une séquence  $m$  de 64 bits (le message utilisé est « IRF—SIC ») ; l'image est divisé en 64 blocs et une matrice pseudo-aléatoire  $W$  de taille 64 est générée à l'aide d'une clé secrète  $k$  qui représente la graine du générateur pseudo-aléatoire de  $\{-1,1\}$ . Pour obtenir la marque du tatouage  $w_i$  du message binaire on utilise l'équation suivante :

$$w_i = m_i \cdot W$$

Avec  $m_i$  est le  $i$ ème composant de message (séquence d'étalement),  $W$  est la séquence pseudo-aléatoire binaire et  $w_i$  est la marque de tatouage sur le  $i$ ème bloc de tatouage. L'image tatouée dans le cas le plus simple, où aucun masque perceptuel n'est utilisé peut être écrite comme suit :

$$(I_w)_i = (I_o)_i + \alpha w_i \tag{3-4}$$

Avec  $\alpha$  une force d'insertion de la marque. Dans notre cas nous avons pris  $\alpha = 1$ .

Si un masque spatial est utilisé, la marque est multiplié élément par élément par le masque JNDEM. L'insertion de la marque  $w_i$  est insérée en utilisant l'équation suivante :

$$(I_w)_i = (I_o)_i + (JND_f)_i .* w_i \quad (3-5)$$

Avec  $JND_f$  le masque final et  $(.*)$  le produit matriciel point par point.

La figure 3-16 présente un exemple, sur l'image Lena, de comparaison entre une marque sans utilisation de masque psychovisuel et avec utilisation de JNDEM. Nous pouvons constater qu'en utilisant un masque la marque s'adapte par rapport aux zones d'insertion, surtout dans les zones sombres et texturées.

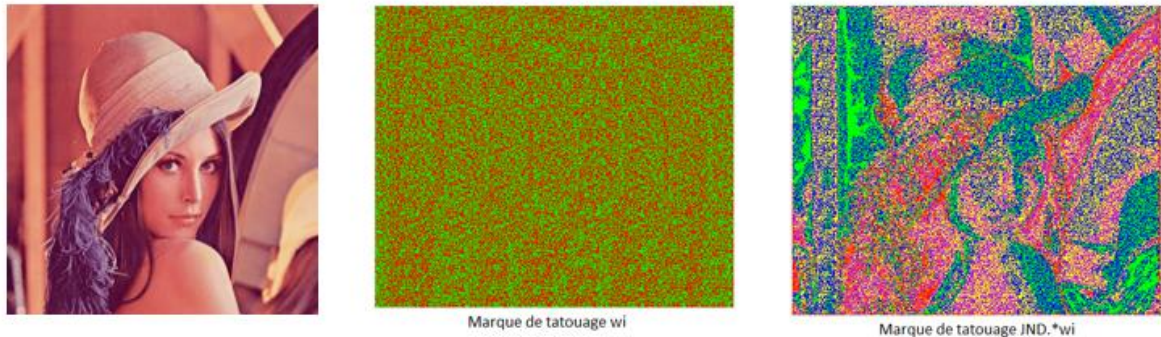


Figure 3-13: Comparaison entre deux marques produites par étalement de spectre : une marque de tatouage spatiale  $w_i$  et une marque de tatouage couplé avec le masque psychovisuel JNDEM. On voit bien que la deuxième marque s'adapte aux zones d'insertion

Dans la suite nous allons tatouer des images de synthèse et des images réelles par la méthode de de tatouage présenté ci-dessus avec des tests plus complets pour des images réelles ainsi que des comparaisons avec d'autres méthodes de tatouage.

### 3.5 Test sur des images de synthèses

Dans cette section, nous examinons les résultats de la méthode proposée sur des images de synthèse. D'abord nous présentons la base d'image utilisée. En outre, l'évaluation de l'imperceptibilité de la méthode a été également comparée à la technique de tatouage mais sans utilisation du masque afin de démontrer les avantages du masque JNDEM proposé dans ce travail.

### 3.5.1 Base expérimentale d'image

Pour évaluer notre JNDEM nous allons le tester d'abord sur deux bases d'images. La base d'image Amsterdam Library of Textures (ALOT)<sup>1</sup> et la base d'image Canadien Multiband Texture (MBT)<sup>2</sup>. ALOT est une collection d'images couleur de 250 textures. La base de données MBT est une collection de 112 images en couleur qui ont deux caractéristiques importantes : leur contenu chromatique riche et leur variation spatiale intra-bande et inter-bande élevée. La résolution des images pour les deux bases ALOT et MBT sont de 768×512 et 640×640 successivement. Les images ont toutes été redimensionnées à 512×512. Cela permet l'utilisation de l'algorithme de tatouage. Un exemple d'images des deux bases est présenté dans les figures 3-12 et 3-13.



Figure 3-14: Exemples d'images de la base expérimentale ALOT.

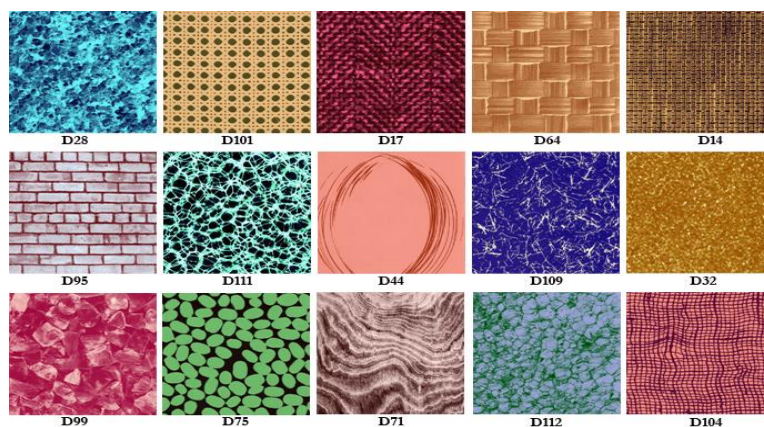


Figure 3-15: Exemples d'images de la base expérimentale MBT.

<sup>1</sup> [http://aloi.science.uva.nl/public\\_alot/](http://aloi.science.uva.nl/public_alot/)

<sup>2</sup> <https://multibandtexture.recherche.usherbrooke.ca/index.html>

Sur la figure 3-16 présente quelques images des bases ALOT et MBT de différentes textures ainsi que le discriminateur de texture proposé. On rappelle que la couleur bleue correspond à des régions uniformes, la couleur verte correspond aux contours et la couleur rouge correspond aux régions texturées. On constate que les différentes régions de l'image sont bien distinguées à partir du discriminateur à d'ondelette de Faber-Schauder à échelle mixée.

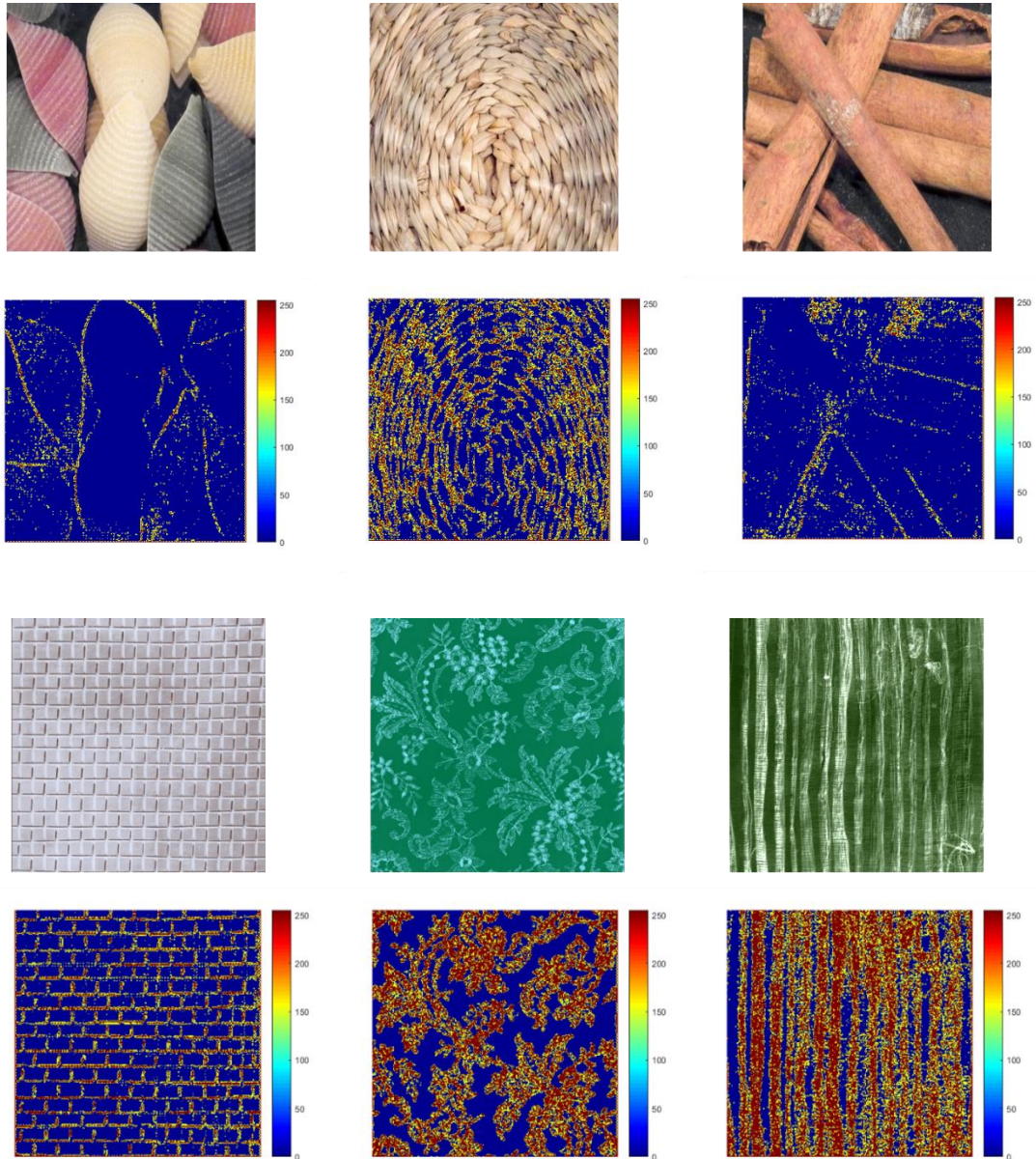


Figure 3-16: Discriminateur de texture de quelques images des deux bases ALOT et MBT.

Dans la figure 3-17 nous avons l'image originale et l'image tatouée sans utilisation de masque ou la marque a été inséré sur toute l'image avec la même force d'insertion (rapport signal à bruit de 40 dB) et l'image tatoué avec masque JNDEM où l'insertion de la marque se fait

d'une manière optimale selon la région d'insertion. Pour évaluer la similarité entre l'image originale et les images tatouées, nous avons utilisé le MSSIM. Cette mesure objective prend en compte la sensibilité de l'œil humain aux changements de structures dans les images. Les tableaux 3-1 et 3-2 présentent les résultats de la mesure de similarité obtenue pour les images tatouées sans masques et les images tatouées avec le masque JNDEM. Idéalement, le MSSIM entre deux image identiques est de un. Les résultats de MSSIM présentées dans les tableaux 3-1 et 3-2 pour les deux méthodes montrent une similarité élevée entre l'image originale et les images tatouées avec masque alors qu'elle est moindre sans masque.

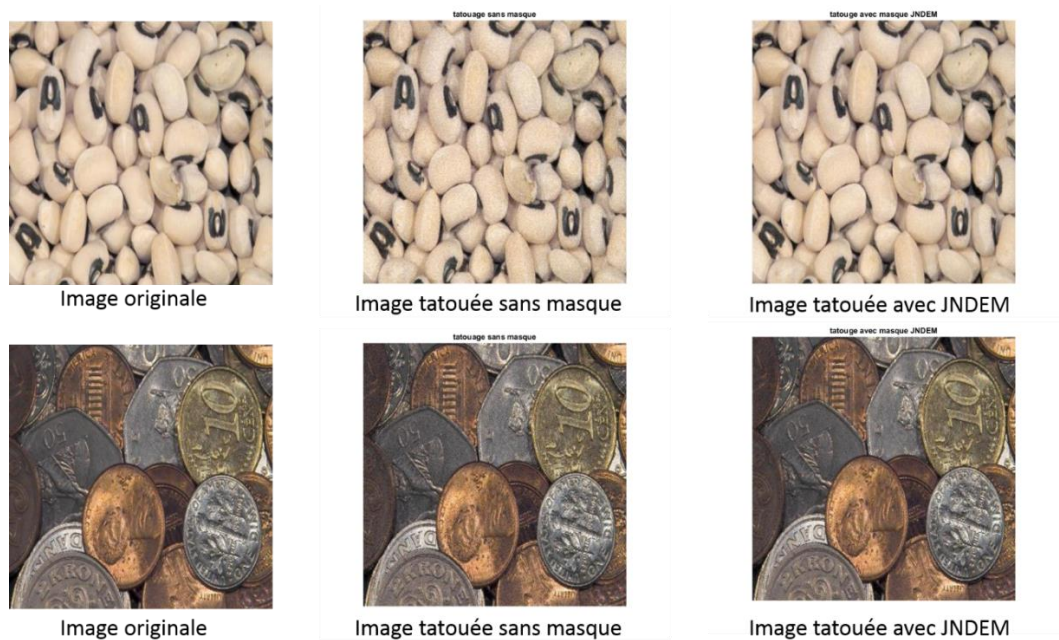


Figure 3-17: Image originale et images tatouée sans utilisation de masque et image tatouée avec masque JNDEM

Métriques objective de la qualité	Tatouage sans masque	Tatouage avec masque JNDEM
Maximum	0.9648	0.9630
Minimum	0.3827	0.6392
Moyenne	0.6152	0.8614

Tableau 3-1: Valeurs du MSSIM pour images tatouées de la base ALOT



Métriques objective de la qualité	Tatouage sans masque	Tatouage avec masque JNDEM
Maximum	0.9772	0.9417
Minimum	0.4623	0.6746
Moyenne	0.7988	0.8745

Tableau 3-2: Valeurs du MSSIM pour images tatouées de la base MBT

### 3.5.2 Evaluation de l'imperceptibilité

Le degré d'imperceptibilité de la marque insérée est important pour s'assurer qu'il n'y a pas de dégradations visuelles relatives au tatouage. Les expériences ont été conduites pour deux cas, le cas des images tatouées sans masque et le cas où le tatouage est basé sur l'utilisation du masque JNDEM. Nous avons calculé la qualité visuelle en utilisant les mesures objectives suivantes (voir chapitre 2) :

- Métriques objectives :
  - Rapport poids bruit maximal pondéré (Weighted Peak Signal-to-Noise Ratio, WPSNR) [89] : Tout comme le PSNR, une valeur élevée de WPSNR indique que l'image est moins déformée. Les valeurs supérieures à 40 dB indiquent une dégradation presque invisible et des valeurs inférieures à 30 dB indiquent une forte dégradation [92]. Pour une image en niveaux de gris de 8 bits, le WPSNR est donné par :

$$WPSNR = 10 \log_{10} \frac{\max(I_o)^2}{\|(I_w - I_o).NVF\|^2} \quad (3-6)$$

avec

$$NVF = \frac{1}{1 + \theta \sigma_{I_o}^2(i, j)}$$

où  $\sigma_x^2(i, j)$  désigne la variance locale de l'image dans une fenêtre centrée sur le pixel avec des coordonnées  $(i, j)$  et  $\theta$  est un paramètre de réglage qui joue le rôle du réglage du contraste.

- Similarité structurelle (SSIM) [83] : une valeur SSIM proche de 1 indique une bonne qualité visuelle :

$$SSIM = \frac{1}{W} \sum_{i=1}^W l(i) \cdot c(i) \cdot s(i), \quad (3-7)$$

$$\text{où } l(i) = \frac{2\overline{x(i).y(i)} + c_1}{\overline{x(i)^2 + y(i)^2} + c_1}, \quad c(i) = \frac{2\sigma_x(i).\sigma_y(i) + c_2}{\sigma_x^2(i) + \sigma_y^2(i) + c_2}, \quad s(i) = \frac{2\sigma_{xy}(i) + c_3}{\sigma_x(i) + \sigma_y(i) + c_3}$$

$$C_1 = (L.K_1)^2, C_2 = (L.K_2)^2, C_3 = \frac{C_2}{2},$$

où  $L$  est la dynamique des niveaux de gris.  $K_1$  et  $K_2$  sont deux constantes fixées par les auteurs à 0.01 et 0.03, respectivement.  $\sigma_x$  et  $\sigma_y$  peuvent être considéré comme des estimations du contraste de  $x$  et  $y$ .

Sur les figures 18 et 19 les résultats des métriques objectives WPSNR et SSIM pour les images tatouées sans utilisation du masque et les images tatouées avec masque JNDEM.

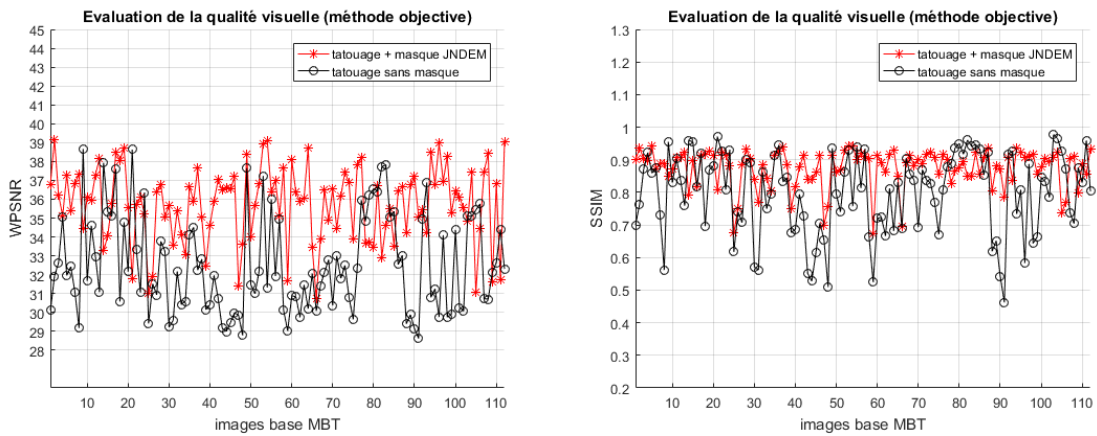


Figure 3-18: Courbe résultat du WPSNR et SSIM de 112 images de la base MBT

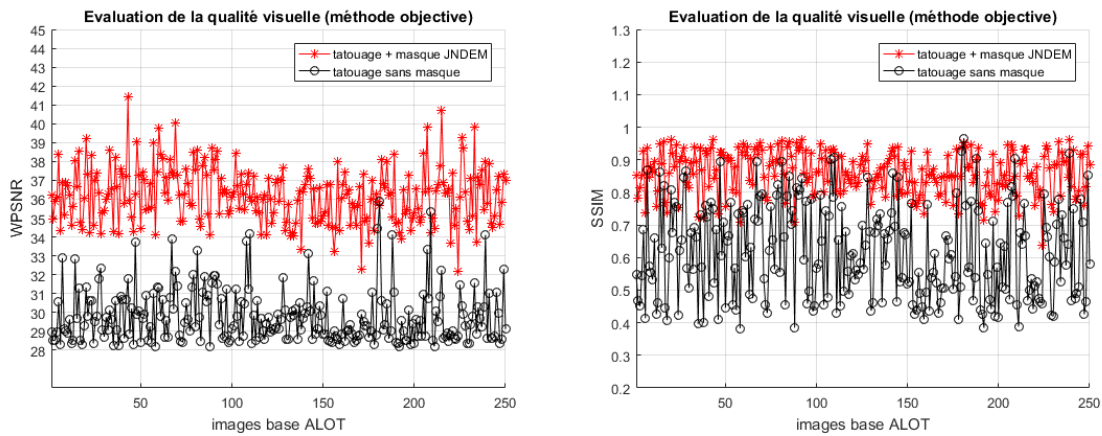


Figure 3-19: Courbe résultat du WPSNR et SSIM de 112 images de la base ALOT

Les résultats trouvés du WPSNR et SSIM des images tatouées avec le masque proposé JNDEM sont meilleurs que ceux de l'autre schéma de tatouage où le masque n'est pas pris en considération. Ceci nous montre clairement l'avantage d'utiliser un masque psychovisuel ou la force d'insertion de la marque se fait d'une manière optimale et adaptée à la zone d'insertion. Ces résultats préliminaires nous conduisent à tester notre JNDEM sur des images

réelles et le comparer à d'autres JND de la littérature et d'évaluer les performances des méthodes.

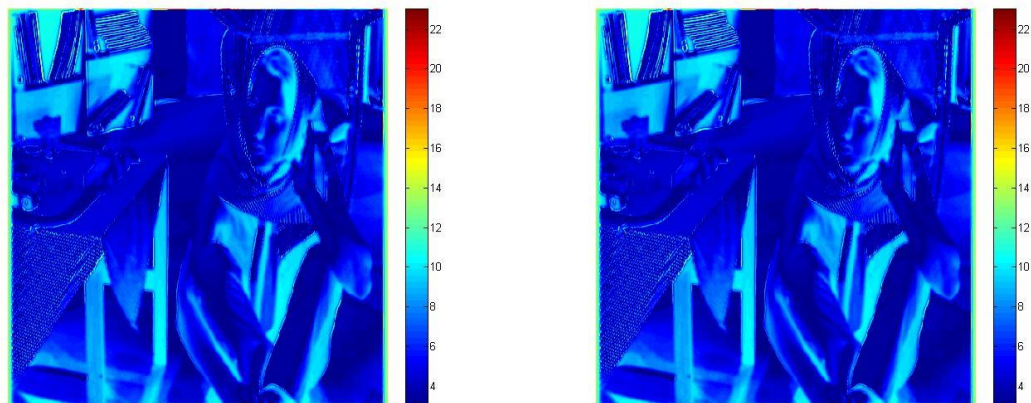
### **3.6 Evaluation et comparaison sur images réelles**

Pour évaluer la performance du tatouage utilisant notre modèle JND à base d'ondelettes à échelles mixées (JNDEM) sur images réelles, nous avons fait une comparaison avec quatre autres modèles JND issu de l'état de l'art :

1. Le modèle JND spatial de Chou et Li [66],
2. Le modèle JND spatial de Qi *et al* [69],
3. Le modèle JND d'ondelette de Barni *et al* [77],
4. Le modèle JND DCT de Watson [70].

Ces méthodes ont été sélectionnées du fait de leur popularité et de leur efficacité. Elles ont été adaptées dans plusieurs travaux récents [93, 94, 95] et sont devenues par conséquent des références dans ce domaine. Par exemple, la méthode de Watson reste l'approche de référence pour définir les masques DCT et les 4 règles de Barni [65] pour la définition des masques psychovisuels sont aussi considérés comme des préalables à tout calcul de JND.

La figure 3-20 représente des exemples pour les différentes implémentations de ces masques de référence pour l'image Barbara.



(a)

(b)

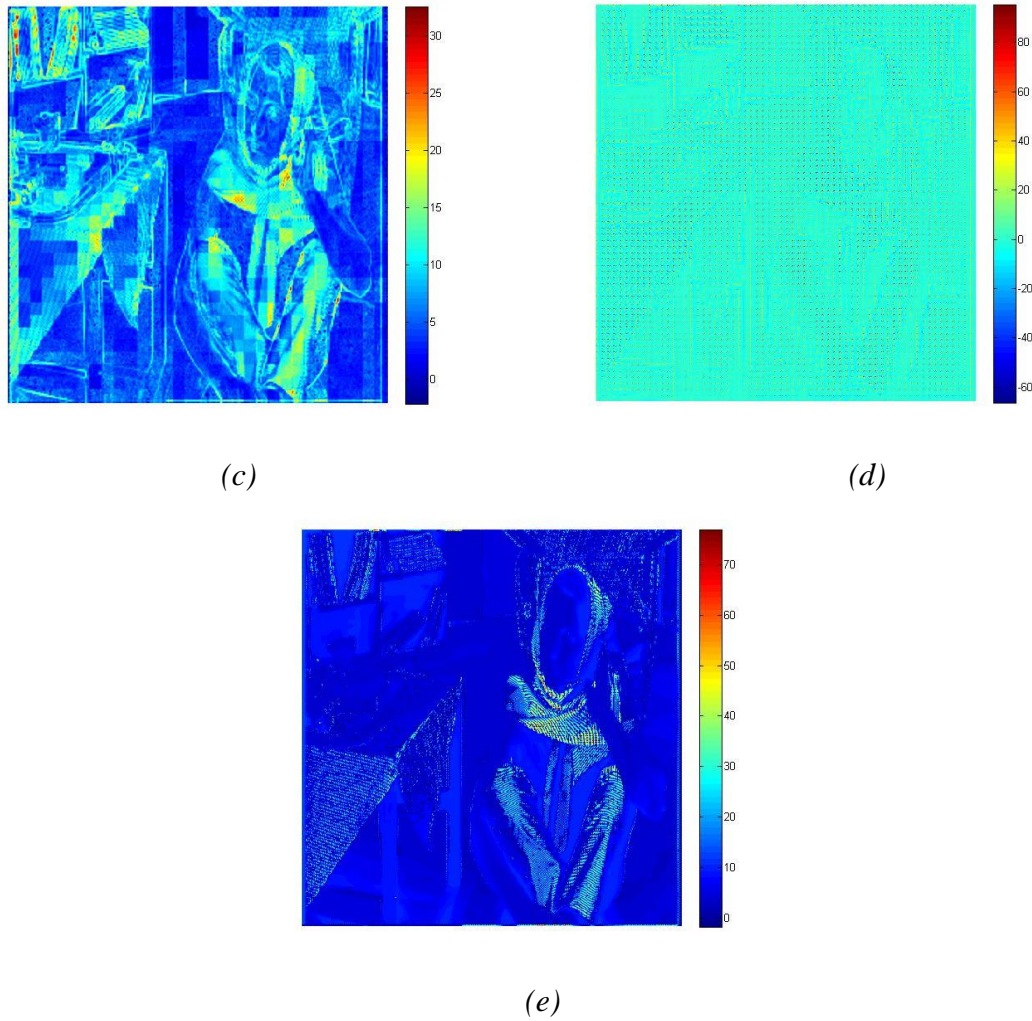
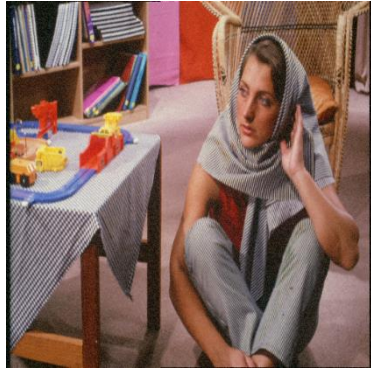


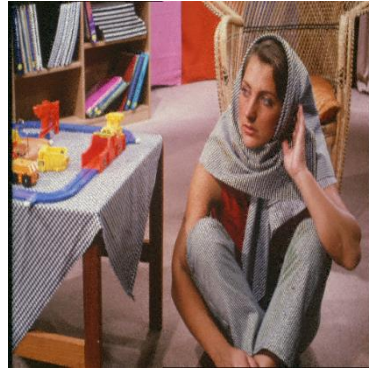
Figure 3-20: Illustrations des différents masques de référence JND et que notre JNDEM sur l'image Barbara : (a) JND de Chou & Li [66], (b) JND de Qi & al [69], (c) JND de Barni & al [77], (d) JND de Watson [70], (e) notre JNDEM

Pour illustrer l'imperceptibilité des 5 modèles JND, l'image originale et les images tatouées par JNDEM et les JND de Chou de Qi de Barni et de Watson ont été présentées dans la figure suivante.

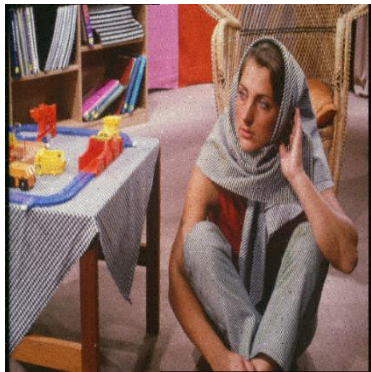




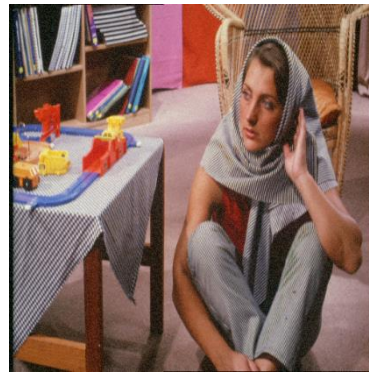
(c)



(d)



(e)



(f)

Figure 3-21: Images originales et images tatouées générées par différents modèles JND : (a) Image originale et (b) (a) JND de Chou & Li [66], (b) JND de Qi & al [69], (c) JND de Barni & al [77], (d) JND de Watson [70], (f) notre JNDEM.

Nous avons utilisé un ensemble de 50 images couleurs naturelles comme indiqué sur la figure 3-22, elles sont toutes de taille 512x512. Pour toutes les images nous avons tatoué uniquement la composante de la luminance (les composants de la chrominance n'ont pas été modifiés).

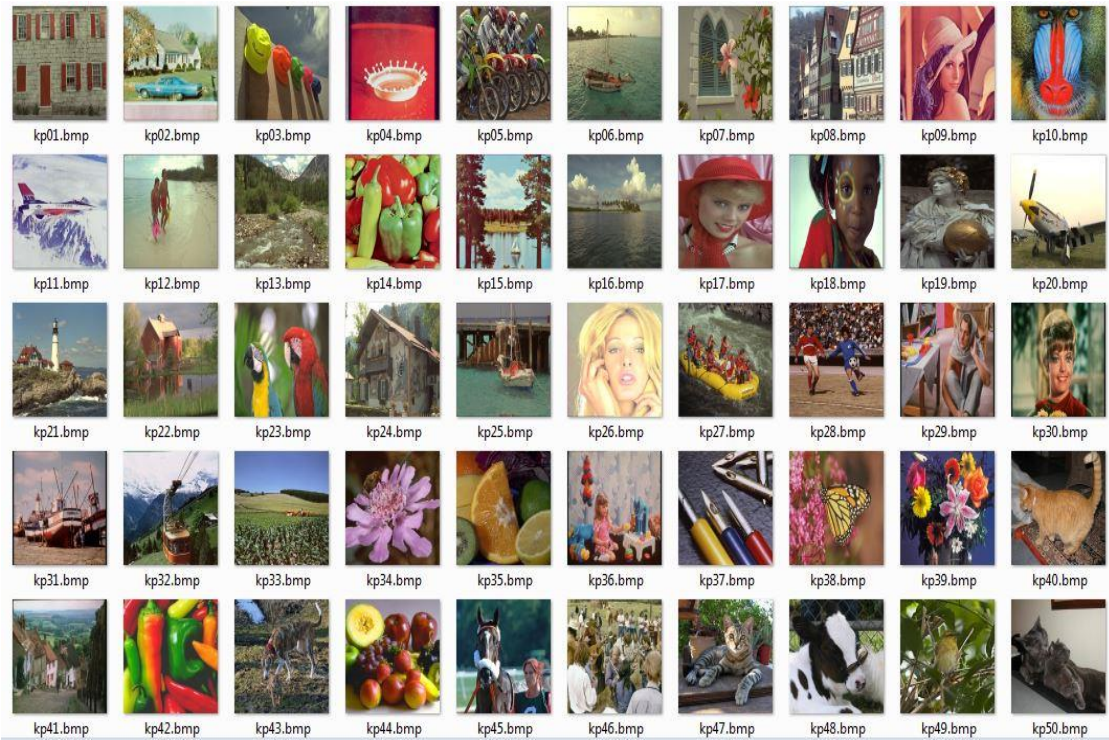


Figure 3-22: Base de 50 images utilisées dans les tests

Nous allons évaluer la qualité visuelle des images tatouées via des métriques objectives. Nous allons aussi effectuer un test subjectif en faisant appel à des êtres humains pour évaluer la qualité visuelle des images tatouées. Enfin, nous allons estimer la robustesse des méthodes.

### **3.6.1 Évaluation des performances**

Dans cette section, nous évaluons la méthode de tatouage JNDEM proposée en termes d'imperceptibilité et de robustesse. Pour l'évaluation de l'imperceptibilité, nous avons utilisé des paramètres de qualité objectifs mais puisque nous utilisons un tatouage avec masque psychovisuel nous avons opté aussi pour l'utilisation de tests subjectifs. Pour l'évaluation de la robustesse nous avons utilisé des attaques courantes.

#### **3.6.1.1 Evaluation de l'imperceptibilité**

Nous avons évalué la qualité visuelle des images tatouées, en utilisant les deux métriques objectives WPSNR et SSIM.

Les figures 3-23 et 3-24 présentent les résultats obtenus pour les deux métriques objectives utilisées.

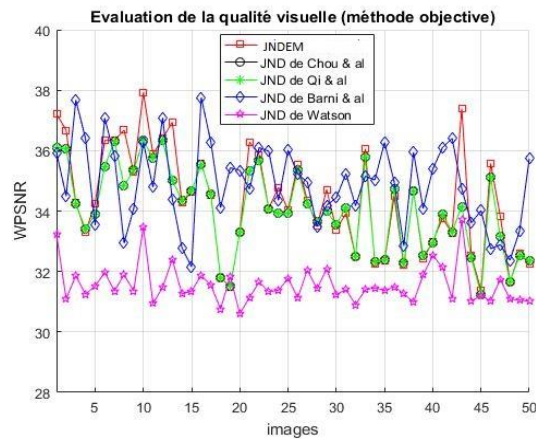


Figure 3-21: Courbe résultat du WPSNR de 50 images tatouées

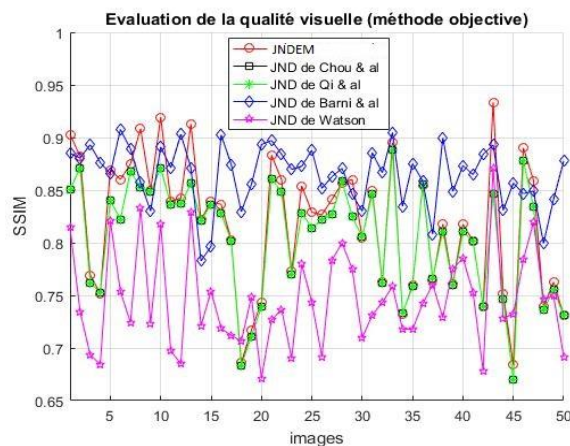


Figure 3-22: Courbe résultat de SSMI de 50 images tatouées

- Métrique subjective : Expérience psychovisuel comparative

D'un autre côté, des tests subjectifs avec des observateurs humains sont utilisés pour évaluer la qualité perceptuelle des images tatouées. Nous avons effectué un test comparatif (décrit au chapitre 2 section 2.3.2.1.2) sur cinquante images (voir figure 13) et 15 observateurs humains (dont des chercheurs en traitement d'image et des observateurs non spécialiste du domaine). La distance de visionnement est de 6 fois la hauteur de l'image et une échelle de notation est fixée à 5 niveaux :

<b>Note</b>	<b>Appréciation</b>	<b>Qualité</b>
<b>5</b>	Imperceptible	excellente
<b>4</b>	Légèrement imperceptible	Bien
<b>3</b>	Légèrement gênant	juste
<b>2</b>	Agaçant	faible
<b>1</b>	Très gênant	mauvaise

En utilisant ce système, les observateurs humains sont invités à classer les images tatouées en fonction de ces cinq échelles. Le classement des résultats de ces expériences donne une mesure quantitative de la perception subjective des images et une note moyenne MOS a été calculée pour chaque image. Tout au long des expériences, l'ordre des images de test est aléatoire et une séance d'explication de l'objectif du test a été faite pour les participants. Un ordinateur HP ProBook 640 GI a été utilisé avec un écran « eDP anti-reflet HD rétro-éclairé par LED 35,56 cm (14 pouces) de diagonale 1,2 plat (1600 x 900). Les résultats de la comparaison du JNDEM avec les autres méthodes se trouvent sur la figure 25.

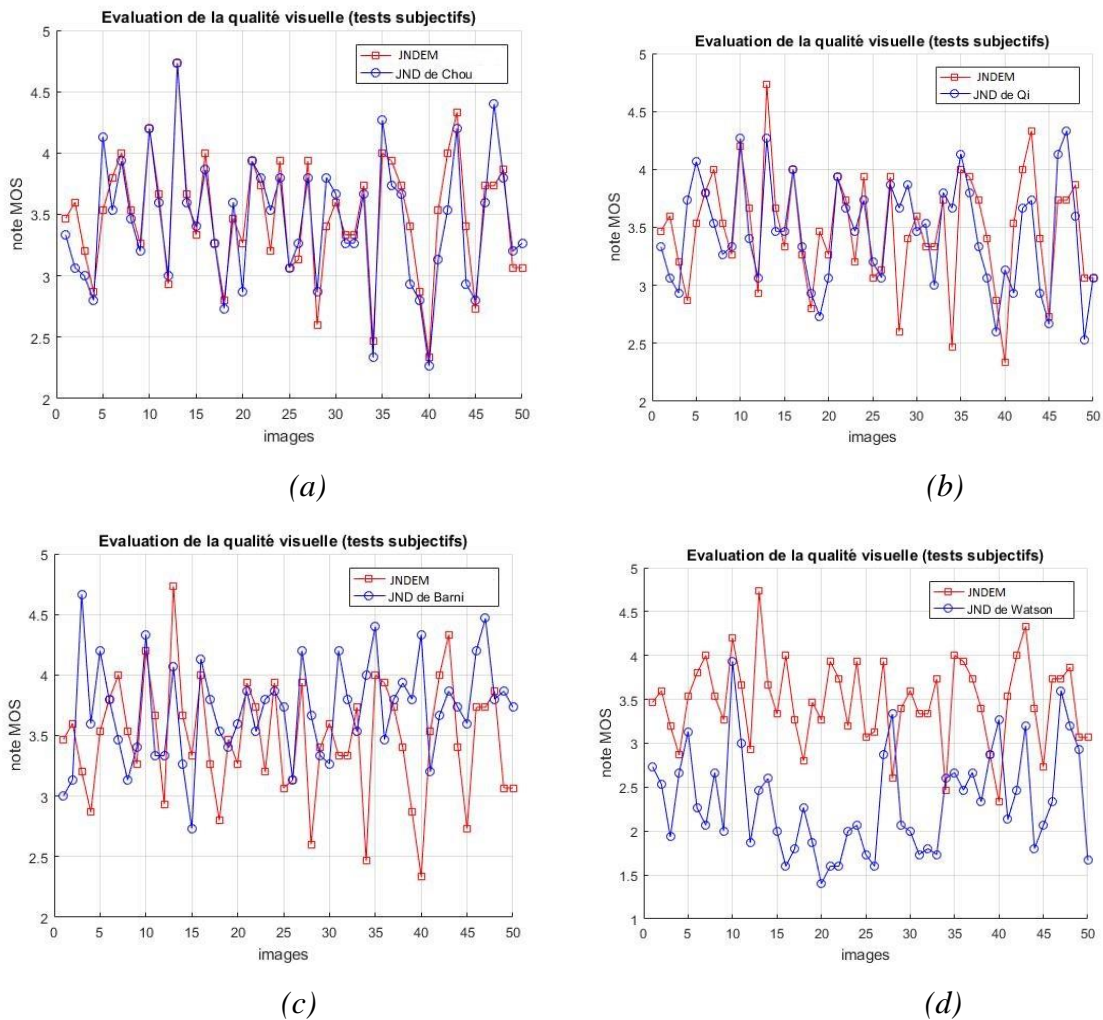


Figure 3-23: Résultat du test subjectif. (a) JNDEM comparé à Chou & al, (b) JNDEM comparé à Qi & al, (c) JNDEM comparé à Barni & al, (d) JNDEM comparé à Watson



Les résultats de la moyenne de 50 images, des métriques WPSNR et SSMI ainsi que la méthode subjective et l'écart type obtenus se trouvent dans le Tableau 3-3, Tableau 3-4 et Tableau 3-5 successivement.

Métriques objective de la qualité	JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
Maximum	37.9237	37.9237	37.7534	33.7089	37.3378
Minimum	31.3786	31.3786	32.1537	30.6029	31.2364
Moyenne	34.3536	34.3536	<b>34.8986</b>	31.5600	34.4153

.Tableau 3-3: Valeurs de WPSNR pour 50 images tatouées

Métriques objective de la qualité	JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
Maximum	0.9335	0.9335	0.9076	0.8712	0.9566
Minimum	0.6838	0.6838	0.7827	0.6710	0.6699
Moyenne	0.8185	0.8185	<b>0.8645</b>	0.7458	0.8416

Tableau 3-4: Valeurs de SSIM pour 50 images tatouées

Métrique subjective de la qualité	Image original	JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
Maximum	5	4.46	4.46	4.46	3.53	4.46
Minimum	3.40	2.13	2.46	2.80	1.60	2.26
Moyenne	4.52	3.32	3.32	<b>3.50</b>	2.20	3.43
Ecart type (Std)	0.32	0.47	0.47	0.43	0.46	<b>0.41</b>

Tableau 3-5: Valeurs MOS pour 50 images tatouées

On peut voir que la méthode de Barni donne la meilleure qualité perceptuelle évaluée au moyen de tests objectifs. Notre JNDEM est également de haute qualité et offre des résultats supérieurs aux JND spatial et Watson. Concernant les tests subjectifs, la méthode de Barni est un peu meilleure que la méthode proposée, alors que d'autres ne sont pas aussi efficaces. Mais notre JNDEM a un écart type meilleur que celui de Barni ce qui prouve qu'il y a une harmonie entre les images tatouées avec le JNDEM. Bien que la méthode de Barni offre la

meilleure qualité visuelle, notre méthode est très proche de celle de Barni, et offre elle aussi de très bons résultats. L'avantage de notre JNDEM et qu'est basé sur un discriminateur de texture qui sépare les différentes zones. Ceci permet d'insérer d'une manière plus adaptée la marque de tatouage. Et aussi en utilisant les différents JND on respecte le seuil de visibilité de la sensibilité de l'œil ce qui se traduit par une imperceptibilité du tatouage.

### **3.6.1.2 Evaluation de la Robustesse**

Dans un deuxième temps, des attaques simulées ont été effectuées pour tester la robustesse du tatouage pour les différents JND. L'insertion des attaques dans le processus du tatouage est expliquée dans le schéma de la figure 3-26. Dans cette section, nous évaluons la performance des algorithmes de tatouage en termes d'extraction du message inséré. En effet nous calculons le taux d'erreur sur les bits (BER) du message extrait, défini comme le rapport entre le nombre de bits du tatouage erronés extraits et le nombre total de bits insérés. La valeur du BER est comprise entre 0 et 1, des faibles valeurs de BER indiquent une bonne robustesse du tatouage :

$$BER = \frac{\text{Nombre de bits d'erreur}}{\text{Nombre total de bits}}$$

Afin de tester la robustesse de la méthode proposée, nous avons effectué différents types d'attaques de signaux ou attaques géométriques, à savoir :

Attaque de bruit	bruit gaussien (AWGN)
Attaque de compression	compression JPEG
Attaque de traitement d'image :	- égalisation de l'histogramme (EH) - ajustement du contraste (AC)
Attaque géométrique	- Rotation - Cropping
Attaque de débruitage	- filtrage moyennneur - filtrage médian

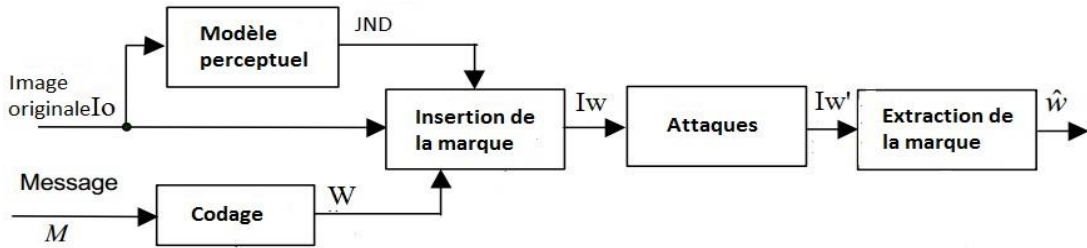


Figure 3-24: Schéma d'un processus de test de tatouage utilisant des modèles perceptuel

### a) Attaque par compression Jpeg

Les résultats du BER de l'extraction de la marque des images tatouées après compression JPEG avec deux facteurs de qualité différente (QF = 80% et 10%) sont présentés sur la figure 3-27. Après une compression Jpeg de facteur qualité 10% les images sont généralement fortement dégradées et des effets de blocs et des artefacts apparaissent nettement. Cependant, La détection est excellente par la méthode JNDEM, le message est extrait et la valeur du BER ne dépasse pas 0.2. Les résultats du JNDEM dépassent les méthodes de tatouages testées de Barni et Watson. La figure 3-28 présente un exemple, sur l'image Barbara, de l'attaque Jpeg par un facteur de qualité QF=10%. Le tableau 3-6 présente les valeurs du BER de 50 images.

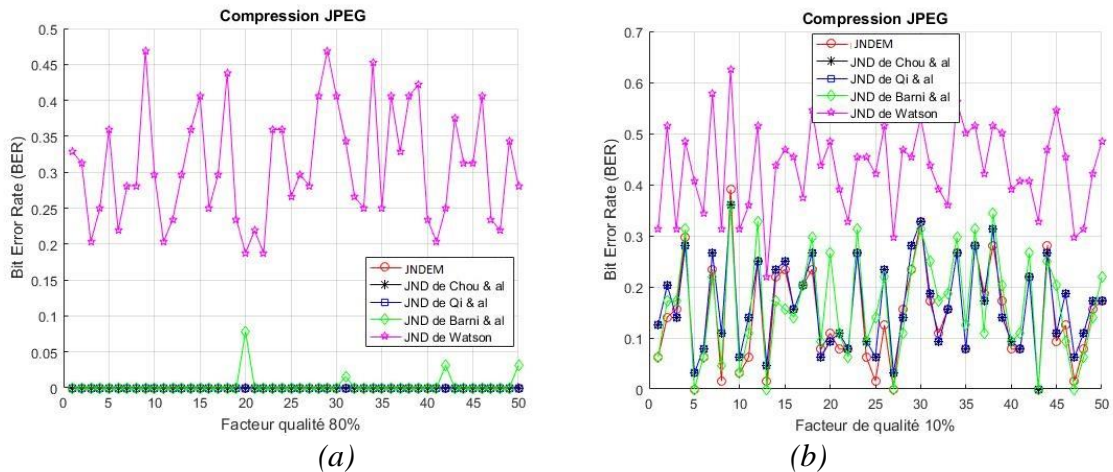


Figure 3-25: Valeurs de détection du BER pour les images tatouées après compression JPEG : (a) 80%, (b) 10%.



IRF—SIC  
 La marque extraite  
 Ber=0

Figure 3-26: Exemple sur l'image Barbara de l'attaque Jpeg avec un QF=10%

		La moyenne du BER de 50 images					
Types d'attaque	Intensité		JND de	JND de	JND de	JND de	JNDEM
			Chou [66]	Qi [69]	Barni [77]	Watson [70]	
JPEG	FQ = 80%	Maximum	0	0	0.0781	0.4688	0
		Minimum	0	0	0	0.1875	0
		Moyenne	0	0	0.0031	0.3091	<b>0</b>
	FQ = 10%	Maximum	0.3594	0.3594	0.3594	0.6250	0.3906
		Minimum	0	0	0	0.2188	0
		Moyenne	<b>0.1628</b>	0.1628	0.1647	0.4306	0.1803

Tableau 3-6: Valeurs BER de 50 images après l'attaque Jpeg

### b) Attaque par bruit gaussien AWGN

Pour la robustesse de l'algorithme de tatouage en ajoutant un bruit gaussien blanc (AWGN) à l'image tatouée nous avons opté pour deux variances de 0.1 et 0.5. Les résultats des valeurs du BER de la marque extraite sont présentés sur la figure 3-29, on peut observer que le détecteur détecte la signature correctement. La figure 3-30 présente un exemple, sur l'image Barbara, de l'attaque AWGN avec une variance de 0.1. Bien que la contamination par un AWGN de variance de 0.5 ait causé aux images un bruitage énorme, le JNDEM reste tout à fait robuste à ce type de traitement. Nous remarquons aussi, d'après le tableau 3-7, que les résultats du BER pour les autres méthodes JND testées sont inférieurs à notre JNDEM.

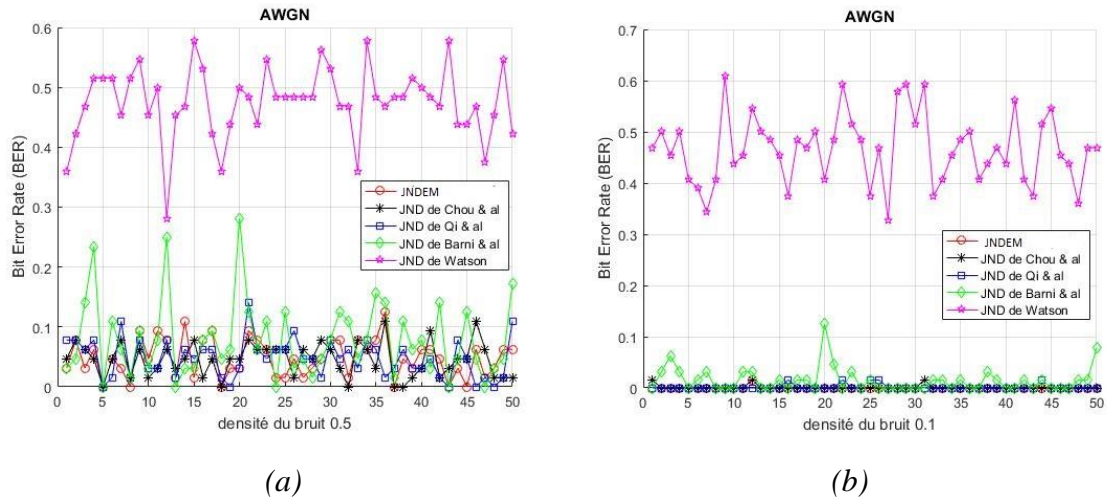


Figure 3-27: Valeurs de détection du BER pour les images tatouées après attaque AWGN : (a) variance 0.1, (b) de variance 0.5.



**IRN-SIC**  
La marque extraite  
Ber=0.0313

Figure 3-28: Exemple sur l'image Barbara de l'attaque AWGN avec une variance  $\sigma = 0.1$

		La moyenne du BER de 50 images					
Types d'attaque	Variance		JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
AWGN	$\sigma = 0.1$	Maximum	0.0156	0.0156	0.1250	0.6094	<b>0.0156</b>
		Minimum	0	0	0	0.3281	<b>0</b>
		Moyenne	0.0016	0.0016	0.0147	0.4656	<b>0.0003</b>
	$\sigma = 0.5$	Maximum	<b>0.1094</b>	<b>0.1094</b>	0.2813	0.5781	0.1250
		Minimum	<b>0</b>	<b>0</b>	0	0.2813	0
		Moyenne	0.0453	0.0478	0.0784	0.4753	<b>0.0432</b>

Tableau 3-7: Valeurs BER de 50 images après l'attaque AWGN

**c) Attaque par rotation**

Pour l'attaque par rotation, nous avons évalué la robustesse du tatouage face à deux angles différents 30 ° et 45 °. Les résultats des valeurs du BER de la marque extraite pour les cinq méthodes sont présentés dans la figure 3-31. Dans ce test, des parties de l'image ont été remplacées par des zéros, ces zones ne contiennent aucune information de la marque ce qui explique la diminution de la robustesse des méthodes si la rotation est importante. La figure 3-32 présente un exemple, sur l'image Barbara, de l'attaque rotation de 45°. D'après le tableau 3-8 nous pouvons conclure que notre algorithme a montré une bonne robustesse et le message est extrait pour une rotation de 30°. Cependant, du fait de sa résistance à de faibles déformations des attaques géométriques, une compensation des attaques géométriques, même assez grossière pourrait être suffisante pour retrouver la marque. Et nous remarquons que pour des degrés de rotation inférieurs à 45° notre méthode a montré des résultats supérieurs par rapport aux autres méthodes.

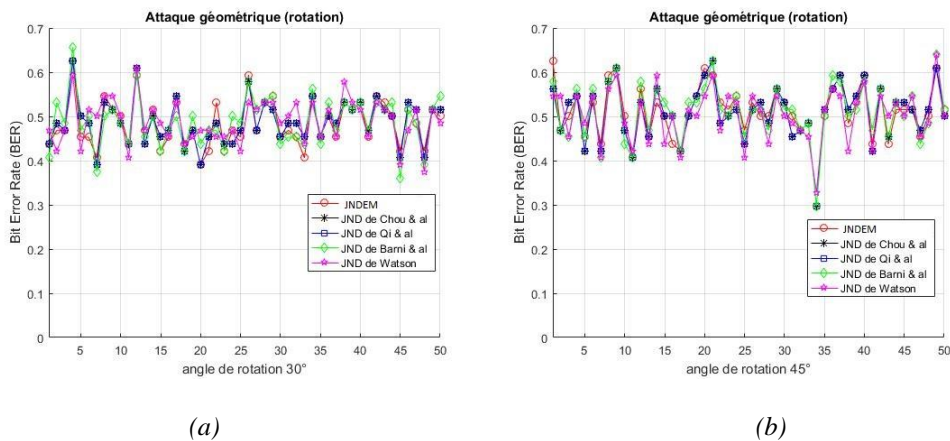


Figure 3-29: Valeurs de détection du BER pour les images tatouées après rotation : (a) de 30°, (b) de 45°.



Figure 3-30: Exemple sur l'image Barbara de l'attaque rotation avec un angle de rotation de 45°

		La moyenne du BER de 50 images					
Types d'attaque	Angle de rotation		JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
Rotation	30°	Maximum	0.6250	0.6250	0.6563	0.6094	0.6250
		Minimum	0.3906	0.3906	0.3594	0.3750	0.3906
		Moyenne	0.4878	0.4878	0.4878	0.4894	<b>0.4869</b>
	45°	Maximum	0.6250	0.6250	0.6406	0.6406	0.6250
		Minimum	0.2969	0.2969	0.2969	0.3281	0.2969
		Moyenne	0.5100	0.5100	0.5119	0.5034	<b>0.5088</b>

Tableau 3-8: Valeurs BER de 50 images après l'attaque rotation

**d) Attaque par cropping**

Pour l'attaque cropping les résultats des valeurs du BER de la marque extraite sont présentés sur la figure 3-33, les images tatouées sont coupées en gardant seulement 3% et 5% de l'image tatouée. La figure 3-34 présente un exemple, sur l'image Barbara, de l'attaque cropping en gardant 3% de l'image. Le tatouage avec le JNDEM a été extrait avec succès avec une attaque cropping de 5% et donne le meilleur résultat comparé aux autres méthodes testées ce qui prouve son efficacité. Ce résultat est obtenu grâce à l'insertion par étalement de spectre qui permet d'avoir une redondance de la marque. Le tableau 3-9 présente les valeurs du BER de 50 images.

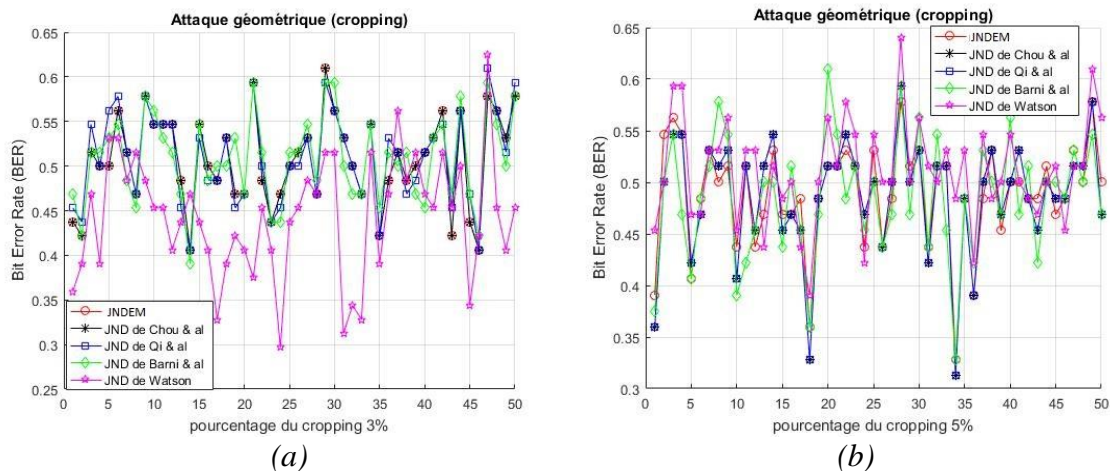


Figure 3-31: Valeurs de détection du BER pour les images tatouées après cropping : (a) de 3%, (b) de 5%.

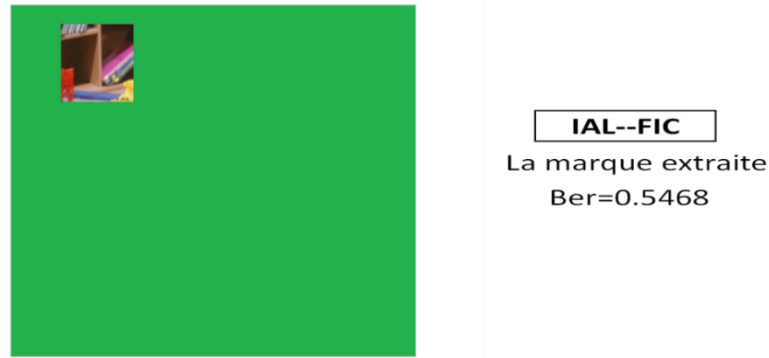


Figure 3-32: Exemple sur l'image Barbara de l'attaque cropping en gardant 3% de l'image

		La moyenne du BER de 50 images					
Types d'attaque	Intensité		JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
Cropping	3%	Maximum	0.5938	0.5938	0.5938	0.6250	0.5781
		Minimum	0.3125	0.3125	0.3906	0.2969	0.3281
		Moyenne	0.4866	0.4866	0.4847	<b>0.4437</b>	0.4884
	5%	Maximum	0.6094	0.6094	0.6094	0.6406	0.6094
		Minimum	0.4063	0.4063	0.3281	0.3906	0.4063
		Moyenne	0.5094	0.5094	0.5075	0.5109	<b>0.5069</b>

Tableau 3-9: Valeurs BER de 50 images après l'attaque cropping

#### e) Attaque par filtrage

Ensuite, les images tatouées ont été filtrées avec un filtre moyenneur et médian de fenêtre 3x3 et 7x7. Les résultats des valeurs du BER de la marque extraite sont présentés sur les figures 3-35 et 3-37 successivement. Les figures 33-6 et 3-38 présentent successivement un exemple, sur l'image Barbara, des attaques de filtrage d'une fenêtre 3x3 nous pouvons voir qu'avec les deux filtres moyenneurs et médian les dégradations de l'image sont assez visibles. Nous estimons que les résultats obtenus sont très satisfaisants, et l'utilisation de ces filtres n'influe pas sur la robustesse de la détection de notre méthode. Même avec une fenêtre de filtre élevée, la marque de tatouage a été extraite pour la méthode JNDEM proposée et les trois autres modèles JND à l'exception du JND DCT. Les résultats obtenus montrent une



robustesse supérieure de la méthode proposée face à ce type d'attaque. En effet notre JNDEM donne les meilleurs résultats pour le filtre moyenneur et aussi pour un filtre médian à partir de 7x7. Les tableaux 3-10 et 3-11 présentent les valeurs du BER de 50 images.

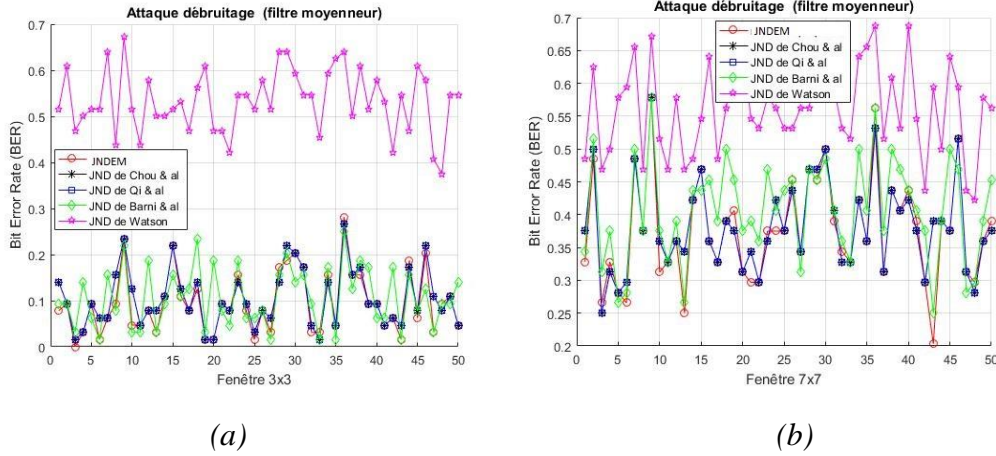


Figure 3-33: Valeurs du BER pour les images tatouées après filtre moyenneur : (a) 3x3, (b) 7x7.



Figure 3-34: Exemple sur l'image Barbara de l'attaque filtre moyenneur d'une fenêtre 3x3.

		La moyenne du BER de 50 images					
Types d'attaque	Fenêtre	JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM	
Filtre moyenneur	3x3	Maximum	0.2656	0.2656	0.2500	0.6719	0.2813
		Minimum	0.0156	0.0156	0.0156	0.3750	0
		Moyenne	0.1047	0.1047	0.1072	0.5334	<b>0.0950</b>
	7x7	Maximum	0.5781	0.5781	0.5781	0.6875	0.5781
		Minimum	0.2500	0.2500	0.2500	0.4219	0.2031
		Moyenne	0.3828	0.3828	0.4069	0.5606	<b>0.3616</b>

Tableau 3-10: Valeurs BER de 50 images après l'attaque filtre moyenneur

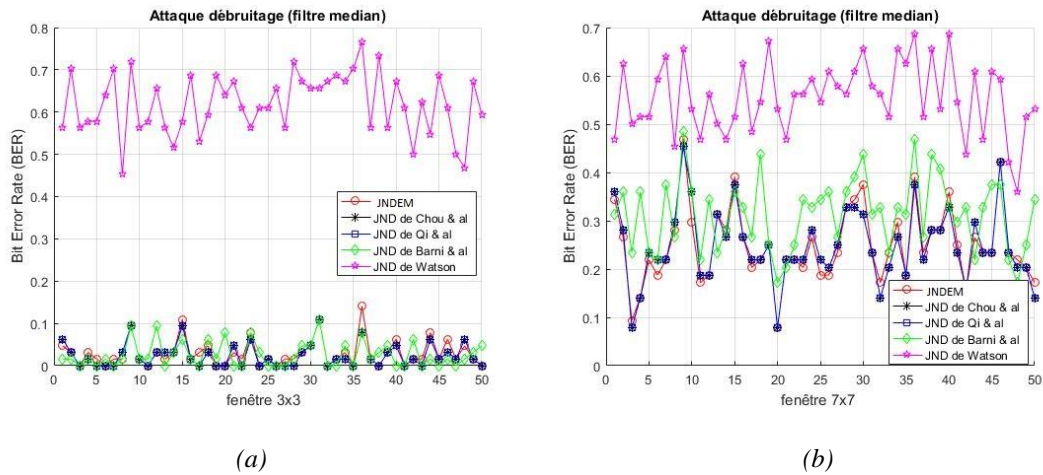


Figure 3-35: Valeurs du BER pour les images tatouées après filtre médian : (a) 3x3, (b) 7x7.



Figure 3-36: Exemple sur l'image Barbara de l'attaque filtre médian d'une fenêtre 3x3.

		La moyenne du BER de 50 images					
Types d'attaque	Fenêtre		JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM
Filtre médian	3x3	Maximum	0.1094	0.1094	0.1094	0.7656	0.1406
		Minimum	0	0	0	0.4531	0
		Moyenne	<b>0.0244</b>	<b>0.0244</b>	0.0259	0.6184	0.0294
	7x7	Maximum	0.4531	0.4531	0.4844	0.6875	0.4688
		Minimum	0.0781	0.0781	0.1719	0.3594	0.0781
		Moyenne	0.2491	0.2491	0.3141	0.5547	<b>0.2344</b>

Tableau 3-11: Valeurs BER de 50 images après l'attaque filtre médian

**f) Attaque par changement d'intensité**

Enfin, le tatouage extrait après égalisation d'histogramme et ajustement de contraste sont présenté sur la figure 3-39. Le BER est égal à zéro pour toutes les images tatouées par les trois méthodes JND testées et la méthode JNDEM proposée ce n'est pas le cas pour la méthode DCT. Les figures 3-40 et 3-41 présentent successivement un exemple, sur l'image Barbara, des attaques par égalisation d'histogramme et ajustement de contraste. Le tableau 3-12 présente les valeurs du BER de 50 images.

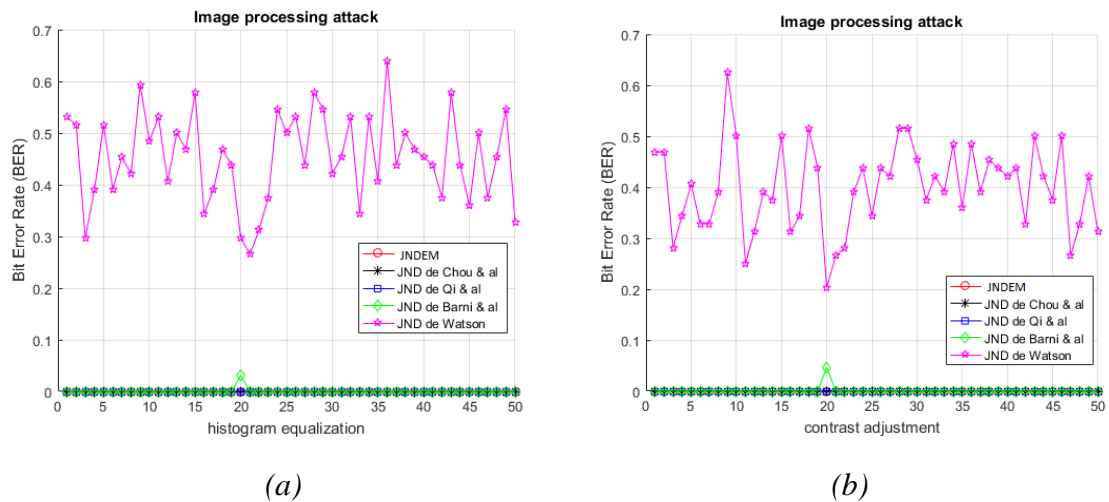


Figure 3-37: Valeurs de détection BER pour les images tatouées après : (a) égalisation d'histogramme et (b) ajustement de contraste.



Figure 3-38: Exemple sur l'image Barbara de l'attaque égalisation d'histogramme.



IRF—SIC

La marque extraite  
Ber=0

Figure 3-39: Exemple sur l'image Barbara de l'attaque ajustement de contraste

Types d'attaque	La moyenne du BER de 50 images					
	JND de Chou [66]	JND de Qi [69]	JND de Barni [77]	JND de Watson [70]	JNDEM	
Egalisation d'histogramme	Maximum	0	0	0	0.6406	0
	Minimum	0	0	0	0.2656	0
	Moyenne	<b>0</b>	<b>0</b>	<b>0</b>	0.4537	<b>0</b>
Ajustement contraste	Maximum	0	0	0	0.6250	0
	Minimum	0	0	0	0.2031	0
	Moyenne	<b>0</b>	<b>0</b>	<b>0</b>	0.3991	<b>0</b>

Tableau 3-12: Valeurs BER de 50 images après les attaques EH et AC

En conclusion, d'après les résultats ci-dessus, nous pouvons voir que notre JNDEM est assez robuste aux attaques et donne un meilleur résultat comparé aux autres JND. Tout au long de ces résultats, il est clair que l'utilisation de modèles perceptuels contribue à améliorer non seulement la transparence mais aussi la robustesse du tatouage. En conclusion, la méthode proposée donne un meilleur compromis entre la qualité visuelle et la résistance aux attaques parmi les méthodes testées. Le succès de notre méthodes est due à l'utilisation du discriminateur qui nous permet d'insérer d'insère la marque d'une manière optimale contrairement aux autres méthodes qui laissent le système décider de l'insertion sans tenir compte de la zone d'insertion. De plus l'insertion en utilisant l'étalement de spectre permet d'avoir une redondance et offre une robustesse supplémentaire.

### **3.4 Conclusion**

Dans ce chapitre, nous avons présenté notre méthode JNDEM basée sur un discriminateur de texture à base d'ondelette de Faber-Schauder à échelles mixées. Ce discriminateur permet de séparer les différentes zones de l'image à savoir, uniforme, texture et contour. Ceci permet d'appliquer des modèles JND spécifiques, pour chaque type de région, considérés comme des références dans la communauté du tatouage numérique. Notre algorithme choisit le masque optimal adapté à chacune des régions à savoir le masque de luminance de Chou pour les régions uniformes et son masque de contours pour les zones de contours et enfin le masque de Qi pour les régions texturées. On a aussi pris en compte la règle 4 de Barni concernant la sensibilité angulaire du SVH.

Nous avons commencé par tester notre JNDEM sur deux bases de texture de la littérature et avons évalué la qualité visuelles des images tatouées avec JNDEM et les avons comparé aux mêmes images mais sans utilisation du masque. Les résultats obtenus ont été très satisfaisants et ont montré l'augmentation au niveau de la qualité visuelle en ajoutant un masque JND à l'algorithme de tatouage.

Nous avons ensuite testé à la fois la qualité visuelle et la robustesse aux attaques de la méthode de tatouage proposée et nous l'avons comparée à quatre autres méthodes de pointe avec différents domaines d'insertion : méthode spatiale de Chou, méthodes spatiale de QI, méthode DCT de Watson et la méthode ondelettes de Barni. Notre méthode donne une haute qualité visuelle et elle très similaire en termes d'imperceptibilité au JND de Barni qui donnent globalement les meilleurs résultats en terme de qualité visuelle. Notre méthode JNDEM est très robuste aux attaques (compression JPEG, filtrages et ajout du bruit) et surpasse les autres modèles JND. Les résultats obtenus montrent que la méthode JNDEM est particulièrement robuste aux dégradations de type traitement d'image et fournit des résultats très encourageants face à ces attaques testées, mais reste cependant peu performant face à des attaques géométriques importantes.

Notre méthode JNDEM est le meilleur compromis entre qualité visuelle et la résistance aux attaques comparée aux 4 autres méthodes testées. Ces résultats s'expliquent par le fait que notre JNDEM combine à la fois les avantages du discriminateur de texture ce qui permet de respecter la sensibilité de l'œil aux différentes régions. Avec note JNDEM on peut insérer

plus dans les zones sombres et très lumineuse et aussi texturées tout en respectant la sensibilité de l'œil humaine aux zones de contours. De plus l'insertion par étalement de spectre offre grâce à son redondance l'alternative de retrouver la marque de tatouage même après de forte dégradation dû aux attaques.

## Chapitre 4

### 4. Nouveau masque de luminance

#### 4.1 Introduction

L'essentiel des travaux dans la littérature sur le tatouage numérique est basé sur le masque de luminance de Chou qui est un modèle conçu pour la compression. Dans ce chapitre nous proposons un nouveau masque de luminance basé sur des expériences subjectives conçues spécifiquement pour le tatouage.

#### 4.2 Test subjectif

Chou a proposé un modèle de luminance adaptée à la compression. Il a effectué des tests subjectifs en utilisant un fond uniforme à l'intérieur duquel se trouvait un carré uniforme de taille 32x32 pixels (figure 4-1-a). On fait varier la luminance du carré jusqu'à ce que le sujet décèle un changement.

Cependant dans le tatouage numérique la marque est considérée comme un bruit qui s'ajoute à l'image. Pour cela nous avons adapté l'expérience de Chou en utilisant un fond uniforme et en ajoutant un bruit sur toute l'image (figure 4-1-b). On fera varier l'intensité du bruit jusqu'à ce que le sujet décèle un changement.

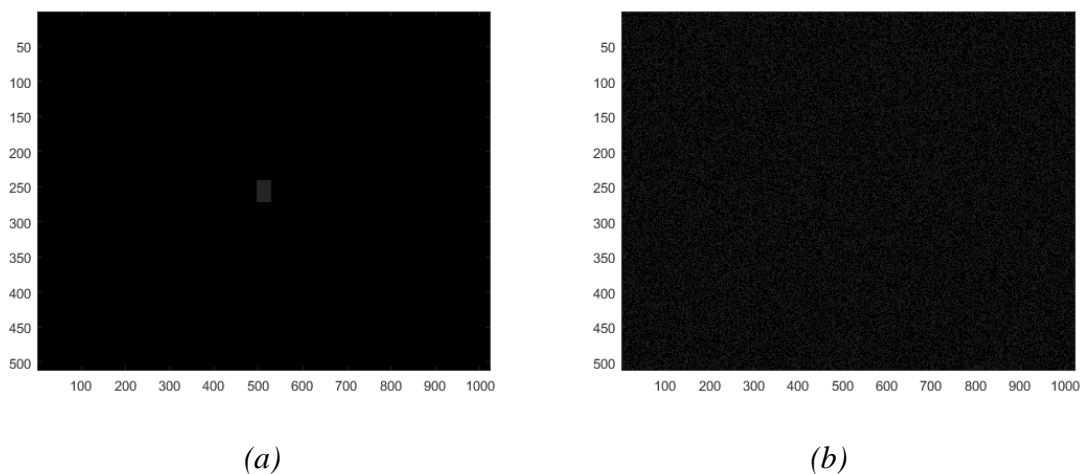


Figure 4-1: (a) L'expérience dans [66], (b) notre expérience.

Nous avons effectués les deux expériences, celle de Chou et notre adaptation, en demandant à des observateurs de noter le seuil de visibilité. Nous avons respecté les mêmes conditions que celles de [66]. Nous avons fait appel à 10 observateurs à qui on a présenté les deux fonds avec le carré et avec le bruit séparément dans une salle avec une lumière faible sur un écran Hp w2207 22 " le contraste et la luminosité de l'écran ont été fixée à 28, une distance de vision de 6 fois la hauteur de l'image est choisie. Une fiche contenant 25 niveaux de gris (de 0 à 255 avec un pas de 8) a été donnée aux observateurs. Pour chaque niveau de gris l'amplitude du bruit est augmentée, pour le carré et pour le bruit. Le seuil de visibilité pour chaque niveau de gris est ainsi déterminé lorsque le bruit ou le carré est juste visible par l'observateur.

Le résultat des expériences se trouve sur la figure 4-2. On voit clairement que l'expérience avec le bruit sur toute l'image donne de meilleur résultat que celle où on a un carré au centre dans le sens où le seuil de visibilité est toujours supérieur. Cela permettra de forcer d'insertion pour le tatouage numérique. On notera sur la figure 4-2 que la pente approximative de la ligne qui relie le seuil de visibilité à la différence de luminance tend à augmenter légèrement lorsque la luminance de fond augmente.

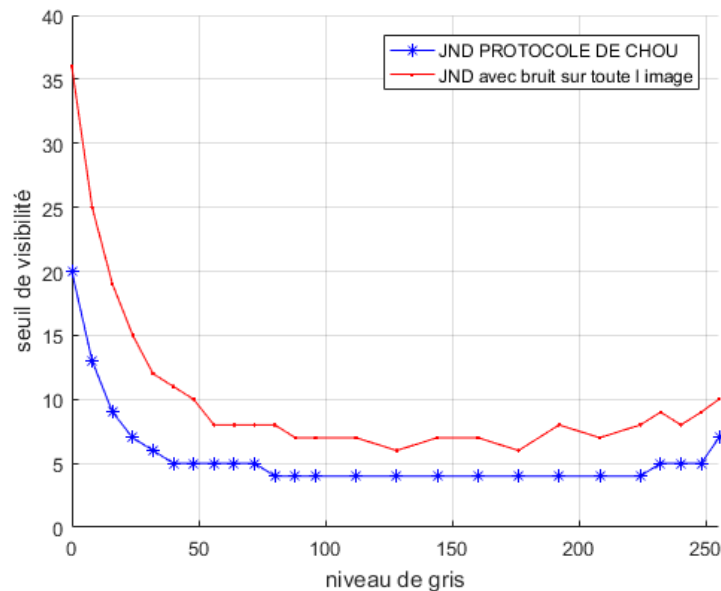


Figure 4-2 : Résultat des tests subjectifs la courbe en bleu la reproduction de l'expérience dans [66], la courbe en rouge notre expérience.



### 4.3 Modélisation de la courbe expérimental

Nous avons modélisé la courbe expérimentale par un polynôme, pour choisir le degré du polynôme nous avons fait plusieurs tests pour trouver le degré qui donnera l'allure la plus proche à la courbe expérimentale. Sur la figure 4-3 quelques graphes avec différent degrés. Les polynômes de degré 6, 7 et 8 donnent l'allure qui ressemble le plus à la courbe expérimentale. Nous avons opté pour un polynôme de degré 7.

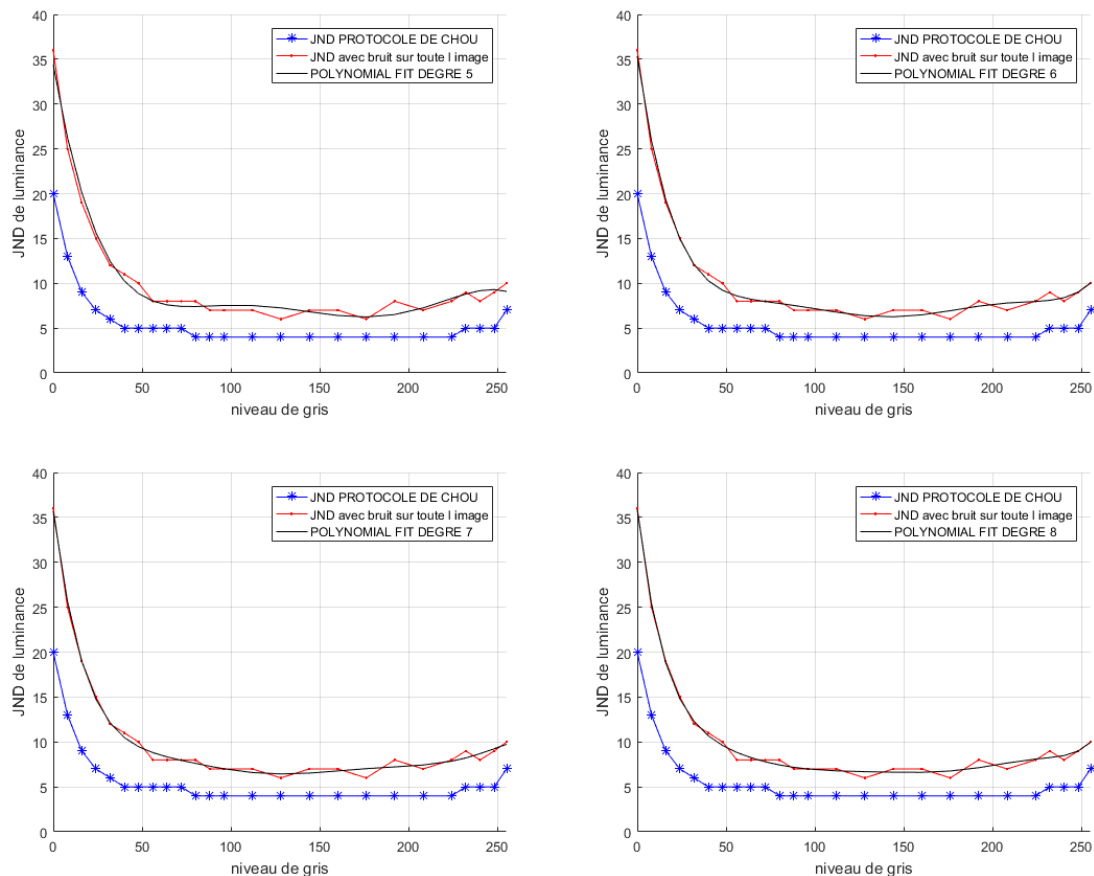


Figure 4-3: La courbe noire représente le polynôme de degré  $n$  avec  $n$  variant de 5 à 8.

Dans [66] figure 4-4-b lorsque le niveau de gris est 0, c'est à dire totalement sombre, le seuil de masquage est de 20. Cela signifie que si on change la valeur de pixel d'origine de 0 à 20, la vision humaine ne fera pas la différence. Lorsque le niveau de gris atteint 255 ce qui indique qu'il est lumineux, le seuil de masquage est 6. Par contre dans notre expérience si le niveau de gris est 0 le seuil de masquage peut atteindre 36 et lorsque le niveau de gris est 255, le seuil arrive à 10.

Le nouveau seuil de visibilité dû à la luminance de fond est donné par fonction  $JND_L$ , dans lequel la relation entre la sensibilité au bruit et la luminance de fond est vérifiée par le test subjectif.

$$JND_L(b_g) = -3.65 \cdot 10^{-14} \cdot b_g^7 + 3.87 \cdot 10^{-11} \cdot b_g^6 - 1.6710^{-8} \cdot b_g^5 + 3.8310^{-6} \cdot b_g^4 + 0.036 \cdot b_g^2 - 1.51 \cdot b_g + 35.66$$

$$b_g(x, y) = \frac{1}{32} \sum_{i=1}^5 \sum_{j=1}^5 I_o(x-3+i, y-3+j) B(i, j),$$

$$B(i, j) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 2 & 0 & 2 & 1 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Avec  $0 \leq x < H$ ,  $0 \leq y < W$ ,  $H$  et  $W$  désignent respectivement la hauteur et la largeur de l'image  $I_o$ .  $b_g(x, y)$  est la luminance moyenne de l'arrière-plan qui est calculée en utilisant le filtre moyennneur  $B$ .

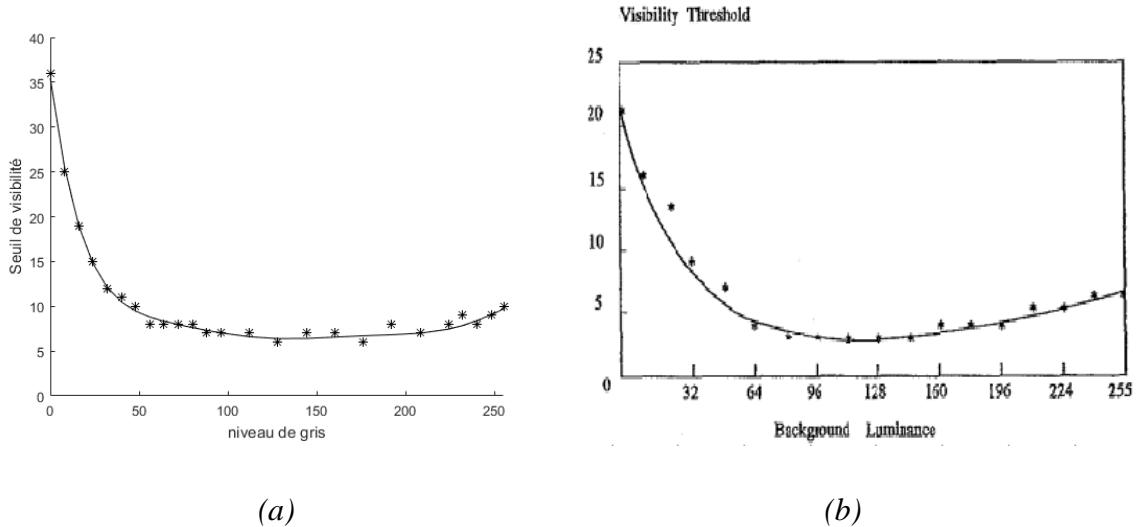


Figure 4-4: (a) Seuils de visibilité expérimentale dus à la luminance de fond. (b) Seuils de visibilité dus à la luminance de l'article [66]

La figure 4-5 présente quelques exemples du masque de luminance de chou tel qu'il est dans son article [66] et le masque de luminance adapté au tatouage numérique que nous avons développé. On voit bien que le masque de luminance que nous avons proposé offre de meilleur résultat au niveau du seuil de visibilité

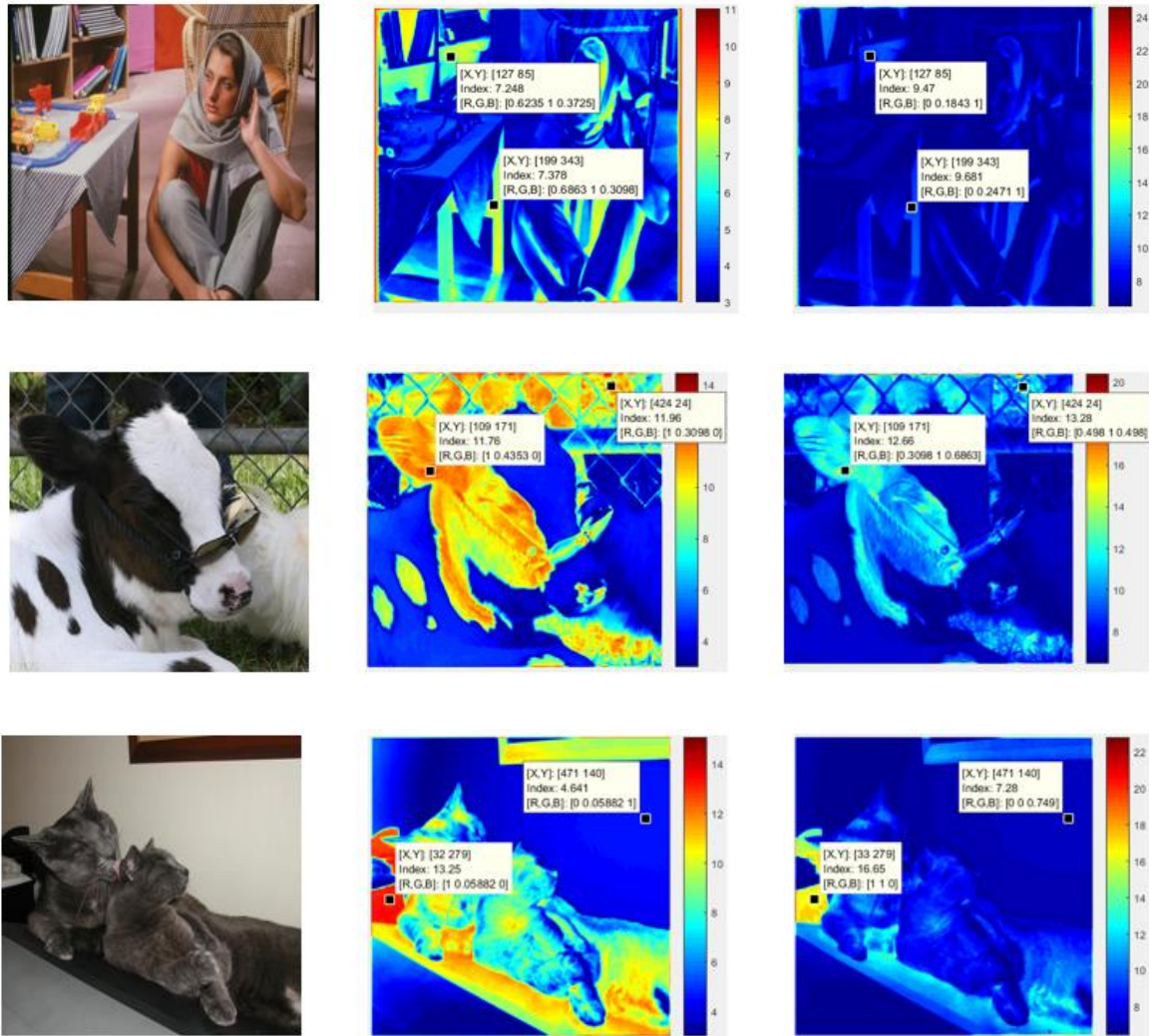


Figure 4-5: Exemple de comparaison entre le masque de luminance de l'article [66] et le masque de luminance que nous avons développé.

## Conclusion

Dans ce chapitre nous avons proposé un nouveau masque de luminance adapté spécifiquement au tatouage numérique. Nous avons effectué des tests subjectifs en faisant appel à des observateurs. La courbe expérimentale pour le test subjectif a été tracée et nous l'avons modélisée par une fonction polynôme. Les résultats obtenus ont été très satisfaisants dans le sens où notre masque de luminance permet plus de force d'insertion dans le tatouage numérique.

## Conclusion générale

Dans cette thèse, nous avons effectué dans un premier temps un examen détaillé de l'état de l'art des méthodes de tatouage d'image. Différents aspects et contraintes ont été abordés surtout pour trouver un compromis entre la robustesse, l'imperceptibilité et la capacité d'insertion. Ceci nous a orienté naturellement vers une approche basée le tatouage psychovisuel qui est actuellement la plus performante et donc a nécessité l'étude des propriétés du système visuel humain (SVH).

Dans le deuxième chapitre, nous avons décrit certaines caractéristiques pertinentes du SVH pour le tatouage comme la sensibilité de l'œil humain à la luminance, au contraste et l'effet de masquage psychovisuel. Cette étude permet de fonder les développements méthodologiques et expérimentaux pour l'élaboration de masques perceptuels garantissant le meilleur compromis entre l'invisibilité et la robustesse.

Dans le troisième chapitre, nous avons présenté notre contribution sur le développement d'un masque psychovisuel basée sur les ondelettes à échelles mixées (JNDEM). C'est une méthode composée de deux parties : dans la première partie nous avons proposé un discriminateur de texture basé sur la densité des coefficients dominants. La seconde étape consiste alors à séparer, en se basant sur un seuil de la densité des coefficients dominants, les différentes zones de l'image à savoir zones uniformes, texturés et de contours. Ces seuils ont été choisis d'après les résultats expérimentaux réalisés sur une base d'images naturelles.

Par la suite nous avons proposé une méthode de tatouage d'images basé notre modèle JNDEM. Notre méthode choisit un masque adapté à chacune des régions discriminée : le masque de luminance de Chou pour les régions uniformes, le masque de contour de Chou pour les zones de contours et enfin le masque de texture de Qi pour les régions texturées. Le modèle prend aussi en compte la sensibilité angulaire du SVH. Enfin, nous avons proposé une méthode de tatouage sur des images couleur avec une insertion multi-bit.

Afin de valider notre méthode de tatouage, nous avons effectué des tests objectifs et subjectifs et une étude de la robustesse. Dans un premier temps, nous avons évalué la qualité visuelle des images de synthèses de textures tatouées. Les résultats obtenus ont été très satisfaisants en termes de WPSNR et de SSIM. Ces premiers résultats nous ont motivé à

approfondir les tests objectifs sur des images naturelles en comparant notre méthode à quatre modèles de la littérature. Nous avons aussi essayé notre méthode en effectuant un test subjectif en faisant appel à des observateurs humains pour l'évaluation de la qualité visuelle. Les résultats obtenus ont montré que notre modèle JNDEM à une très bonne qualité visuelle en termes de tolérance à la distorsion et de qualité perceptuelle. Les tests de la robustesse aux attaques standards comme la compression JPEG, le filtrage, cropping, le bruit et le filtrage, montrent que la méthode de tatouage proposée résiste avec succès aux opérations de traitement de signal sévères mais ils sont relativement fragiles vis-à-vis des attaques de désynchronisation. De plus elle est robuste aux attaques. Ce dernier point est d'une grande importance pour les applications réelles. En conclusion, la méthode proposée est le meilleur compromis entre la qualité visuelle et la résistance aux attaques comparée aux quatre autres méthodes testées.

Dans le quatrième chapitre, comme perspective à notre travail, nous avons proposé un nouveau masque de luminance adapté au tatouage numérique. Nous avons effectué des tests subjectifs en faisant appel à des observateurs humains. La courbe expérimentale obtenue a été modélisée et pourra remplacer le masque de luminance de Chou qui a été développé pour la compression des images numériques

## Perspectives

La méthode de tatouage JNDEM présentée dans ce mémoire de thèse peut être améliorée.

Pour le masque de luminance proposé dans le chapitre 4, son utilisation dans le modèle JNDEM proposé serait susceptible de conduire à une plus grande efficacité tout en garantissant une bonne imperceptibilité.

Comme autre perspective, nous visons à développer un nouveau masque de texture adapté au tatouage. Il se base sur la Décomposition Empirique Multimodale (EMD pour Empirical Mode Decomposition en anglais). L'EMD postule que tout signal réel oscillant se décompose en une composante fortement oscillante, et une moyenne locale qui contient les variations lentes du signal. Cette dernière se décompose elle-même, ce qui conduit à une décomposition récursive multi-échelles. Une fois le masque de texture développé nous pourrions l'utiliser avec le masque de luminance pour les tester en tatouage numérique. Nous pensons que cet axe de recherche est très prometteur. C'est dans cette direction que nous orienterons nos prochains travaux.

Le système de décision a démontré son efficacité mais des pistes d'amélioration sont possibles. D'une part, les bornes de discrimination des 3 masques ont été déterminées de manière empirique. D'autre part, le système de décision actuel reste simpliste dans le sens où un seul masque est appliqué à la fois sur la base d'un seuillage binaire. Une autre perspective d'amélioration serait de travailler à partir d'un système de décision basé sur l'agrégation des masques. Chaque masque apporterait sa contribution, et la combinaison serait plus représentative de la réalité. De nombreuses techniques d'agrégation existent, celles basées sur les systèmes d'inférence floue sont bien adaptées à ce type de problématique. Dans la continuité, un réseau de neurones "neuro-flou" pourrait être développé. L'apprentissage supervisé devrait permettre, à partir de l'expertise existante, d'optimiser à la fois les bornes de discrimination des 3 masques et les coefficients d'agrégation, l'ensemble permettant d'améliorer la performance globale du système de décision.

---

## Bibliographie

- [1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans Image Process*, vol. 6, pp. 1673-87, 1997
- [2] M. El-Hajji, "La sécurité d'images par le tatouage numérique dans le domaine d'ondelettes," Université Ibnou Zohr, Faculté des Sciences, Agadir, 2012.
- [3] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital watermarking and steganography*: Morgan Kaufmann, 2007.
- [4] K. Tanaka, Y. Nakamura, K. Matsui , "Embedding secret information into a dithered multi-level image," *IEEE Military Communications Conference*, vol. 1, pp. 216-220, September-october 1990.
- [5] W. Bender, D. Gruhl et N. Morimoto , "Techniques for data hiding," *SPIE conference on Storage and Retrieval for Image and Video Databases III*, vol. 2420, pp. 164-173, Février 1995.
- [6] I. Pitas, T. Kaskalis, "Applying signature on digital image," *Workshop on nonlinear signal and image processing, Neo marmaras IEEE*, pp. 460-463, June 1995.
- [7] W. Lu, H. Lu, "Robust watermarking based on sub-sampling and nonnegative matrix factorization", *Informatica*, Vol. 19, No. 4, pp. 555- 566, December 2008.
- [8] P. Bas, "Méthode de tatouage d'images fondées sur le contenu," 2000.
- [9] R. RIAD, "Tatouage robuste d'images imprimées." Thèse de doctorat. Université d'Orléans 2015.
- [10] F. Ros, J. Borla, F. Leclerc, R. Harba, N. Launay, "Watermarking for Plastic Card Supports," *9ème Conférence Maghrébine sur Les Technologies de L'Information, MCSEAI*, 2006.
- [11] V. Solachidis and L. Pitas, "Circularly symmetric watermark embedding in 2-D DFT domain," *IEEE Trans Image Process*, vol. 10, pp. 1741-53, 2001.
- [12] V. Licks and R. Jordan, "On digital image watermarking robust to geometric transformations," in *International Conference on Image Processing*, 2000, pp. 690-693.

- 
- [13] A. Poljicak, L. Mandic, and D. Agic, "Discrete Fourier transform–based watermarking method with an optimal implementation radius," *Journal of Electronic Imaging*, vol. 20, pp. 033008-033008-8, 2011.
- [14] R. Riad, H. Douzi, M. El-hajji, R. Harba, F. Ros, "A Print-Scan Resilient Watermarking based on Fourier Transform and Image Restoration". *International Journal of Computer Applications*, Published by Foundation of Computer Science (FCS), NY, USA. Vol. 128(15), pp.13-17, October 2015.
- [15] K.R. RAO, P. YIP, 'Discrete cosine transform: algorithms, advantages, applications," *Academic Press Inc*, 1990.
- [16] A. B. Watson and J. Solomon, "Model of visual contrast gain control and pattern masking," *Journal of the Optical Society of America*, vol. 14, p. 2379–2391, September 1997.
- [17] E. Kosh, J. Zhao, "Towards Robust and Hidden Image Copyright Labeling," *IEEE Workshop on Nonlinear Signal and Image Processing*, October 1995.
- [18] M. A. Suhail and M. S. Obaidat, "Digital watermarking-based DCT and JPEG model," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, pp. 1640-1647, 2003.
- [19] Wei, G. U. O., and Q. I. Wenfa. "An Improved Print-Scan Resistant Watermarking Scheme." *Applied Mechanics & Materials* 731 (2014).
- [20] S. A. Parah, J. A. Sheikh, N. A. Loan, and G. M. Bhat, "Robust and blind watermarking technique in DCT domain using inter-block coefficient differencing." *Digital Signal Processing* 53 (2016): 11-24.
- [21] Y. Hu, Z. Wang, H. Liu, G. Guo. "A Geometric Distortion Resilient Image Watermark Algorithm Based on DWT-DFT." *Journal Of Software*, Vol. 6(9): pp. 1805-1812. 2011.
- [22] L. Quan, and AI. Qingsong. "A combination of DCT based and SVD based watermarking," *ICSP proceedings of IEEE International conference on signal processing*, pp. 873-876, 2004.
- [23] T. Stathaki, and P. Dafas, "Digital Image Watermarking Using Block-Based Karhunen-Loeve Transform." *Proceedings of the 3rd International Symposium (ISPA), Rome, Italy*, pp. 1072–1075, September 18-20, 2003.



- 
- [24] V.Sanathi, and P. Arulmozhivarman. "Hadamard transform based adaptive visible/invisible watermarking scheme for digital images." *Journal of Information Security and Applications* 18.4 (2013): 167-179.
- [25] A.Mansouri, A. Mahmoudi Aznaveh, and F. Torkamani Azar. "SVD-based digital image watermarking using complex wavelet transform." *Sadhana* 34.3 (2009): 393-406.
- [26] J. J. K. Ò. Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," *Signal Processing*, vol. 66, pp. 303-317, 5/28/ 1998.
- [27] J. Zhao, "Robust Image Watermarking Algorithm Based on Radon and Analytic Fourier-Mellin Transforms," *Open Automation and Control Systems Journal*, vol. 7, pp. 1071-1074, 2015.
- [28] K. LOUKHAOUKHA, "Tatouage numérique des images dans le domaine des ondelettes basé sur la décomposition en valeurs singulières et l'optimisation multi-objective."Thèse de doctorat. Université Laval2010.
- [29] S. G. MALLAT, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation." *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 11(7):674-693, juillet 1989.
- [30] W. SWELDENS: The Lifting Scheme, "A new philosophy in biorthogonal wavelet constructions." In *SPIE conference on Wavelet Applications in Signal and Image Processing III*, vol. 2569, p. 68-79, San Diego, USA, juillet 1995.
- [31] SWELDENS, Wim. The lifting scheme, "A construction of second generation wavelets." *SIAM journal on mathematical analysis*, 1998, vol. 29, no 2, p. 511-546.
- [32] H. Douzi, "Base d'ondelettes de Faber\_Schauder et applications au traitement d'images," PhD thesis, University of Paris 6, 2001
- [33] M. El Hajji, H. Douzi, R. Harba, D. Mammass, and F. Ros, "New image watermarking algorithm based on mixed scales wavelets," *J. Electron. Imaging* 21(1), 1–7, 2012
- [34] S. Basheera, D. B. Prakash, and Naganjaneyulu, "Blind Medical Image Watermarking Technique for Secure Recovery of Hidden Data." In *Advances in Digital Image Processing and Information Technology* (pp. 185-192). Springer Berlin Heidelberg2011.

- 
- [35] A. Ahmad, G. R. Sinha, and N. Kashyap, "3-Level DWT Image Watermarking against Frequency and Geometrical Attacks." *International Journal of Computer Network and Information Security (IJCNIS)*, 6(12), 58, 2014.
- [36] P. Maneesha, K. Yogendra, "Digital Watermarking Algorithm for Embedding Color Image using Two Level DWT." *International Journal of Computer Applications (IJCA)*. (pp. 19- 24), 116(13), 2015
- [37] S. Voloshynovskiy, S. Pereira, V. Iquise, T. Pun, "Attack modelling: Towards a second generation watermarking benchmark," *Proc. Signal Processing*, vol. 81, p. 1177–1214, 2001.
- [38] C.-Y. Lin and S.-F. Chang, "Distortion Modeling and Invariant Extraction for Digital Image Print-and-Scan Process," *ISMIP 99*, December 1999.
- [39] M. Kutter and F. A. P. Petitcolas, "Fair benchmark for image watermarking systems," in *Security and Watermarking of Multimedia Contents*, pp. 226-239, 1999.
- [40] S. Craver, N. Memon, B.-L. Yeo and M. Yeung, "Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks, and implementations.," *IEEE Journal on Selected Areas in Communications (Special issue on Copyright and Privacy Protection)*, vol. 16, n° 14, pp. 473-586, May 1998.
- [41] N. Nikolaidis, V. Solachidis, A. Tefas, V. Arguriou, and I. Pitas, "Benchmarking of still image watermarking methods: Principles and state of the art," *Proc. of Electronic Imaging and the Visual Arts (EVA2002)*, 2002.
- [42] Dong, C., Li, J., & Chen, Y. W. (2012, May). "A DWT-DCT Based Robust Multiple Watermarks for Medical Image." In *Photonics and Optoelectronics (SOPO), 2012 Symposium on* (pp. 1-4). IEEE.
- [43] A. Dhaou, S. Jabra, and E. Zagrouba. "Nouvelle approche robuste de tatouage des images médicales basée sur les ROI et RONI A new robust approach of Medical image watermarking based on ROI and RONI."
- [44] A. Khan Kamran, S.A. Malik "A high capacity reversible watermarking approach for authenticating images: exploiting down-sampling, histogram processing, and block selection," *Inf. Sci.*, 256 (2014), pp. 162–183

- [45] H. H. Ralaivao, P. A. Randriamantsoa, and T. Raminoson. "S curisation des images par combinaison de tatouage et de cryptage." MADA-ETI, ISSN 2220-0673, Vol.1, 2016
- [46] L.R. Haddada, B. Dorizzi, and N. E. Ben Amara. "A combined watermarking approach for securing biometric data." *Signal Processing: Image Communication* 55 (2017): 23-31.
- [47] V. Itier, W. Puech, and A. G. Bors, "Cryptanalysis aspects in 3-D watermarking." In *Image Processing (ICIP), 2014 IEEE International Conference on* (pp. 4772-4776). IEEE, October 2014.
- [48] N. Heydari, and B. Minaei-Bidgoli, "Reduce Energy Consumption and Send Secure Data Wireless Multimedia Sensor Networks Using A Combination of Techniques For Multi-Layer Watermark And Deep Learning". *International Journal of Computer Science and Network Security (IJCSNS)*, 17(2), 98, 2017.
- [49] Uchida, Y., Nagai, Y., Sakazawa, S., & Satoh, S. I.. "Embedding Watermarks into Deep Neural Networks." arXiv preprint arXiv: 1701.04082, 2017.
- [50] R. A. Young, "Oh say, can you see? the physiology of vision," in *Proc. SPIE Human Vision, Visual Processing and Digital Display*, San Jose, CA, vol. 1453, pp. 92–123,1991.
- [51][https://www.google.com/search?q=coupe+lat%C3%A9rale+oeil+humain&newwindow=1&source=lnms&tbm=isch&sa=X&ved=0ahUKEwjYp5n48rvZAhUDVxQKHcMLCEMQ\\_AUICigB&biw=1366&bih=637#imgrc=upFO9DvvIoPsCM](https://www.google.com/search?q=coupe+lat%C3%A9rale+oeil+humain&newwindow=1&source=lnms&tbm=isch&sa=X&ved=0ahUKEwjYp5n48rvZAhUDVxQKHcMLCEMQ_AUICigB&biw=1366&bih=637#imgrc=upFO9DvvIoPsCM):
- [52] <https://www.bioinformatics.org/oeil-couleur/dossier/retine.html>
- [53] L. Rastislav, ed. "Perceptual digital imaging: methods and applications." CRC Press, 2012
- [54] <http://www.rennes.supelec.fr/ren/perso/jweiss/tv/perception/image52.gif>
- [55] G. T. Fechner, "Elemente der psychophysic," *Ch. XVI, Breitkopf und H rtel, Leipzig*, vol. II, 1860.
- [56] A. Michelson, "Studies in optic," *Chicago Press*, 1927.
- [57] P. Moon and D. E. Spencer, "The visual effect of nonuniform surrounds," *Journal of the Optical Society of America*, vol. 35, pp. 233–248, 1945

- [58] M. K. Kundu and S. K. Pal, "Thresholding for edge detection using human psychovisual phenomena," *Pattern Recognition Letter*, vol. 4, pp. 433–441, 1986.
- [59] F. Kretz, "Subjectively optimal quantification of pictures," *IEEE Transactions on Communications*, vol.25, pp. 1288–1292, 1975.
- [60] K. Belkacem-Boussaid and A. Beghdadi, "A new image smoothing method based on a simple model of spatial processing in the early stages of human vision," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, February 2000.
- [61] E. Peli, "Contrast in complex images," *Journal of Opt. Soc. Am. A*, vol. 7, no. 10, pp. 2032–2040, 1990.
- [62] R. G. Robson, "Spatial and temporal contrast sensitivity functions of the visual system," *Journal of Optical Society of America*, vol. 56, pp. 1411–1442, 1966.
- [63] P. G. J. Barten, "Evaluation of subjective image quality with the square-root integral method," *Journal of the Optical Society of America A: Optics, Image Science and Vision*, vol. 7, no. 10, pp. 2024–2031, 1990.
- [64] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *Journal of the Optical Society of America*, vol. 70, pp. 1458–1471, 1980.
- [65] M. Barni, and F. Bartolini, eds. "watermarking systems engineering: enabling digital assets security and other applications." CRC Press, 2004.
- [66] C.H. Chou, and Y.C. Li, "A perceptually tuned subband image coder based on the measure of Just noticeable-distortion profile," *IEEE Trans. Circ. Syst. Video Technol.* 5(6), 467–476, 1995.
- [67] X. Zhang, W. Lin, Z. Lu, E. Ong, and S. Yao, "Motion-compensated residue pre-processing in video coding based on just-noticeable-distortion profile," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 742–750, June 2005.
- [68] X. Zhang, W. Lin, Z. Lu, E. Ong, and S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Processing: Image Communication*, vol. 20, no. 7, pp. 662–680, August 2005.

- [69] Q. Huiyan, Z. Dong and Z. Jiying, "Human visual system based adaptive digital image watermarking," *Sig. Process.* 88(1), 174–188, 2008
- [70] A. B. Watson, "Dct quantization matrices visually optimized for individual images," in *Proc. SPIE Int. Conf. Human Vision, Visual Processing, and Digital Display*, 1993, vol. IV, pp. 202–216.
- [71] H. A. Peterson, A. J. Ahumada Jr., and A. B. Watson, "An improved detection model for dct coefficient quantization," in *SPIE Proceedings*, 1993, vol. 1913, pp. 191–201.
- [72] H. A. Peterson, A. J. Ahumada Jr., and Andrew B. Watson, "The visibility of dct quantization noise," *Society for Information Display Digest of Technical Papers XXIV*, pp. 942–945, 1993.
- [73] A. J. Ahumada Jr. and H. A. Peterson, "Luminance-model-based dct quantization for color image compression," *Proceedings of the SPIE, Human Vision, Visual Processing, and Digital Display III*, 1992.
- [74] X. Zhang, W. Lin, and P. Xue, "Improved estimation for just-noticeable visual distortion," *Signal Processing*, vol. 85, no. 4, pp. 795–808, Avril 2005.
- [75] Z. Liu, L. J. Karam, and A. B. Watson, "Jpeg2000 encoding with perceptual distortion control," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1763–1778, 2006.
- [76] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1164–1175, August 1997.
- [77] M. Barni, and F. Bartolini, Franco, and PIVA, Alessandro. "Improved wavelet-based watermarking through pixel-wise masking." *Image Processing, IEEE Transactions on*, vol. 10, no 5, p. 783-791, (2001)
- [78] P.Bang Nguyen, A. Beghdadi and M. Luong, "Perceptual watermarking using a new Just-Noticeable-Difference model," *Sig. Proc. Image Commun.* 28(10), 1506–1525, 2013
- [79] "UIT-T P.910 subjective video quality assessment methods for multimedia applications," International Telecommunication Union-Telecommunication sector, 1999.

- [80] "UIT-R BT.500-11 methodology for the subjective assessment of the quality of television pictures," International Telecommunications Union - Radiocommunication sector, 2002.
- [81] A. Chetouani, "Vers un système d'évaluation de la qualité d'image multi-critères." Thèse de doctorat 2010, Paris; Université Paris-Nord-Paris XIII.
- [82] Durant. "Méthodes d'évaluation subjective de la qualité audiovisuelles pour applications multimédias." Recommandation UIT-T P.911.
- [83] Z. Wang, A. C. Bovik, H. R. Sheikh, and P. E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*.
- [84] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, no. 4, pp. 81–84, March 2002.
- [85] S. Pereira, S. Voloshynovskiy, M. Madueño, S. Marchand-Maillet, and T. Pun, "Second generation benchmarking and application oriented evaluation," in *Proceedings of Third International Workshop on Information Hiding*, Pittsburgh, PA, U.S.A., April 2001, pp. 219–239.
- [86] S. Pereira, "Robust digital image watermarking," Ph.D. thesis, University of Geneva, Swiss, 2000.
- [87] R. Shnayderman, E. Gusev, and A. M. Eskicioglu, "An svd-based gray-scale image quality measure for local and global assessment," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 422–429, February 2006.
- [88] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. of the 37th IEEE Asilomar Conf. on Signals, Systems, and Computers*, pp. 1398–1402, 2003.
- [89] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," *In Information Hiding*, Springer Berlin Heidelberg, pp. 211-236, January 2000.

- [90] H. Douzi, D. Mammass, and F. Nouboud, "Faber-Schauder wavelet transform, application to edge detection and image characterization," *J. Math. Imag. Vis.* 14(2), 91–101, 2001
- [91] M. Amar, R. Harba, H. Douzi, F. Ros, M. El Hajji, R. Riad and K. Gourrame. "Perceptual Image Watermarking based on a Mixed-scale Wavelet Representation." *International Journal of Computer Applications* 172(8):1-9, August 2017
- [92] A. Cheddad, J. Condell, K. Curran, and M. Kevitt, "Digital image steganography survey and analysis of current methods," *Signal Process*, Vol 90, pp. 727-752, 2010
- [93] Yang, X. K., et al. "Just noticeable distortion model and its applications in video coding." *Signal Processing: Image Communication* 20.7: 662-680, 2005.
- [94] A. Pramila, K. Anja and S. Tapio. "Toward an interactive poster using digital watermarking and a mobile phone camera." *Signal, Image and Video Processing* 6.2: 211-222, 2012.
- [95] C. Naornita, "A new pixel-wise mask for watermarking." *Proceedings of the 9th workshop on Multimedia & security*. ACM, 2007

## Bibliographie de l'auteur

### *Revue:*

**M. Amar**, R. Harba, H. Douzi, F. Ros, M. El-Hajji, R. Riad and K. Gourrame "Perceptual Image Watermarking based on a Mixed-scale Wavelet Representation." *International Journal of Computer Applications* 172(8):1-9, August 2017.

### *Chapitre de livre*

**M. Amar**, R. Harba, H. Douzi, F. Ros, M. El-hajji, R. Riad, et K. Gourrame "A JND Model Using a Texture-Edge Selector Based on Faber-Schauder Wavelet Lifting Scheme," *Lecture Notes in Computer Science*, edited by Springer. Vol. 9680, pp. 328-336, 2016.

K. Gourrame, H. Douzi, R. Harba, F. Ros, M. El-hajji, R. Riad, et **M. Amar** "Robust Print-cam Image Watermarking in Fourier Domain," *Lecture Notes in Computer Science*, edited by Springer. Vol. 9680, p. 356-365, 2016.

### *Communications*

**M. Amar**, R. Harba, H. Douzi, F. Ros, M. El-hajji, R. Riad, et K. Gourrame "A JND Model Using a Texture-Edge Selector Based on Faber-Schauder Wavelet Lifting Scheme," *International Conference on Image and Signal Processing*. Springer International, pp. 328-336, Jun 30, 2016 - Jul 1, 2016. Trois rivières Québec Canada.

**M. Amar**, R. Harba, H. Douzi, F. Ros, M. El-Hajji et R. Riad "Perceptual models based on wavelet for digital Watermarking." Agadir School on Numerical Methods in Engineering, 19 - 21 October 2015, Agadir, Maroc.

**M. Amar**, H. Douzi et M. El-Hajji "Masques psychovisuels à base d'ondelette pour le watermarking." 2ème édition des journées scientifiques de la faculté des sciences d'Agadir JDFSA-Mai 2014 Agadir, Maroc.

### *Revue en cours d'évaluation*

K. Gourrame, H. Douzi, R. Harba, R. Riad, F. Ros, **M. Amar** et M. El-hajji, "Image watermarking for smartphones," *Multimedia Tools and Applications* (Soumis).



## Meina AMAR

### Masque psychovisuel à base d'ondelettes pour le Watermarking

**Résumé :** Le tatouage d'images (ou watermarking) consiste à insérer de l'information de copyright, ou d'autres messages de vérification, invisible à l'œil humain. Parmi toutes les approches possibles, le tatouage psychovisuel est le plus performant. Il consiste à utiliser les connaissances du Système Visuel Humain (SVH) pour en déduire la force du tatouage de manière optimale. Dans cette étude, nous avons proposé une nouvelle méthode de tatouage d'image psychovisuel qui combine les avantages des ondelettes à celle du domaine spatial par l'utilisation d'une représentation en ondelettes à échelle mixée (JNDEM). En considérant la densité des coefficients d'ondelettes dominants, notre système est capable de différencier les zones uniformes, celles incluant des contours, et enfin celles texturées. Nous avons sélectionné les méthodes efficaces de tatouage psychovisuel adaptée à chaque zone, à savoir celle de Chou pour les zones uniformes, de nouveau celle de Chou pour les contours, et finalement celle de Qi pour les textures. Nous avons aussi pris en compte la sensibilité du SVH qui varie avec l'orientation locale de l'image. Nous avons comparé notre JNDEM à quatre autres méthodes de tatouage de la littérature en termes de qualité visuelle et de robustesse. La méthode JNDEM conduit à une grande qualité visuelle mesurée par des tests objectifs et subjectifs. De plus elle est robuste aux attaques. Ce dernier point est d'une grande importance pour les applications réelles. En conclusion, la méthode proposée est le meilleur compromis entre la qualité visuelle et la résistance aux attaques comparée aux quatre autres méthodes testées.

**Mots-clés :** Masque psychovisuel, Système visuel humain, tatouage numérique, ondelette, échelles mixées

### Wavelet perceptual model for digital watermarking

**Abstract:** Watermarking consists in inserting copyright information, or other verification messages, invisible to the human eye. Among all the possible approaches, perceptual watermarking is the most efficient. Perceptual image watermarking consists in using knowledge of the Human Visual System (HVS) to choose the strength of the watermark. In this study, we propose a new perceptual image watermarking method (JNDEM) that combines the advantages of both the wavelet domain and the spatial domain since a mixed-scale wavelet representation is applied. By considering the density of the dominant wavelet coefficients, our scheme is able to differentiate uniform, edge and texture areas. We selected effective Just Noticeable Difference (JND) models from the literature, i.e. luminance and edge masking developed by Chou, and texture masking developed by Qi. We also took into account the HVS sensitivity which varies with the orientation of the image activity. We have compared our JNDEM to four other JND models from the literature in terms of visual quality and of robustness to attacks. Results show that the proposed method yields a high visual quality and is very robust against attacks. This last point is of great importance for real applications. In conclusion, the proposed method is the best trade-off between visual quality and resistance to attacks among the four other tested methods.

**Keywords:** perceptual models; Human Visual System; digital watermarking; wavelet; lifting scheme

Laboratoire IRF-SIC, Faculté des Sciences d'Agadir Université Ibn Zohr  
BP 8106 - Cité Dakhla, 80000 Agadir, Maroc.

Laboratoire PRISME – Pôle IRAuS, Axe Image et Vision Polytech'Orléans,  
12 Rue de Blois, BP 6744 45067 Orléans Cedex 2 France.

