



HAL
open science

Modélisation Psychovisuelle pour le tatouage des images

Florent Autrusseau

► **To cite this version:**

Florent Autrusseau. Modélisation Psychovisuelle pour le tatouage des images. Traitement du signal et de l'image [eess.SP]. Université de Nantes, 2002. Français. NNT: . tel-00616511

HAL Id: tel-00616511

<https://theses.hal.science/tel-00616511>

Submitted on 22 Aug 2011

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.

UNIVERSITÉ DE NANTES

ÉCOLE DOCTORALE STIM

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DES MATÉRIAUX

Année 2002

Thèse de Doctorat de l'Université de Nantes

Spécialité Electronique

Option : Traitement d'images

Présentée et soutenue publiquement par

Florent Autrusseau

Le 7 novembre 2002.

à l'École polytechnique de l'université de Nantes

**TATOUAGE D'IMAGES FONDÉ SUR LA MODÉLISATION
DU SYSTÈME VISUEL HUMAIN ET SUR LA TRANSFORMATION MOJETTE**

Jury

Rapporteurs :

JM. CHASSERY

H. MAITRE

Examineurs :

D. BARBA

A. SAADANE

JP. GUÉDON

M. KUNT

P. NGUYEN

Directeur de thèse : Dominique BARBA

Laboratoire : IRCCyN

Table des matières

| | |
|--|-----------|
| Glossaire | 7 |
| Introduction Générale | 9 |
| I État de l’art | 13 |
| 1 Le Système Visuel Humain | 17 |
| 1.1 Introduction | 18 |
| 1.2 Bases anatomiques et physiologiques du S.V.H. | 20 |
| 1.2.1 Anatomie de l’œil | 20 |
| 1.2.1.1 l’œil | 20 |
| 1.2.1.2 La rétine | 21 |
| 1.2.1.3 La fovéa | 22 |
| 1.2.1.4 Les photo-récepteurs | 23 |
| 1.2.2 Les voies optiques, le chiasma | 25 |
| 1.2.3 Le corps genouillé latéral | 26 |
| 1.2.4 Le cortex visuel | 27 |
| 1.3 Modélisation du SVH | 29 |
| 1.3.1 Introduction | 29 |
| 1.3.2 Perception de la luminance | 29 |
| 1.3.3 Définition du contraste | 31 |
| 1.3.4 Fonction de sensibilité au contraste (CSF) | 32 |
| 1.3.5 Masquage | 33 |
| 1.3.6 Modèles Mono-canaux | 34 |
| 1.3.7 Modèles Multi-canaux | 36 |
| 1.3.8 Modèle proposé | 42 |
| 1.3.8.1 Filtres Dom | 43 |
| 1.3.8.2 Filtres Fan | 45 |

| | | |
|-----------|---|------------|
| 1.4 | Conclusion | 47 |
| 2 | La transformation Mojette | 49 |
| 2.1 | Introduction | 50 |
| 2.2 | Transformation Mojette directe | 50 |
| 2.3 | Transformation Mojette inverse et notion de restructibilité | 52 |
| 2.3.1 | Ensemble suffisant de projections pour la reconstruction | 52 |
| 2.3.2 | Notion de fantôme et théorème de restructibilité partielle | 54 |
| 2.3.3 | Algorithme de la transformée Mojette inverse | 57 |
| 2.4 | Opérateur d'addition | 60 |
| 2.5 | Propriétés | 63 |
| 2.6 | Applications | 67 |
| 2.7 | Conclusion | 67 |
| 3 | Introduction au tatouage d'images | 69 |
| 3.1 | Introduction | 70 |
| 3.2 | historique | 70 |
| 3.3 | Les applications | 71 |
| 3.4 | Les attaques | 73 |
| 3.5 | Schéma global de tatouage | 75 |
| 3.6 | Le choix de l'espace | 76 |
| 3.6.1 | Méthodes spatiales | 78 |
| 3.6.2 | Méthodes fréquentielles | 80 |
| 3.6.3 | Méthodes utilisant des critères visuels | 82 |
| 3.6.4 | Mesures de distorsions | 89 |
| 3.6.4.1 | Calculs d'erreurs | 89 |
| 3.6.4.2 | Méthodes subjectives | 91 |
| 3.6.4.3 | Méthodes objectives | 93 |
| 3.7 | Les techniques d'extraction et de détection | 94 |
| 3.8 | Tatouage et cryptographie : liens et différences | 95 |
| 3.9 | Conclusion | 96 |
| II | Le tatouage d'images | 99 |
| 4 | Le tatouage basé sur des critères psychovisuels | 103 |
| 4.1 | Introduction | 104 |
| 4.2 | Création d'un masque perceptuel spatial | 104 |

| | | |
|----------|---|------------|
| 4.2.1 | Sélection des sites | 105 |
| 4.2.2 | Détermination des forces | 107 |
| 4.3 | Applications du masque | 111 |
| 4.3.1 | Le domaine DCT | 111 |
| 4.3.2 | Le domaine des ondelettes | 113 |
| 4.3.3 | Une application spatiale du masque | 115 |
| 4.3.4 | Expérimentations | 126 |
| 4.4 | Le tatouage psychovisuel | 129 |
| 4.4.1 | Insertion de la marque | 129 |
| 4.4.2 | Processus de détection | 131 |
| 4.4.3 | Expérimentations | 133 |
| 4.5 | Conclusion | 140 |
| 5 | Tatouage Mojette | 141 |
| 5.1 | Introduction | 142 |
| 5.2 | Utilisation des fantômes dans une méthode de tatouage | 142 |
| 5.2.1 | Introduction | 142 |
| 5.2.2 | Schéma de tatouage | 143 |
| 5.2.2.1 | Insertion des marques | 144 |
| 5.2.2.2 | Extraction de la marque | 145 |
| 5.2.3 | Gestion des clefs | 147 |
| 5.2.4 | Conclusion | 151 |
| 5.3 | Exploitation de la redondance : tatouage de flux Mojette | 152 |
| 5.3.1 | Introduction | 152 |
| 5.3.2 | Brouillage des plans binaires | 152 |
| 5.3.3 | Reconstruction des plans binaires | 157 |
| 5.3.4 | Conclusion | 161 |
| 5.4 | Une méthode de tatouage spatial adaptable à la stéganographie | 162 |
| 5.4.1 | Introduction | 162 |
| 5.4.2 | Algorithme de protection des droits d'auteurs | 162 |
| 5.4.3 | Algorithme de Stéganographie | 164 |
| 5.4.4 | Résultats | 166 |
| 5.4.5 | Conclusion | 177 |
| 5.5 | Conclusion | 177 |

| | | |
|----------|--|------------|
| 6 | Une méthode alliant masques psychovisuels et transformation Mojette | 179 |
| 6.1 | Introduction | 180 |
| 6.2 | Choix de l'espace de tatouage | 180 |
| 6.3 | Création du masque | 185 |
| 6.4 | Schéma d'insertion | 187 |
| 6.5 | Schéma de détection | 190 |
| 6.6 | Expérimentations | 195 |
| 6.7 | Conclusion | 203 |
| 7 | Contribution au projet Aquamars | 205 |
| 7.1 | Introduction | 206 |
| 7.2 | Contexte général du projet | 206 |
| 7.3 | Description des travaux | 207 |
| 7.3.1 | Détail des sous-projets | 207 |
| 7.3.2 | Contribution de l'équipe IVC | 208 |
| 7.4 | Conclusion | 209 |
| | Conclusion Générale | 213 |
| | Perspectives | 217 |
| | Annexes | 218 |
| A | Psychovisuel | 219 |
| B | Mojette | 221 |
| C | Tatouage | 223 |
| D | Attaques | 227 |
| | Bibliographie | 231 |
| | Production scientifique | 243 |

Glossaire

Modélisation Psychovisuelle

DCP : Décomposition en canaux perceptuels.

SVH : Système Visuel Humain.

CSF : Contrast Sensitivity function.

JND : Seuil de perception (Just Noticeable Difference).

BF : Basses fréquences spatiales.

HF : Hautes fréquences spatiales.

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

PSNR : Peak Signal to Noise Ratio.

TNL : Transformations Non Linéaires (fonction Gamma).

Transformation Mojette

TM : Transformation Mojette.

TMI : Transformation Mojette Inverse.

ES2P : Élément structurel à 2 pixels.

QoS : Quality of Service.

Proj : Projection Mojette.

Traitement du signal

FFT / IFFT : Transformation (Inverse) de Fourier rapide (Fast Fourier Transform).

DFT / IDFT : Transformation Discrète (Inverse) de Fourier (Discrete Fourier Transform).

DCT / IDCT : Transformation (Inverse) en cosinus discret (Discrete Cosinus Transform).

DWT / IDWT : Transformation (Inverse) en Ondettes (Discrete Wavelet Transform).

Tatouage d'images

IO : Image Originale (Image de référence).

IT : Image Tatouée.

IS : Image "stéganographie".

NG : Niveau(x) de Gris.

WM : Watermark (Filigrane).

Kpu : Clef Publique.

Kpr : Clef Privée.

Introduction Générale

Le traitement des images numériques a permis l'apparition récente des nouvelles technologies de l'information et de la communication. Cette discipline comporte différentes catégories, comme la compression, la transmission, le codage, l'évaluation de la qualité, la cryptographie et bien d'autres. Le tatouage d'images, domaine plus récent que ceux cités ci-dessus, est apparu au début des années 1990, tout d'abord dans une optique de protection des droits d'auteurs des images numériques. Les images numériques étant aisément duplicables à l'identique, le problème de la propriété intellectuelle se pose inévitablement. Pour établir un parallèle avec un autre domaine où les problèmes de propriété intellectuelle sont présents, citons le domaine artistique, et plus précisément la peinture : un tableau de maître ne peut jamais être copié exactement à l'identique, un oeil expérimenté peut le plus souvent trouver des différences entre une peinture originale et une copie. Les images numériques sont facilement duplicables au bit près, d'où la nécessité d'insérer un copyright que seule la personne reconnue comme propriétaire ou utilisateur autorisé sera apte à extraire.

Cependant, le tatouage ne se borne pas à la protection des droits d'auteurs, bien que cette application soit une des plus connues. La stéganographie, ou l'art de camoufler de l'information dans un support, est aussi un domaine très intéressant et très répandu actuellement. La principale différence entre la protection de droits d'auteurs et la stéganographie réside dans le fait qu'en stéganographie, le support importe peu, l'essentiel étant le message inséré à l'intérieur, qui, pour plus de sécurité est habituellement lui même chiffré. Pour une application de droits d'auteurs, c'est ce support même qui est protégé par l'information qu'il contient.

Il est aussi intéressant de noter l'existence de différentes applications un peu moins répandues, mais auxquelles le tatouage d'images, appelé parfois de son nom anglo-saxon *watermarking* (action d'insérer un filigrane), peut aussi être associé : l'indexation, l'authentification des données, etc ...

Nous montrerons que la cryptographie est présente dans certaines applications du tatouage, que ce soit dans les aspects de gestion des clefs ou dans le cryptage même du copyright.

Si des liens existent entre ces différents domaines, ils sont néanmoins mis en œuvre dans des applications

très différentes. La protection des droits d'auteurs a pour objectif l'insertion robuste d'informations concernant l'auteur du document ou de ses ayants-droits. La stéganographie, permet de transmettre la plus grande quantité d'information possible noyée dans un support anodin, cette information doit être invisible et indétectable pour toute personne non autorisée. La cryptographie permet d'établir une liaison sécurisée entre deux personnes en rendant la communication incompréhensible pour une tierce personne. Comme nous le constaterons par la suite, nos travaux nous ont amenés à mettre en œuvre une technique de stéganographie ainsi qu'une technique de cryptographie, ces dernières seront détaillées dans ce mémoire.

L'application la plus largement étudiée dans ce manuscrit est la protection des droits d'auteurs. Comme nous le verrons par la suite, c'est la plus exigeante et la plus complète, dans le sens où elle doit respecter de nombreuses contraintes, autant en terme de robustesse que d'invisibilité. L'objectif principal de nos travaux est d'exploiter un modèle du Système Visuel Humain (SVH), afin de permettre une optimisation des méthodes de tatouage. Une modélisation du SVH permet l'optimisation conjointe de l'invisibilité et de la robustesse de la méthode de tatouage. Une étude des différentes attaques envisageables lors d'un processus de tatouage doit aussi être mise en œuvre afin de définir un espace permettant de faire face à ces attaques, tout en respectant la contrainte d'invisibilité de la marque.

Comme nous pourrons le constater dans l'état de l'art, il existe de très nombreuses méthodes de tatouage d'images, utilisant tous types de transformations. Nous nous sommes concentrés dans ce mémoire sur deux aspects fortement développés au sein de l'équipe IVC : la modélisation psycho-visuelle et la transformation Mojette. Il semble exister certains points communs entre ces deux domaines comme par exemple la représentation angulaire et radiale de l'information image. Par contre, rien n'avait été réalisé pour relier ces deux champs de recherche avant ce travail.

Ce document est présenté en deux parties principales. Dans la première partie, nous détaillons certains aspects connexes au domaine du tatouage des images : la modélisation du système visuel humain ainsi que la transformation Mojette. Dans cette partie, nous dressons un état de l'art des techniques actuelles permettant le tatouage des images.

L'étude de ces différents aspects est prise en compte dans la seconde partie, où nous présenterons les techniques de tatouages mises en œuvre lors de cette thèse. Nous détaillerons tout d'abord diverses méthodes de tatouage exploitant des critères psychovisuels, afin d'optimiser conjointement l'invisibilité et la robustesse de la marque. Dans le but d'explorer certains aspects connexes à la protection des droits d'auteurs ou d'apporter certaines solutions face aux attaques géométriques, nous détaillerons ensuite certaines méthodes de tatouage exploitant les propriétés de la transformation Mojette. Enfin, l'optimisation des aspects perceptuels et de

la transformation Mojette sera effectuée dans une dernière méthode de tatouage dont le but est d'offrir de très bons résultats autant en terme d'invisibilité qu'en termes de robustesse vis-à-vis d'un large champ de dégradations.

Première partie

État de l'art

De nombreux outils sont exploitables pour mener à bien l'élaboration d'un mécanisme de tatouage d'images. Parmi les études menées au sein de l'équipe IVC, deux axes de recherches sont susceptibles d'apporter des éléments nouveaux au domaine du tatouage d'images. Ces deux études portent sur la modélisation psychovisuelle d'une part et sur la transformation Mojette d'autre part. Nous avons donc souhaité adapter ces deux "outils" au domaine du tatouage.

La première partie de ce mémoire est donc consacrée à l'étude de ces deux outils, mais aussi à l'étude des techniques de tatouage existantes. Un tel état de l'art pourra alors nous permettre de définir les atouts ainsi que les lacunes des principales méthodes de tatouage.

Différentes études du système visuel humain ont été menées au sein du laboratoire. Les objectifs principaux étaient la modélisation du SVH pour la visualisation des images en niveaux de gris, le codage d'images ou encore l'étude des dégradations et l'évaluation de la qualité. L'utilisation du modèle psychovisuel dans une application de codage d'images visait à supprimer une certaine quantité d'information de l'image. L'adaptation du modèle visuel dans une application de tatouage se fait non pas par la recherche de l'information susceptible d'être supprimée, mais plutôt par la recherche du seuil de visibilité. L'obtention de ce seuil de perception permet d'optimiser conjointement les processus d'insertion et de détection de la marque en obtenant le meilleur compromis entre l'invisibilité et la robustesse de la marque. Un des principaux avantages de l'utilisation d'un modèle psychovisuel pour le tatouage des images réside dans le fait que l'invisibilité de la marque est garantie sans toutefois devoir passer par des tests contraignants d'évaluation de la qualité des images tatouées.

Une étude du SVH permettant de retenir certains éléments exploitables dans le domaine du tatouage des images est conduite dans le premier chapitre. Différentes modélisations du SVH y sont détaillées et plus particulièrement le modèle développé au laboratoire. 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.

Certaines des propriétés de la transformation Mojette, telles que les notions de redondance ou de fantômes, peuvent être exploitées dans le but de "camoufler" une quantité importante d'information. L'utilisation de la transformation Mojette peut permettre la mise en œuvre d'une technique de tatouage des données, mais peut

tout aussi bien garantir la transmission de ces données. Une caractéristique intéressante de la transformation Mojette est sa capacité à caractériser l'image originale par plusieurs ensembles de données (projections) distinctes. L'étude d'une seule de ces projections n'est pas suffisante à elle seule pour définir les données originales, une certaine quantité de projections est nécessaire pour reconstituer l'image.

Puis nous détaillerons les propriétés principales de la transformation *Mojette* (chapitre 2). Ici encore, certaines caractéristiques intéressantes seront retenues, la redondance ou encore la notion de fantômes Mojette sont susceptibles d'être exploitées dans une méthode de tatouage.

Le troisième chapitre sera consacré à une revue des principaux résultats obtenus depuis dix ans dans le domaine du tatouage des images. Les techniques majeures utilisées seront disséquées afin de comprendre les atouts et les faiblesses inhérentes à ce domaine du traitement des images. Afin d'établir ces points positifs ou négatifs pour une technique de tatouage donnée, il faut être capable de classifier aussi le type d'attaques auxquelles peuvent être soumises les images tatouées. Enfin, nous analyserons l'impact du choix de l'espace de tatouage choisi (pour inférer des bonnes propriétés concernant la transformation Mojette). La problématique de l'évaluation de la qualité de l'image tatouée (donc de la définition d'une distance) sera enfin abordée.

Chapitre 1

Le Système Visuel Humain

Sommaire

| | | |
|------------|--|-----------|
| 1.1 | Introduction | 18 |
| 1.2 | Bases anatomiques et physiologiques du S.V.H. | 20 |
| 1.2.1 | Anatomie de l'œil | 20 |
| 1.2.2 | Les voies optiques, le chiasma | 25 |
| 1.2.3 | Le corps genouillé latéral | 26 |
| 1.2.4 | Le cortex visuel | 27 |
| 1.3 | Modélisation du SVH | 29 |
| 1.3.1 | Introduction | 29 |
| 1.3.2 | Perception de la luminance | 29 |
| 1.3.3 | Définition du contraste | 31 |
| 1.3.4 | Fonction de sensibilité au contraste (CSF) | 32 |
| 1.3.5 | Masquage | 33 |
| 1.3.6 | Modèles Mono-canaux | 34 |
| 1.3.7 | Modèles Multi-canaux | 36 |
| 1.3.8 | Modèle proposé | 42 |
| 1.4 | Conclusion | 47 |

1.1 Introduction

Certaines applications du traitement d'images numériques nécessitent une modélisation du système visuel humain. Parmi ces applications, les plus enclines à une telle étude sont celles de la compression d'images avec pertes, de l'évaluation de la qualité et, plus récemment le domaine du tatouage d'images.

Pour ce dernier domaine il y a nécessité d'utiliser un modèle du Système Visuel Humain (SVH) pour les applications orientées vers la protection de droits d'auteurs. Le besoin de produire un filigrane (un message invisible) est effectivement très important. Bien que l'invisibilité et la robustesse soient inversement proportionnelles (antinomiques), il n'en reste pas moins qu'un tatouage visible est évidemment aisément localisable, il est donc aussi aisément effaçable. Nous pouvons observer figure 1.1(a) une image marquée d'un copyright visible que nous avons "attaqué" à l'aide d'un utilitaire disponible gratuitement sur Internet : GIMP (GNU Image Manipulation Program). Le résultat, présenté figure 1.1(b) montre clairement la fragilité des méthodes de tatouage visible. Aucune connaissance a priori sur le tatouage d'images n'est nécessaire pour réaliser une telle expérience, donc, à plus forte raison, un "*pirate confirmé*" peut facilement venir à bout de toute méthode de tatouage visible.



FIG. 1.1 – (a) : Image marquée d'un copyright visible, (b) Suppression du copyright

Ces tatouages visibles comportent d'autres inconvénients. Ils doivent être réversibles, afin d'éviter de conserver intégralement l'image originale. Le filigrane doit également être éparpillé sur la totalité de l'image pour pouvoir résister à des découpages d'images (appelés cropping), dont un exemple est donné figure 1.2(a) et 1.2(b).



FIG. 1.2 – (a) : Image marquée d'un copyright visible, (b) Suppression du copyright par découpage

Comme nous le verrons par la suite (section 3.3), les caractéristiques principales requises au tatouage d'images en fonction des différentes applications sont la capacité, la robustesse, l'invisibilité, la sécurité, une faible complexité d'insertion et enfin, une faible complexité d'extraction. Concernant la protection des droits d'auteurs, les caractéristiques requises sont essentiellement l'invisibilité, la robustesse ainsi que, dans une moindre mesure, la capacité de la marque insérée. Si une faible capacité d'information peut parfois être suffisante pour décrire l'auteur de l'image ou ses ayants-droits, un problème de détection peut toutefois survenir dans les schémas de tatouage utilisant des marques trop courtes. Nous pouvons trouver dans [25] une étude portant sur les probabilités de fausse alarme et sur l'utilité d'utiliser des marques de longueur importante, lors du processus de détection. Les différentes attentes des schémas de tatouage selon les applications, sont détaillées par Fridrich dans [37].

Dans un schéma de codage, les modèles du système visuel humain (S.V.H.) permettent de définir l'information la moins pertinente d'une image, (information dont la suppression ne produirait aucune gêne pour les observateurs). Dans un contexte de tatouage, l'étude du système visuel doit au contraire permettre de définir l'information pertinente de l'image qui ne sera pas modifiée dans la plupart des attaques, dont les schémas de codage, sans toutefois générer d'artéfacts visibles. Le modèle visuel permet donc l'obtention d'une carte représentant les sites de l'image rendant toute modification imperceptible mais néanmoins robuste. Cette carte psychovisuelle, de même taille que l'image originale, doit donc définir pour tout pixel si ce dernier est modifiable, mais doit aussi déterminer la force maximale applicable en ce site sans causer de dégradations

visibles.

L'œil et le système visuel ont fait l'objet de nombreuses études depuis très longtemps. Le sujet est très vaste et très complexe.

Des études ont été menées au laboratoire, dans le but de modéliser le système visuel humain dans des applications de codage [115], [14], ou d'évaluation de la qualité [15] [22]. Nous présenterons dans ce chapitre, les bases de la perception visuelle, nous détaillerons pas à pas le chemin optique qui va du champ visuel jusqu'au cortex visuel primaire, partie du cerveau qui traite ces informations. Puis nous présenterons différentes modélisations du Système Visuel Humain, dont celle que nous avons retenue dans le cadre des différents schémas de tatouage présentés dans la partie II.

1.2 Bases anatomiques et physiologiques du S.V.H.

1.2.1 Anatomie de l'œil

1.2.1.1 l'œil

l'œil, représenté figure 1.3 peut être considéré comme un récepteur de signaux lumineux. Il focalise ces signaux pour en former une image qu'il devra alors transmettre au cerveau au moyen du nerf optique.

Le cristallin, sorte de lentille imparfaite, se comporte comme une lentille de focale variable, qui forme l'image de l'objet observé sur la rétine, couche photosensible qui tapisse le fond de l'œil.

La pupille, commandée par l'iris, joue le rôle d'un diaphragme à réglage automatique dont le diamètre s'ajuste en fonction du niveau d'éclairement reçu.

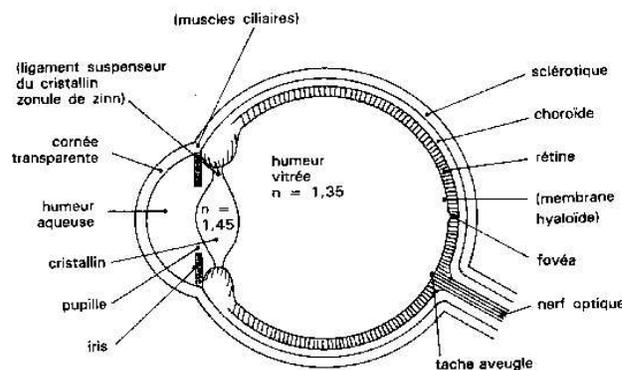


FIG. 1.3 – Coupe schématique de l'œil

1.2.1.2 La rétine

La rétine est la couche neurosensorielle, elle occupe une superficie d'environ 12.5 cm^2 . Elle transforme la lumière reçue en signaux électriques qui sont transmis vers le cortex via le nerf optique.

La rétine présente d'importantes différences régionales, ayant des conséquences fonctionnelles. Le centre de la rétine est dévolu à la vision des détails et à la vision colorée (photopique), alors que le pourtour est dévolu à la vision achromatique (scotopique) et plus globale.

On trouve plusieurs types de cellules nerveuses au niveau rétinien : les cellules de la voie directe (photorécepteurs -cônes et bâtonnets-, cellules bipolaires et cellules ganglionnaires) et celles des voies indirectes (cellules horizontales et 3 types de cellules amacrines), les cellules ganglionnaires, les cellules bipolaires, les récepteurs et les cellules pigmentaires (voir figure 1.4).

Il est important de noter que la lumière doit traverser l'ensemble des couches rétiniennes avant d'atteindre les photorécepteurs, sauf au niveau de la fovéa. En regard du centre optique de l'œil se trouve la fovéa, (appelée aussi tâche jaune), qui ne contient que des cônes. A ce niveau chaque cône est connecté à une seule cellule bipolaire et une seule cellule ganglionnaire. L'acuité visuelle est maximale.

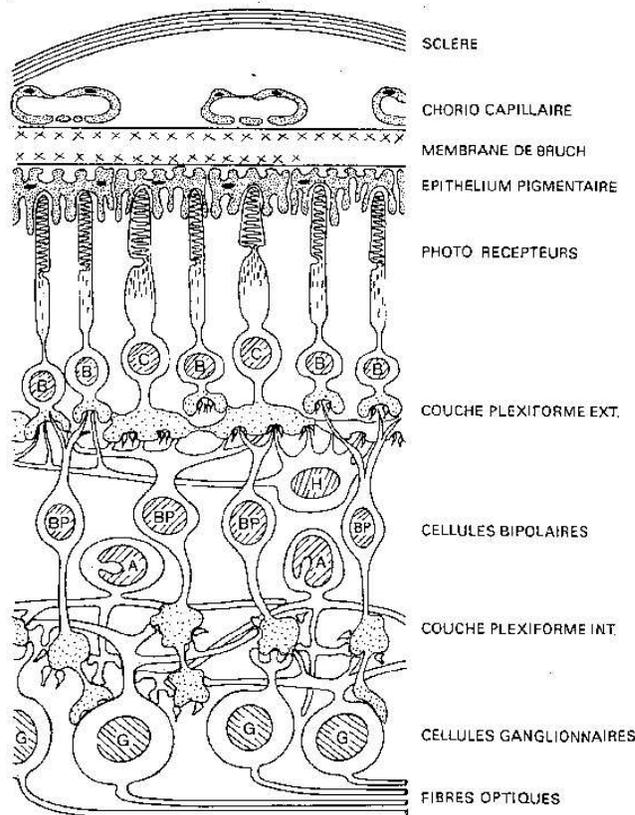


FIG. 1.4 – Structure de la rétine

Au centre optique la fovéa forme une dépression, au niveau de laquelle le faisceau lumineux arrive directement sur les cônes.

La papille optique, point de départ du nerf optique, se situe du côté nasal, à environ 4 mm de la fovéa, il n'y a pas de photorécepteurs à cet endroit qui est donc insensible à la lumière et porte donc le nom de tâche aveugle.

1.2.1.3 La fovéa

La fovéa est située juste dans l'axe visuel de l'œil, à un endroit où la rétine est très mince car on n'y rencontre que des cônes, les autres cellules visuelles étant rejetées sur le côté; il n'y a là que 2 500 cônes, mais chacun d'eux est relayé par une, et une seule, cellule associative. Cette région, de ce fait, assure le maximum d'acuité visuelle. Dès que l'on s'éloigne de la fovéa, l'acuité visuelle baisse rapidement, d'où la nécessité, pour voir net, d'orienter le globe oculaire en direction de l'objet fixé; cette fonction est dévolue aux muscles oculomoteurs insérés sur la capsule du globe. Du côté temporal, le champ visuel s'étend jusqu'à 95 degrés du point de

fixation. Il n'atteint que 60 degrés du côté nasal, 50 degrés vers le haut, et 80 degrés vers le bas.

1.2.1.4 Les photo-récepteurs

Il existe deux catégories de photorécepteurs, de formes différentes, les cônes et les bâtonnets. Leur densité locale varie en fonction de l'angle α où ils sont situés par rapport à l'axe optique figure 1.5.

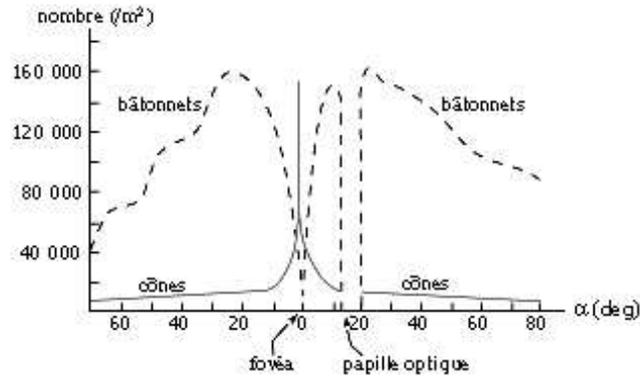


FIG. 1.5 – Distribution des cônes et des bâtonnets

- Les bâtonnets, en nombre extrêmement élevé (75 à 150 millions), se répartissent surtout en dehors de la fovéa. Ils ont un diamètre moyen de 2,5 à 3 microns et sont distants de 10 à 20 microns; comme, de plus, de nombreux bâtonnets sont reliés à une même fibre nerveuse, la résolution est faible, ce qui correspond à une vision globale de l'image. Par contre leur sensibilité est très bonne, ce qui explique qu'ils interviennent essentiellement en faible lumière (vision crépusculaire ou vision scotopique). Leur sensibilité spectrale est assez constante en fonction de la longueur d'onde; comme il n'existe qu'un seul type, la vision des couleurs n'est pas assurée par les bâtonnets.
- Les cônes, beaucoup moins nombreux (6 à 7 millions) se trouvent essentiellement dans la fovéa. Ils sont plus petits (1 à 2 microns) et aussi plus serrés (distance moyenne de 2,5 à 10 microns); il n'y a que quelques cônes reliés à une même fibre nerveuse, d'où une bonne résolution. Par contre les cônes sont peu sensibles, figure 1.6. et interviennent essentiellement en vision diurne normale (vision photopique). En traitement d'images, on peut admettre que l'on est presque toujours dans ce cas, donc seule la réponse des cônes est à considérer.

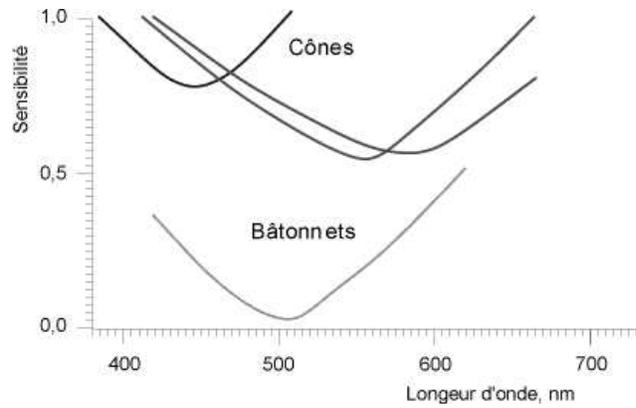


FIG. 1.6 – sensibilité spectrale des cônes et des bâtonnets

Il existe en fait trois types de cônes dont les sensibilités spectrales, dues aux pigments qu'ils contiennent, sont différentes. La figure 1.7 montre les courbes d'absorption de ces pigments en fonction de la longueur d'onde de la lumière; les maxima se situent respectivement dans le bleu (430 nm), le vert (555 nm) et le jaune-vert (585 nm).

Ce sont les cônes qui permettent la vision des couleurs.

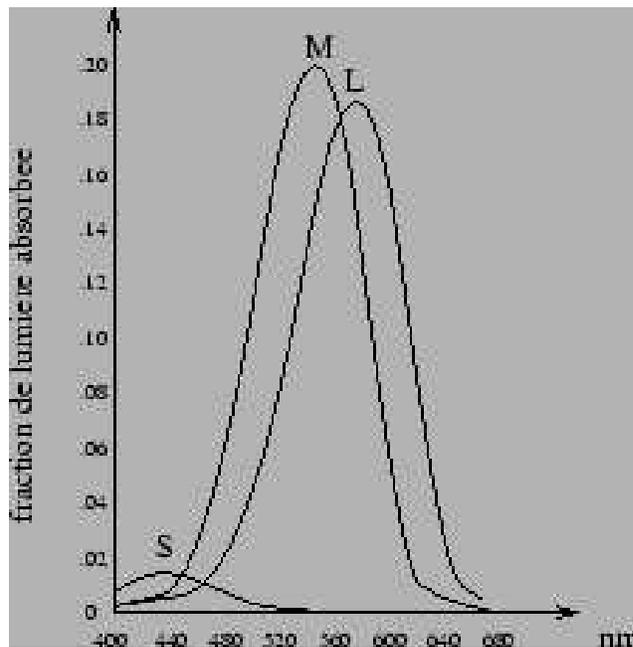


FIG. 1.7 – Courbes d'absorption relative des pigments des 3 types de cônes

1.2.2 Les voies optiques, le chiasma

On appelle voie optique (ou voies visuelles), figure 1.8, la succession de neurones qui commence dans l'épaisseur de la rétine et se termine sur le cortex visuel primaire. Le nerf optique débute au niveau de la rétine, s'éloigne du globe oculaire et se termine au niveau d'une structure en forme de X, appelée le chiasma optique, où se croisent une partie des fibres de chacun des deux nerfs. Après le chiasma, la voie se continue par la bandelette optique, passe par différents centres nerveux, et se termine sur le lobe occipital (postérieur) de l'hémisphère cérébral du même côté.

Après avoir franchi le chiasma optique, certaines fibres vont vers une structure appelée tubercule quadrijumeau antérieur, dont le rôle serait particulièrement important, notamment pour la perception impliquant des mouvements oculomoteurs, tandis que d'autres atteignent un relais appelé Corps Genouillé Latéral (C.G.L.) où s'effectue la vision en profondeur. De là, ce que l'on nomme les radiations optiques conduisent ensuite les influx nerveux dans des zones du cortex situées à l'arrière des lobes occipitaux : les aires visuelles. Celles-ci sont responsables de la perception des objets dans l'espace, de leurs relations les uns aux autres, ainsi que de la perception de la lumière et de l'ombre.

Un objet situé du côté externe du champ visuel d'un œil (du côté droit pour l'œil droit, gauche pour l'œil gauche) se projette du côté interne de la rétine, du fait du principe d'inversion des images. Le message nerveux emprunte les fibres internes du nerf optique, croise la ligne médiane au niveau du chiasma, et se termine dans l'hémisphère opposé (gauche pour l'œil droit, droit pour l'œil gauche).

Si l'objet est situé du côté interne du champ visuel, l'image se forme du côté externe de la rétine, le message ne croise pas la ligne médiane et se termine dans l'hémisphère du même côté. Au total, l'hémisphère droit reçoit la moitié gauche du champ visuel de chaque œil, et l'hémisphère gauche reçoit la moitié droite.

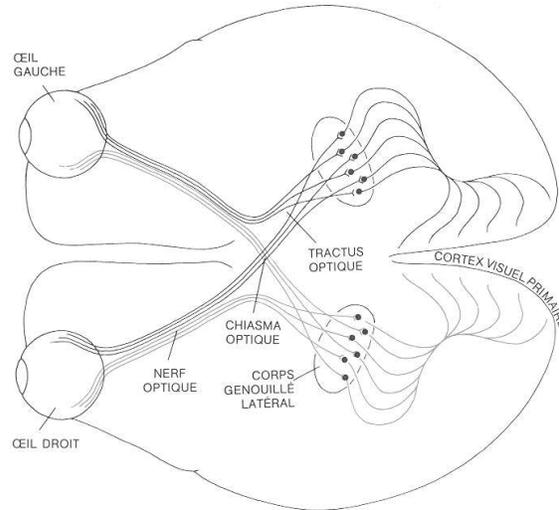


FIG. 1.8 – Les voies visuelles

En résumé, pour chaque œil, l'information visuelle d'un objet perçu du côté nasal est projetée via le chiasma optique, passe par le C.G.L. situé dans l'hémisphère opposé, avant d'être transmise au cortex visuel primaire. A l'opposé, pour tout objet perçu du côté temporal, l'information est transmise par le nerf optique vers le C.G.L. situé dans le même hémisphère, puis au cortex visuel primaire.

1.2.3 Le corps genouillé latéral

Les axones des cellules ganglionnaires de la rétine convergent pour former le nerf optique qui se projette sur plusieurs structures dont le corps genouillé via le chiasma optique.

Nous avons vu que les informations provenant d'une part de l'hémisphère temporale de l'œil droit (hémichamp gauche de la scène visuelle) et d'autre part de l'hémisphère nasale de l'œil gauche (hémichamp gauche de la scène visuelle) se rejoignent sans se mélanger. C'est à partir des deux corps genouillés latéraux que l'information de chaque œil est projetée et mélangée sur les aires visuelles corticales. C'est cette fusion d'information qui est en partie responsable de notre perception du relief d'une scène visuelle. Le C.G.L. est la principale structure relais à travers laquelle les messages visuels sont retransmis de la rétine, via 90% des fibres du nerf optique, vers les régions du cortex cérébral situées au niveau des pôles occipitaux des hémisphères cérébraux.

1.2.4 Le cortex visuel

Le cortex visuel est la dernière étape du processus de la perception visuelle, avant le traitement des informations par le cerveau. Il est responsable de la pensée, du langage, de la volonté. Il analyse des sensations, des déclenchements de mouvements volontaires. Cette couche de neurones est responsable de toutes les fonctions supérieures de l'organisme.

Cet organe est divisé en deux parties :

- l'hémisphère gauche recueille, traite les informations et commande les mouvements du côté droit du corps.
- l'hémisphère droit assure cette même fonction mais pour la partie gauche du corps.

Au sein du cortex, on peut donc trouver deux hémisphères.

Cependant chacun peut être divisé en quatre continents ou lobes :

- Le lobe frontal réfléchit et donne des ordres aux muscles.
- Le lobe pariétal se situe au sommet de la tête. Il interprète les sensations reçues par la peau.
- Le lobe occipital se situe, lui, à la base du crâne. Il analyse les images. Il contient pas moins d'une trentaine d'aires visuelles.
- Le lobe temporal décrypte les sons et participe à la mémoire.

Ces quatre continents se retrouvent par symétrie dans chaque hémisphère. Néanmoins certaines aires sont spécifiques à une hémisphère. C'est le cas pour les centres de reconnaissance des mélodies et de compréhension des émotions contenues dans la voix qui sont localisés dans l'hémisphère droit ou pour les centres de calcul, de compréhension des phrases, chiffres, symboles et de production du langage qui se trouvent dans l'hémisphère gauche.

Le lien entre ces deux hémisphères se fait grâce au corps calleux. Même si chaque partie du cerveau est concentrée sur une tâche précise, ses acteurs collaborent pour déclencher les comportements nécessaires.

Le cortex visuel est composé d'une trentaine de régions différentes ou aires corticales, schématisées figure 1.9.

Toutes ces parties travaillent ensemble pour élaborer une image unique et nette. Mais on ne peut définir et comprendre le rôle que de quelques-unes.

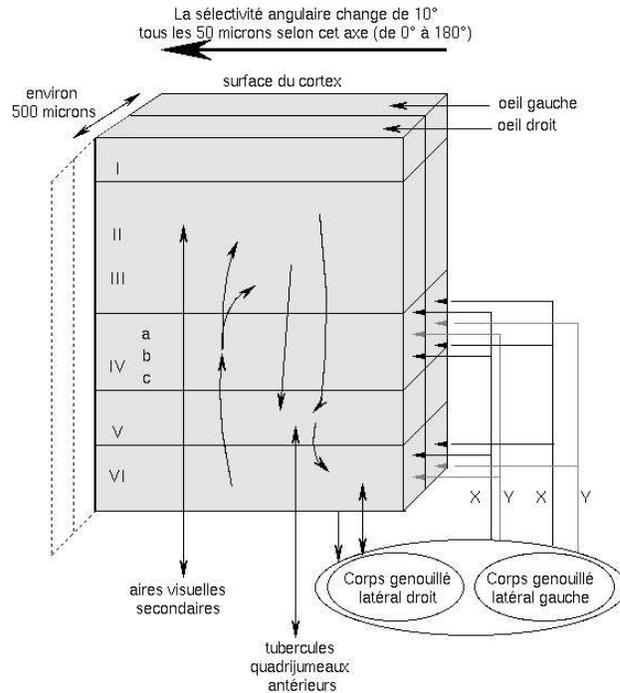


FIG. 1.9 – Représentation schématique d'une coupe, perpendiculaire à la surface du cortex, de l'aire visuelle primaire

La première, V1, est l'aire visuelle primaire. C'est la partie la plus importante car sans elle, nous serions aveugles. Les informations venant de la rétine par l'intermédiaire du nerf optique, arrivent sur elle. En effet chaque partie de la rétine a son reflet parfaitement "géographique" sur le cortex visuel primaire. Cette partie analyse les informations recueillies sur la forme, la couleur, le mouvement, et les distribue ensuite aux autres aires appartenant au cortex visuel (telles V2, V3, V4, V5).

L'aire V2, reçoit beaucoup d'informations et a pour but de trier plus finement les informations transmises par V1. Elle traite à la fois des contours tels que les formes géométriques, des orientations (horizontale et verticale), des textures (lisse, rugueux...) et des couleurs.

En même temps que l'aire V2, les autres aires reçoivent également des informations de l'aire V1, les traitent sur le même rythme mais chacune dans sa spécialité.

L'aire V3 analyse les formes en mouvement et apprécie les distances.

L'aire V4 s'occupe du traitement des couleurs et des formes immobiles.

Enfin, l'aire V5 s'occupe de la perception des mouvements tels que la direction et la vitesse.

La plupart de la trentaine d'aires visuelles identifiées sont situées dans la région occipitale (arrière) du

cerveau. Cependant certaines occupent les régions pariétales (côtés et voûte). Elles auraient pour but la localisation des objets. Les aires visuelles temporales, seraient, quand à elles, dévolues à l'identification. Néanmoins, à ce jour, il reste difficile de savoir comment toutes ces régions coordonnent leur savoir-faire et construisent des images nettes.

1.3 Modélisation du SVH

1.3.1 Introduction

Les disciplines de la compression ou de l'évaluation de la qualité ont permis la mise en œuvre de plusieurs études approfondies du S.V.H. Nous pouvons trouver dans [80], une description exhaustive, tant des aspects anatomiques du S.V.H., que des modèles visuels de la littérature, ainsi qu'une étude des méthodes d'évaluation de la qualité ou des schémas de compression.

Dans cette section, nous allons présenter respectivement la perception de la luminance (paragraphe 1.3.2), la sensibilité au contraste (paragraphe 1.3.4), les effets de masquage (1.3.5), ainsi que les différents modèles visuels de la littérature (paragraphe 1.3.6, 1.3.7 et 1.3.8).

1.3.2 Perception de la luminance

La compréhension du processus de la perception visuelle est un élément à considérer lors de la modélisation du S.V.H. 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. Plusieurs processus contribuent à la perception d'une large dynamique de la luminance [$3.10^{-2}cd/m^2$ à $3.10^5cd/m^2$]. Alors que les cônes nécessitent une certaine luminosité, les bâtonnets réagissent déjà à partir de très basses luminances. Cette sensibilité aux très basses luminances est garantie également par une intégration spatiale effectuée au niveau des cellules ganglionnaires.

Une des premières modélisations quantitatives de la perception de la luminance a été énoncée par Weber. La loi de Weber est modélisée par :

$$\frac{\Delta L}{L} = C^{te} \quad (1.1)$$

Cette loi indique que si sur un fond uniforme de luminance L , dite d'adaptation, on superpose un stimulus type médaillon de luminance $\Delta L + L$ (figure 1.10), le rapport $\frac{\Delta L}{L}$ est pratiquement constant dans un large

domaine de luminosité.

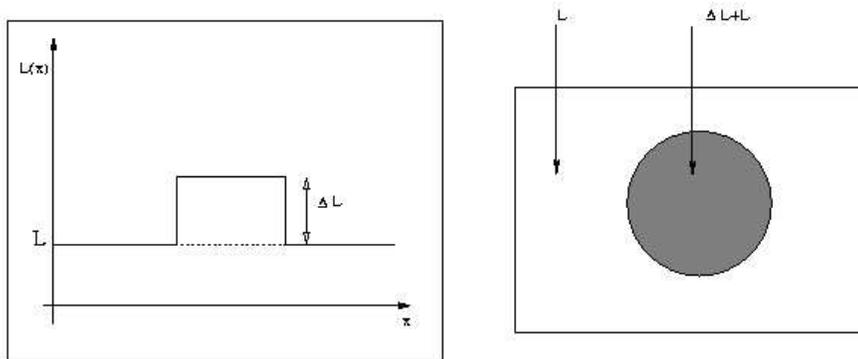


FIG. 1.10 – l'expérience de Weber

Ces mesures donnant des seuils de détection en fonction de la luminance de fond, quand celle ci est suffisamment élevée, ont été confirmées dans l'intervalle des luminances moyennes par Sakmann et Creutzfeld [109] et retenue par Sakrison [110].

Une légère incohérence en basses luminances a été corrigée par Moon et Spencer [79] dans [115], donnant lieu à l'expression suivante :

$$\frac{\Delta L}{L} = \left(\frac{C_{\infty}}{L} \right) (0,456 + \sqrt{L})^2 \quad (1.2)$$

Au seuil de détection, et lorsque la luminance augmente, le rapport $\frac{\Delta L}{L}$ tend vers la constante de Weber C_{∞} . Cette constante dépend de la géométrie et de la taille du stimulus.

La loi de Weber, complétée par Moon et Spencer, montre que le S.V.H. comporte une transformation de type logarithmique telle que la représente la figure 1.11.

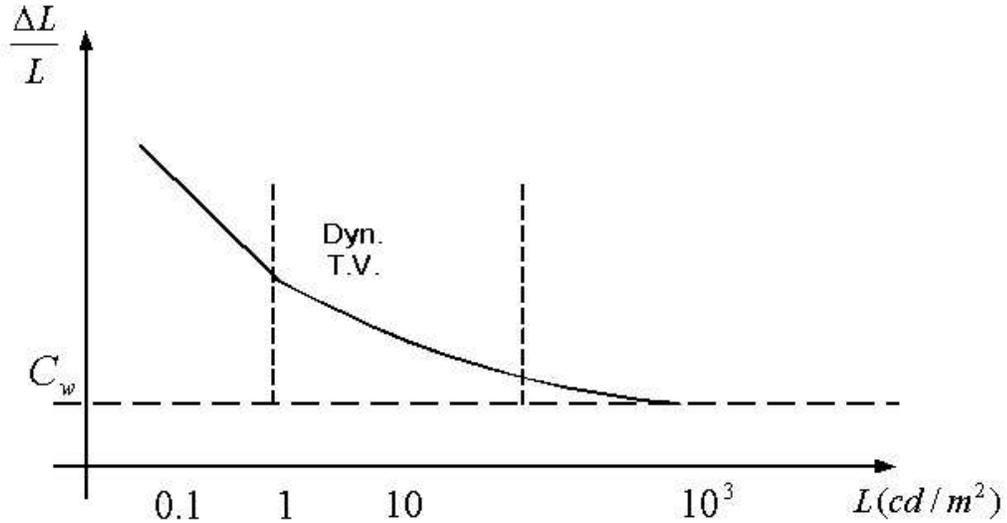


FIG. 1.11 – Loi de Moon et Spencer

1.3.3 Définition du contraste

Le contraste est le rapport entre l'intensité locale d'une image et son intensité moyenne. Il existe un nombre important de définitions du contraste, adaptées à des stimuli plus ou moins complexes. Par exemple, dans le cas de stimuli sinusoïdaux, la formule de Michelson est généralement adoptée :

$$C = \frac{L_{max} - L_{min}}{2L_{moy}} \quad (1.3)$$

Le contraste local à bande limitée proposé par E. Peli [88] est beaucoup plus complexe et reflète le fait très important que la perception d'un détail de l'image dépend aussi de son environnement local. Le calcul du contraste local suppose une décomposition de l'image en sous-bandes visuelles et est défini par le rapport entre la luminance locale d'une sous-bande donnée et la luminance moyenne locale relative à ce canal, c'est à dire, pour le site considéré (m,n) , la somme des luminances de toutes les sous-bandes radiales inférieures à la sous-bande.

$$C_i(m, n) = \frac{L_i(m, n)}{\sum_{k=0}^{i-1} L_k(m, n)} \quad (1.4)$$

où i représente le i^{eme} canal radial et $\sum_{k=0}^{i-1} L_k(m, n)$, le signal basses fréquences relatif à la i^{eme} bande.

1.3.4 Fonction de sensibilité au contraste (CSF)

Afin de déterminer la sensibilité au contraste du S.V.H., plusieurs études ont été menées.

Une des premières expériences sur le comportement fréquentiel est celle de Campbell et Robson [23], ils ont étudié la détection par un observateur de mires sinusoïdales verticales fixes du type :

$$L(x) = L(1 + c \sin(2\pi fx)) \quad (1.5)$$

où L est la luminance de fond et à déterminer la valeur minimale du contraste “ c ” permettant de détecter la fréquence f . Le contraste obtenu est appelé “contraste seuil” ou “seuil de visibilité”. l’inverse de ce contraste représente la sensibilité visuelle à la fréquence f .

Une des caractéristiques les plus importantes de la création des modèles visuels concerne la décroissance de la sensibilité pour les hautes fréquences spatiales. Ce phénomène est modélisé par la fonction de sensibilité au contraste (CSF).

Plusieurs CSF achromatiques sont détaillées par Barten dans [11].

La courbe de la CSF, (figure 1.12) est usuellement assimilée à la fonction de transfert d’un filtre spatial linéaire en basses et hautes fréquences.

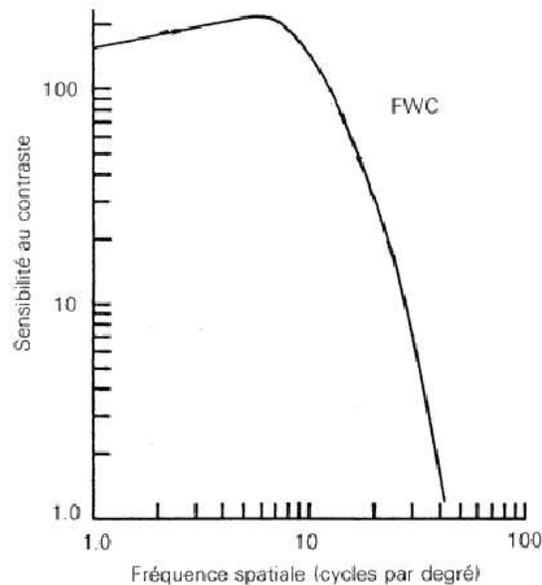


FIG. 1.12 – Fonction de sensibilité au contraste

La sensibilité visuelle est maximale pour des fréquences spatiales moyennes (de l'ordre de $3cy/d^\circ$), elle

décroit pour les fréquences plus élevées. Ce filtrage provient du comportement passe-bas de la lentille et de la pupille, ainsi qu'au nombre limité de cellules rétinienne par unité de longueur.

1.3.5 Masquage

En accord avec la modélisation perceptuelle ci dessus, les signaux ayant des caractéristiques voisines sont traités par les mêmes canaux visuels et suivent donc le même cheminement de l'œil jusqu'au cortex. Il apparaît en fait que de tels signaux interagissent entre eux et sont soumis à des effets non linéaires. l'effet non linéaire le plus considéré est l'effet de masquage. Ce dernier traduit la variation du seuil de détection d'un stimulus due à la présence d'un signal, qualifié de signal masquant, ayant des caractéristiques voisines et un niveau plus fort. Etudier l'effet de masquage revient donc à modéliser la variation du seuil de détection d'un stimulus en fonction des caractéristiques du signal masquant.

l'effet de masquage a fait l'objet de plusieurs études et différents modèles ont été proposés dans la littérature [69], [35], [53], [117]. Le modèle largement utilisé [122], traduisant cet effet non linéaire s'écrit lorsque le stimulus et le signal masquant ont les mêmes caractéristiques :

$$C_0^m(f, \theta, L) = C_0(f, \theta, L) \max \left[1, \left| \frac{C_m(f, \theta)}{C_0(f, \theta, L)} \right|^\alpha \right] \quad (1.6)$$

- $C_0^m(f, \theta, L)$ est le contraste seuil du stimulus de caractéristiques (f, θ) en présence du signal masquant et présenté seul sur un fond de luminance L .
- $C_0(f, \theta, L) = \frac{1}{CSF(f, \theta, L)}$ est le contraste seuil du stimulus de caractéristiques (f, θ) présenté seul sur un fond de luminance L .
- $C_0(f, \theta)$ est le contraste du signal masquant de mêmes caractéristiques (f, θ) que le stimulus.
- α est un paramètre dépendant de (f, θ) et varie typiquement entre $0.6 \leq \alpha \leq 1.1$

Le contraste seuil $C_0^m(f, \theta, L)$ varie en fonction de la variation des caractéristiques f_m, θ_m du signal masquant. Pour prendre en compte cette dépendance, une fonction de pondération est utilisée et le nouveau modèle s'écrit

$$C_0^m(f, \theta, L) = C_0(f, \theta, L) \max \left[1, \left| k \left(\frac{f_m}{f}, \theta_m - \theta \right) \frac{C_m(f_m, \theta_m, L_m)}{C_0(f, \theta, L)} \right|^\alpha \right] \quad (1.7)$$

où $k \left(\frac{f_m}{f}, \theta_m - \theta \right)$ est la fonction de pondération modélisée généralement par une gaussienne :

$$k\left(\frac{f_m}{f}, \theta_m - \theta\right) = \exp\left[-\left(\frac{\log_2^2\left(\frac{f_m}{f}\right)}{\sigma_f^2} + \frac{(\theta_m - \theta)^2}{\sigma_\theta^2}\right)\right] \quad (1.8)$$

Le paramètre σ_f représente la largeur de bande radiale, définie à mi hauteur, elle est de l'ordre de 2.5 octaves à $1cy/d^\circ$ et de 1.5 octaves à $16cy/d^\circ$ avec une décroissance linéaire. La largeur de bande angulaire σ_θ est de l'ordre de 30 degrés à $1cy/d^\circ$ et de 15 degrés à $16cy/d^\circ$.

1.3.6 Modèles Mono-canaux

En se basant sur les phénomènes présentés dans les sections précédentes, différentes modélisations du S.V.H. sont apparues.

Ces modèles ont été élaborés sur la base du contenu fréquentiel des images, l'information spectrale pouvant inscrire une plus large gamme de variabilité d'images.

Nous pouvons observer deux catégories principales de modèles du SVH, les modèles mono-canaux et les modèles multi-canaux.

Le modèle proposé par Schade [111], avait pour objectif, la prédiction du SVH pour la conception d'écrans de télévision. Ce modèle suppose que l'image neuronale peut être obtenue par une transformation linéaire et invariante par translation de l'image rétinale.

Ce modèle est adapté à des stimuli simples (mires horizontales ou verticales), mais est trop simpliste pour des signaux complexes.

Mannos et Sakrison [74], ont proposé un modèle basé sur les travaux psychophysiques de Campbell et Robson, décrits précédemment (section 1.3.4)

A partir des résultats obtenus sur les mires horizontales, Mannos et Sakrison proposent une modélisation de la CSF :

$$CSF(f) \approx \frac{1}{m_s} \approx c \left(\frac{f}{f_0}\right) \exp\left(-\frac{f}{f_0}\right) \quad (1.9)$$

où $c(f)$ est une fonction linéaire correspondant au processus d'inhibition latérale engendrée par l'interaction neuronale au niveau de la rétine et des couches supérieures.

Pour simplifier la modélisation du SVH, les auteurs proposent un modèle isotropique :

$$CSF(f) = a \left(b + \frac{f}{f_0}\right) \exp\left(-\frac{f}{f_0}\right)^c \quad (1.10)$$

a, b, c et f_0 représentent les paramètres du modèle.

l'expression

$$CSF(f) = 2.6 (0.0192 + 0.114f) \exp(-0.114f)^{1.1} \quad (1.11)$$

a été utilisée par les auteurs, dans le cadre de l'évaluation de la qualité des images dégradées, pour pondérer le spectre de l'image d'erreur. Bien que cette modélisation soit largement admise, diverses études sont menées pour prendre en compte les différents paramètres expérimentaux. Ainsi la C.S.F. proposée par Barten [10] tient compte de la variation de la luminance de fond L et des conditions d'observation. Cette C.S.F. s'écrit

$$CSF(f, L) = a(L, f, \omega) f \sqrt{1 + 0.06 \exp[b(L)f]}. \exp -b(L)f \quad (1.12)$$

où L représente la luminance de fond et est exprimée en candelas par mètre carré.

f représente la fréquence radiale et s'exprime en cycles par degrés. Les paramètres du modèle $a(L, f, \omega)$ et $b(L)$ sont donnés par

$$a(L, f, \omega) = \frac{540 (1 + 0.7/L)^{-0.2}}{1 + \frac{12}{\omega(1+f/3)^2}} \quad (1.13)$$

$$b(L) = 0.3 (1 + 100/L)^{0.15} \quad (1.14)$$

l'angle solide ω , exprimé en degrés, est donné par $\omega = \frac{180 \sqrt{A}}{\pi D}$ avec A taille de l'image et D distance d'observation.

Cette C.S.F. a été modifiée pour rendre compte du caractère non isotropique de la sensibilité visuelle :

$$CSF(f, L, \theta) = a(L, f, \omega) f \sqrt{1 + 0.06 \exp[b(L)f]}. \exp [-b(L)f\Gamma(\theta)] \quad (1.15)$$

avec $\Gamma(\theta) = 1 - 0.079 [\cos(4\theta) - 1]$

Hall et Hall [50] proposent une adaptation du modèle de Mannos et Sakrison où le filtre passe-bande est remplacé par un filtre passe-bas suivi d'un passe haut. Une non linéarité représentant la sensibilité des cellules de la rétine étant placée entre ces deux filtres.

Ce modèle est mieux adapté au comportement du SVH, où le passe-bas correspond à la formation de l'image sur la rétine, et le passe haut correspond à la formation de l'image neuronale.

Limb propose dans [71] un schéma de mesure objective de la qualité des images. L'auteur propose un

modèle complet de la vision humaine, incluant un filtrage de l'erreur par un filtre passe-bas et une pondération par une fonction de masquage.

Ce modèle, adapté à la mesure de qualité, bien qu'assez complet, n'est pas totalement satisfaisant et ne donne pas une variance minimale pour tous les types de distorsions.

Bien que souvent utilisée, ces modélisations "mono-canal" ne suffisent pas toutefois à expliquer le comportement du S.V.H. vis à vis de stimuli complexes [110] [43]. Il est admis [118] que le S.V.H. utilise pour l'analyse des signaux d'entrée un ensemble de canaux dont chacun est sensible à une orientation et à une fréquence spatiale donnée.

Les caractéristiques de ces canaux, séparables dans une représentation polaire, ont fait l'objet de plusieurs études [19], [1], [94].

Bien que notre préoccupation principale soit ici le traitement des images fixes monochromes, nous pouvons toutefois citer une métrique de qualité d'images couleurs, proposée par Faugeras [34], ou encore, une métrique de qualité adaptée aux séquences vidéo [73].

1.3.7 Modèles Multi-canaux

Les modèles mono-canaux, généralement assez simples à mettre en œuvre, ne permettent toutefois pas la prédiction des formes complexes.

Certaines expériences psychophysiques [23] [107], mettent en évidence la structure multirésolution du SVH.

De nombreux modèles de vision multi-canaux ont été développés, un point commun à la plupart de ces modèles est l'utilisation en première étape, d'un filtre passe bande (modélisant la CSF).

Le découpage fréquentiel diffère souvent en fonction des auteurs, Sakrison [110] propose 12 fréquences radiales pour 9 orientations, Watson [123] propose 8 fréquences radiales et 8 orientations, pour Zetsche [129] 5 fréquences radiales et 6 orientations sont nécessaires, Comes [26] emploie 4 fréquences radiales et 4 orientations et enfin, Lubin [72] propose 7 fréquences radiales pour 4 orientations.

Marion présente dans [76] un modèle classique du comportement du SVH pour des images monochromes, représenté sur la figure 1.13.

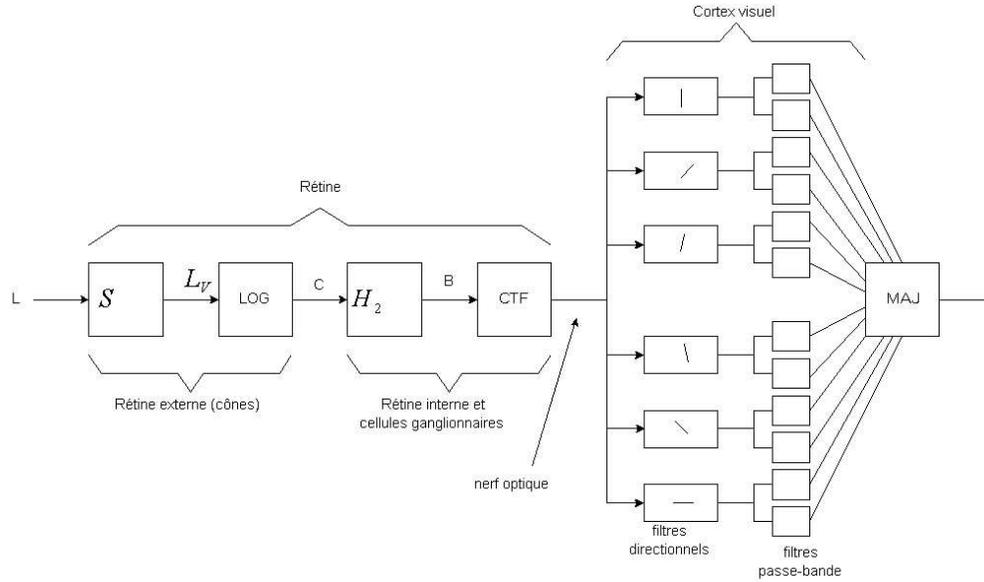


FIG. 1.13 – Modèle type du SVH

La rétine interne, les cellules ganglionnaires, le corps genouillé latéral, ainsi que dans une certaine mesure, le cortex visuel sont pris en considération dans ce modèle visuel.

La fonction de transfert isotrope $H(\rho)$ modélise l'ensemble de la rétine et des cellules ganglionnaires. Le cortex est modélisé par un ensemble de filtres directionnels, puis une batterie de filtres passe-bande mesurent l'intensité de l'excitation dans chacune des directions. La sortie du schéma permet la sélection du signal le plus important.

Karunasekera et Kingsbury présentent dans [62], [63] et [61], une métrique d'évaluation de la qualité des images, basée sur une étude du SVH face à différentes dégradations. Le schéma global est représenté sur la figure 1.14. L'image d'erreur (différence entre l'image de référence et l'image dégradée), est décomposée en images de contours selon quatre directions (-45° , 0° , 45° et 90°), ainsi qu'en une bande de base (basses fréquences). Chacune de ces images subit une fonction de masquage, décomposé en deux parties. Un premier masquage dû à l'activité de l'image pour chaque direction sélectionnée est mis en œuvre, un second masquage dû à la luminance moyenne de l'image est appliqué.

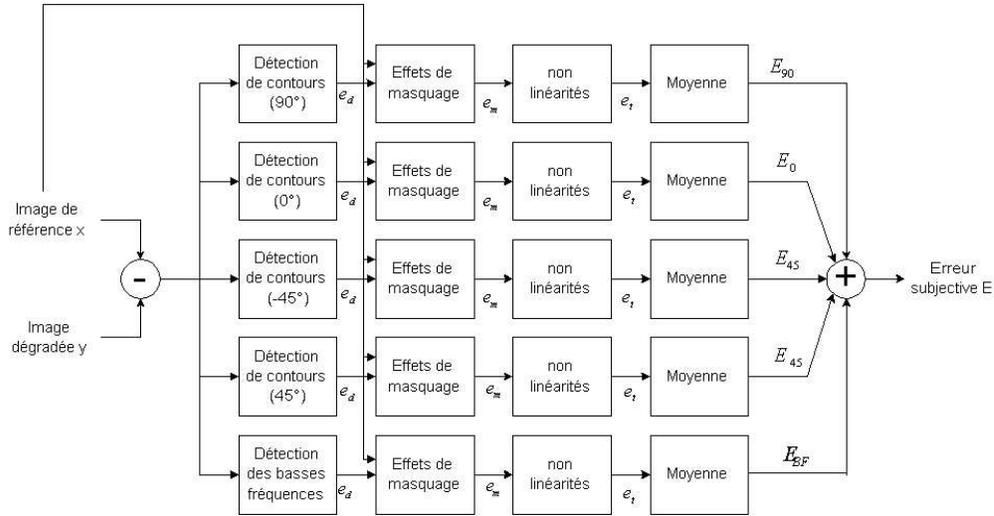


FIG. 1.14 – Modèle développé par Karunasekera et Kingsbury

Pour toute image de référence x , l'activité directionnelle A_θ est calculée à partir d'un filtre directionnel et d'un filtre passe-bas. Chaque image de contour directionnel e_d fournit une image d'activité e_a , selon l'équation 1.16.

$$e_a = \frac{e_d}{\left(1 + \left(\frac{A_\theta}{a_0}\right)^{r_a}\right)} \quad (1.16)$$

Chacune de ces images représentant l'activité directionnelle subit alors le masquage dû à la luminance moyenne de l'image e_m .

$$e_m = \frac{e_a}{\left(1 + \left(\frac{B}{b_0}\right)^{r_b}\right)} \quad (1.17)$$

où B représente la luminance moyenne de l'image.

Une fois l'effet de masquage modélisé, une transformation non linéaire est calculée pour chaque angle directionnel θ selon l'équation 1.18

$$e_{t\theta} = (e_{m\theta})^{p_0} \quad (1.18)$$

Puis, finalement, le calcul de l'erreur finale pour chaque direction, s'effectue par calcul de la moyenne de chaque image d'erreur $e_{t\theta}$ et l'erreur finale est donnée par la somme des quatre erreurs directionnelles et de l'erreur représentant les basses fréquences.

Le Visible Difference Predictor [31], [32] de S. Daly est probablement la métrique de distorsion d'images la plus connue. Cet algorithme fournit une carte d'erreurs entre une image de référence et une version dégradée quelconque. Cette carte d'erreur indique la probabilité de détection d'une différence visuelle en tout point de l'image. Le modèle de vision exploité dans le VDP, inclut une fonction non linéaire, modélisant l'adaptation du système visuel à la luminance, une CSF en deux dimensions est employée pour caractériser la sensibilité fréquentielle du SVH. Une transformation de Fourier est donc nécessaire entre la non linéarité et la CSF, cette dernière est spécifiée en cycles par degrés.

$$f_{visuel} = \frac{\pi}{180} \frac{1}{2 \times \arctan\left(\frac{D_w}{D_v}\right)} cy/d^\circ \quad (1.19)$$

où D_w représente la taille de l'image sur l'écran et D_v , la distance d'observation. Habituellement, on considère que la distance de visualisation est proportionnelle à la taille de l'image, par exemple, selon les recommandations du CCIR [24], la distance d'observation est normalisée à 6 fois la hauteur de l'écran ($D_v = 6 \times D_w$).

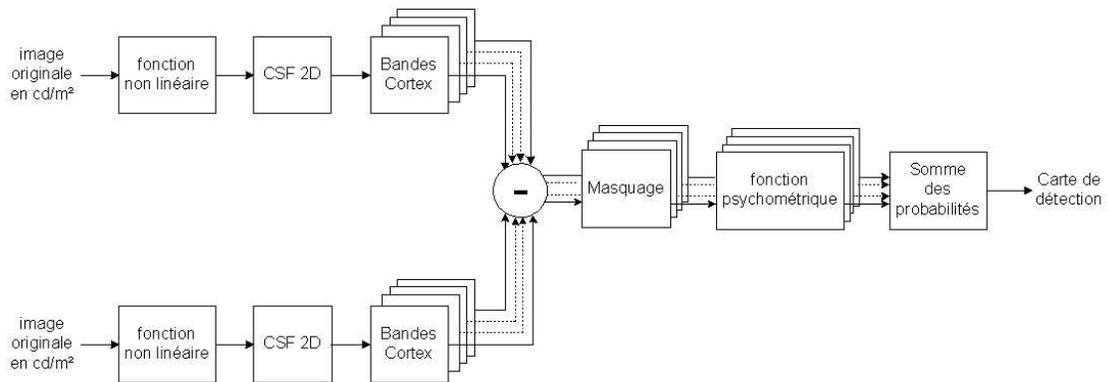


FIG. 1.15 – Le 'Visible Difference Predictor' (VDP) de Scott Daly

Le VDP, figure 1.15, nécessite, outre l'image de référence et l'image dégradée, plusieurs paramètres d'entrée, tels que la distance de visualisation, permettant la conversion des fréquences spatiales normalisées en fréquences spatiales en cycles par degrés. Certaines caractéristiques de l'écran (fonction gamma) sont nécessaires pour convertir les niveaux de gris de l'image en luminances perçues.

La CSF utilisée par Daly est très complète, car elle dépend non seulement de la fréquence spatiale, mais aussi de l'orientation, du niveau d'adaptation à la luminance, de la taille de l'image en degrés visuels, de la distance d'observation et de l'excentricité du stimulus. L'expression de la CSF utilisée est la suivante :

$$CSF(u, v) = P.Min \left\{ CSF_1 \left(\frac{\rho}{r_a \cdot r_b \cdot r_\theta}, L_A, I^2 \right), CSF_1 (\rho, L_A, I^2) \right\} \quad (1.20)$$

$$CSF_1(\rho, L_A, I^2) = \left(\left(3.23 (\rho^2 I^2)^{-0.3} \right)^5 + 1 \right)^{-1/5} \cdot 0.9 \rho A \exp -0.9 \rho B \quad (1.21)$$

où A et B sont des fonctions du niveau de l'intensité lumineuse (L_A), $\rho = \frac{\sqrt{u^2+v^2}}{I}$ est la fréquence radiale, $\theta = \arctan \frac{u}{v}$ représente l'orientation, I est la taille de l'image

Daly a utilisé les filtres cortex, développés par Watson [122] lors de la construction de la décomposition fréquentielle. Chaque filtre cortex est le produit de deux filtres passe-bande, permettant la sélection de fréquences radiales, ainsi que de certaines orientations.

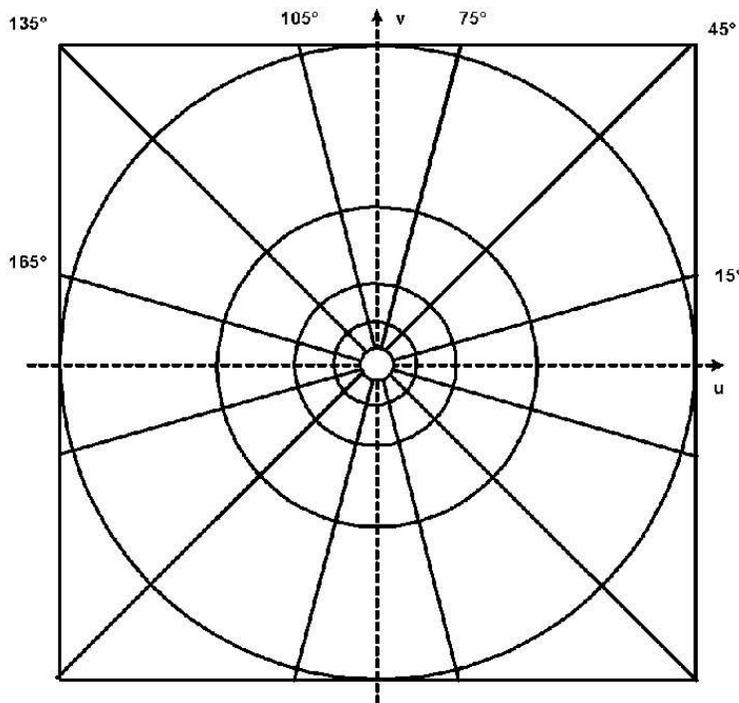


FIG. 1.16 – Décomposition fréquentielle utilisée par Daly

L'étape suivante du VDP consiste à convertir chacune de ces 31 images en unité de contraste, chaque image spatiale correspondant à un filtre cortex (une transformation de Fourier inverse est alors nécessaire), que nous appellerons sous-bande, est normalisée par la valeur moyenne de l'image cortex de la bande de base.

L'objectif de la fonction de masquage est de fournir, pour chaque site (m, n) de chaque sous-bande (i, j) , l'élévation du seuil de visibilité $T_{i,j}(m, n)$.

$$T_{i,j}(m,n) = \left(1 + \left(k_1 \cdot (k_2 \cdot |c_{i,j}(m,n)|)^2 \right)^b \right)^{\frac{1}{b}} \quad (1.22)$$

où $|c_{i,j}(m,n)|$ représente le contraste du site (m,n) de la sous-bande (i,j) .

Une carte d'erreurs perceptuelles est donc calculée pour chaque sous-bande selon l'équation 1.23

$$\Delta C_{i,j}(m,n) = \frac{c_{i,j}^O(m,n) - c_{i,j}^D(m,n)}{\min [T_{i,j}^O(m,n), T_{i,j}^D(m,n)]} \quad (1.23)$$

où $c_{i,j}^O(m,n)$ représente le contraste de l'image originale et $c_{i,j}^D(m,n)$ celui de l'image dégradée.

Une sommation des 31 cartes d'erreur, permet alors l'obtention d'une unique carte de probabilité de détection.

Enfin, la fonction psychométrique indique la probabilité de détection d'un signal c normalisé en contraste en présence d'un masque α . La formule représentant cette fonction psychométrique est donnée par l'équation 1.24.

$$P(c) = 1 - \exp \left[- \left(\frac{c}{\alpha} \right)^\beta \right] \quad (1.24)$$

α et β sont des paramètres expérimentaux et $P(c)$ la probabilité de détection.

La probabilité de détection dans la bande k,l pour le pixel i,j est donnée par :

$$P_{k,l}(i,j) = 1 - \exp \left(- \left(\frac{\Delta C_{k,l}(i,j)}{T_e^{k,l}(i,j)} \right)^\beta \right) \quad (1.25)$$

Finalement, la sommation des probabilités est calculée par l'équation 1.26

$$P_T [i,j] = 1 - \prod_{k=1,K;l=1,L} (1 - P_{k,l} [i,j]) \quad (1.26)$$

où K et L sont les nombres des fréquences radiales et d'orientations.

Si la modélisation psychovisuelle utilisée dans ce schéma opère dans l'espace de Fourier, il est néanmoins intéressant de noter la possibilité d'adapter le VDP à d'autres domaines. Bradley [17], propose une adaptation du VDP au domaine des ondelettes.

Ces travaux ont été le fondement de nombreuses métriques d'évaluation de la qualité.

1.3.8 Modèle proposé

Dans cette section, nous présentons le modèle multi-canaux, mis en œuvre au sein du laboratoire, ainsi que la Décomposition en Canaux Perceptuels (DCP) qu'il utilise. Les filtres employés dans la DCP sont détaillés, ainsi que les paramètres utilisés. Ce modèle est établi d'après des tests psychophysiques, effectués sur un ensemble d'observateurs par H. Sénane [115]. Cette modélisation du SVH se base sur le modèle présenté par Watson [123], cependant, le modèle présenté ici a été affiné et présente certaines différences.

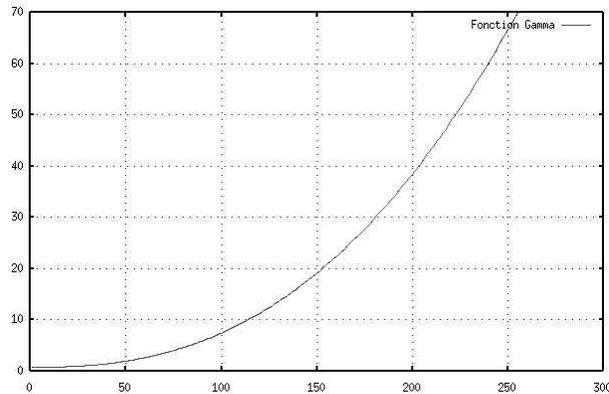


FIG. 1.17 – fonction Gamma de l'écran

La modélisation du système visuel doit tout d'abord passer par une modélisation de la perception des luminances, nous devons donc établir un modèle traduisant les valeurs numériques des niveaux de gris de l'image en luminances perçues. Cette transformation, représentée figure 1.17 est appelée "fonction gamma" car la luminance issue de la majorité des écrans est une fonction de la forme :

$$L = Min + Max \left(\frac{N}{255} \right)^\gamma \quad (1.27)$$

Où L est la luminance perçue pour tout pixel de l'image, Min représente la luminance minimale affichable par l'écran (en Candela par mètre carré : Cd/m^2), Max est la luminance maximale affichable par ce même écran et N le niveau de gris du pixel correspondant.

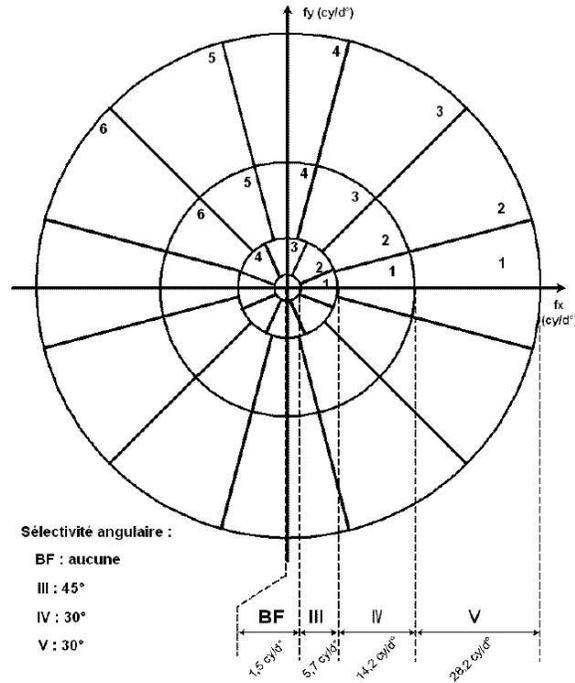


FIG. 1.18 – Décomposition en canaux perceptuels

Les filtres utilisés dans la Décomposition en canaux perceptuels (DCP) sont analogues aux filtres *cortex* de Watson [123]. Ils ont toutefois été adaptés au pavage fréquentiel du S.V.H. figure 1.18 qui n'est pas de type dyadique en fréquences radiales et ne possèdent pas une sélectivité angulaire constante vis-à-vis de la bande de fréquences radiales. Ces filtres sont obtenus par produit de filtres définissant la sélectivité radiale (appelés filtres *Dom*) et de filtres caractérisant la sélectivité angulaire (appelés filtres *Fan*).

Les sections suivantes sont dévolues à la présentation des différents filtres analytiques conduisant à l'élaboration d'une telle décomposition.

1.3.8.1 Filtres Dom

Les filtres *Dom* passe-bandes radiaux sont construits à partir du filtre passe-bas à symétrie circulaire, appelé filtre *Mesa*. Ce dernier est modélisable dans le domaine des fréquences spatiales par la convolution d'une fonction porte de valeur unité à l'intérieur du cercle de rayon $\beta/2$, $\beta \in [0, 1]$ et de valeur nulle à l'extérieur avec une gaussienne d'écart type $\sigma = \frac{\beta}{2 \cdot \gamma \cdot \sqrt{2\pi}}$. Cet écart type dépend directement du paramètre β afin d'avoir une pente de bande de transition diminuant lorsque la fréquence de coupure diminue, c'est à dire quand β diminue. Un choix correct du couple (β, γ) permet alors au filtre *Mesa* d'avoir une bande passante

maximale tout en ayant une bande de transition suffisamment douce et un gain quasiment nul aux fréquences spatiales $u = \pm 1/2$ et $v = \pm 1/2$. Cette dernière condition est importante pour la construction des couronnes (filtres *Dom*). L'expression analytique du filtre *Mesa* $M_0(u, v)$ est :

$$M_0(u, v) = \left(\frac{\gamma}{f_0} \right)^2 . e^{-\pi((\omega \cdot \gamma)/f_0)^2} \oplus \Pi \left(\frac{\omega}{2 \cdot f_0} \right) \quad (1.28)$$

où $f_0 = \beta/2$, $\omega^2 = u^2 + v^2$ et $\Pi \left(\frac{\omega}{2 \cdot f_0} \right)$ représente la fonction porte de hauteur unité, de diamètre $2 \cdot f_0$ et centré sur l'origine.

γ , figure 1.19 est un paramètre de raideur de l'atténuation, lié à l'écart type σ_0 de la gaussienne par

$$\sigma_0 = \frac{1}{\sqrt{2\pi}} \frac{f_0}{\gamma} \quad (1.29)$$

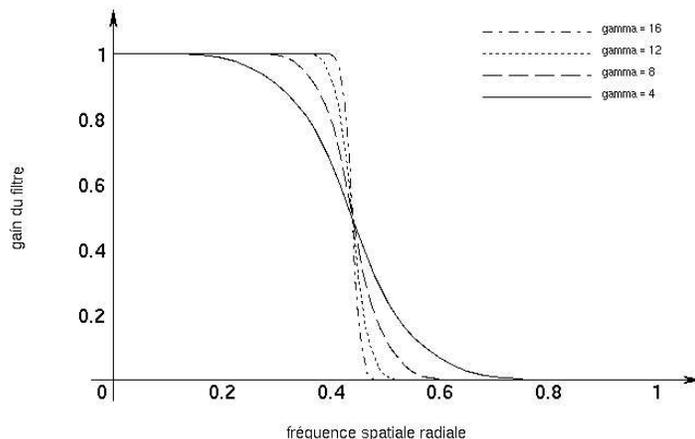
Le filtre *Mesa* d'indice i s'écrit donc

$$M_i(u, v) = M_0 \left(\left(\prod_{j=1}^{i-1} s_j \right) u, \left(\prod_{j=1}^{i-1} s_j \right) v \right) \quad (1.30)$$

où s est un facteur d'échelle.

Un ensemble de K filtres *Mesa* peut donc être généré à partir du seul filtre *Mesa* d'ordre 0 en réduisant à chaque fois la fréquence de coupure du filtre obtenu d'un facteur s .

La transition et le positionnement des filtres doivent être effectués de façon à pouvoir reconstruire exactement le spectre initial. Le chevauchement de ces filtres doit donc fournir un gain unitaire en tout point de la DCP.

FIG. 1.19 – influence du paramètre γ sur la transition du filtre

Des filtres *Mesa-homothétiques*, qui sont en fait les versions homothétiques du filtre *Mesa* de base, sont alors définis.

$$M_k(u, v) = M_0(s_k \cdot u, s_k \cdot v) \quad (1.31)$$

où s_k est le rapport d'homothétie.

Watson avait choisi $s_k = 2^k$ pour obtenir une décomposition dyadique. Dans le cas présent, la décomposition doit respecter la partition spectrale rappelée par la figure 1.18.

Un filtre *Dom* est donc engendré par une simple différence entre deux filtres *Mesa* d'indices consécutifs, (et donc de fréquence de coupure radiale différente)

$$Dom_i(u, v) = M_{i-1}(u, v) - M_i(u, v) \quad (1.32)$$

1.3.8.2 Filtres Fan

L'évolution du gain de ces filtres selon les directions angulaires est également décrite par l'action d'un flou gaussien sur un filtre angulaire idéal. Selon la direction horizontale cette évolution est donnée par

$$M'_0(u, v) = H(v) \otimes \gamma_b \exp(-\pi \gamma_b^2 v^2) \quad (1.33)$$

où

- $H(v)$ est le filtre 'échelon' qui sépare le domaine fréquentiel en deux sous domaines de gain unité pour l'un

et nul pour l'autre.

- γ_b est un nouveau paramètre de raideur.

Compte tenu que le filtre échelon n'a pas de variation selon l'axe u , la convolution peut s'écrire dans ce cas :

$$M'_0(u, v) = \int_{-\infty}^v \gamma_b \exp(-\pi\gamma_b^2 \omega^2) d\omega = F(\gamma_b, v) \quad (1.34)$$

L'expression du filtre échelon ayant une orientation θ est alors

$$M'_\theta(u, v) = F(\gamma_b, (v \cdot \cos \theta - u \cdot \sin \theta)) \quad (1.35)$$

Le filtre Fan correspondant à la k^i eme direction est finalement donné par

$$Fan_k(u, v) = M'_{\theta_k}(u, v) - M'_{\theta_{k+1}}(u, v) \quad (1.36)$$

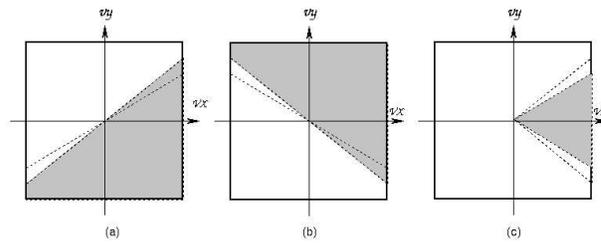


FIG. 1.20 – construction d'un filtre Fan

Nous pouvons observer figure 1.20, les trois étapes menant à la construction d'un filtre Fan.

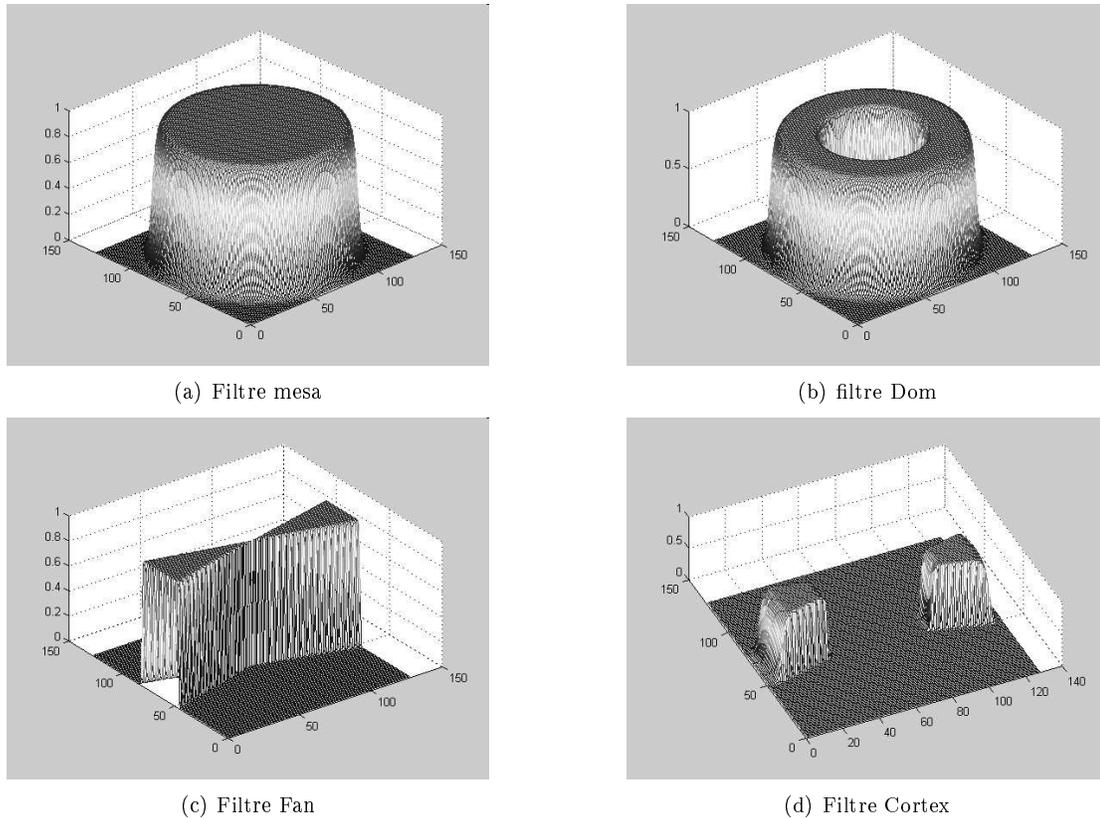


FIG. 1.21 – Étapes de construction d'un filtre analytique

Les différentes étapes de construction des filtres cortex détaillés précédemment sont illustrées dans la figure 1.21. Nous apercevons donc figure 1.21(a) un filtre MESA, en 1.21(b), un filtre DOM, la figure 1.21(c) présente un filtre Fan et enfin, un filtre Cortex est présenté figure 1.21(d).

1.4 Conclusion

Nous avons vu dans ce chapitre un descriptif détaillé du système visuel humain, la transmission de l'information visuelle depuis le champ visuel jusqu'au cortex à été étudiée étape par étape. Ceci nous à permis d'introduire différents modèles du S.V.H. découlant de ces observations psychophysiques. Ces modèles visuels sont, pour la plupart, utilisés dans des applications d'évaluation de la qualité ou de codage, mais ils sont cependant susceptibles d'être exploités dans d'autres applications prenant en compte des critères visuels approfondis, telles que les applications d'amélioration de la qualité des images.

Concernant le domaine du tatouage des images, l'optimisation des méthodes de tatouage d'images, c'est

à dire, l'obtention du meilleur compromis entre l'invisibilité et la robustesse, peut être mis en œuvre par l'utilisation de modèles visuels.

Chapitre 2

La transformation Mojette

Sommaire

| | | |
|------------|--|-----------|
| 2.1 | Introduction | 50 |
| 2.2 | Transformation Mojette directe | 50 |
| 2.3 | Transformation Mojette inverse et notion de restructibilité | 52 |
| 2.3.1 | Ensemble suffisant de projections pour la reconstruction | 52 |
| 2.3.2 | Notion de fantôme et théorème de restructibilité partielle | 54 |
| 2.3.3 | Algorithme de la transformée Mojette inverse | 57 |
| 2.4 | Opérateur d'addition | 60 |
| 2.5 | Propriétés | 63 |
| 2.6 | Applications | 67 |
| 2.7 | Conclusion | 67 |

2.1 Introduction

La transformée Mojette est une transformée de Radon [101] discrète exacte définie pour des angles de projections spécifiques. Cette transformée et son inverse stable ont été définies par Guédon en 1995 [46] [45]. Le domaine transformé 2-D d'une image $f(x, y)$ est un ensemble de projections où chaque élément (dénoté bin comme en tomographie) correspond à la somme des valeurs de pixels le long de la ligne de projection.

La transformée de Radon est utilisée pour deux types de propriétés. La première est de pouvoir reconstruire une information image à partir de projections : c'est la tomographie. Les exemples de tomographes sont nombreux en imagerie médicale (scanner, IRM, ...), mais sont aussi utilisés en astronomie ou sismologie. La seconde utilité de la transformée de Radon est de représenter une image 2D via sa version fréquentielle polaire /radiale (théorème de la tranche centrale). Dans ce contexte, la transformée de Radon peut être utilisée au sein d'un schéma de traitement d'images. Il est alors impératif d'utiliser une version discrète et dont l'approximation après transformation inverse soit la meilleure. Ce point est toujours délicat avec la transformée de Radon qui n'est pas un problème bien-posé comme la transformée de Fourier par exemple. Ceci explique en grande partie l'utilisation de la version discrète exacte appelée Mojette.

La transformation Mojette est, de par la redondance qu'elle introduit, utilisée essentiellement dans les domaines de la transmission, de la qualité de service, de la description multiple ou dans des applications de serveurs distribués.

Nous allons présenter dans ce chapitre les différents aspects de la transformation Mojette. Nous détaillerons le principe de la transformation Mojette directe dans la section 2.2. La section 2.3 sera dévolue au processus de la transformation Mojette inverse. Nous y étudierons certains aspects tels que la notion de fantôme Mojette ou la notion de restructibilité. Nous constaterons ensuite (section 2.4), les possibilités offertes par la transformation Mojette d'utiliser différents opérateurs mathématiques. Certaines propriétés inhérentes à cette transformation, telles que le théorème de la tranche centrale, la "Mojette spline", ou encore la "Mojette Verbert" seront données dans la section 2.5. Enfin, nous verrons les différentes applications de la transformation Mojette en traitement d'images (2.6).

2.2 Transformation Mojette directe

La transformée Mojette ne donne un ensemble de projections $proj(m)$ d'une image ou région $f(k, l)$ que pour des angles θ de la forme $\tan \theta = \left(\frac{q}{p}\right)$ où p et q sont des entiers relatifs premiers entre eux. De plus, pour respecter la π -périodicité de l'opérateur, seules les valeurs positives de q sont définies par l'équivalence

$$(p, -q) = (-p, q).$$

Cette transformée est linéaire. Elle définit pour chaque angle de projection par l'opérateur $M_{p,q}$:

$$M_{p,q}f(k, l) = proj(p, q, b) = proj_{p,q}(b) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(k, l)\Delta(b + kq - lp), \quad (2.1)$$

où $\Delta = \begin{cases} 1 & \text{si } b = 0 \\ 0 & \text{si } b \neq 0 \end{cases}$, et où le pixel (k,l) appartient à la région. La transformée Mojette $M_I f(k, l)$

correspond à l'ensemble de I projections $M_{p_i,q_i}f(k, l)$ pour différents angles de projection indexés par i : $M_I f = \{proj_{p_i,q_i} = M_{p_i,q_i}f, i \in [1...I]\}$.

Un exemple est présenté sur la figure 2.1. La valeur d'un bin est donc égale à la sommation des valeurs de pixels de la ligne $b = lp - kq$. La différence majeure avec la transformée de Radon est donc l'échantillonnage des projections qui dépend de l'angle, de plus, seuls les angles discrets sont permis. Cela induit un nombre de bins sur une projection (p, q) qui dépend à la fois de la forme de la région et de l'angle. Par exemple, pour une région rectangulaire $P \times Q$, la projection (p, q) est composée de B bins avec $B(p, q) = (Q - 1)|p| + (P - 1)|q| + 1$.

La complexité algorithmique de la transformée Mojette d'une région composée de N pixels représentée par un ensemble de I projections est $O(IN)$. La transformation Mojette est donc linéaire dans le nombre de pixels et linéaire dans le nombre de projections.

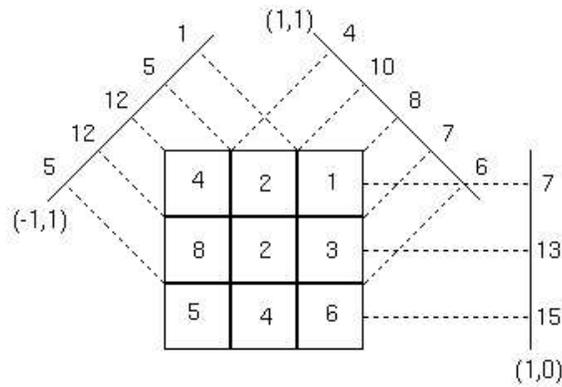


FIG. 2.1 – La transformation Mojette

La figure ci-dessus illustre le principe d'échantillonnage de la transformation Mojette pour une image de taille 3×3 et pour l'ensemble de directions $\{ (1, 0), (1, 1), (-1, 1) \}$. La figure 2.8 présente quelques exemples de projections Mojette pour l'image naturelle "bateau" (cf. Annexe C).

2.3 Transformation Mojette inverse et notion de restructibilité

Une fois la définition de la transformation Mojette présentée, le premier problème posé consiste à savoir à partir de quel ensemble de projections la reconstruction est possible. Son corollaire consiste à savoir de quelle part d'information image on dispose lorsque la condition précédente n'est pas remplie.

2.3.1 Ensemble suffisant de projections pour la reconstruction

Le premier résultat sur les conditions de reconstruction d'une région à partir d'un ensemble suffisant de projections pour l'inversion de la transformation Mojette a été trouvé dans Katz [64] dans un contexte différent. Afin de généraliser, une image $f(k, l)$ de support G convexe est dit restructible par une ensemble de I projections $M_I f$ lorsque $f(k, l)$ peut être obtenue à partir des informations des I projections par la transformation Mojette inverse. Le résultat trouvé dans Katz peut être réécrit comme l'équivalence des deux propositions :

Lemme de Katz (1979)

il y a équivalence entre les deux propositions

i) $f(k, l)$ définie sur une région G rectangulaire $P \times Q$ peut être reconstruite à partir de l'ensemble $M_I f$

ii)

$$P \leq P_I = \sum_{i=1}^I |p_i| \quad (2.2)$$

ou

$$Q \leq Q_I = \sum_{i=1}^I |q_i| \quad (2.3)$$

Ce résultat a été étendu d'une manière indépendante par Normand et Guédon [83] en utilisant la morphologie mathématique et le concept d'élément structurant à deux pixels (noté ES2P). Chaque direction de projection correspond à un élément structurant $\{O, (p, q)\}$ de deux pixels : l'origine de la grille et le point (p, q) . Dans la suite, la région R est définie comme la suite de dilatations par chaque ES2P de l'ensemble $S_I = \{O, (p_i, q_i), i \in [1 \dots I]\}$. Comme la dilatation morphologique par un ES2P consiste en la réunion ensembliste de l'image d'entrée et de sa translatée par le vecteur (p, q) , la forme R peut ainsi être construite par duplication.

Théorème 1 (Normand-Guédon 1996) [83]

La région R est la plus petite région non-restructible i.e. n'importe quelle région qui ne contient pas

entièrement R peut être reconstruite à partir de S_I . Une autre façon d'exprimer ce résultat est l'équivalence entre ces deux propositions :

- i) $f(k, l)$ définie sur le convexe G peut être reconstruite à partir de l'ensemble $M_I f$,
- ii) R construite par dilatations à partir de S_I ne peut pas être incluse dans G ,

Le corollaire est également intéressant car plus pratique à utiliser.

corollaire :

Dans les conditions de définition du théorème précédent, on a l'équivalence

- i) l'érosion de G par R est l'ensemble vide,
- ii) G peut être reconstruite à partir de l'ensemble $M_I f$.

La figure 2.2 ci-dessous illustre le lemme de Katz présenté précédemment, deux cas de figure sont possibles pour une image de taille 3×3 . La figure (a) représente un ensemble non reconstructible, alors que pour la figure (b), nous constatons que l'ensemble des directions permet la reconstruction de l'image. En (a) nous avons : $\sum |p| = 2$ et $\sum |q| = 2$ alors qu'en (b) : $\sum |p| = 3$ et $\sum |q| = 3$.

Il est important de noter que sur la figure ci-dessous, seules les valeurs positives de p respectent la notion représentée.

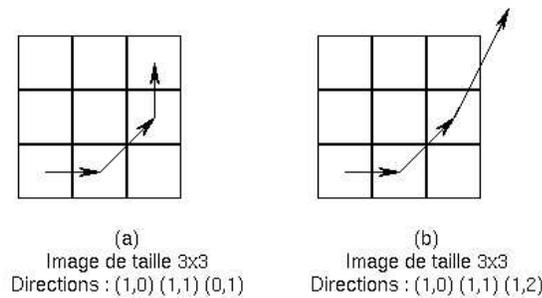


FIG. 2.2 – Notion de reconstructibilité selon le lemme de Katz

La figure 2.3 illustre la notion de reconstructibilité selon le théorème 1.

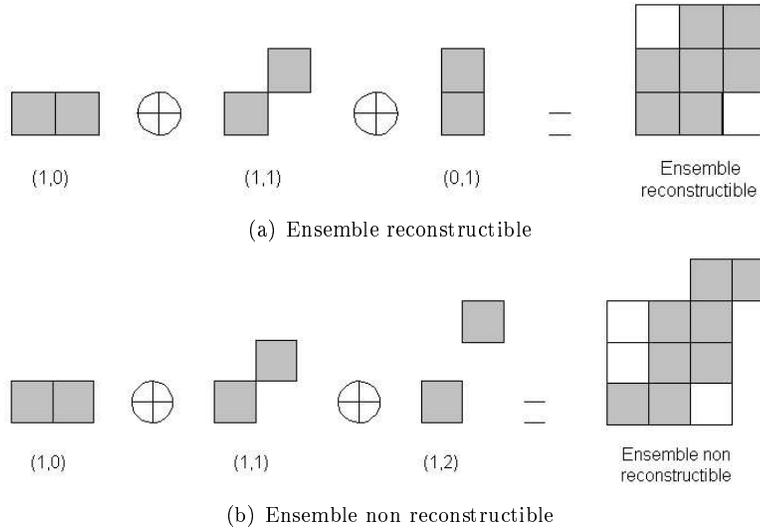


FIG. 2.3 – Notion de restructibilité selon le théorème 1

Dans le premier cas (figure 2.3(a)) la dilatation des éléments structurant produit un support pouvant être inclus dans l'image (de taille 3×3) : l'ensemble de directions est dit restructible. Dans le second cas (figure 2.3(b)), la dilatation des éléments structurant fournit un support de dimensions supérieures à l'image : l'ensemble de direction est non restructible.

2.3.2 Notion de fantôme et théorème de restructibilité partielle

Lorsque l'érosion de G par R n'est pas l'ensemble vide on ne peut pas reconstruire (le déterminant de la matrice associée est nul). Par contre l'espace nul associé à un ensemble insuffisant de projections peut être construit à partir de vecteurs de base dénotés fantômes car correspondant à des formes dans l'image qui se projettent sur le vecteur nul pour les projections de I .

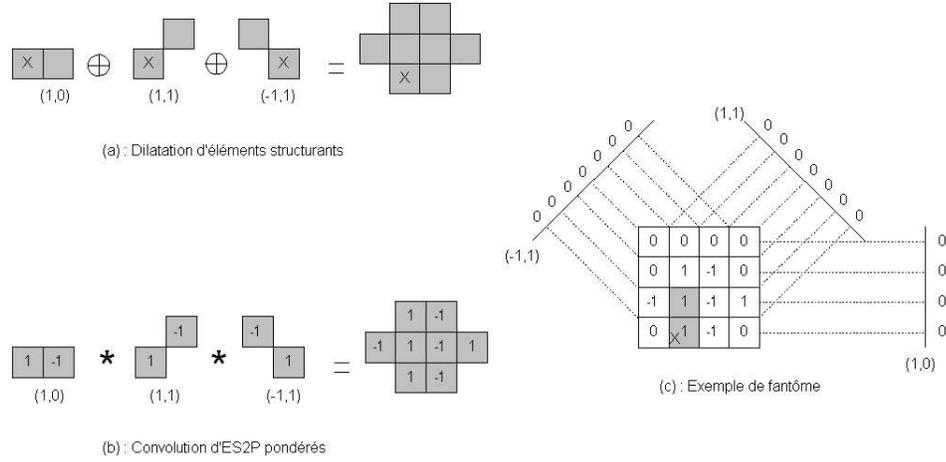


FIG. 2.4 – Exemple de fantôme Mojette

Un fantôme élémentaire est constitué par la suite des convolutions d'éléments structurants à deux pixels valués par $+1$ et -1 , ce qui permet d'obtenir zéro sur les projections correspondantes aux directions fournies par les éléments structurants. Un fantôme élémentaire est représenté sur la figure 2.4(c). L'élément structurant de départ est positionné sur le pixel contenant un repère (une croix). Nous pouvons constater que ce même fantôme peut tout aussi bien être décalé d'un pixel vers le haut. Les deux pixels de l'image permettant de positionner le fantôme sont représentés en gris sur la figure. Le résultat des dilatations d'éléments structurants est représenté en (a), alors que (b) représente la convolution d'éléments structurants pondérés. Le nombre d'inconnues est le cardinal de l'ensemble érodé.

Nous pouvons aisément déterminer pour tout fantôme basé sur un ensemble de directions définies noté $\{p_{F_i}, q_{F_i}\}$, le nombre d'emplacements possibles pour ce fantôme dans l'image. Notons Q le nombre de lignes de l'image et P le nombre de colonnes. On a alors :

$$\left(Q - \sum_i |q_{F_i}|\right) \times \left(P - \sum_i |p_{F_i}|\right) = N_{Pos}, \quad (2.4)$$

où N_{Pos} représente le nombre d'emplacements possibles du fantôme dans l'image.

Nous verrons par la suite, lors d'une application de tatouage d'images (chapitre 5.2) l'importance de N_{Pos} sur la sécurité du schéma de tatouage.

On place alors le fantôme élémentaire f_i où toute combinaison linéaire de celui-ci sur chaque pixel de l'érodé de sorte à obtenir à chaque fois une forme contenue dans la région de départ et un vecteur de base de l'espace nul (un seul pixel de l'érodé à 1 et les autres à zéro). Un exemple est donné sur la figure 2.4 où

l'érodé contient deux pixels. On note R_I la région reconstruite par I sous la contrainte d'un érodé nul (ce qui donne par application du théorème 1, une inversion possible), et a_i la valeur (inconnue) du pixel i de l'érodé. En écrivant l'égalité entre la région de départ et sa décomposition en somme directe sur l'espace projeté (en mettant l'érodé à zéro) et sur l'espace nul on obtient le théorème suivant :

Théorème 2 (Philippé - Guédon 1997) [93]

$$R = R_I + \sum_{i=1}^c a_i f_i \quad (2.5)$$

où $c = \text{Card}(\text{erode}_I(R))$ c'est à dire le nombre de pixels de l'érodé de la région R par les différentes projections de l'ensemble S_I .

Si l'on érode le support donné dans la figure 2.4 par les éléments structurants à deux pixels de directions $(1,0)$, $(1,1)$, $(-1,1)$, la région érodée est composée des deux pixels que nous avons représentés en grisé sur la figure 2.4. Si nous supposons nulles les valeurs de l'érodé [44], la reconstruction Mojette devient alors possible avec ces trois directions données. Selon l'équation 2.5, nous pouvons décomposer l'image de la façon décrite sur la figure 2.5.

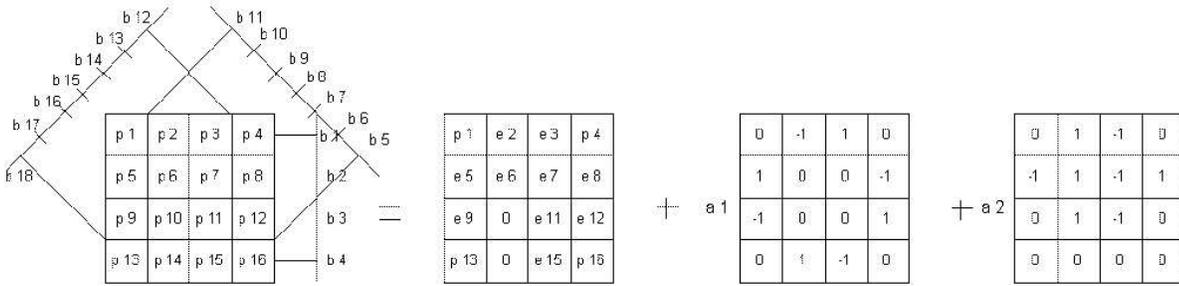


FIG. 2.5 – Décomposition de la reconstruction Mojette

où

$$\begin{aligned} e_{15} &= b_4 - p_{13} - p_{16} & e_{11} &= b_3 - e_9 - e_{12} \\ e_9 &= b_{17} & e_2 &= b_{13} - e_7 - e_{12} \\ e_5 &= b_{16} - e_{15} & e_3 &= b_1 - p_1 - e_2 - p_4 \\ e_7 &= b_8 - p_4 - p_{13} & e_8 &= b_7 - e_4 \\ e_{12} &= b_6 - e_{15} & e_6 &= b_2 - e_5 - e_7 - e_8. \end{aligned}$$

2.3.3 Algorithme de la transformée Mojette inverse

La transformation Mojette inverse correspond à un algorithme simple et rapide, comme explicité dans la thèse de N. Normand [82]. Une fois connue la forme de la région convexe, une version binaire (pixel à 1 pour la forme et 0 pour le fond) est calculée ainsi que ses projections. Cela donne explicitement le nombre de pixels correspondant à un bin. Une simple itération sur le nombre de pixel à reconstruire est réalisée en utilisant à chaque fois une correspondance univoque pixel -bin (si cela n'est pas possible, la région ne peut être reconstruite par l'ensemble des projections comme déterminé dans le théorème 1). La valeur du bin est rétroprojetée dans le pixel et soustraite de toutes les projections. Le nombre de pixels contribuant aux bins correspondants est décrémenté de un.

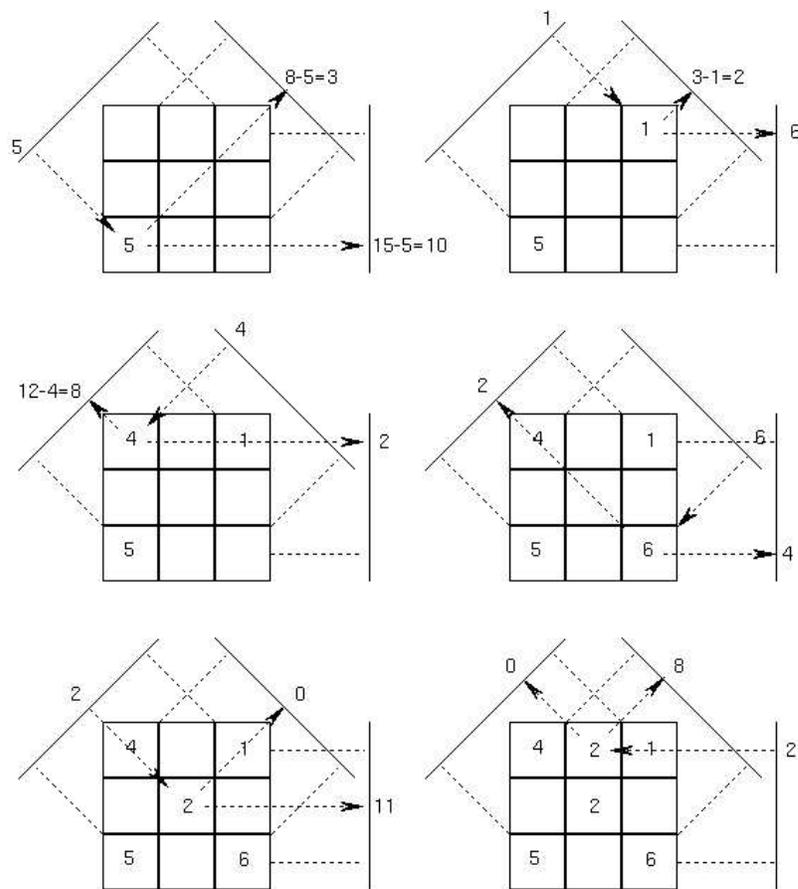


FIG. 2.6 – La transformation Mojette inverse

Procédant ainsi, on érode la forme qui reste à reconstruire du pixel rétroprojeté. La complexité algorith-

mique de la transformée inverse d'une région de N pixels et d'un jeu de I projections est également de $O(IN)$ ce qui donne une transformation équilibrée [83].

Une autre propriété importante de la transformation Mojette inverse est, dans un contexte de transmission, l'utilisation des projections dès leur réception pour débiter la reconstruction (en mettant à zéro les projections non reçues et en additionnant ces projections aux mises à jour effectuées).

La transformation Mojette inverse, bien que très simple à comprendre selon un exemple visuel simple, tel que celui présenté sur la figure 2.6, est toutefois un peu plus complexe à mettre en œuvre algorithmiquement.

Pour une image rectangulaire de taille $(P \times Q)$, l'algorithme de Mojette directe, trouvé dans [44] est donné ci-dessous.

Algorithm 1 Algorithme de la Mojette Directe

```

pour  $lig = 1$  à  $Q$  faire
  pour  $col = 1$  à  $P$  faire
     $m \leftarrow p \times lig - q \times col$  // calcul du bin m correspondant
     $proj+ \leftarrow image(lig, col)$  // incrémentation du bin
  fin pour
fin pour

```

L'algorithme de la transformation Mojette inverse, légèrement plus complexe, est décrit ci-dessous.

Algorithm 2 Algorithme de la Mojette Inverse

```

Initialisation:  $image \leftarrow -1$ 
pour  $compteur = 1$  à  $P \times Q$  faire
   $(i, m) \leftarrow Phase_1$  // chercher correspondance univoque
   $(lig, col) \leftarrow Phase_2(i, m)$  // chercher élément d'image correspondant
   $Phase_3$  // mise à jour sur chaque projection incrémentation du bin
fin pour // l'image est reconstruite

```

Détaillons à présent les trois phases utilisées dans cet algorithme.

L'objectif de la première phase, est de chercher les bins à correspondance univoque (bins n'étant issus que d'un unique pixel).

Algorithm 3 $Phase_1$

```

// chercher correspondance univoque
 $pastrouvé \leftarrow vrai$ ,  $pasfini \leftarrow vrai$ 
tant que  $pastrouvé$  et  $pasfini$  faire
  on parcourt les projections  $proj(i, m)$ 
fin tant que
// Ici, on a trouvé  $(i, m)$  ou bien le support n'est pas restructurable

```

Lors de la seconde phase, on 'rétroprojette' le bin à correspondance univoque dans l'image à reconstruire.

Algorithm 4 *Phase₂*

```

// reconstruisons le pixel correspondant : il n'existe qu'un seul élément de image dont la valeur est -1 sur
la ligne  $m = p_i \times lig + q_i \times col$ 
 $col \leftarrow 1$ ,  $pastrouvé \leftarrow \text{vrai}$ 
tant que  $pastrouvé$  faire
  si  $(m - q_i \times col)$  est divisible par  $p_i$  (donne  $lig$ ) et  $(\text{image}(lig, col) = -1)$ 
  alors  $pastrouvé \leftarrow \text{faux}$ 
  sinon  $col \leftarrow +1$ 
  fin si
fin tant que

```

La dernière phase permet, pour chaque bin 'rétroprojeté', de mettre à jour le bin correspondant au pixel dans toutes les autres projections.

Algorithm 5 *Phase₃*

```

0: // mise à jour des projections
pour  $i \leftarrow 1$  à  $I$  faire
   $m \leftarrow p_i \times lig - q_i \times col$  // calcul du bin  $m$  correspondant
   $proj(i, m) \leftarrow \text{image}(lig, col)$  // décrémentation du bin
   $proj\#(i, m) \leftarrow 1$  // décrémentation du # éléments
fin pour

```

La figure 2.7 ci-dessous, présente les différentes étapes de l'algorithme de la transformation Mojette inverse.

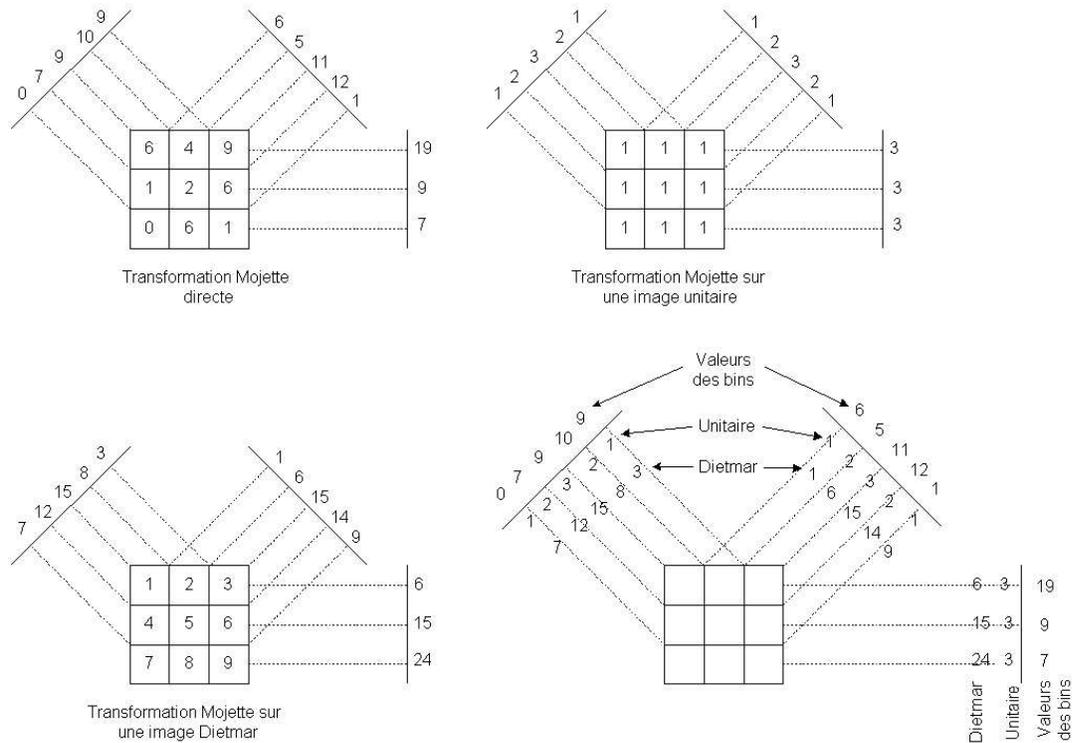


FIG. 2.7 – Étapes de l'algorithme de transformation Mojette inverse

Nous constatons que deux transformations Mojette directes sont nécessaires pour la transformation Mojette inverse. Une transformation Mojette sur une image unitaire (représentant la forme du support) permet l'obtention du nombre de pixels traversés. La transformation Mojette sur une image baptisée "Dietmar" permettra de définir les pixels où devront être rétroprojetés les bins.

En balayant tous les bins disponibles, lorsqu'un bin correspondant à l'image unitaire est égal à un, cela signifie que l'on peut rétroprojeter la valeur du bin dans le pixel dont la position est donnée par le bin correspondant au Dietmar.

2.4 Opérateur d'addition

Nous avons donc pu constater dans le paragraphe 2.2, que la transformation Mojette effectue plusieurs projections d'une image en deux dimensions vers autant de supports en une dimension. Les projections Mojette sont en fait des additions des niveaux de gris de l'image, il est donc envisageable pour ce même opérateur Mojette, de remplacer les additions par tout autre opérateur mathématique commutatif.

Dans le but "d'alléger" les projections obtenues, il est donc possible d'effectuer une transformation en

modulo des bins.

La figure 2.8(a), représente la projection selon la direction $(1, 0)$ de l'image 'bateau', alors que la figure 2.8(b) représente la projection modulo 256 selon la même direction. Les figures 2.8(c) et 2.8(d) représentent respectivement les projections en niveaux de gris et modulo 256 de l'image bateau selon la direction $(0, 1)$. Enfin, les figures 2.8(e) à 2.8(h), représentent les projections selon les directions $(1, 1)$ et $(1, 2)$.

Il est à noter que les figures 2.8(a), 2.8(c), 2.8(e) et 2.8(g) ont été redimensionnées, car elles ont des maximas allant de 22548 Niveaux de Gris (NG) pour la direction $(1, 2)$ à 51127 NG pour la direction $(0, 1)$, alors que les projections en modulo 256 sont évidemment limitées à 256 niveaux.

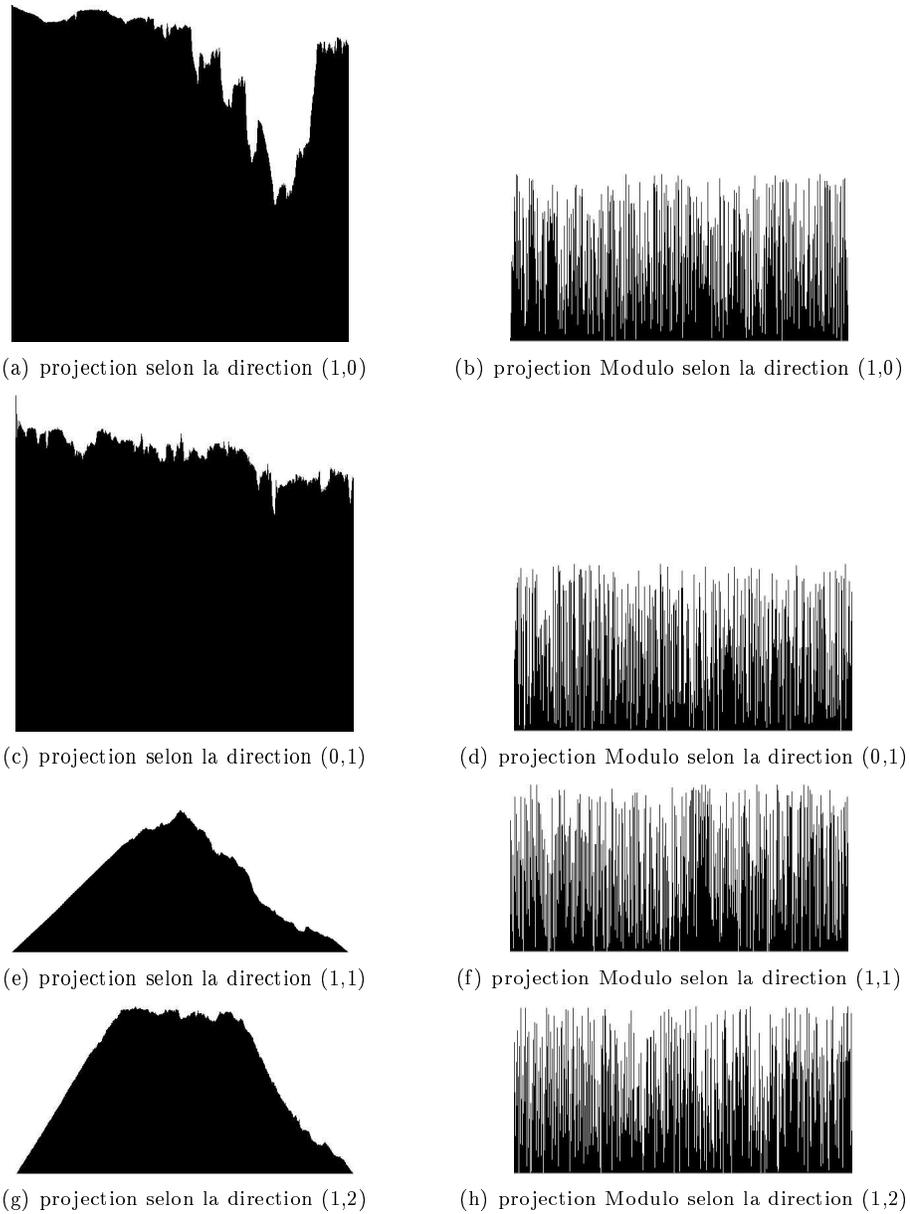
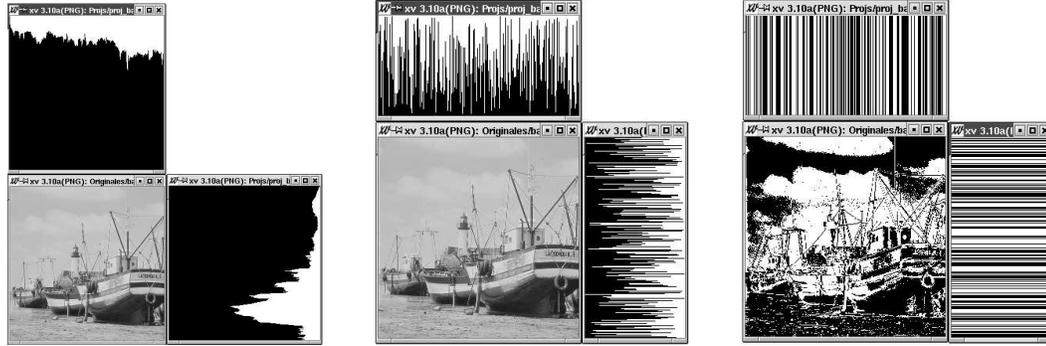


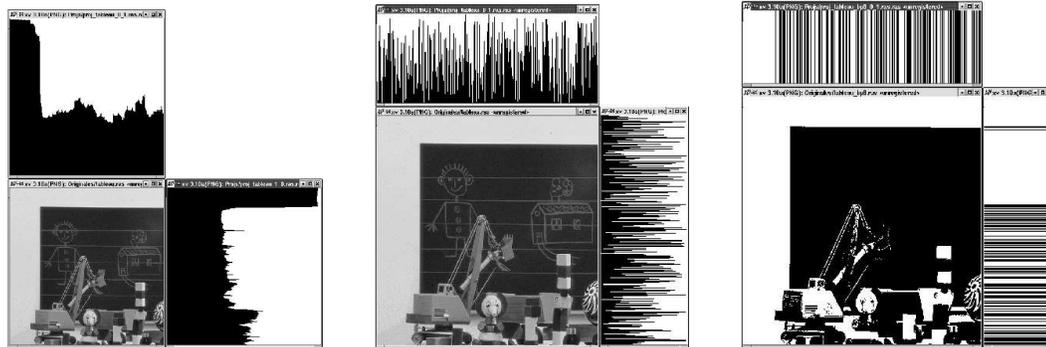
FIG. 2.8 – Projections Mojette de l'image 'bateau' selon différentes directions

Il est donc envisageable d'effectuer une transformation Mojette sur une image binaire où éventuellement sur les différents plans binaires d'une image en niveaux de gris, à l'aide de l'opérateur OU exclusif (XOR).

La figure 2.9 illustre pour les images bateau et tableau, les trois opérateurs cités ci-dessus. Les figures 2.9(c) et 2.9(f) ont été calculées sur le plans binaire de poids le plus important des images 'bateau' et 'tableau'.



(a) projection de 'bateau' selon les directions (1,0) et (0,1) (b) projection Modulo 256 de 'bateau' selon les directions (1,0) et les directions (1,0) et (0,1) (c) projection XOR de 'bateau' selon les directions (1,0) et (0,1)



(d) projection de 'tableau' selon les directions (1,0) et (0,1) (e) projection Modulo 256 de 'tableau' selon les directions (1,0) et (0,1) (f) projection XOR de 'tableau' selon les directions (1,0) et (0,1)

FIG. 2.9 – Projections Mojette des images 'bateau' et 'tableau' selon les directions (1,0) et (0,1)

2.5 Propriétés

Nous allons présenter dans cette section quelques propriétés de la transformation Mojette.

La transformation Mojette étant une version discrète spécifique de la transformation de Radon, le théorème de la tranche centrale est aussi applicable dans le *domaine Mojette*.

Rappelons ce théorème, qui opère dans le domaine continu :

soit la transformée de Fourier :

$$F(u, v) = \int \int f(x, y) \exp(-2\pi i (ux + vy)) dx dy \quad (2.6)$$

pour la ligne $v = 0$ du spectre de Fourier, on a :

$$\begin{aligned}
F(u, 0) &= \int \left[\int f(x, y) dy \right] \exp(-2\pi i (ux)) dx \\
&= \mathcal{F}(p_{\theta=0}(t))
\end{aligned} \tag{2.7}$$

où $p_{\theta}(t)$ représente une projection de l'image suivant un angle θ . Et pour une ligne passant par l'origine du spectre de Fourier avec un angle θ on a :

$$\begin{aligned}
F(r \cos \theta, r \sin \theta) &= \int \int f(x, y) \exp(-2\pi i (xr \cos \theta + yr \sin \theta)) dx dy \\
&= \int \left[\int f(t \cos \theta - s \sin \theta, s \cos \theta + t \sin \theta) ds \right] \exp(-2\pi i r t) dt \\
&= \mathcal{F}(p_{\theta}(t))
\end{aligned} \tag{2.8}$$

où le changement de variable suivant a été utilisé :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} t \\ s \end{pmatrix} \tag{2.9}$$

En conclusion, à une ligne du spectre de Fourier d'orientation θ correspond la Transformée de Fourier de la projection de l'image selon cet angle.

La Mojette, telle que nous l'avons présentée dans l'équation 2.1 et décrite par la suite, pourrait tout aussi bien être appelée "Mojette-Dirac". Ici, seulement les valeurs de l'image se trouvant sur une grille discrète, sont prises en compte lors de la projection. Le noyau d'interpolation est une fonction de Dirac de noyau cardinal $nc(x) = \delta(x)$. L'équation 2.1 peut alors se réécrire :

$$f_{\delta}(x, y) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(k, l) \cdot \delta(x - k\Delta) \cdot \delta(y - l\Delta) \tag{2.10}$$

La Mojette-spline [47] est définie pour d'autres noyaux d'interpolation.

L'interpolateur constant par morceaux est défini par l'équation 2.11 :

$$nc_0(x) = \begin{cases} 1 & , si & , |x| < \frac{1}{2}, \\ \frac{1}{2} & , si & , |x| = \frac{1}{2}, \\ 0 & sinon \end{cases} \tag{2.11}$$

Rappelons l'équation 2.1 :

$$M_{p,q}f(k, l) = proj(p, q, b) = proj_{p,q}(b) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(k, l)\Delta(b + kq - lp), \quad (2.12)$$

avec le noyau d'interpolation défini ci-dessus, cette équation devient :

$$proj_0(t, \theta) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(k, l) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} nc_0(x - k\Delta)nc_0(y - l\Delta)\delta(t - x \cdot \cos \theta + y \cdot \sin \theta)dx dy \quad (2.13)$$

Comme nous l'avons vu section 2.1, la contrainte de 'discrétisation' Mojette impose des angles θ tels que $\tan \theta = \left(\frac{q}{p}\right)$, ce qui conduit à :

$$proj_0(t, \theta) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(k, l) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} nc_0(x - k\Delta)nc_0(y - l\Delta)\delta(t \cdot h - x \cdot p + y \cdot q)dx dy \quad (2.14)$$

Pour un angle θ donné, soit pour $noyau_0(t, \Delta, p, q)$:

$$noyau_0(t, \Delta, p, q) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} nc_0(x - k\Delta)nc_0(y - l\Delta)\delta(t \cdot h - x \cdot p + y \cdot q)dx dy \quad (2.15)$$

Soit :

$$noyau_0(t, \Delta, p, q) = nc_0\left(\frac{p - t \cdot h}{q} - k \cdot \Delta\right) \times nc_0\left(\frac{x \cdot p - t \cdot h}{q} - l \cdot \Delta\right) \quad (2.16)$$

Enfin, la Mojette-Haar est donnée par l'équation 2.17 :

$$M_{p,q}^0 f(k, l) = proj_0(p, q, b) = proj_{p,q}(b) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(k, l)trapeze(b + kp - lq) \quad (2.17)$$

avec

$$trapeze(b + k \cdot p - l \cdot q) = \begin{cases} 1, & \text{si } b \leq \frac{|p| - |q|}{2}, \\ \frac{|p| + q - 2b}{2 \cdot \min(|p|, |q|)}, & \text{si } \left(\frac{|p| - |q|}{2}\right) \leq b \leq \frac{|p| + q}{2} \\ 0 & \text{sinon} \end{cases} \quad (2.18)$$

Une autre forme de transformation Mojette, appelée 'Mojette Verbert' [119], peut permettre dans certains cas de compresser des images binaires. Cette forme de Mojette calcule pour un nombre de directions restreint

au maximum, la somme des pixels de valeur '1' traversés par les directions de projection. Ici, lorsque la valeur d'un bin est égale à '0' ou au nombre de pixels traversés, la rétroprojection peut s'opérer sur une ligne entière. Cette dernière sera composée uniquement de '0' dans le premier cas ou uniquement de '1' dans le second.

La figure 2.10 illustre ce principe pour une image binaire de taille 4×4 et les directions de projection $(1, 1)$ et $(1, -1)$. Les quatre bins entourés suffisent à eux seuls, à reconstruire les sept pixels représentés en grisé.

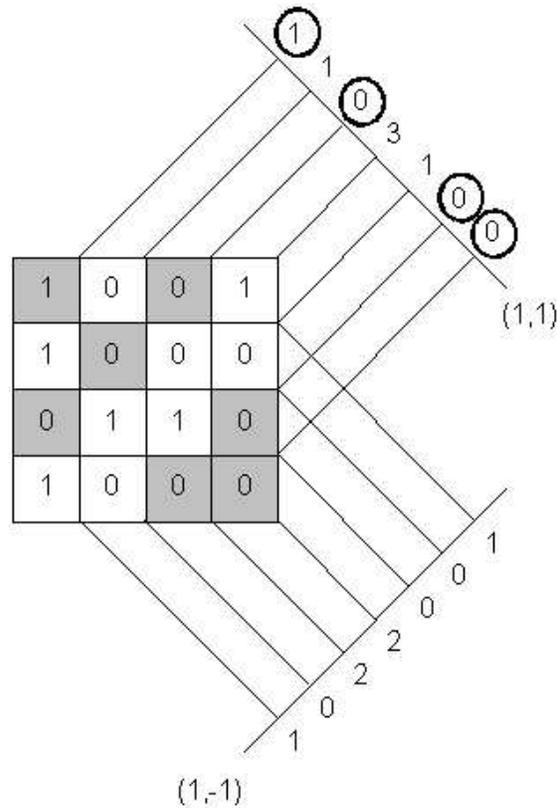


FIG. 2.10 – Principe de la Mojette 'Verbert'

Nous pouvons noter sur cet exemple et avec ce procédé que, outre les quatre bins entourés, la prise en compte des sept bins de la projection $(1, -1)$ suffisent à reconstruire l'image. Donc, ici, onze bins suffisent à reconstruire seize pixels.

Moyennant l'utilisation d'un codage de type Huffman ou RLE (Run Length Encoding), les onze bins nécessaires à la reconstruction peuvent être codés de façon à comprimer l'information représentant l'image.

2.6 Applications

De part la redondance qu' elle introduit, la transformation Mojette est principalement utilisée dans des applications de transmission d'information[48], ou de qualité de service. La redondance permet dans un schéma de transmission d'information par le biais de réseaux informatiques, de garantir une bonne réception malgré d'éventuelles pertes du réseau ou les additions de bruits classiques dans de telles applications. Les caractéristiques détaillées dans ce chapitre permettent à la transformation Mojette d'être employée dans des applications de code correcteur d'erreur. Certaines études sont en cours sur l'exploitation de la transformation Mojette en n-dimensions [86] [120]. Ne nombreuses études mettant en jeu la transformation Mojette ont permis son adaptation à diverses applications tel que la diffusion de comptes rendus médicaux multimédias, la mise en œuvre d'un protocole de transmission ionosphérique avec protection des flux. La transformation Mojette a aussi permis l'élaboration d'une application de commerce électronique assurant l'intégrité des données et gérant les aspects de confidentialité. Des études portant sur l'élaboration d'une base de données distribuée et sécurisée sont actuellement en cours. De plus, certaines disciplines issues du traitement des images, tel que l'analyse de textures où la compression des données peuvent être assurées par la transformation Mojette. Ces différentes applications sont détaillées par Guédon dans [44]. De plus, nous allons voir par la suite, différentes applications de tatouage d'images, mais aussi de stéganographie ainsi que de cryptographie, utilisant la transformation Mojette.

2.7 Conclusion

La transformation Mojette est une transformation réversible dont la caractéristique principale est la redondance qu'elle introduit. Cette redondance explique les nombreuses exploitations de la Mojette dans des applications de transmission, de qualité de service , de description multiple [85], où de systèmes d'information distribuée [49]. Outre la redondance, une caractéristique très intéressante qu'offre la transformation Mojette est la notion de fantômes, que nous exploiterons par la suite pour "cacher" de l'information dans des images.

Chapitre 3

Introduction au tatouage d'images

Sommaire

| | | |
|------------|---|-----------|
| 3.1 | Introduction | 70 |
| 3.2 | historique | 70 |
| 3.3 | Les applications | 71 |
| 3.4 | Les attaques. | 73 |
| 3.5 | Schéma global de tatouage | 75 |
| 3.6 | Le choix de l'espace | 76 |
| 3.6.1 | Méthodes spatiales | 78 |
| 3.6.2 | Méthodes fréquentielles | 80 |
| 3.6.3 | Méthodes utilisant des critères visuels | 82 |
| 3.6.4 | Mesures de distorsions | 89 |
| 3.7 | Les techniques d'extraction et de détection | 94 |
| 3.8 | Tatouage et cryptographie : liens et différences | 95 |
| 3.9 | Conclusion | 96 |

3.1 Introduction

Le tatouage d'image ou watermarking est apparu au début des années 1990, dans le but de parer le piratage des documents multimédia.

Le tatouage consiste donc à insérer une information numérique, éventuellement cryptée, dans une image, une vidéo, une séquence sonore ou même un texte numérisé. Les applications du tatouage d'images sont vastes. La littérature nous fournit diverses approches de tatouage, tant sur des images monochromes que sur des images couleurs, bien que ce dernier cas n'en soit qu'à ses balbutiements. Il existe différentes méthodes de tatouage de séquences vidéo et le domaine du tatouage de signaux audio est également assez exploité.

Nous allons présenter dans ce chapitre, un bref historique des méthodes de tatouage, nous présenterons les différentes applications que recouvrent le terme générique de "tatouage d'images", un inventaire complet des attaques à envisager sera ensuite détaillé. Dans la section 3.6 le problème du choix des espaces potentiels sera étudié, enfin les principales méthodes existantes seront analysées dans les paragraphes 3.6.1 et 3.6.2.

3.2 historique

Le tatouage est utilisé sur n'importe quel support numérique, nous pouvons trouver tout aussi bien du tatouage de séquences vidéos [126], [78], [52], [51], de flux vidéo [81] ou audio [65], [16] que du tatouage de texte tel que la méthode, proposée par Brassil et al. [2], qui présente une méthode de tatouage par décalages imperceptibles de mots ou de lignes du texte. Cependant, les méthodes de tatouage de texte souffrent encore d'un problème important de robustesse. Nous pouvons aussi trouver dans [20],[21], une méthode de tatouage de partitions musicales par décalages des notes. Un article regroupant différents médias est celui de Bender et al. [16].

Les travaux effectués dans le cadre de cette thèse, portent sur l'insertion de copyrights numériques dans des images fixes monochromes. Le tatouage des images monochromes peut être étendu aux images couleurs, par l'application de la méthode sur la composante achromatique. Toutefois, une étude psychovisuelle de la perception des images couleurs permettrait une meilleure optimisation du tatouage. Nous verrons que la restriction de la discipline du tatouage aux images monochromes, permet tout de même un vaste champ d'investigations.

Concernant l'insertion du filigrane, différentes approches sont étudiées pour tatouer une image, les premières méthodes d'insertion s'opéraient dans le domaine spatial. Le tatouage spatial consiste la plupart du temps, à modifier les bits les moins significatifs d'une image (appelés LSB), ou à modifier les valeurs des niveaux de

gris de certains pixels pré-sélectionnés en s'aidant ou non de masques pour « camoufler » le filigrane. Comme nous le constaterons par la suite le tatouage 'spatial' assure aisément l'invisibilité mais pose des problèmes de détection, notamment pour des attaques géométriques en provoquant une dé-synchronisation de la marque. Il est bien évidemment possible de concevoir un schéma de 're-synchronisation' de la marque, c'est à dire une compensation des déformations géométriques, afin de rétablir l'image avant de procéder à la détection de la marque. Cette solution n'est pas réellement évidente car aucun indice ne nous permet de savoir dans quelle position était l'image lors de l'insertion.

Une solution qui s'impose naturellement pour éviter ce manque de robustesse, sans avoir à se soucier de la 're-synchronisation', est l'exploitation des domaines transformés. Là encore les solutions ne manquent pas. Nous trouvons donc du tatouage de coefficients DCT, DFT, ou ondelettes, certains auteurs utilisent la transformation de Fourier-Mellin et d'autres exploitent la théorie de l'étalement spectral (Spread Spectrum)[95].

Certaines études portant sur la recherche de liens ou sur des comparaisons entre les différents domaines ont été menées dans [98] pour les domaines DCT et ondelettes.

Quant à l'aspect psychovisuel du tatouage, du fait de sa complexité, il n'est pas encore étudié de façon approfondie. La plupart des schémas de tatouage utilisant une sélection de sites, se basent sur des critères empiriques.

Concernant l'extraction du copyright, là encore plusieurs solutions s'offrent à nous. L'extraction peut être :

- aveugle dite aussi publique : le tatoueur n'a recours à aucune information
- semi aveugle : utilisation d'un masque par exemple
- non-aveugle, dite privée : on a recours à l'image originale ainsi qu'au filigrane original.

Évidemment, l'extraction de la marque doit s'avérer fructueuse, après avoir fait subir à l'image des traitements tels que la compression (JPEG), des déformations géométriques, différentes sortes de filtrages, de la quantification, l'addition de bruits, le tatouage multiple, un découpage de l'image et bien d'autres manipulations.

Nous verrons dans ce chapitre, les différentes applications du tatouage d'images, mais aussi une liste complète des attaques envisageables, ainsi qu'une étude bibliographique des schémas de tatouage existants.

3.3 Les applications

Le tatouage d'images numériques peut être employé pour diverses applications, comme nous venons de le voir, la protection des droits d'auteurs en est une d'importance. Cependant d'autres applications, telles que le fingerprinting, ou la stéganographie sont à considérer. Voici donc les principales caractéristiques des différentes

applications du tatouage d'image :

- **La protection de droits d'auteurs** est donc l'application la plus complexe et la plus fréquemment étudiée, ici, la marque doit être invisible mais aussi la plus résistante possible, car il est utilisé pour parer le piratage de données. C'est essentiellement cette application que nous étudions ici.
- **Le fingerprinting** (ou empreinte) est utilisé dans le but de suivre les copies illégales des documents numériques. Le propriétaire peut insérer différents filigranes dans les copies de l'image. Ceci permet, lorsque l'on trouve une copie illégale de déterminer l'identité du client ayant transgressé la licence.
- **La stéganographie** sert à transmettre un message secret 'camouflé' dans un support numérique. Cette technique nécessite donc une parfaite invisibilité, mais aussi de solides notions de cryptographie, cependant la robustesse n'est pas ici une qualité importante.
- **l'authentification des données** (ou tatouage fragile) permet de savoir si l'image a subi des malversations et si tel est le cas, certaines informations pointant la localisation des dégradations peuvent être extraites. Comme son nom l'indique, cette méthode de tatouage ne doit pas être robuste, mais bien au contraire la marque insérée doit être la plus fragile possible pour être effacée dès lors qu'une malversation apparaît.
- D'autres applications telles que l'indexation peuvent tout aussi bien être assurées par le tatouage d'images.

Fridrich a résumé dans [37] les différentes applications du tatouage d'images ainsi que les exigences requises pour chaque application en matière de robustesse, invisibilité, capacité ...

Le tableau 3.1 ci-dessous, présente ces exigences.

| Exigences | | | | | | Applications |
|-----------|---|---|---|----|----|--|
| C | R | I | S | CI | CE | |
| 7 | 2 | 7 | 7 | 2 | 2 | Stéganographie |
| 1 | 7 | 4 | 7 | 2 | 2 | Protection des droits d'auteurs |
| 3 | 7 | 4 | 7 | 1 | 6 | Fingerprinting |
| 4 | 5 | 4 | 1 | 2 | 7 | Addition d'informations (sous-titres, ...) |
| 2 | 6 | 5 | 7 | 7 | 2 | Protection de l'intégrité |
| 2 | 5 | 4 | 7 | 1 | 7 | Contrôle de copies (DVD) |
| 2 | 6 | 4 | 7 | 1 | 7 | Browsers intelligents, information automatique de copyright, ... |

TAB. 3.1 – Exigences en fonction des différentes applications

où C signifie Capacité, R : Robustesse, I : Invisibilité, S : Sécurité, CI : Complexité d'insertion et CE : Complexité d'extraction. Les valeurs données dans la colonne exigence représentent le degré requis dans chacune des catégories citées ci-dessus, en fonction de l'application. Ces valeurs sont numérotées de 1 à 7 dans un ordre

croissant d'exigence.

Les contraintes fixées dans le cas des méthodes de protection des droits d'auteurs sont très bien détaillées par Bartolini et al. dans [13]. Les auteurs détaillent les méthodes symétriques et asymétriques de tatouage, mais aussi les schémas d'extraction aveugle ou semi-aveugle. La notion de clef publique / clef privée est explicitée, ainsi que la réversibilité du tatouage.

3.4 Les attaques.

Afin d'améliorer les performances des différents schémas de tatouages existants, certains travaux portant sur les aspects de la robustesse des schémas ont été conduits. Nous pouvons trouver certains utilitaires opérant de nombreuses attaques, tels que Stirmark, créée à Cambridge [68], [92] et [91] le plus utilisé et le plus éprouvé de tous, mais aussi Checkmark, de l'Université de Genève [77] ou encore UnZign. Un bref descriptif de ces algorithmes peut nous aider à comprendre les différentes attaques à craindre sur les schémas de tatouage :

- **Stirmark** est écrit en langage C par Fabien Petitcolas et Martin Kutter, les attaques proposées par cet algorithme sont assez dévastatrices et touchent principalement aux problèmes de (dé-)synchronisation des schémas de détection (d'extraction). Effectivement, Stirmark effectue 137 attaques dont 100 déformations géométriques. Une liste exhaustive de ces attaques est détaillée dans [68], certaines images résultantes de Stirmark sont présentées dans l'Annexe D.
- **Checkmark**, écrit en langage Matlab, utilise plutôt des attaques propres au contenu de l'image, telles que du codage avec perte, du filtrage passe bas, ...
- **UnZign**, un logiciel de dégradations d'images créé en langage C sous Unix, fût disponible sur Internet à l'adresse <http://www.altern.org/watermark>, mais n'est actuellement plus maintenu.

La modélisation des attaques est une étape très importante de la mise en œuvre d'un algorithme de tatouage. l'algorithme doit être conçu à partir de deux principes de base énoncés précédemment, l'invisibilité et la robustesse. Ces deux qualités étant très corrélées, il est indispensable d'étudier leurs interactions.

Les attaques les plus fréquemment rencontrées sont scindées en deux catégories :

Les attaques intentionnelles, mises en oeuvre le plus souvent par des pirates confirmés ou tout au moins par des personnes ayant de solides notions en traitement des images.

Les attaques naturelles, dues aux différents traitements que peuvent subir une image, imputables par exemple à des problèmes de transmission ou d'affichage de l'image. les attaques naturelles sont beaucoup moins violentes que les précédentes.

Il est donc évident que les schémas de tatouage doivent résister au moins aux attaques naturelles.

Nous récapitulons ci-dessous la plupart des attaques répertoriées dans la littérature [68]. Comme nous le verrons par la suite, la robustesse des schémas de tatouage varie en fonction des types d'attaques appliquées et du domaine d'insertion choisi.

Les attaques naturelles

- Le codage avec perte (JPEG, JPEG 2000, SPIHT, ...).
- l'addition de bruits (bruit gaussien, bruit blanc, bruit rose ...).
- Les modifications de la palette.
- La modification de luminance.
- Le rehaussement de contraste.
- La correction gamma.

Les attaques intentionnelles

- Les filtrages (passe bas, passe haut, médian).
- Les modifications d'histogramme.
- l'étalement des pixels.
- La pixelisation (sous échantillonnage suivi d'un sur-échantillonnage).
- La quantification des niveaux de gris (où des couleurs).
- La combinaison impression-scannerisation.
- Le tatouage multiple.
- Les déformations géométriques : la rotation, la translation, le découpage, le zoom, le patchwork (ajout d'images), les changements d'échelle, l'effet miroir, l'effet de mosaïque, la suppression de lignes (et/ou de colonnes).
- Toutes les combinaisons possibles de ces attaques.

Outre les attaques propres au contenu de l'image, il faut aussi prendre en considération les attaques dites "de procédure" [29]. Lors de ce type de malversations, surnommée "IBM attack", un pirate crée une image qu'il supposera originale (I') par soustraction d'un faux filigrane (W') à la "vraie" image tatouée (I) où une marque (W) est déjà insérée. Le malfaiteur peut ainsi utiliser sa clef pour extraire son (faux) filigrane (W') de l'image (I') en prouvant que cette dernière lui appartient.

Une solution possible à ce type d'attaque est l'insertion dans l'image d'une signature (fonction de hachage) [41], [40], [38], ou alors, d'avoir recours à un tiers de confiance pour stocker différents paramètres et permettant de prendre en compte la date de dépôt.

3.5 Schéma global de tatouage

Nous avons vu précédemment la nécessité de robustesse des schémas de tatouage dans le cadre de la protection des droits d'auteurs. Effectivement, dans le but de s'appropriier l'image, un pirate n'aura d'autre alternative que de la dégrader, afin de supprimer la marque. De plus, selon les applications envisagées, le filigrane pourra être extrait de l'image ou seulement détecté.

Dans la suite, nous appellerons clef, un paramètre qui s'applique sur une fonction à une variable qui est un message, on notera $f(message, clef) = f_{clef}(message)$.

Nous représentons figure 3.1 le processus général d'un schéma de tatouage comportant l'utilisation d'une clef de cryptage des données ainsi qu'une clef d'insertion. Nous pouvons observer que l'image originale est soumise à un algorithme de tatouage. L'utilisation d'une ou plusieurs clefs est fortement souhaitable lors du processus d'insertion. L'image tatouée ainsi obtenue peut subir des attaques, intentionnelles ou naturelles, comme nous l'avons vu précédemment. En fonction de l'application, deux schémas sont envisageables à la réception de l'image potentiellement tatouée : un schéma d'extraction de la marque ou un schéma de détection. Les schémas opérant une extraction de la marque sont exploités dans les applications où la récupération de l'image originale est souhaitable (imagerie médicale). Les schémas utilisant un processus de détection de la marque sont utilisés dans le but de déterminer la présence du filigrane dans l'image (schéma typique de protection des droits d'auteurs). Pour chacune de ces options, les clefs utilisées lors de l'insertion doivent alors être utilisées pour détecter ou extraire la marque.

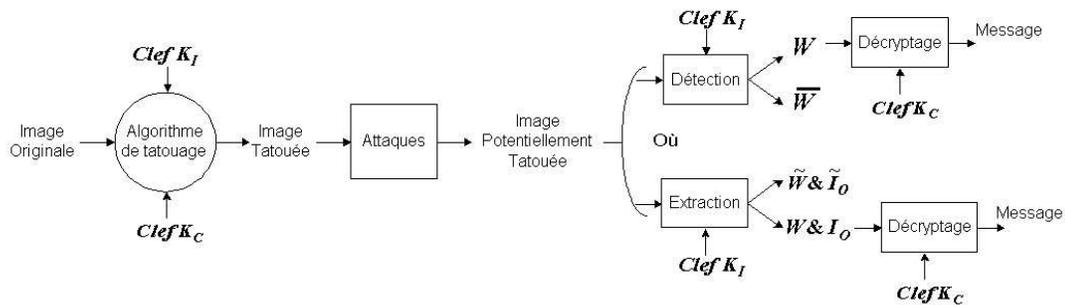


FIG. 3.1 – Schéma global de tatouage

La figure 3.2 présente dans le détail, les différentes étapes d'un schéma de tatouage. Nous constatons que la marque, issue d'un système cryptographique, codée à l'aide d'une clef K_C , est insérée dans l'image, ou dans une représentation transformée de l'image, par le biais d'une clef, notée K_I . Bien que l'utilisation d'un masque perceptuel ne soit pas automatique, elle est fortement souhaitable afin d'octroyer la force maximale au filigrane,

ou comme nous le verrons par la suite, afin de sélectionner les composantes de l'image (bits, fréquences, ...) permettant une meilleure détection.

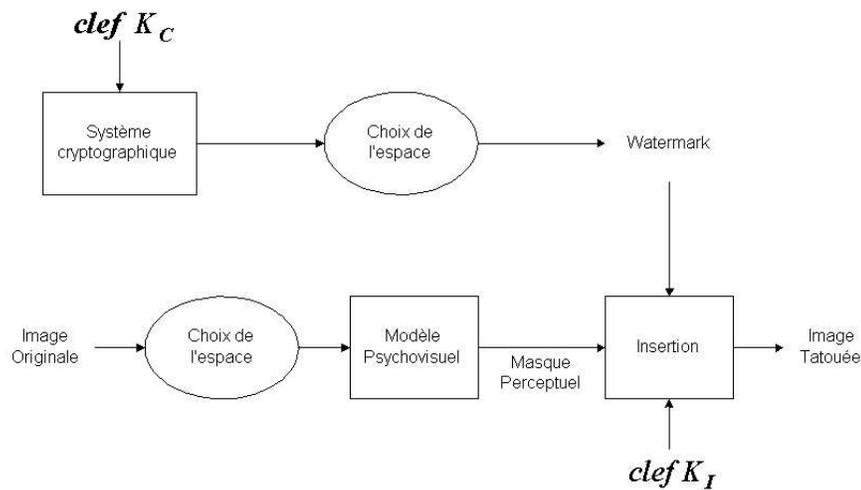


FIG. 3.2 – Schéma usuel de tatouage

La plupart des techniques de tatouage peuvent tout aussi bien se résumer par l'équation suivante [4] :

$$I_W(x, y) = I_O(x, y) + M_s(x, y) \times M_f(x, y) \times W(x, y), \quad (3.1)$$

où $I_W(x, y)$ représente l'image tatouée ou une représentation dans un espace transformé, $I_O(x, y)$ l'image originale ou sa représentation transformée, $M_s(x, y)$ un masque représentant les sites tatouables, $M_f(x, y)$ est un masque visuel de pondération et $W(x, y)$ le filigrane.

Nous devons noter que dans la figure 3.2 ou encore dans l'équation 3.1, l'utilisation d'un modèle visuel est optionnelle et peut tout aussi bien être remplacée par une sélection empirique des sites ou des forces de la marque.

3.6 Le choix de l'espace

Une des étapes la plus importante de la mise en œuvre d'un algorithme de tatouage est probablement le choix d'un espace de tatouage. Les premiers algorithmes de tatouage d'images [113] proposaient d'insérer un filigrane binaire directement dans les bits les moins significatifs de l'image appelés LSB pour 'Less Significant Bits'.

Si ce type d'approche garantit essentiellement l'invisibilité de la marque, l'aspect de la robustesse est

néanmoins peu assuré. Effectivement, la plupart des attaques des images numériques, modifient en premier lieu les LSB de l'image.

Une expérience simple montre la vulnérabilité des LSB. Effectuons la différence pixel par pixel entre une image de référence et sa version dégradée par un codage JPEG de facteur de qualité $q=75\%$, puis décomposons cette image de différence en plans binaires. Nous constatons, figure 3.3 (Les pixels noirs

représentent les bits de valeur nulle et les blancs les bits de valeur unitaire) que les plans binaires de poids faibles (LSB) ont subi les plus importantes modifications.

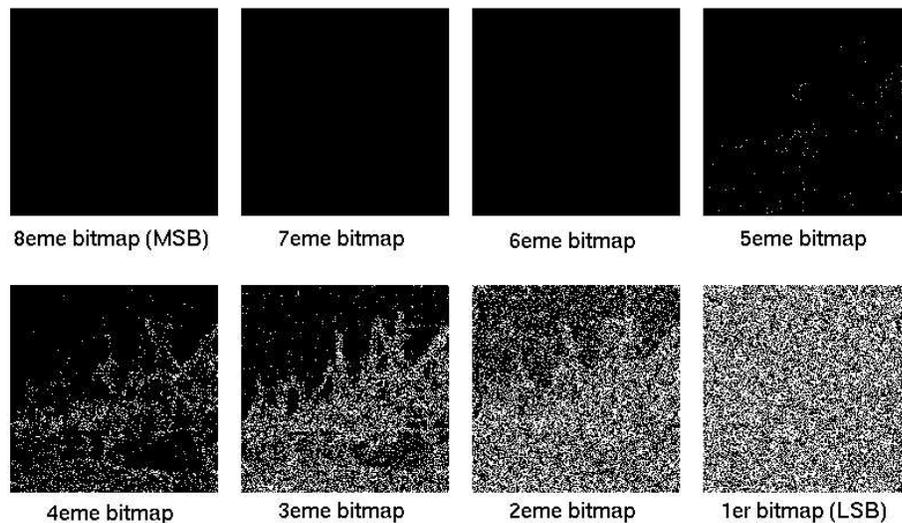


FIG. 3.3 – Plans binaires de l'image de différences

La même expérience, réalisée avec quatre images différentes, est reconduite avec différentes autres dégradations.

Le tableau 3.2 ci-dessous résume les différents résultats obtenus :

| | PB 1 | PB 2 | PB 3 | PB 4 | PB 5 | PB 6 | PB 7 | PB 7 |
|-------------------------|---------|---------|---------|---------|---------|---------|---------|--------|
| bateau JPEG75 | 50.0118 | 34.8262 | 18.8793 | 6.3262 | 0.3158 | 0.0003 | 0.0000 | 0.0000 |
| bateau Fil Med 2x2 | 49.8451 | 36.1248 | 22.6802 | 13.8904 | 7.5504 | 3.2176 | 0.8777 | 0.0583 |
| bateau Fil Gauss 3x3 | 49.7798 | 34.9994 | 20.9972 | 12.6541 | 5.5206 | 1.2886 | 0.0114 | 0.0000 |
| singe Fil Med 3x3 | 49.9847 | 47.5257 | 42.4648 | 29.1648 | 12.3798 | 4.1160 | 0.5649 | 0.0484 |
| singe Sharpening | 49.9931 | 49.2126 | 47.8820 | 45.3876 | 39.3112 | 29.6512 | 22.1599 | 0.9426 |
| singe Fil Gauss 3x3 | 50.1762 | 47.9125 | 43.6195 | 31.0611 | 17.2832 | 4.4231 | 0.0286 | 0.0000 |
| iba Fil Med 3x3 | 49.8329 | 44.4282 | 33.1146 | 17.7997 | 7.6705 | 2.3048 | 0.5455 | 0.0522 |
| iba Fil Med 4x4 | 49.8744 | 46.6449 | 39.6480 | 30.0701 | 19.4488 | 10.3195 | 3.1047 | 0.6305 |
| iba Sharpening | 49.9008 | 49.0520 | 46.9242 | 43.1774 | 34.4348 | 22.4090 | 8.8085 | 0.4661 |
| tableau Fil Med 2x2 | 50.0106 | 41.6210 | 26.7589 | 10.8436 | 4.5383 | 1.9592 | 0.8991 | 0.0984 |
| tableau Sharpening | 49.9534 | 48.4081 | 45.2075 | 38.0954 | 23.9250 | 10.2466 | 2.3578 | 0.0267 |
| tableau FilGauss 3x3 | 50.1201 | 41.7644 | 24.9679 | 9.3284 | 3.6975 | 0.9647 | 0.0240 | 0.0000 |

TAB. 3.2 – Pourcentages de pixels modifiés pour différentes attaques en fonction des plans binaires

Le tableau 3.2 confirme la faiblesse des plans binaires de poids faible face à des attaques simples, telles que des filtrages ou des rehaussement de contraste.

Les méthodes de tatouage insérant la marque dans un espace transformé présentent en général de meilleurs résultats en terme de robustesse, ainsi que d'invisibilité, étant donné que les modèles visuels sont souvent mis en œuvre dans des domaines transformés, tels que Fourier, DCT, ondelettes.

Cependant, les espaces envisageables sont nombreux et souvent applicables à certains types d'attaques spécifiques. Un tatouage DCT [9], [97] ..., est adapté au codage JPEG. Un tatouage opéré dans le domaine des ondelettes [114], [59] [67] [66] peut résister plus aisément au codage JPEG 2000, un tatouage spatial de l'image par exemple ne sera pas adapté aux déformations géométriques.

3.6.1 Méthodes spatiales

Le tatouage d'image a débuté dans le domaine spatial, les premières méthodes de tatouage effectuaient des modifications binaires. Par exemple Van Schyndel [113] proposait trois méthodes d'insertion dans les plans

binaires de l'image : Une méthode de substitution des LSB, une méthode d'addition binaire, et une méthode, appelée "patchwork". Cette dernière propose de choisir un ensemble n de paires de pixels (a_i, b_i) et d'augmenter pour chaque paire, la luminosité de a_i d'un pas prédéfini, en diminuant simultanément la luminosité de b_i . La modification des niveaux de gris a été largement exploitée lors des balbutiements du tatouage d'images. Gordon Braudaway présentait dans [18] une méthode simple basée sur la modification de l'intensité des pixels. La marque est constituée d'un tableau numérique de même dimensions (I, J) que l'image originale, la valeur de chaque élément de ce tableau est distribuée aléatoirement pour chaque position (i, j) des pixels, de sorte que : $1 \geq w(i, j) \geq (1 - 2\beta)$ avec $1 \leq i \leq I$; $1 \leq j \leq J$; $0 \leq \beta \leq 0.1$, où β est un paramètre de poids de la marque. Puis, la luminosité de l'image originale est simplement multipliée par celle du filigrane pour tout pixel. La robustesse du filigrane réside dans son imprévisibilité. l'aspect cryptographique est donc pris en compte par l'auteur qui propose l'utilisation de l'algorithme DES (Data Encryption Standard).

Les travaux de Bender et al. [16] sont reconnus dans le domaine du tatouage spatial. Certaines contraintes sont à envisager pour toute méthode de tatouage :

- Le signal hôte ne doit pas être perceptiblement dégradé.
- Les données doivent être insérées directement dans le média et non dans une en-tête.
- Les données insérées doivent être robustes face à différentes attaques, malveillantes ou naturelles.
- Un codage asymétrique des données insérées est souhaitable.
- Un codage correcteur d'erreurs doit être envisagé afin d'assurer l'intégrité des données.
- Les données insérées doivent être étendues sur toute l'image de façon à ce que l'examen d'un seul échantillon permette de les récupérer.

La méthode dite "Patchwork" développée par Bender et al. est basée sur un procédé statistique pseudo-aléatoire. Le principe est de sélectionner de façon pseudo-aléatoire, deux patches A et B sur une image, puis d'augmenter la luminosité du patch A tout en diminuant simultanément celle du patch B . Les auteurs ont étendu ce principe simple à différentes formes de patches, en permettant de concentrer l'énergie de la marque dans certaines bandes de fréquences prédéfinies. L'algorithme de détection nécessite de stocker une quantité importante d'information. Les positions des pixels tatoués doivent être connues afin de détecter la marque. De plus, la quantité d'information insérée dans l'image est relativement faible. Cette méthode a été étendue à une méthode de tatouage par blocs de textures, les patches précédents sont à présent remplacés par des zones de textures semblables. La détection dans ce dernier cas s'effectue par auto-corrélation de l'image tatouée.

Les techniques de tatouage opérant dans le domaine spatial comportent l'avantage de fournir une capacité de tatouage importante mais permettent aussi de garantir aisément l'invisibilité. Outre ces avantages, les

techniques de tatouage spatial montrent parfois des lacunes importantes en termes de robustesse de la marque vis-à-vis des attaques. Ces techniques de tatouages se montrent souvent peu robustes face aux attaques géométriques.

3.6.2 Méthodes fréquentielles

Comme nous venons de le voir, le principal inconvénient des méthodes de tatouage spatial est la faible robustesse face aux attaques et notamment face aux attaques géométriques.

Les techniques de tatouage spatial telles que celles que nous venons de découvrir peuvent tout aussi bien être adaptées aux domaines fréquentiels.

Les domaines transformés les plus fréquemment exploités pour les applications de tatouage d'image sont le domaine de Fourier et le domaine DCT (Transformation en Cosinus Discret). L'arrivée récente dans le monde de l'imagerie numérique, du codage JPEG 2000, méthode de codage exploitant le domaine des ondelettes (DWT : Discrete Wavelet Transform), va probablement permettre l'augmentation des études de tatouage dans le domaine des ondelettes, domaine actuellement peu exploité.

Ó Ruanaidh et al. ont présenté dans [102] une méthode de tatouage de la phase de la DFT. Ici, les auteurs ont opté pour la phase du spectre, car l'insertion d'une marque dans les composantes les plus importantes de l'image permet d'accroître la robustesse du schéma. Une attaque opérant dans la phase du spectre dégraderait rapidement la qualité de l'image. Une seconde raison justifiant le tatouage de la phase du spectre provient de la théorie de la communication : la modulation de la phase possède une meilleure immunité au bruit que la modulation d'amplitude.

Cependant, plusieurs techniques de tatouage exploitent la modulation d'amplitude de la DFT. L'invariance du spectre en translations ou décalages a motivé certains auteurs à tatouer l'amplitude du spectre de Fourier [54],[55],[89],[84].

Une méthode assez répandue de tatouage d'images, extensible aux signaux audio voire aux séquences vidéo est celle présentée par Cox et al. [28],[27]. Cette méthode, applicable dans le domaine DCT "full frame" ou DFT (Discrete Fourier Transform), consiste à insérer un bruit dans le contenu fréquentiel de l'image. Les techniques d'étalement spectral sont employées. La communication par étalement spectral [95] (Spread Spectrum Communications) a pour principe la transmission d'un signal à bande étroite sur une bande plus large, de façon à ce que l'énergie du signal, présente dans chaque fréquence soit indécélable. Le filigrane est alors "étalé" uniformément sur le spectre, de telle sorte que son amplitude soit faible et donc indétectable en tout point. Dans le cas de la méthode de Cox et al., la marque à insérer dans l'image représente une distribution

Gaussienne, pondérée par une force constante. Considérons un filigrane $X = x_1, \dots, x_n$ et les coefficients du spectre original $V = v_1, \dots, v_n$. Nommons les coefficients tatoués $V' = v'_1, \dots, v'_n$ et la force α constante pour tous les coefficients. Trois méthodes de pondération ont été étudiées :

- $v'_i = v_i + \alpha x_i$
- $v'_i = v_i(1 + \alpha x_i)$
- $v'_i = v_i(e^{\alpha x_i})$

Si l'on note X^* la marque extraite, la détection du filigrane s'effectue selon l'équation :

$$Sim(X, X^*) = \frac{X * .X}{\sqrt{X * .X^*}}. \quad (3.2)$$

La principale innovation apparue dans ces travaux est le choix de l'insertion de la marque non pas dans les composantes perceptuellement insignifiantes de l'image mais au contraire dans les composantes significatives.

Wu et Liu [127], proposent une méthode de tatouage par blocs DCT. L'insertion de la marque s'effectue ici directement dans les coefficients DCT, après quantification, par le biais d'une Look Up Table (LUT). Une LUT composée uniquement de '1' ou de '0' est créée de façon aléatoire. Chaque coefficient DCT (par bloc 8*8), est mis en correspondance avec cette table afin d'insérer la marque. Ce schéma de tatouage permet l'insertion de formes binaires dans les coefficients DCT mais comporte un inconvénient qui est la nécessité de stockage de la LUT lors de la procédure d'extraction. Cette méthode n'est pas applicable en l'état dans des blocs DCT correspondant à des régions uniformes de l'image, où la marque pourrait causer des dégradations perceptibles. Les auteurs ont alors mis en œuvre la technique dite du 'back-up embedding', c'est à dire, l'insertion ne s'effectue plus dans les zones de niveau de gris constant de l'image, mais dans des zones comportant une relation géométrique avec le bloc à tatouer.

Une autre méthode de tatouage de blocs DCT est proposée par Hsu et Wu dans [57] et [58]. Ici, le filigrane est une forme binaire visuellement reconnaissable. L'insertion s'effectue dans une bande moyenne fréquence pour chaque bloc DCT de taille 8×8 . Une image de taille réduite, composée de la concaténation des bandes moyennes fréquences de chaque bloc DCT est formée. Le nombre de coefficients à stocker dans cette imagerie doit être choisi de façon à obtenir une imagerie de même taille que le filigrane.

Piva et al. [97] proposent une méthode d'insertion dans les coefficients DCT. Ici, La DCT est calculée sur toute l'image, et non pas par blocs. Les coefficients sont ordonnés en zig-zag, puis les $L + M$ premiers coefficients sont sélectionnés. On insère la marque sur les coefficients allant de L à $L + M$. Afin de garantir l'invisibilité de la marque, une pondération prenant en compte les caractéristiques locales de l'image est alors appliquée. L'image originale et l'image tatouée sont additionnées pixel par pixel, en fonction d'un coefficient de

pondération. Une nouvelle image tatouée est alors obtenue. Le processus de détection s'effectue en extrayant les coefficients potentiellement tatoués (de L à $L + M$) puis en calculant la corrélation entre ces coefficients et un filigrane différent. La comparaison de cette corrélation avec un seuil prédéfini (test d'hypothèse) permet de savoir si la marque est présente ou non.

3.6.3 Méthodes utilisant des critères visuels

Certains schémas de tatouage opérant dans le domaine DCT ont utilisé des masques perceptuels, tel que Bartolini et al. dans [12], Kankanhalli et al. dans [60] ou Podilchuk et Zeng présentant une méthode psychovisuelle plus aboutie dans [99].

Dans [12], Bartolini et al. proposent trois masques perceptuels distincts, basés sur des considérations différentes : Un premier masque nommé M_s , basé sur le modèle théorique du S.V.H. proposé par Western et al. [125], un second masque M_σ basé sur le calcul de la variance locale et enfin, un troisième masque M_{bp} basé sur des considérations heuristiques, prenant en compte un filtrage de Sobel de l'image, des filtrages médian et un filtre passe bande.

Les coefficients DCT d'une image I sont modifiés par l'insertion du filigrane w_{ij} selon la loi suivante :

$$y_j = x_j + \alpha \times w_{ij}|x_j|, \quad (3.3)$$

x_j représentant les coefficients DCT originaux, y_j représentent les coefficients tatoués par la marque w_{ij} pondérée par un paramètre α . Ici, les coefficients tatoués sont substitués aux coefficients originaux, puis la DCT inverse fournit l'image tatouée I' . L'inconvénient principal de cette méthode réside dans le fait que toute l'étendue du domaine DCT est tatouée, aucune distinction n'est faite entre les zones uniformes ou les contours de l'image. Les auteurs ont donc opté pour la création d'une seconde image tatouée I'' , basée sur l'image originale I et sa version tatouée I' .

$$I'' = (1 - M)I + M \times I'. \quad (3.4)$$

Le masque M ayant des valeurs comprises entre 0 et 1, il donne une mesure de l'insensibilité au bruit pour chaque pixel.

Sachant que $W = I' - I$, l'équation 3.4 peut se réécrire de la façon suivante :

$$I'' = I + M \times W. \quad (3.5)$$

Afin d'accroître la robustesse de l'algorithme, les auteurs ont choisi de concentrer autant que possible le contenu spectral des images de masquage dans les faibles fréquences. En observant l'équation 3.5, nous constatons qu'en masquant le filigrane, son énergie dépend aussi de l'image de masquage et en particulier de la valeur de α dont la valeur moyenne est :

$$\bar{\alpha} = \alpha \times M. \quad (3.6)$$

Les auteurs ont créé trois images de masquage, un masque M_s basé sur des critères visuels, un second masque M_σ basé sur les considérations suivantes :

Les régions de forte activité sont insensibles aux perturbations, les contours sont plus sensibles aux distortions que les régions texturées, et enfin, les régions les plus sombres ou les plus claires sont moins sensibles au bruit.

Le masque M_σ est construit en calculant pour chaque point de l'image, la variance locale à travers une fenêtre W de taille 9×9 .

$$M_\sigma(x, y) = \frac{1}{M_{max}} \sum_{(i, j) \in W} [I(i, j) - \mu_W(i, j)]^2, \quad (3.7)$$

où $\mu_W(i, j)$ est la moyenne locale calculée dans la fenêtre W centrée sur le pixel (x, y) et $M_{max} = max_{m, n} M_\sigma(m, n)$ est la valeur maximale de la variance locale sur toute l'image.

La principale faiblesse de ce masque tient du fait qu'il ne différencie pas les régions texturées des contours.

Un second masque basé sur des considération heuristiques M_{bp} a été créé. Ce dernier est basé sur l'hypothèse que le filigrane sera moins perceptible dans les zones de l'image où les fréquences du filigrane sont présentes. Différents filtres sont utilisés lors de la création de ce masque. Un filtrage passe-bande est utilisé pour extraire la gamme de fréquences du filigrane, un filtrage de Sobel est utilisé dans le but de mettre en évidence les contours de l'image, et enfin deux seuillages ayant pour objectif l'extraction des zones les plus claires et les plus sombres de l'image. Les quatre images obtenues sont alors ajoutées afin de n'obtenir qu'un seul et unique masque perceptuel.

Les auteurs confirment dans cet article l'apport du masquage en termes de robustesse de la marque. Cependant, selon eux, le masque basé sur une modélisation du système visuel humain n'offre pas de meilleurs résultats que les masques basés sur des considérations heuristiques. Ils en déduisent que les modèles théoriques effectuent une modélisation correcte du comportement de l'oeil lors du processus de vision humaine, mais restent peu performants dans la modélisation de l'intégration de l'information effectuée par le cerveau.

Kankanhalli et al. proposent dans [60] un processus de création de masque J.N.D. basé sur la transformation en cosinus discret calculée par blocs de taille 8×8 . Les auteurs se basent sur les hypothèses suivantes : Les contours contribuant principalement à la perception de l'image, ils possèdent la plus faible sensibilité au bruit. Les zones uniformes sont très sensibles aux variations. Les régions texturées ont un très fort niveau de sensibilité au bruit. Enfin, les régions de niveau de gris moyen possède un faible J.N.D. comparé aux autres zones.

Pour chaque bloc DCT, un processus de classification des textures est mis en œuvre, les blocs comportant la plus forte activité subissent alors une séparation des contours et des textures.

Appelons $X_{i,j}$ le coefficient DCT positionné au pixel (i, j) , l'énergie des coefficients AC comportant les variations de fréquences de l'image (contours, textures) se présente sous la forme suivante :

$$E_{AC} = \log \sum_{i,j} (X_{ij})^2 - (X_{00})^2, \quad (3.8)$$

où X_{00} représente le coefficient DC du bloc DCT. Si G est la valeur maximale des niveaux de gris, l'énergie maximale du bloc est donnée par :

$$M_{max} = \log \left(\frac{G}{2} \right)^2. \quad (3.9)$$

Une normalisation permet l'obtention d'un "degré de texture" pour le bloc b :

$$R_b = \frac{E_{AC}}{E_{max}}. \quad (3.10)$$

Un indice de distorsion $I_b \in 1..8$ (intervalle choisi de façon empirique) est alloué pour chaque bloc en fonction de la valeur de R_b . Une forte valeur de I_b implique une faible tolérance aux distorsions.

Afin de déterminer si un bloc comporte des contours ou des textures, la mesure des fluctuations pour tout point x est donnée par le gradient ∇_x . L'amplitude des fluctuations est comparée à un seuil prédéfini λ .

Un bloc comportant une forte valeur de R_b comporte un contour si :

$$Card\{x|x \text{ est un pixel et } ampl(\nabla_x) > \lambda\} < C, \quad (3.11)$$

La valeur du JND est allouée pour chaque bloc b de la façon suivante :

$$\hat{J}(b) = K \times I_b, \quad (3.12)$$

où $K \times I_b$ varie entre 3 et 12 pour I_b variant de 1 à 8, $\hat{J}(b)$ est la valeur JND allouée pour tous les pixels

du bloc b . Ces valeurs ont été définies de façon empirique.

Une dernière correction portant sur la sensibilité à la luminance est apportée au calcul des valeurs JND. Pour chaque pixel (i, j) du bloc b :

$$J(i, j, b) = \hat{J}(b) + L(i, j, b), \quad (3.13)$$

avec

$$L(i, j, b) = \left[(128 - I(i, j, b))^2 / \alpha \right]. \quad (3.14)$$

α représentant une constante prédéfinie et $I(i, j, b)$ la valeur en niveaux de gris des pixels de l'image.

Un filigrane binaire subit plusieurs traitements (étalement et modulation par une séquence pseudo-aléatoire) avant d'être amplifié par les valeurs du JND. La forme binaire est donc étalée puis dupliquée de façon à couvrir toute l'image. Le signal modulé $w_i = J(i, j, b).b_i.p_i$ est enfin ajouté au canal de luminance de l'image I_i fournissant l'image tatouée :

$$\hat{I}_i = I_i + J(i, j, b).b_i.p_i \quad (3.15)$$

où b_i représentent les bits "étalés" du filigrane et p_i la séquence pseudo-aléatoire.

Dans un tel schéma, la détection du filigrane s'effectue par une sélection des formes binaires dupliquées dans l'image puis par le calcul de la corrélation entre l'image tatouée et ces marques modulées et étalées.

Cette méthode exploite surtout des critères empiriques pour modéliser la perception humaine. Un masque perceptuel est établi dans le domaine DCT par blocs 8×8 . Certaines études sont nécessaires sur les fondements de la méthode détaillée ici : nous conduirons une telle étude dans la section 4.3.3.

Dans [33], Delaigle et al. présentent une méthode de tatouage utilisant une modélisation du S.V.H. Le contraste relatif à une gamme de fréquence et à une orientation donnée est obtenu par le calcul de l'énergie locale, extraite par un filtre de Gabor. Les caractéristiques du filtre de Gabor sont adaptées à chaque composante perceptuelle. L'effet de ce filtrage est assimilable au filtrage réalisé par le cortex visuel.

Le critère de masquage est défini de la façon suivante :

un bruit est masqué pour tout pixel (x, y) si

$$\forall(f_0, \theta_0), E_{mask.(f_0, \theta_0)}(x, y) \geq E_{noise.(f_0, \theta_0)}(x, y), \quad (3.16)$$

f représentant les fréquences, θ les orientations, E représente l'énergie extraite par les filtres de Gabor.

Les auteurs supposent une décomposition en canaux visuels comportant cinq fréquences radiales et quatre orientations.

Les filtres de Gabor appliqués sur le spectre de Fourier sont définis par l'équation 3.17

$$G_{(f_0, \theta_0)}(f, \theta) = \exp \left[- \left(\frac{\log^2(f/f_0)}{F^2(f_0)} + \frac{(\theta - \theta_0)^2}{\Theta^2(f_0)} \right) \right] \quad (3.17)$$

Les auteurs ont choisi de limiter le filigrane aux composantes perceptuelles horizontales, seuls les filtres horizontaux ($\theta_0 = 0$) seront alors utilisés.

Le filigrane est composé ici d'une séquence pseudo-aléatoire. Cette dernière subit alors un nombre suffisant de décalages afin de couvrir toute l'image. Chaque séquence est composée de 31 symboles, chaque symbole est inséré dans un rectangle de dimensions 8×32 pixels de l'image. Une image de dimensions 512×512 comporte donc 33 séquences. Une séquence est modulée pour une fréquence secrète positionnée dans une bande de fréquences horizontale et moyenne fréquence.

Pour satisfaire le critère perceptuel défini précédemment, l'amplitude de la séquence doit être inférieure à l'énergie locale de la sous-bande considérée.

Un des inconvénients issus de cette méthode de tatouage est l'insuffisance du masquage près des contours de l'image. Le filigrane est donc plus facilement perceptible autour des contours alors que les zones texturées permettent l'insertion d'une marque d'amplitude plus élevée. Les auteurs utilisent une détection de contours, suivie d'un filtrage morphologique afin de permettre l'insertion de la marque dans les zones texturées tout en évitant de marquer les contours isolés.

La détection de la marque s'effectue par le calcul de la corrélation entre les séquences extraites et les séquences originales décalées. Le choix du coefficient de plus forte amplitude permet de définir la présence du filigrane. La robustesse de ce schéma a été testée avec succès face à diverses attaques telles que la résistance à l'addition de bruit, le codage JPEG ou encore le filtrage passe-bas. Cependant, il semble qu'un tel procédé n'offre pas une forte résistance aux attaques géométriques, car ce schéma n'envisage pas la re-synchronisation du filigrane.

Podilchuk et Zeng ont présenté dans [99] une technique de tatouage basée sur des modèles visuels développés dans le contexte de la compression d'images. Ce schéma est une amélioration basée sur les critères psychovisuels de la méthode proposée par Cox et al. détaillée précédemment.

Ici, les auteurs proposent un processus de tatouage applicable aux domaines DCT et DWT. A la différence des travaux de Swanson et al. [116] où la pondération fréquentielle était identique pour tous les blocs de

la DCT, Podilchuk et Zeng présentent une méthode de tatouage dépendante des caractéristiques locales de l'image.

Tout comme dans les travaux de Cox, les auteurs pensent que la marque doit être insérée dans les composantes "perceptuellement significatives" de l'image. Si le filigrane était inséré dans les composantes inutiles de l'image, la plupart des schémas de codage suffiraient à le supprimer.

Un filigrane w_i de longueur n_I dépendante de l'image, est inséré dans l'image tout en respectant la contrainte :

$$\frac{1}{n_I} \sum_{i=0}^{n_I-1} (a_{I,i} w_i)^2 < P_I, \quad (3.18)$$

où P_I représente la puissance maximale d'un filigrane imperceptible. Le poids maximal permis, noté $a_{I,i}$, ainsi que la longueur n_I permettent de déterminer un seuil JND local. Ces travaux se sont basés sur des modèles perceptuels exploités pour le domaine de la compression. Les trois principales caractéristiques du système visuel humain exploitées ici sont : la sensibilité fréquentielle, la sensibilité aux luminances et l'effet de masquage.

Les deux techniques de tatouage présentées ici peuvent s'écrire :

$$X_{u,v}^* = \begin{cases} X_{u,v} + J_{u,v} w_{u,v} \text{ si } X_{u,v} > J_{u,v} \\ X_{u,v} \text{ sinon,} \end{cases} \quad (3.19)$$

où $X_{u,v}$ représentent les coefficients fréquentiels de l'image originale $(x_{i,j})$. $X_{u,v}^*$ représentent les coefficients de l'image tatouée, $w_{u,v}$ est le filigrane formé d'une distribution Gaussienne de valeur moyenne nulle et de variance unitaire et enfin, $J_{u,v}$ représente la masque perceptuel défini pour chaque coefficient.

Une légère modification de l'équation 3.19 permet l'insertion du filigrane dans les coefficients perceptuellement significatifs :

$$X_{u,v}^* = \begin{cases} X_{u,v} + J_{u,v} w_{u,v} \text{ si } X_{u,v} > J_{u,v} \text{ et } \frac{J_{u,v}}{X_{u,v}} < T_{JND} \\ X_{u,v} \text{ sinon,} \end{cases} \quad (3.20)$$

où T_{JND} est un seuil défini de façon empirique.

Les travaux de Peterson [90] proposent des matrices de quantification de la luminance ainsi que des composantes chromatiques, en fonction des conditions de visualisation. Ce modèle prend en compte la sensibilité aux fréquences, mais ne prend pas en compte la sensibilité aux luminances ni les effets de masquage. Les auteurs ont exploité les seuils de quantification définis par Watson dans [124] prenant en considération la luminance

locale, ainsi que le masquage. Ces seuils sont adaptés par Podilchuk et Zeng selon les travaux de Safranek [108], qui permettant l'obtention de seuils perceptuels (JND), par un contrôle des quantificateurs locaux.

L'insertion peut se produire directement dans un flux JPEG par le biais d'une décompression partielle.

La fonction de sensibilité à la luminance utilisée par Watson dans [124] est donnée par la formule suivante :

$$t_{u,v,b}^L = t_{u,v}^F \left(\frac{X_{0,0,b}}{\bar{X}_{0,0}} \right)^a, \quad (3.21)$$

où $X_{0,0,b}$ est le coefficient DC de la DCT pour le bloc b , $\bar{X}_{0,0}$ est le coefficient DC correspondant à la luminance moyenne et a est un paramètre contrôlant la sensibilité à la luminance. La sensibilité fréquentielle du modèle est notée $t_{u,v}^F$.

Le seuil de masquage du contraste est défini par :

$$t_{u,v,b}^C = \text{Max} \left[t_{u,v,b}^C, |X_{u,v,b}|^{w_{u,v}} (t_{u,v,b}^L)^{1-w_{u,v}} \right], \quad (3.22)$$

$w_{u,v}$ est un nombre compris entre 0 et 1, fixé ici de façon empirique à 0,7. Ces seuils de masquage sont utilisés pour déterminer la localisation ainsi que la force maximale du filigrane. Ce dernier est composé d'une séquence de nombres réels suivant une distribution gaussienne de moyenne nulle et de variance unitaire.

L'insertion du filigrane s'effectue de la façon suivante :

$$X_{u,v,b}^* = \begin{cases} X_{u,v,b} + t_{u,v,b}^C w_{u,v,b} \text{ si } X_{u,v,b} > t_{u,v,b}^C \\ X_{u,v,b} \text{ sinon} \end{cases} \quad (3.23)$$

où $X_{u,v,b}$ représentent les coefficients DCT du bloc b , $X_{u,v,b}^*$ représentent les coefficients DCT tatoués. $w_{u,v,b}$ est la séquence du filigrane et enfin, $t_{u,v,b}^C$ représentent les seuils JND détaillés précédemment.

Concernant le modèle visuel adapté au tatouage des coefficients ondelettes, la procédure d'insertion est identique, ainsi que le filigrane inséré, cependant, les auteurs se sont basés sur les travaux de Watson dans [121] pour la création du masque perceptuel.

Le principal point faible de ce schéma est la nécessité de conserver l'image originale lors du processus de détection. Le schéma de détection consiste à calculer la corrélation entre la marque extraite (différence entre l'image originale et l'image reçue) et la marque originale, le résultat de cette corrélation est alors comparé à un seuil prédéfini, afin de confirmer la présence du filigrane.

Le schéma de détection par corrélation normalisée peut s'exprimer

$$w_{s,u,v,b}^* = X_{u,v,b} - \hat{X}_{u,v,b}^* \quad (3.24)$$

$$w_{u,v,b}^* = \frac{w_{s,u,v,b}^*}{t_{u,v,b}^C} \quad (3.25)$$

$$\rho_{ww^*} = \frac{w^* \cdot w}{\sqrt{E_w \cdot E_w}} \quad (3.26)$$

$w_{s,u,v,b}^*$ représente le filigrane reçu modulé par le masque JND et potentiellement dégradé. $w_{u,v,b}^*$ représente le filigrane reçu et ρ_{ww^*} est le coefficient de corrélation normalisé entre les deux signaux w et w^* .

La dernière étape du processus de détection est la comparaison de la corrélation avec un seuil prédéfini.

$$\begin{aligned} \rho_{ww^*} > T_\rho & \quad \text{Filigrane } w \text{ détecté} \\ \rho_{ww^*} \leq T_\rho & \quad \text{Filigrane } w \text{ non détecté} \end{aligned} \quad (3.27)$$

Cette méthode de tatouage est très intéressante compte tenu des aspects perceptuels mis en œuvre. Elle comporte cependant quelques défauts, comme la nécessité de conserver l'image originale lors du processus de détection, ou encore quelques lacunes majeures en termes de robustesse. Par exemple, la suppression de lignes ou de colonnes, ou encore une rotation, modifieraient totalement les masques JND par blocs DCT et ne permettraient pas de synchronisation de la marque pour la détection du filigrane.

3.6.4 Mesures de distorsions

L'évaluation de la robustesse d'un schéma de tatouage ne peut se faire que par la mise en œuvre de tests de détection ou d'extraction vis-à-vis du plus grand nombre possible d'attaques. Nous avons vu précédemment (section 3.4) une liste exhaustive des attaques répertoriées dans Stirmark ou Checkmark.

3.6.4.1 Calculs d'erreurs

Concernant l'invisibilité, son évaluation peut passer par de nombreuses métriques de distorsion, Kutter et Petitcolas en présentent un grand nombre dans [68]. Ces dernières sont représentées dans le tableau 3.3 :

| | |
|---|---|
| métriques de différences | |
| Différence maximale | $MD = \max_{m,n} I_{m,n} - \tilde{I}_{m,n} $ |
| Différence absolue moyenne | $AD = \frac{1}{MN} \sum_{m,n} I_{m,n} - \tilde{I}_{m,n} $ |
| Différence absolue moyenne Normalisée | $NAD = \sum_{m,n} I_{m,n} - \tilde{I}_{m,n} / \sum_{m,n} I_{m,n} $ |
| Erreur quadratique moyenne | $MSE = \frac{1}{MN} \sum_{m,n} (I_{m,n} - \tilde{I}_{m,n})^2$ |
| Erreur quadratique moyenne Normalisée | $NMSE = \sum_{m,n} (I_{m,n} - \tilde{I}_{m,n})^2 / \sum_{m,n} I_{m,n}^2$ |
| Norme L^P | $L^p = \left(\frac{1}{MN} I_{m,n} - \tilde{I}_{m,n} ^p \right)$ |
| Laplacien de l'erreur quadratique moyenne | $LMSE = \sum_{m,n} (\nabla^2 I_{m,n} - \nabla^2 \tilde{I}_{m,n})^2 / \sum_{m,n} (\nabla^2 I_{m,n})^2$ |
| Rapport signal à bruit | $SNR = \sum_{m,n} I_{m,n}^2 / \sum_{m,n} (I_{m,n} - \tilde{I}_{m,n})^2$ |
| Rapport signal à bruit maximal | $PSNR = MN \max_{m,n} I_{m,n}^2 / \sum_{m,n} (I_{m,n} - \tilde{I}_{m,n})^2$ |
| Fidélité | $IF = 1 - \sum_{m,n} (I_{m,n} - \tilde{I}_{m,n})^2 / \sum_{m,n} I_{m,n}^2$ |
| métriques de corrélation | |
| Intercorrélation Normalisée | $NC = \sum_{m,n} I_{m,n} \tilde{I}_{m,n} / \sum_{m,n} I_{m,n}^2$ |
| Corrélation Qualité | $CQ = \sum_{m,n} I_{m,n} \tilde{I}_{m,n} / \sum_{m,n} I_{m,n}$ |
| autres | |
| Contenu structurel | $SC = \sum_{m,n} I_{m,n}^2 / \sum_{m,n} \tilde{I}_{m,n}^2$ |
| Rapport signal à bruit (global Sigma) | $GSSNR = \sum_b \sigma_b^2 / \sum_b (\sigma_b - \tilde{\sigma}_b)^2$ où $\sigma_b = \sqrt{\frac{1}{P} \sum_{blockb} I_{m,n}^2 - \left(\frac{1}{P} \sum_{blockb} I_{m,n} \right)^2}$ |
| Rapport signal à bruit (Sigma) | $SSNR = \frac{1}{P} \sum_b SSNR_b$ où $SSNR_b = 10 \log_{10} \frac{\sigma_b^2}{(\sigma_b - \tilde{\sigma}_b)^2}$ |
| Ratio sigma / erreur | $SE R_b = \frac{\sigma_b^2}{\frac{1}{P} \sum_{blockb} (I_{m,n} - \tilde{I}_{m,n})^2}$ |
| similarités d'histogrammes | $HS = \sum_{c=0}^{255} f_I(c) - f_{\tilde{I}}(c) $ où $f_I(c)$ est la fréquence relative au NG c |

TAB. 3.3 – Comportement sur le spectre des attaques

Dans ce tableau, $I(m, n)$ représente le pixel (m, n) de l'image I et $\tilde{I}(m, n)$ représente l'image dégradée (tatouée).

Les métriques de différences, appelées aussi mesures de distances, sont globalement assez simples, les plus fréquemment utilisées pour des applications de tatouage sont l'erreur quadratique moyenne ou encore le rapport signal à bruit.

Comme nous le constatons sur la figure 3.4 ci-dessous, pour deux images dont le rapport signal à bruit (PSNR) est similaire, les dégradations occasionnées sont très différentes.



(a) image 'bateau' ayant subi un filtrage passe bas de noyau 3×3



(b) image 'bateau' codée en JPEG $q=10\%$

FIG. 3.4 – dégradations fournissant un PSNR similaire

L'image 3.4(a) a subi un filtrage passe bas de noyau 3×3 et fournit un PSNR de 33.75, alors que l'image 3.4(b), qui a subi un codage JPEG de facteur de qualité $q = 10\%$ présente des dégradations perceptuellement plus gênantes et obtient un PSNR de 32.76. Le PSNR, bien que fréquemment utilisé ne semble donc pas être approprié pour juger de façon fiable la qualité des images tatouées.

Les métriques de corrélation sont légèrement plus complexes : elles permettent de calculer la corrélation entre les deux images. De telles métriques ne mesurent plus la différence entre deux images, mais plutôt la ressemblance des images.

Si ces métriques représentent grossièrement une mesure de dégradation des images elles ne sont cependant pas adaptées au système visuel humain. Plusieurs études, basées sur le principe du "*Visible Difference Predictor*" de Daly [31], effectuées au sein de l'équipe IVC, par N. Bekkat et P. Le Callet [15], [22], sont beaucoup plus approfondies et prennent en compte le modèle du S.V.H. détaillé précédemment (cf. section 3.6.4.3).

3.6.4.2 Méthodes subjectives

La méthode la plus fiable d'évaluation de la qualité est la mise en place de tests psychovisuels de mesure subjective de la qualité.

Ces mesures subjectives sont établies selon un protocole expérimental précis [24]. Il existe deux méthodes principales d'évaluation subjective de la qualité : les méthodes différentielles et les méthodes sans référence. La méthode qui semble être la plus stable est appelée méthode à double stimulus. Ici, l'ordre de présentation des

images est le suivant : l'image originale est présentée, suivie d'une version dégradée. Cette séquence se répète avec différentes dégradations. Une représentation de cette séquence est donnée figure 3.5.

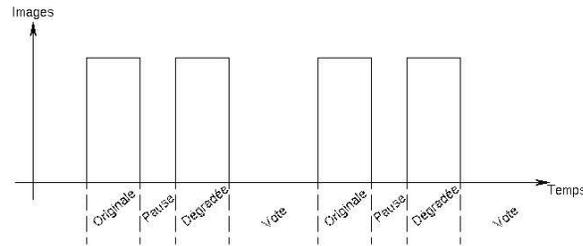


FIG. 3.5 – Séquence de la méthode à double stimulus

L'observateur doit juger la qualité de chaque image dégradée, comparativement à la qualité de l'image originale, qui revient de façon récurrente. Pour l'évaluation, l'observateur doit se baser sur l'échelle suivante :

| Note | Qualité | Dégradations |
|------|-------------|--------------------------------|
| 5 | excellente | imperceptibles |
| 4 | bonne | perceptibles mais non-gênantes |
| 3 | assez bonne | légèrement gênantes |
| 2 | médiocre | gênantes |
| 1 | mauvaise | très gênantes |

TAB. 3.4 – Échelle d'évaluation de la méthode à double stimulus

Une seconde méthode d'évaluation de qualité fréquemment utilisée est la méthode comparative. L'observateur doit noter la qualité d'une image comparativement à la précédente, qui n'est plus obligatoirement l'image originale. Cette séquence peut être envisagée avec ou sans image de référence. Un schéma illustrant cette méthode est présenté figure 3.6.

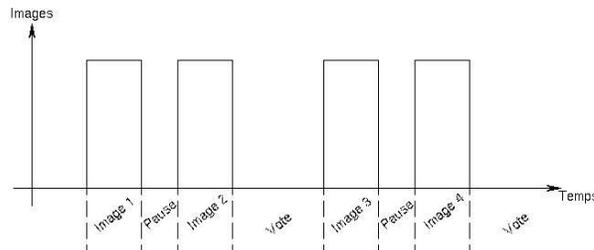


FIG. 3.6 – Séquence de la méthode comparative

L'échelle d'évaluation est donc adaptée à la séquence :

| Note | Evaluation |
|------|------------------------|
| -3 | beaucoup moins bonne |
| -2 | moins bonne |
| -1 | légèrement moins bonne |
| 0 | identique |
| 1 | légèrement mieux |
| 2 | mieux |
| 3 | beaucoup mieux |

TAB. 3.5 – Échelle d'évaluation de la méthode comparative

3.6.4.3 Méthodes objectives

Nous avons vu en introduction, section 1.1 que certaines modélisations du SVH sont exploitées dans le but d'évaluer la qualité des images. Nous venons de voir que certaines normes de visualisation ont été mises en oeuvre afin de soumettre des séquences d'images dégradées à un ensemble d'observateurs, en vue d'obtenir une note moyenne de qualité (Recommandations CCIR [24]).

De nombreuses études ont été réalisées dans le but d'automatiser ces évaluations fastidieuses et contraignantes. Comme nous l'avons vu au chapitre 1, il existe certains schéma, tels que le VDP de Scott Daly [31], susceptibles de produire une carte de probabilité de détection des erreurs par le SVH. Il est alors envisageable à partir de cette carte de probabilités, d'opérer un cumul des erreurs afin de n'obtenir qu'une note d'évaluation objective de la qualité.

Nachida Bekkat [15] et Patrick Le Callet [22] ont basé leurs travaux sur la modélisation du SVH présentée section 1.3.8 et ont étudié différentes pondérations possibles des cartes d'erreur obtenues pour chaque sous-bande. Un pooling fréquentiel des 17 sous-bandes de la DCP permet l'obtention d'une carte d'erreurs et enfin, un pooling spatial de cette carte permet de ne retenir qu'une note représentant la qualité de l'image dégradée comparativement à l'image de référence. Cette note objective doit représenter la meilleure corrélation possible avec la note moyenne de qualité subjective fournie par un ensemble d'observateurs.

Chacune de ces trois méthodes d'évaluation de la qualité peuvent être utilisées pour valider l'invisibilité des schémas de tatouages. La première méthode, l'utilisation de métriques empiriques de mesures d'erreurs, bien que faciles à mettre en oeuvre, ces métriques ne sont pas toujours très fiables.

La méthode de mesure subjective de qualité doit respecter un protocole expérimental précis et rigoureux, de nombreuses conditions assez contraignantes doivent être respectées. Cependant, si un nombre suffisamment important d'observateurs est disponible, les résultats sont très fiables.

La troisième méthode, de mesure objective semble être un bon compromis entre la simplicité de mise en œuvre de la première et la complexité de la seconde.

Toutefois, il est important de noter que l'utilisation préalable d'un modèle psychovisuel lors du processus d'insertion garantit l'invisibilité et permet d'éviter la mise en place d'évaluation *a posteriori* de la qualité.

3.7 Les techniques d'extraction et de détection

Les techniques de détection les plus fréquemment employées sont le coefficient de corrélation, la fonction d'intercorrélation, ou comme nous le verrons par la suite, la fonction d'intercorrélation pondérée. La fonction de similarités exploitée dans [27] et [28] est aussi souvent utilisée. La distance de Hamming peut aussi être envisagée en substitution de la corrélation dans des schémas d'insertion de séquences binaires.

De nombreuses méthodes nécessitent la conservation de l'image originale voire de l'image tatouée lors du processus d'extraction ou de détection. Si cette option permet en général de meilleurs résultats, elle n'est tout de même pas recommandée du fait de l'importance de la quantité d'information à stocker et n'évite en rien les attaques dites 'de procédure' comme l'attaque 'IBM'.

Pour certaines méthodes de tatouage, il peut être nécessaire d'effectuer une sélection préalable des pixels ou des coefficients potentiellement tatoués, avant de procéder aux techniques de détection citées ci-dessus. Par exemple, lors d'un processus de tatouage utilisant un masque psychovisuel [60] [8], les sites tatoués devront être sélectionnés avant de subir la technique adéquate de détection.

Nous proposerons dans le chapitre 5, une technique assez innovante de détection combinant la notion de projections Mojette et la fonction d'intercorrélation normalisée.

Quant aux techniques d'extraction, elles sont toujours dépendantes de la méthode d'insertion choisie. L'extraction du filigrane peut être suivie d'un processus de détection afin de confirmer la présence de la marque dans les données extraites.

Parallèlement à ces méthodes de détection ou d'extraction de filigranes, certaines études portent sur la différence entre les schémas de tatouage dit public ou privé. Craver présente dans [30] un algorithme de tatouage, ainsi qu'un processus dit "à divulgation nulle" [112] permettant de prouver la connaissance d'une information sans la révéler. L'intérêt ici est de créer un schéma de tatouage dans lequel la présence du filigrane peut être vérifiée sans révéler suffisamment d'information pour permettre la suppression de la marque.

Supposons un transfert de données mettant en jeu quatre personnages : Alice, une cliente de banque, Bob représentant la banque et gérant une base de données des clients, Carl, un employé de banque ayant besoin d'informations pour fournir le service souhaité et enfin Eve, observant les transactions et souhaitant pirater

une transaction. De nombreuses fraudes sont possibles entre ces personnages. Craver utilise une méthode de tatouage proposée par Pitas dans [96], puis étudie deux méthodes de transmission à divulgation nulle. La première méthode exploitée utilise les graphes isomorphes (changement de représentation entre graphes), permettant à Alice de révéler à Bob l'isomorphisme entre deux graphes. La seconde méthode, appelée "logarithme discret" permettant à Alice de dévoiler un logarithme à Bob afin de prouver son identité. L'intérêt principal de cette étude est de permettre à une personne de s'identifier auprès d'une autorité sans pour autant dévoiler suffisamment d'informations permettant à une tierce personne de pirater ces informations.

3.8 Tatouage et cryptographie : liens et différences

Si les domaines du tatouage et de la cryptographie sont très liés [3] [36] [103], il existe cependant une différence importante qui réside dans le fait que des données cryptées doivent être illisibles pour une personne non autorisée. Dans ce cas, les données en question sont le média numérique lui-même (l'image, la séquence vidéo ou audio ...), alors que dans un contexte de tatouage, les données tatouées doivent paraître originales mais l'information "camouflée" dans ces données (le filigrane) doit être illisible pour toute personne non autorisée.

Pour résumer, dans un schéma de cryptographie, l'image est rendue illisible et irrécupérable pour ceux ne détenant pas la clé adéquate, alors que pour le tatouage, l'image doit être de meilleure qualité possible, mais les données insérées doivent être irrécupérables.

Le domaine de la stéganographie est apparenté à celui de la cryptographie, les deux disciplines permettent d'échanger des messages avec un correspondant sans que d'autres personnes ne puissent en prendre connaissance. En cryptographie habituelle, la sécurité repose sur le fait que le message ne sera sans doute pas compris. Avec la stéganographie, la sécurité est double, d'une part, l'information à transmettre est cryptée, d'autre part, ce message ne sera probablement pas détecté, car il est noyé dans un support anodin et de dimensions nettement supérieures au message.

Le principe de base du tatouage est d'insérer dans un média numérique, une quantité d'information nécessaire à identifier le propriétaire ou les ayants droit.

Le schéma de tatouage présenté section 5.3 montre bien les liens étroits qui existent entre tatouage et cryptographie.

Les fonctions de hachage sont largement utilisées dans les schémas de sécurité. Ces dernières ont pour but de vérifier l'authenticité des données. Effectivement, une fonction de hachage représente une concaténation du document, si ce concentré du support a été modifié, cela signifie que les données reçues ont subi une malversation.

Ces empreintes, issues du domaine de la cryptographie, sont largement utilisées dans les schémas de tatouage, notamment par Fridrich et al. dans [38] [39] [40] [41] [42].

Certaines méthodes de tatouage insèrent directement dans l'image sa propre version cryptée [100]. Ici, l'insertion d'un bit s'effectue spatialement dans les plans binaires, par blocs 8×8 de l'image. L'insertion d'une valeur de la marque se fait par augmentation ou diminution des 64 pixels du bloc, de façon à obtenir pour la valeur moyenne du bloc, la valeur souhaitée d'insertion.

3.9 Conclusion

Nous avons présenté dans ce chapitre, certaines méthodes existantes de tatouage d'images, il existe à ce jour une très grande variété de schémas de tatouage.

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. Effectivement, la plupart des attaques, qu'elles soient intentionnelles ou naturelles, peuvent être assez dévastatrices. Pour s'adapter à la majorité des attaques, il est nécessaire d'étudier les différents domaines proposés par le traitement du signal.

Concernant les aspects psychovisuels, certains auteurs estiment qu'une sélection empirique des sites peut parfois être suffisante, du fait de leur complexité. Cependant ils peuvent être étudiés de façon approfondie dans le but d'optimiser une méthode de tatouage comme nous le ferons par la suite .

Nous avons présenté dans cette première partie deux des principaux axes de recherche développés au sein de l'équipe IVC : L'étude du système visuel humain et la transformation Mojette.

L'étude de ces deux aspects a été effectuée dans l'optique d'apporter des solutions aux principales exigences des techniques de tatouages détaillées dans le troisième chapitre.

Comme nous l'avons vu, l'étude des bases anatomiques et physiologiques du système visuel humain a permis la création de différents modèles perceptuels. Nous avons étudié les principaux modèles de la littérature, puis le modèle psychovisuel établi au sein de l'équipe IVC a été détaillé. Certains aspects semblent parfaitement adaptés à la problématique du tatouage des images. Parmi ces aspects, nous pouvons trouver les notions de seuil de détection, de contraste local à bande limitée ou encore de fonction de sensibilité au contraste. Ces notions, ainsi que l'effet de masquage ont été prises en compte lors de l'élaboration du modèle psychovisuel.

Dans le but de retenir certains aspects exploitables lors de la création d'un schéma de tatouage, nous avons étudié les principales propriétés de la transformation Mojette. Ces dernières sont d'une part le partage de l'information en différentes projections, mais aussi la redondance introduite ou encore la notion de fantôme.

Les différentes applications du tatouage des images ont été abordées, parmi lesquelles nous pouvons citer la stéganographie, la protection des droits d'auteurs, le tatouage fragile ou encore l'authentification. Nous avons pu constater la diversité des exigences requises en fonction de ces applications. Concernant les applications de protection des droits d'auteurs, une étude des attaques envisageables a été conduite, dans le but de fournir un algorithme de tatouage résistant au maximum de malversations possibles. La problématique du choix des espaces de tatouage a été abordée, permettant ainsi de définir les principaux avantages et inconvénients pour chaque espace transformé. Une telle étude nous permet d'affirmer que les méthodes de tatouage spatial souffrent d'un certain manque de robustesse, mais permettent toutefois l'insertion d'une très grande quantité d'informations tout maintenant la marque l'invisibilité. Les méthodes fréquentielles semblent présenter quant à elles de meilleurs résultats en terme de robustesse mais sont beaucoup plus complexes à mettre en œuvre si l'on souhaite prendre en compte des critères perceptuels. Une étude détaillée de l'élaboration de masques perceptuels opérant dans le domaine de Fourier ou dans le domaine DCT a été établie. Certains aspects connexes au tatouage des images tel que la mesure des distorsions introduites par la marque, ou encore la

notion de clef ont été abordés. Dans un dernier temps, nous avons présenté de façon succincte la cryptographie, ainsi que les fonctions de hachage, outils mis en œuvre par la suite.

Deuxième partie

Le tatouage d'images

Nous allons à présent procéder d'une part à l'exploitation des critères visuels et d'autre part à l'exploitation de la transformation Mojette dans divers schémas complets de tatouage. L'optimisation conjointe de ces deux aspects (psychovisuel et Mojette), passera par une étude séparée. Ces travaux seront donc décomposés en trois étapes distinctes.

Tout d'abord, différentes techniques de tatouage mettant à profit l'étude du SVH seront mises en œuvre (chapitre 4). Nous étudions d'une part les différentes étapes nécessaires à l'élaboration d'un masque perceptuel ainsi que sa possible adaptation dans différents domaines transformés. Le comportement du masque lors de son adaptation dans ces différents domaines pourra nous amener à y apporter certaines modifications, en vue d'améliorer sa robustesse. La validation du masque obtenu sera ensuite testée dans un schéma complet de tatouage spatial. D'autre part une technique de tatouage opérant directement l'insertion de l'information de droits d'auteurs dans les sous-bandes fournies par la décomposition en canaux perceptuels sera détaillée.

Ensuite, les principales propriétés de la transformation Mojette seront exploitées lors de la réalisation de plusieurs schémas de tatouage (chapitre 5). Nous essaierons d'exploiter la structure discrète de l'opérateur Mojette pour optimiser la robustesse de la marque insérée. Une telle structure discrète se prête forcément à des explosions combinatoires susceptibles d'être exploitées lors d'une utilisation en cryptographie ou en stéganographie. Nous proposerons tout d'abord une méthode de protection de droits d'auteurs adaptée à l'imagerie médicale, incluant la vérification de l'intégrité. Cette méthode doit prendre en compte les différentes exigences issues du traitement des images médicales comme la réversibilité. Nous verrons ensuite qu'il existe des liens très étroits unissant le tatouage d'image à des fins de protection de droits d'auteurs à certaines disciplines telles que la stéganographie, ou la cryptographie. Nous procéderons donc à la mise en œuvre d'une méthode de tatouage Mojette, adaptable à chacune de ces disciplines. Finalement, les atouts de ces deux études seront concentrées dans l'élaboration d'un schéma de tatouage exploitant conjointement les aspects perceptuels et la transformation Mojette (chapitre 6). Dans un dernier chapitre (chap 7), nous présenterons quelques travaux réalisés avec des partenaires du monde de l'industrie et de la recherche pour montrer l'aspect pratique des recherches engagées.

Chapitre 4

Le tatouage basé sur des critères psychovisuels

Sommaire

| | | |
|------------|--|------------|
| 4.1 | Introduction | 104 |
| 4.2 | Création d'un masque perceptuel spatial | 104 |
| 4.2.1 | Sélection des sites | 105 |
| 4.2.2 | Détermination des forces | 107 |
| 4.3 | Applications du masque | 111 |
| 4.3.1 | Le domaine DCT | 111 |
| 4.3.2 | Le domaine des ondelettes | 113 |
| 4.3.3 | Une application spatiale du masque | 115 |
| 4.3.4 | Expérimentations | 126 |
| 4.4 | Le tatouage psychovisuel | 129 |
| 4.4.1 | Insertion de la marque | 129 |
| 4.4.2 | Processus de détection | 131 |
| 4.4.3 | Expérimentations | 133 |
| 4.5 | Conclusion | 140 |

4.1 Introduction

Nous avons vu dans la section 3.1 l'importance de l'invisibilité ainsi que de la robustesse du filigrane. Pour satisfaire ces deux contraintes antagonistes, qui en imposent une troisième qui est la capacité d'insertion, les schémas de watermarking s'orientent de plus en plus vers l'exploitation des modèles du S.V.H. 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 S.V.H. ainsi que ses capacités de masquage pour assurer aussi bien la transparence que la robustesse du filigrane.

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 n'est plus de supprimer ou de modifier les composantes inutiles, mais au contraire de modifier judicieusement les composantes utiles, qui ne seront pas altérées par les schémas de compression par exemple.

L'utilisation d'un modèle psychovisuel pour le tatouage ne doit pas permettre l'insertion du copyright dans les composantes les moins significatives visuellement, mais au contraire, l'insertion doit s'effectuer uniquement au seuil de visibilité, afin que toute attaque dégrade immédiatement la qualité de l'image.

Dans ce chapitre, nous étudierons l'élaboration d'un masque perceptuel (section 4.2), nous verrons que la création de tels masques passe obligatoirement par la sélection de sites propices, ainsi que par la détermination de forces maximales permises. Puis nous testerons (section 4.3) la validité du masque obtenu dans un schéma de tatouage simple par modification des plans binaires de l'image. Enfin, dans la section 4.4, nous présenterons un schéma complet de tatouage des contrastes de l'image. La création du masque, ainsi que le schéma de tatouage dans les contrastes, exploitent pleinement l'étude du système visuel humain présenté au chapitre 1.

4.2 Création d'un masque perceptuel spatial

Certains auteurs [12] [60], ont opté pour la création de masques perceptuels dans le but d'optimiser le processus d'insertion du filigrane dans l'image. Ces masques perceptuels souvent appelés masques J.N.D. pour 'Just Noticeable Difference' permettent de définir la valeur maximale des niveaux de gris modifiables dans l'image sans occasionner d'artefacts visibles.

Deux étapes doivent précéder l'insertion d'un filigrane dans l'image : La sélection des sites les plus propices à l'invisibilité puis la détermination des forces maximales dans ces sites sans créer d'artefacts visibles. Ces

informations peuvent donc être contenues dans un masque dit J.N.D.

Deux traitements préliminaires sont effectués sur l'image originale $I^0(m, n)$ exprimée en niveaux de gris. Le premier consiste en une transformation permettant de passer de l'image en niveaux de gris, perceptuellement sans signification, à l'image des luminances affichées. Cette transformation correspond à la fonction gamma de l'écran utilisé. Le deuxième traitement transforme les luminances objectives affichées par l'écran en luminances subjectives $L_{subj}(m, n)$ perçues par l'observateur. La transformation en luminances subjectives proposé par Mannos et Sakrison [74] est retenue, cette dernière est normalisée par la valeur de la luminance maximale affichée à l'écran :

$$L_{subj} = L_{obj_{max}} \times \left(\frac{L_{obj}}{L_{obj_{max}}} \right)^{\frac{1}{3}}. \quad (4.1)$$

La sélection des sites propices au tatouage est effectuée sur l'image des luminances ainsi obtenues. Cette sélection, décrite ci-dessous, exploite la Décomposition en Canaux Perceptuels détaillée figure 1.18.

Pour chacun des sites retenus, un modèle de visibilité des erreurs, également discuté, permet de déterminer la force maximale $\Delta L_{i,j}^0(m, n)$ au delà de laquelle la marque insérée engendre des dégradations visibles. Cette force maximale est ensuite utilisée pour pondérer le filigrane avant son insertion. Après reconstruction, les transformations non linéaires inverses permettent de calculer l'image tatouée en niveaux de gris.

4.2.1 Sélection des sites

La sélection des sites exploite la décomposition en canaux perceptuels donnée section 1.3.8. Comme nous l'avons vu précédemment, cette décomposition utilise 4 bandes radiales appelées couronnes. La couronne BF (basses fréquences) est non sélective en orientation. Les largeurs de bande des couronnes III, IV, V, sont respectivement de 1.9 octaves, 1.3 octaves et 1 octave. Chacune de ces couronnes est sélective en orientation. Cette sélectivité angulaire dépend de la fréquence spatiale et vaut 45 degrés pour la couronne III et 30 degrés pour les couronnes IV et V.

Le choix des sous-bandes les plus propices au tatouage est conditionné par le compromis invisibilité-robustesse. Alors que la modification des seules basses fréquences spatiales engendre en général des dégradations facilement perceptibles, celles des hautes fréquences spatiales génère un tatouage qui reste non robuste puisqu'un simple filtrage passe bas permet de l'éliminer. Pour tirer avantage de tout le contenu de l'image et satisfaire les critères d'invisibilité et de robustesse, nous pensons qu'il est nécessaire de retenir des sites dans chacune des couronnes III, IV, et V. Pour chaque couronne, différentes orientations sont donc possibles, dans ces travaux, dans le but d'exploiter les différentes orientations offertes par la DCP, nous avons retenu les sites

des sous-bandes (III,1) (couronne III orientation 1), (IV,3) et (V,4).

Ce choix, fait ici pour alléger l'algorithme tout en sélectionnant différentes orientations afin de garantir une optimisation du nombre de sites retenus, peut également se faire en fonction du contenu de l'image, ce point sera abordé dans la section 6.3. Nous constaterons par le biais de différentes méthodes de tatouage, que si le choix de trois orientations différentes permet de maximiser le contenu du masque perceptuel, il n'apporte cependant pas d'optimisation du point de vue de la robustesse.

Une fois choisies, les sous-bandes sont extraites moyennant l'utilisation de filtres cortex. L'avantage de ces filtres, décrits section 1.3.7, réside dans le fait que leur réponse impulsionnelle, tout comme le champ réceptif de la plupart des cellules du cortex visuel humain, ressemble à une fonction de Gabor 2-D. Ces filtres sont le résultat du produit entre les filtres qui définissent la sélectivité radiale (appelés filtres Dom) et les filtres qui caractérisent la sélectivité angulaire (appelés filtres Fan).

Comme nous avons pu le constater au chapitre précédent, différentes directions d'investigation sont envisageables pour la sélection des sites. Dans un premier temps, nous avons opté pour une méthode de sélection des zones de fréquences moyennes, ainsi qu'une sélection des contours de l'image pour la création d'un masque J.N.D. Nous détaillerons par la suite les avantages et inconvénients d'un tel choix et nous décrirons l'évolution du masque perceptuel en fonction des expérimentations.

La figure 4.1 présente l'algorithme que nous avons développé dans [7] pour la création d'un masque J.N.D.

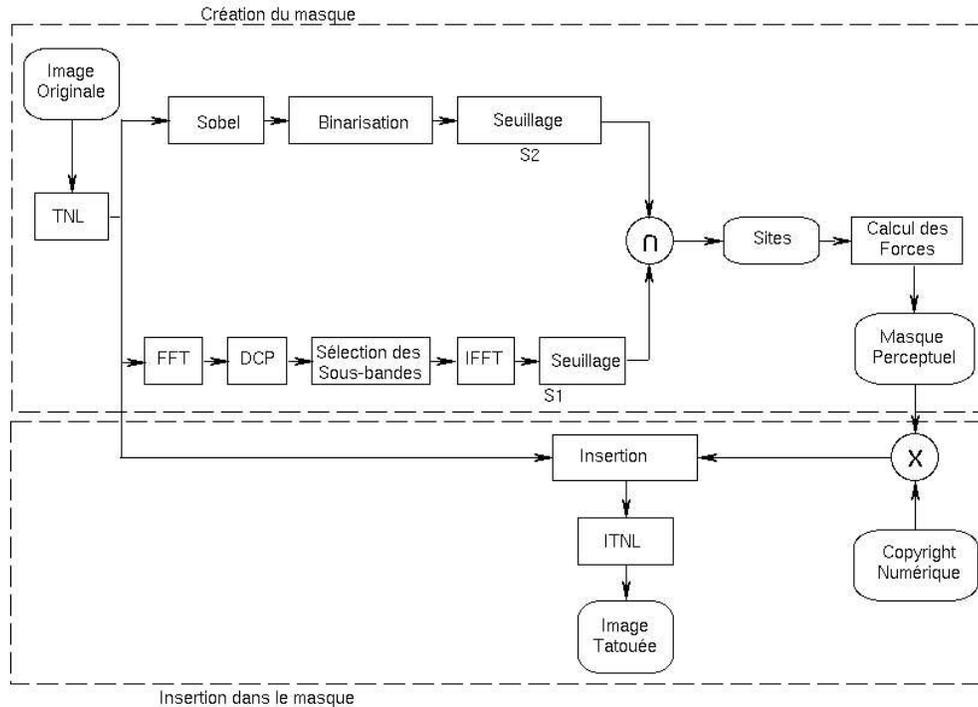


FIG. 4.1 – Création d'un masque J.N.D.

Le schéma de création du masque est décomposé en deux étapes distinctes : La détection des contours de l'image et l'extraction des sous-bandes propices au tatouage.

La détection des contours s'effectue par un filtrage de Sobel sur l'image, suivi d'une binarisation de l'image obtenue.

L'extraction des sous-bandes s'effectue en trois étapes principales, la transformation de Fourier de l'image, permet l'application de la DCP sur le spectre, puis la sélection et extraction des sous-bandes propices.

L'intersection entre ces sites définis psychovisuellement et les contours de l'image nous permettent donc d'obtenir les sites du masque J.N.D. dépendants de l'image traitée.

4.2.2 Détermination des forces

Une fois les sites définis, la force maximale de chacun de ces sites doit être calculée. Pour ce faire, nous avons vu à la section 1.3.4 que la définition du contraste pour des signaux complexes est loin d'être évidente. Effectivement, pour des signaux simples du type incrément ou signal sinusoïdal, le contraste est en général bien défini soit par la formule de Michelson $C = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$, soit par la fraction de Weber $c = \frac{\Delta L}{L}$. Pour les signaux complexes, différentes définitions sont données dans la littérature. Le contraste quadratique moyen

est utilisé dans [87] et [104] pour comparer le contraste de deux images. Ce contraste est défini en tout point (m, n) pour une image de taille $M \times N$ par l'équation 4.2

$$c_{rms} = \left(\frac{1}{(N-1)(M-1)} \sum_{m=1}^M \sum_{n=1}^N \left(\frac{L(m, n) - L_{moy}}{L_{max}} \right)^2 \right)^{1/2}, \quad (4.2)$$

où $L(m, n)$ représente la luminance du point (m, n) , L_{moy} et L_{max} représentent respectivement la luminance moyenne et la luminance maximale de l'image.

Une autre définition du contraste est donnée dans [56]. Il est défini dans le domaine de Fourier et pour chaque bloc s'écrit

$$c(u, v) = 2 \frac{A(u, v)}{DC}, \quad (4.3)$$

où u et v représentent les fréquences spatiales, $A(u, v)$ l'amplitude du spectre de l'image et DC représente la valeur moyenne. Ces définitions et d'autres moins connues présentent toutefois des inconvénients. Ainsi, pour la première définition ci-dessus, l'inconvénient majeur est qu'elle ne prenne en compte ni la structure spatiale de l'image ni son contenu fréquentiel. Si ce dernier point est pris en compte dans la deuxième définition, il n'en demeure pas moins que, pour cette dernière, la localisation spatiale reste très grossière.

Pour une bonne définition du contraste, il semble nécessaire que les résolutions spatiale et fréquentielle soient directement liées. Une définition répondant naturellement à ce critère est celle du contraste local à bande limitée, développé par Peli dans [88]. Ce contraste est d'une part local pour exprimer le fait que la sensibilité d'un observateur aux changements de luminance varie avec la luminance moyenne locale. Il est d'autre part à bande limitée car la perception des dégradations dépend directement de leur localisation spectrale.

Ce contraste est donné en tout pixel de position (m, n) , par l'équation :

$$c_i(m, n) = \frac{L_i(m, n)}{\sum_{k=0}^{i-1} L_k^i(m, n)}, \quad (4.4)$$

où l'indice i représente le i^{eme} canal radial et le dénominateur représente le signal basses fréquences relatif au i^{eme} canal radial.

Cette définition du contraste traduit le fait que la perception d'un détail au sein d'une image dépend de la valeur moyenne de son environnement local et pour lequel l'extension spatiale est adaptée à la résolution de ce détail. Pour une décomposition telle que celle de la figure 1.18 présentée section 1.3.8, où les canaux sont à la fois sélectifs radialement et angulairement, le contraste local à bande limitée a été modifié pour s'écrire :

$$c_{i,j}(m,n) = \frac{L_{i,j}(m,n)}{\sum_{k=0}^{i-1} \sum_{l=0}^{card(l)} L_{k,l}^i(m,n)}, \quad (4.5)$$

Où $L_{i,j}(m,n)$ et $c_{i,j}(m,n)$ représentent respectivement la luminance et le contraste au point (m,n) du i^{eme} canal radial et j^{eme} secteur angulaire. $Card(l)$ est le nombre de secteurs angulaires dans le k^{eme} canal radial.

Le dénominateur représentant en fait la luminance moyenne locale correspondant à tous les canaux de support spectral inférieur à celui du i^{eme} canal. L'expression du contraste peut donc se mettre sous la forme

$$c_{i,j}(m,n) = \frac{L_{i,j}(m,n)}{\overline{L_{i,j}}(m,n)} \quad (4.6)$$

où $\overline{L_{i,j}}(m,n)$ est la luminance moyenne locale au point (m,n) ; la localité étant relative à la sous-bande radiale i .

Pour déterminer la force du tatouage dans notre travail, le contraste local à bande limitée est donc choisi. Des études menées au laboratoire ont permis d'établir des lois de quantification du contraste local à bande limitée. La notion de contraste seuil $\Delta C_{i,j}$ a été définie pour permettre de fixer un pas de quantification maximal applicable aux contrastes d'une image sans générer d'artefacts visibles. Ce contraste seuil est donné par un modèle de visibilité des dégradations, qui est établi à partir des résultats d'un ensemble de tests psychophysiques, ces tests sont détaillés respectivement dans [15] et [115].

Ces valeurs de contraste seuil ont été déterminées lorsque l'effet de masquage intra canal seul (interaction de signaux appartenant à une seule sous-bande) est pris en compte. En fonction des caractéristiques locales, ce modèle s'écrit :

$$\Delta C_{i,j} = \frac{E_{i,j}}{L_0} \left(\frac{\Delta f_i}{f_{0,i}} \right)^{k_i} \quad (4.7)$$

où $\Delta C_{i,j}$ est le contraste seuil de la sous-bande i,j ,

$E_{i,j}$ est l'énergie de la sous-bande (i,j) ,

L_0 est la luminance de l'écran utilisé pendant les tests,

Δf_i est la largeur de bande radiale i ,

$f_{0,i}$ est la fréquence centrale de la sous-bande i ,

et k_i est une constante qui dépend de la sous-bande i .

La variation maximale $\Delta L_{i,j}(m,n)$ (ou force maximale) que l'on peut s'autoriser sur la luminance en chaque sous-bande (i,j) et en chaque site (m,n) avant que les dégradations deviennent visibles peut être donnée en première approximation par :

$$\Delta L_{i,j}(m,n) = \Delta C_{i,j} \times \bar{L}_{i,j}(m,n) \quad (4.8)$$

Où $\Delta C_{i,j}(m,n) = \Delta C_{i,j} = \text{constante}$ en tout point (m,n) . $\Delta C_{i,j}$ est, comme nous l'avons vu, mesurée à partir de tests psychophysiques pour chaque sous-bande i,j et $\bar{L}_{i,j}(m,n)$ est la luminance moyenne locale à la sous-bande i,j et en tout point (m,n) .

Cette variation, fonction de la luminance moyenne locale, permet donc d'adapter la force du filigrane aux caractéristiques locales du site considéré.

Traduite en niveaux de gris, cette erreur maximale donne le nombre maximal de bits à modifier pour que la dégradation générée reste en dessous du seuil de perception.

| Couronne | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|-----------|--------|--------|--------|--------|--------|--------|
| BF (I+II) | 0.5 | | | | | |
| III | 0.0034 | 0.004 | 0.0034 | 0.004 | | |
| IV | 0.0066 | 0.010 | 0.010 | 0.0066 | 0.010 | 0.010 |
| V | 0.026 | 0.04 | 0.04 | 0.026 | 0.04 | 0.04 |

TAB. 4.1 – Valeurs du plus petit pas de quantification par sous-bande proposés par N. Bekkat dans [15]

Les tableaux 4.1 et 4.2 présentent respectivement les résultats des premiers seuils de quantification obtenus pour chaque canal dans [15] et [115].

| Couronne | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|-----------|--------|--------|--------|--------|--------|--------|
| BF (I+II) | 0.001 | | | | | |
| III | 0.005 | 0.006 | 0.005 | 0.006 | | |
| IV | 0.01 | 0.015 | 0.01 | 0.015 | 0.015 | |
| V | 0.04 | 0.06 | 0.06 | 0.04 | 0.06 | 0.06 |

TAB. 4.2 – Valeurs du plus petit pas de quantification par sous-bande proposés par H. Sénane dans [115]

Pour valider le masque obtenu, le schéma représenté figure 4.1 effectue une insertion de la marque directement en niveaux de gris dans l'image. Cette même insertion peut toutefois s'effectuer de différentes façons, soit par addition en tout point des sites, de la force du masque psychovisuel dans l'image, soit par addition et soustraction respectivement d'un pixel sur deux, dans le but de compenser chaque modification par les modifications du voisinage. Un observateur effectuant la comparaison de l'image obtenue avec l'image de référence détecte dans le premier cas un rehaussement global des niveaux de gris, sans toutefois que cela procure de gêne, tandis que pour la seconde alternative, aucune différence n'est perceptible selon la méthode d'évaluation subjective à "double stimulus" décrite dans le paragraphe 3.6.4.2.

4.3 Applications du masque

Dans cette section, nous présentons les différents domaines dans lesquels le masque a été testé. Nous constaterons des résultats variables en terme de robustesse en fonction des domaines d'insertion. Cependant, l'invisibilité de la marque est toujours assurée par l'utilisation d'un tel masque perceptuel. De plus, l'utilisation de masques JND permet d'éviter l'utilisation de critères d'évaluation de qualité (complexes à mettre en œuvre pour les critères objectifs et contraignants pour les critères subjectifs), tels que ceux que nous avons présentés dans le chapitre 3.

4.3.1 Le domaine DCT

L'efficacité de ce masque perceptuel a tout d'abord été testée dans le domaine D.C.T. [128]. Dans ces travaux, des tests expérimentaux ont été conduits sur l'estimation de la capacité de tatouage des images. Les performances du masque que nous proposons ont été testées et comparées à un masque empirique opérant une sélection des coefficients moyenne fréquence dans des blocs DCT 8×8 , comme faisaient Hsu et al. dans [57] et [58].

L'adaptation du masque perceptuel spatial au domaine DCT s'est effectuée en trois étapes :

- Le masque perceptuel a tout d'abord été divisé en blocs de taille 8×8 ,
- Pour chaque bloc, un décompte des pixels dépassant un seuil prédéfini (et ajustable) est effectué,
- Si le nombre de pixels tatouables du bloc est supérieur à 15, le bloc est considéré comme tatouable. Ce nombre de 15 ayant été obtenu par des tests expérimentaux.

Une fois le masque perceptuel spatial adapté au domaine DCT, l'étude de Byron Yu portait sur la comparaison des deux masques en termes de capacité des images. Cette comparaison s'est effectuée selon la procédure expérimentale suivante :

- Fixer une valeur de pas de quantification JPEG,
- Débuter avec un filigrane de faible longueur, ayant une force suffisamment faible pour obtenir une parfaite invisibilité et une détection fiable à 100 %,
- Augmenter la force de la marque jusqu'à ne plus satisfaire les deux critères.

Les résultats obtenus par le masque empirique créé à l'ENST, pour une force d'insertion fixée à 10, sont donnés sur le tableau 4.3. Les images de test utilisées sont fournies dans l'Annexe C

| Image | Dimensions | Capacité (bits) | Capacité (bits/pixel $\times 10^{-3}$) |
|-----------|------------------|-----------------|---|
| eglise | 512 \times 512 | 355 | 1.35 |
| bureau | 256 \times 192 | 50 | 1.02 |
| bateau | 720 \times 576 | 967 | 2.33 |
| chambre | 512 \times 512 | 301 | 1.15 |
| singe | 512 \times 512 | 366 | 1.40 |
| engrenage | 768 \times 512 | 255 | 0.65 |
| water | 512 \times 512 | 10 | 0.04 |
| franck | 720 \times 576 | 284 | 0.68 |
| iba | 720 \times 576 | 780 | 1.88 |
| enfant | 720 \times 576 | 882 | 2.13 |

TAB. 4.3 – Capacité expérimentale du masque empirique à 100% de détection

Les mêmes tests ont été reconduits avec le masque perceptuel, et ont fournis les résultats présentés dans le tableau 4.4

| Image | Dimensions | Force | Capacité (bits) | Capacité (bits/pixel $\times 10^{-3}$) |
|-----------|------------------|-------|-----------------|---|
| eglise | 512 \times 512 | 25 | 595 | 2.27 |
| bureau | 256 \times 192 | 20 | 51 | 1.04 |
| bateau | 720 \times 576 | 10 | 115 | 0.28 |
| chambre | 512 \times 512 | 20 | 97 | 0.37 |
| singe | 512 \times 512 | 5 | 1160 | 4.43 |
| engrenage | 768 \times 512 | 25 | 569 | 1.45 |
| water | 512 \times 512 | 0 | 0 | 0 |
| franck | 720 \times 576 | 0 | 0 | 0 |
| iba | 720 \times 576 | 837 | 1 | 2.02 |
| Enfant | 720 \times 576 | 647 | 2 | 1.56 |

TAB. 4.4 – Capacité expérimentale du masque perceptuel à 100% de détection

Nous pouvons constater que les résultats fournis par le masque perceptuel sont encourageants pour les images comportant des contours ou des zones texturées, telles que 'singe', 'église' ou encore 'engrenage', mais ils sont cependant parfois insatisfaisant pour les images présentant de nombreuses zones uniformes ('bureau' ou 'frank') ou pour les images d'aspect sombre, dont la moyenne des niveaux de gris est faible telles que 'chambre' ou 'water'. Nous verrons par la suite que la prise en considération de ces résultats nous a amenés à conduire d'autres tests et à modifier le processus de création du masque perceptuel.

Outre les résultats très encourageant obtenus dans le domaine DCT, une étude portant sur l'adaptation du masque perceptuel au domaine des ondelettes (DWT pour 'Discrete Wavelet Transform') a été réalisée.

4.3.2 Le domaine des ondelettes

Nous pouvons trouver dans [75], deux méthodes de tatouage exploitant le masque perceptuel décrit ci-dessus. Une première approche consistait à effectuer un post-traitement sur une image tatouée à l'aide d'un procédé de tatouage des coefficients ondelettes. Ici, une image était très fortement tatouée, puis subissait un seuillage afin de maintenir les valeurs des pixels tatoués en-deçà des valeurs fournies par le masque. Si cette technique garantissait l'invisibilité de l'image tatouée, elle n'améliorait cependant pas la robustesse, car le 'filtrage' de l'image tatouée par le masque, était assimilable à une attaque et réduisait donc la capacité de détection du schéma.

Au vu de ces résultats, cette méthode a donc été délaissée au profit d'une seconde méthode permettant l'optimisation de la force de la marque à appliquer pour les coefficients à tatouer. Le principe de cette seconde méthode, consiste à représenter spatialement le filigrane préalablement inséré dans le domaine DWT.

L'équation de reconstruction d'une image I décomposée selon une base de paquets d'ondelettes prédéfinie B , est la suivante :

$$I(i, j) = \sum_{(p,k,l) \in B} \sum_{m,n} C_{p,k,l}(m, n) W_{p,k,l,m,n}(i, j) \quad (4.9)$$

où $C_{p,k,l}(m, n)$ est un paquet correspondant à un noeud de l'arbre binaire de décomposition. Chacun de ces paquets contient l'information relative à toute l'image dans la bande de fréquence (k, l) et dont la taille est déterminée par le niveau de résolution p .

Les ondelettes W étant séparables, elles respectent l'équation suivante :

$$W_{p,k,l,m,n}(i, j) = \frac{1}{2^{-p}} W_k(2^{-p_i} - m) W_l(2^{-p_j} - n) \quad (4.10)$$

où (i, j) représente la position des pixels, les paramètres m et n représentent les transitions spatiales et p est le niveau de résolution du paquet.

Pour simplifier les équations, nous noterons k pour décrire non seulement le niveau de résolution, mais aussi la bande de fréquence. L'équation 4.9 devient alors :

$$I(i, j) = \sum_{k \in B} \sum_{m,n} C_k(m, n) W_{k,m,n}(i, j) \quad (4.11)$$

Après la procédure de tatouage, la meilleure base devient B^T . Notons $B^T = SB_{mod}^T \cup SB_{nonmod}^T$, nous séparons les noeuds des coefficients modifiés de ceux n'ayant pas subi de modification. L'équation de reconsti-

tution devient alors :

$$I(i, j) = \sum_{k \in SB_{mod}^T} \sum_{m, n} C_k(m, n) W_{k, m, n}(i, j) + \sum_{k \in SB_{nonmod}^T} \sum_{m, n} C_k(m, n) W_{k, m, n}(i, j) \quad (4.12)$$

La différence entre l'image tatouée et l'image originale est donnée par :

$$I^T(i, j) - I(i, j) = \sum_{k \in SB_{mod}^T} \sum_{m, n} C_k^T(m, n) - C_k(m, n) W_{k, m, n}(i, j) \quad (4.13)$$

L'équation 4.13 peut alors s'écrire :

$$I^T(i, j) - I(i, j) = \sum_{k \in SB_{mod}^T} (\alpha_k - 1) \sum_{m, n} C_k(m, n) W_{k, m, n}(i, j) \quad (4.14)$$

où $\sum_{m, n} C_k(m, n) W_{k, m, n}(i, j)$ représente l'image obtenue par inversion de la décomposition en base de paquets d'ondelettes de l'unique paquet k , toutes les contributions relatives aux autres paquets ayant été annulées. Si nous notons I_k , l'image obtenue par reconstruction du seul paquet k , $I_k(i, j) = \sum_{k \in SB_{mod}^T} C_k(m, n) W_{k, m, n}(i, j)$, nous pouvons réécrire l'équation 4.14 :

$$I^T(i, j) - I(i, j) = \sum_{k \in SB_{mod}^T} (\alpha_k - 1) I_k(i, j) \quad (4.15)$$

Ceci fournit l'expression des modifications de l'image dans le domaine spatial et permet donc l'application du modèle psychovisuel.

Le critère perceptuel est vérifié si :

$$|(I(i, j) - I^T(i, j))| \leq F(i, j) \quad (4.16)$$

Ceci implique la condition suivante :

$$\forall(i, j), \left| \sum_{k \in SB_{mod}^T} (\alpha_k - 1) I_k(i, j) \right| \leq F(i, j) \quad (4.17)$$

La maximisation des coefficients α_k tout en respectant l'équation 4.17, permet l'optimisation de la méthode de tatouage.

La réalisation pratique d'une telle méthode est assez complexe et demande un temps de calcul assez important. Pour alléger ces calculs, la longueur du filigrane a donc été réduite. Les résultats en termes d'invisibilité

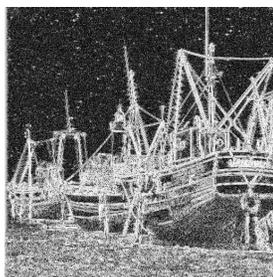
ont été assez encourageants, mais cependant les performances en termes de robustesse s'en trouvent dégradées.

Nous avons par ailleurs testé dans [8] les qualités de ce masque lors d'un processus de tatouage et d'extraction dans les bits les moins significatifs appelés LSB (Less Significant Bits). Le paragraphe suivant est dédié à la présentation de ces travaux.

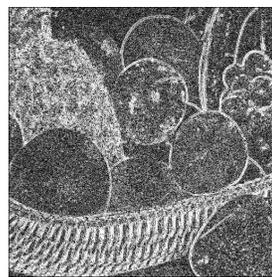
4.3.3 Une application spatiale du masque

Des tests nous ont permis de constater que le tatouage des contours d'une image procurait un masque J.N.D. de forte amplitude, mais ne garantissait que très peu la robustesse. Effectivement, une expérience simple le prouve : pour réaliser une étude du comportement des attaques sur les images, nous prenons tout d'abord un ensemble d'images de contenus variés. Nous dégradons ces images à l'aide de différents algorithmes (JPEG, filtrage médian, passe-bas, ...), puis nous créons les images résultant de la différence en valeur absolue entre les versions originales et les versions dégradées, nous constatons que les pixels appartenant aux contours de l'image sont ceux qui ont subi les modifications les plus importantes.

La figure 4.2 illustre bien cette faiblesse des contours vis-à-vis de différentes attaques. Une égalisation d'histogramme appliquée sur ces images de différences permet de visualiser les pixels (blancs) subissant les dégradations les plus importantes.



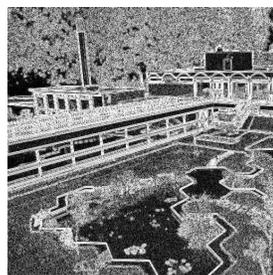
(a) bateau : filtrage 3x3



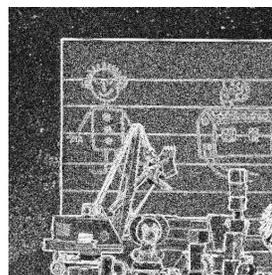
(b) fruit : rehaussement de contraste
75%



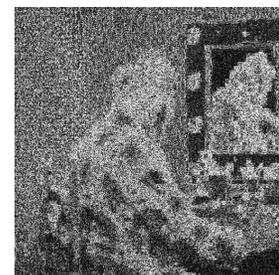
(c) lena : Pixelisation coef=2



(d) iba : filtrage 5x5



(e) tableau : filtrage 3x3



(f) clown : JPEG, q=75%

FIG. 4.2 – Comportement des contours des images en fonction de diverses attaques

Nous constatons donc sur la figure 4.2 que les six images, ayant subi des attaques, présentent des écarts de niveaux de gris plus importants sur les contours que sur les zones homogènes. Cependant, il est important de noter que toutes les images ou toutes les dégradations ne se comportent pas de la même façon. Le codage JPEG peut par exemple introduire des effets de blocs, quant aux déformations géométriques, non représentées ici, n'attaquent pas les contours, mais posent surtout des problèmes de désynchronisation de la marque.

Le schéma présenté figure 4.1 fut donc modifié et la détection de contours a été remplacée par une détection de zones uniformes [8]. Nous avons réalisé, sur plusieurs images, des tests de robustesse de la marque face à différentes attaques, respectivement dans les contours, dans les zones uniformes et sur toute l'image.

Cette détection de zones uniformes de l'image s'effectue simplement par le calcul de la différence de chaque pixel avec ses 4 voisins, si ces différences sont nulles le pixel considéré appartient bien à une zone uniforme.

Les détections des zones uniformes correspondant aux images utilisées sur la figure 4.2 sont présentées ci-dessous :

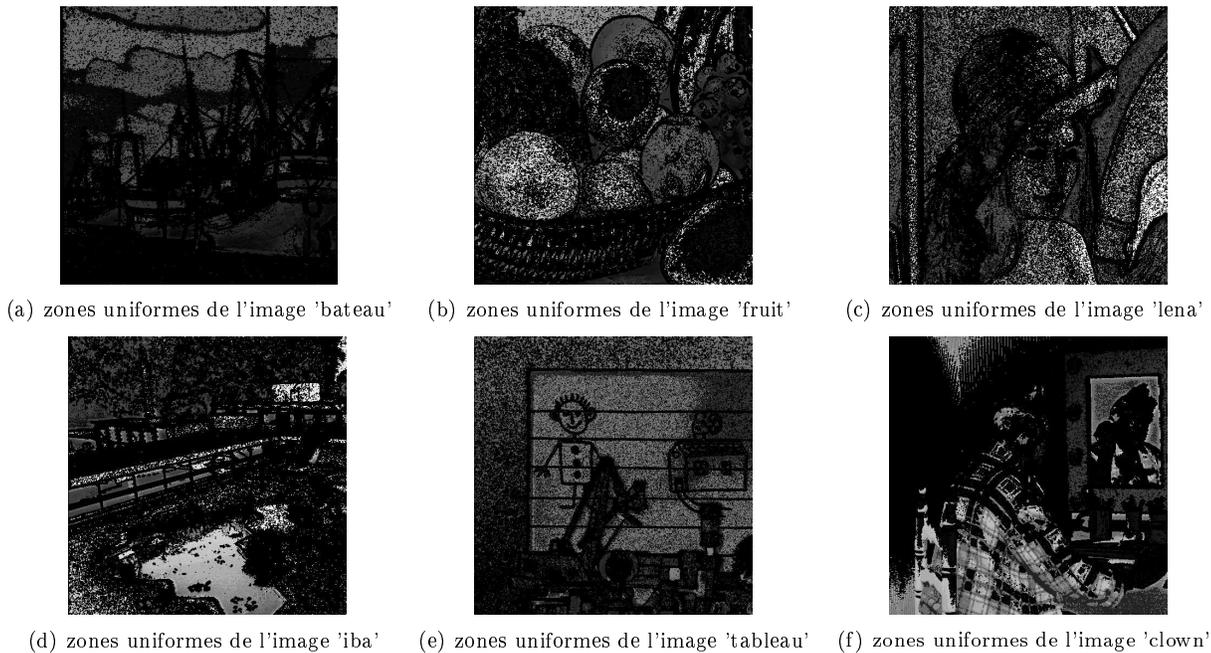


FIG. 4.3 – Images représentant les zones uniformes des images de test

Dans la figure 4.3, les pixels noirs représentent les zones de l'image n'ayant pas été retenues par le filtrage de l'image, ce sont les zones "non-uniformes". Les pixels les plus clairs représentent les zones uniformes des images selon le processus de sélection décrit précédemment.

L'allocation de la force maximale de tatouage dans ces zones uniformes se fait selon la loi de Weber. Pour

tout pixel de position (m, n) , la force maximale permise est donnée en luminances par l'équation :

$$\Delta L(m, n) = L(m, n) \times \Delta C \quad (4.18)$$

où ΔC représente la constante de Weber et vaut 0.02.

La fonction d'intercorrélation normalisée, équation 4.19 a été appliquée entre la marque binaire originale et la marque extraite.

$$r(d) = \frac{\sum_i [(x(i) - m_x) * (y(i - d) - m_y)]}{\sqrt{\sum_i (x(i) - m_x)^2} \sqrt{\sum_i (y(i - d) - m_y)^2}} \quad (4.19)$$

où $x(i)$ et $y(i)$ représentent les signaux d'entrée à corrélérer, m_x et m_y représentent leurs moyennes respectives et d représente un paramètre de décalage.

Les images 4.4(d) à 4.4(i) présentent les résultats des fonctions d'intercorrélation pour l'image 'bateau' tatouée respectivement dans les contours, les zones uniforme ou sur toute l'étendue de l'image. Ces fonctions d'intercorrélation sont calculées entre le filigrane original préalablement stocké et le filigrane extrait de l'image tatouée et attaquée par des filtrages passe-bas de noyau 3×3 et 5×5 .

Dans les courbes présentant l'intercorrélation, l'axe des abscisses correspond aux différents décalages des données notées d dans l'équation 4.19. Dans le cas présent, l'intercorrélation étant normalisée, l'axe des ordonnées est obligatoirement compris entre -1 et 1 .

Pour les images 4.4(a) à 4.7(f), nous avons inséré, dans l'image 'bateau', une marque binaire dans les contours, les zones uniformes ou sur toute l'image, dans le troisième LSB pour chaque pixel tatouable. Un tel tatouage ne dégrade pas visiblement les images, mais permet des résultats de détection très différents en fonction des zones sélectionnées.

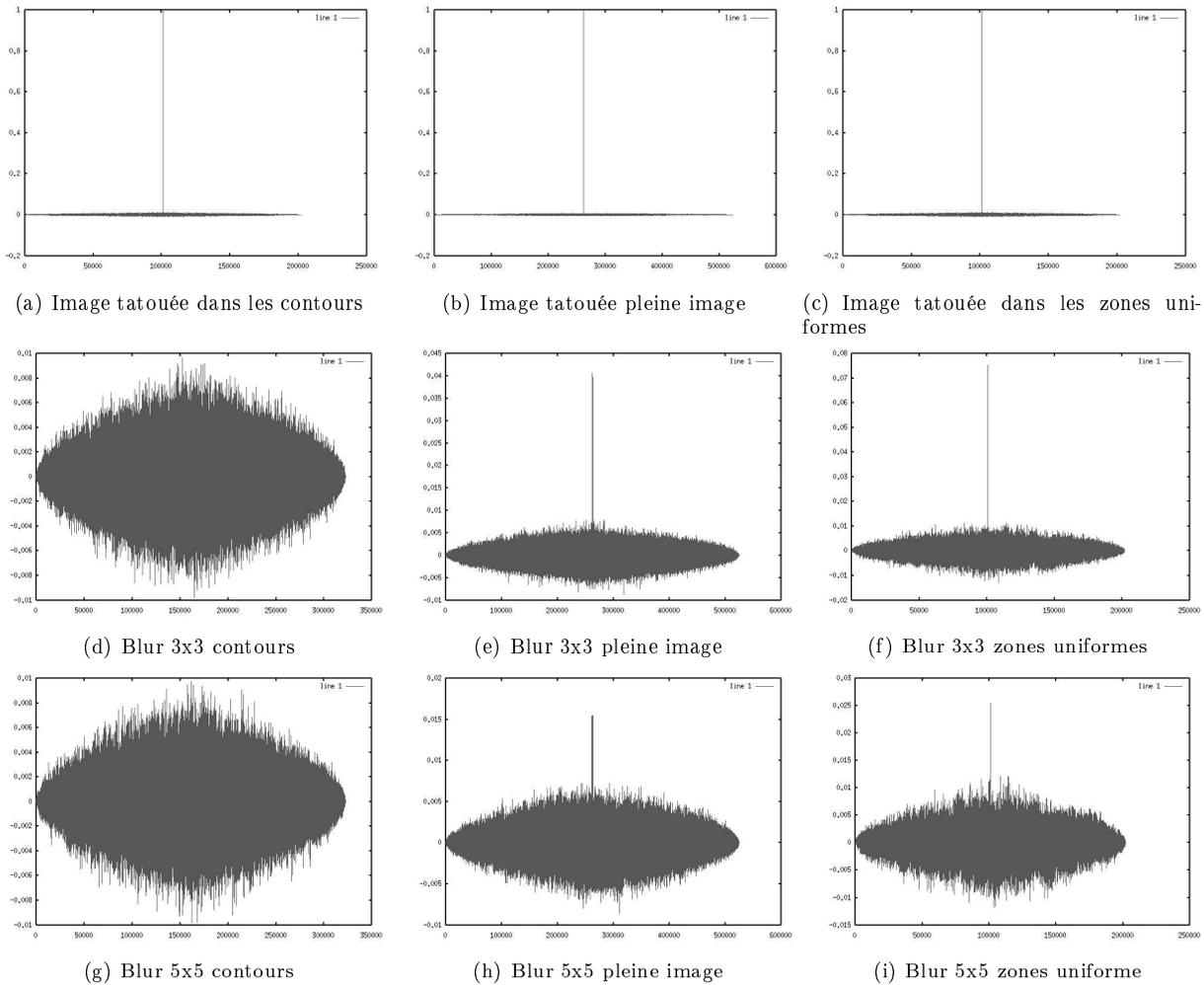


FIG. 4.4 – Intercorrélations dans le cas de filtrages

Nous pouvons clairement observer sur les résultats d'intercorrélation représentées figure 4.4, que la détection d'une marque insérée dans les contours de l'image est bien moins robuste face aux filtrages passe-bas qu'une marque insérée dans les zones uniformes ou sur toute l'étendue de l'image.

Une autre série de résultats, encore pour l'image 'bateau' et concernant le codage JPEG, sont présentés sur la figure 4.5. Les résultats du codage JPEG fournis, ont été obtenus avec un facteur de qualité fixé respectivement à $q=30\%$ et $q=40\%$, dans ces deux cas, le processus se trouve au seuil de détection pour un tatouage dans les zones uniformes. Pour des facteurs de qualité inférieurs, aucune des trois zones de tatouage ne permettrait de détection fructueuse. Nous pouvons observer la variation du pouvoir de détection en fonction des zones d'insertion. Encore une fois, les résultats de détection de la marque sont meilleurs pour une insertion dans les zones uniformes que dans les contours, ou dans toute l'image.

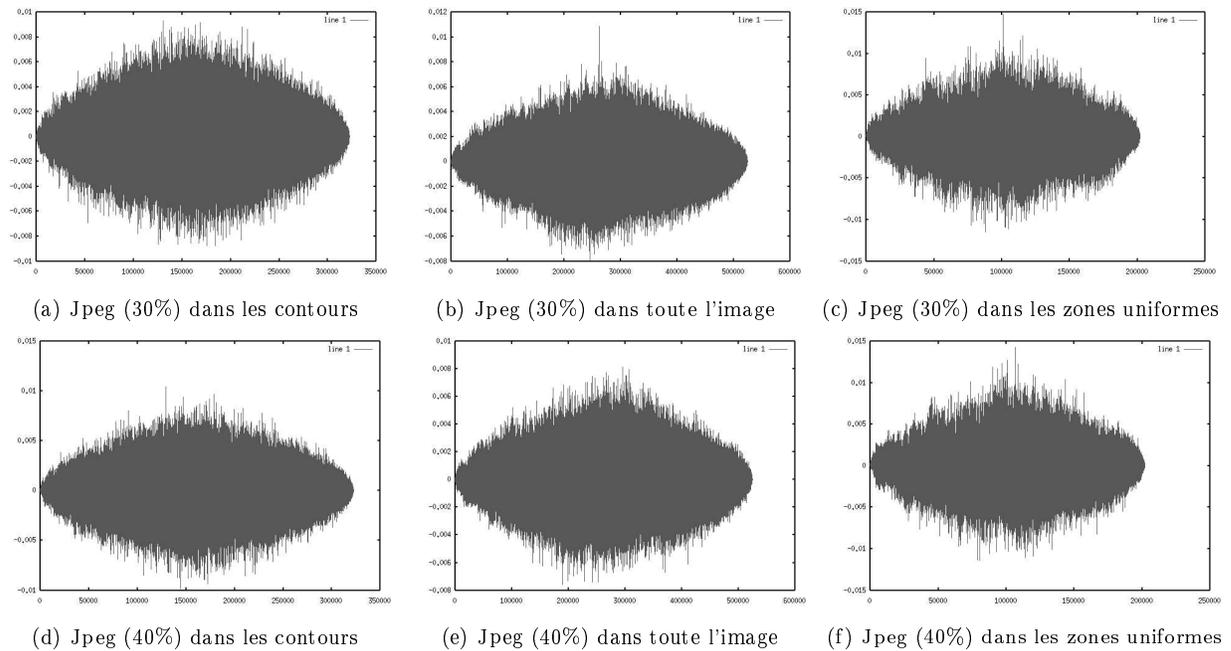


FIG. 4.5 – Intercorrélations dans le cas de codage JPEG

La figure 4.6 présente, toujours pour l'image 'bateau', les résultats de l'intercorrélation dans le cas d'addition de bruits dans l'image. Contrairement aux attaques présentées sur les figures 4.4 et 4.5, cette attaque est totalement indépendante de l'image considérée. Cette dernière fournit donc globalement des résultats similaires, bien que les fonctions d'intercorrélation concernant la détection dans les zones uniformes soient légèrement plus performantes.

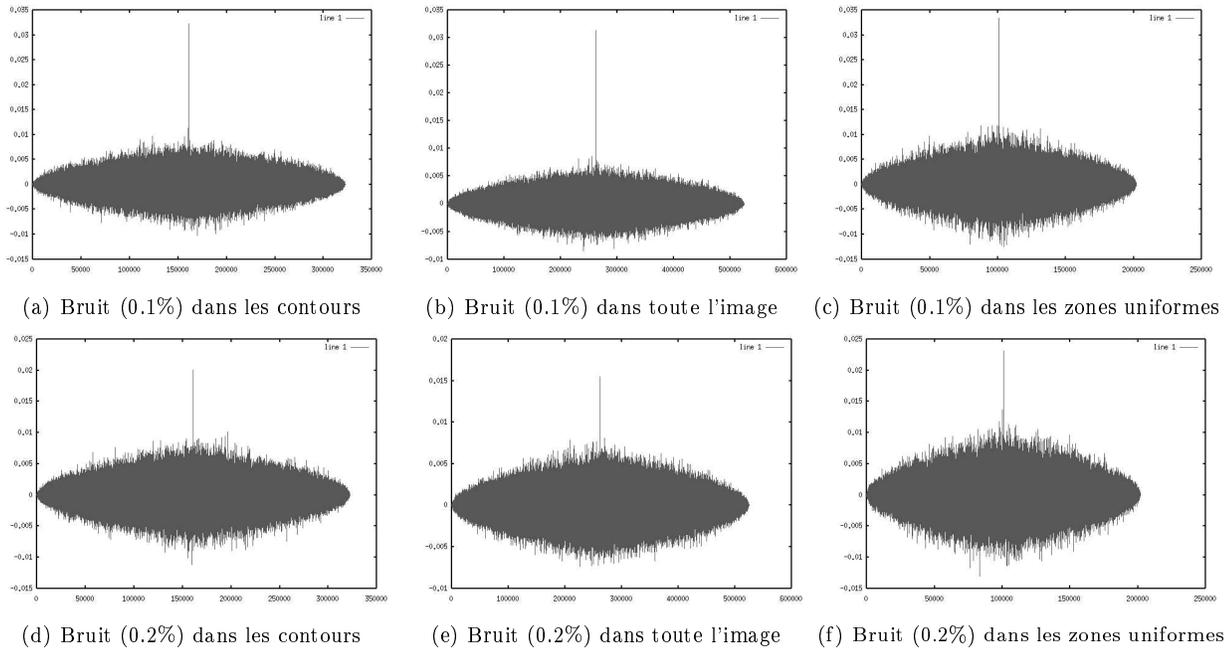


FIG. 4.6 – Intercorrélations dans le cas d'addition de bruits

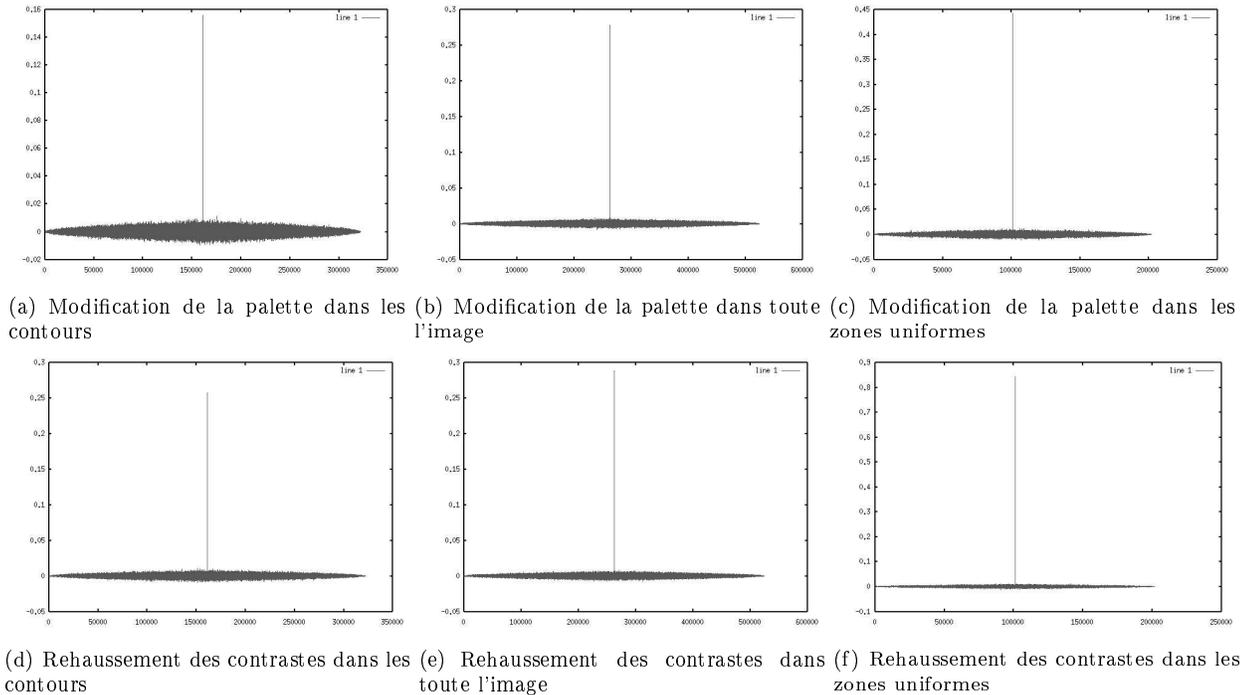


FIG. 4.7 – InterCorrélation dans le cas de modifications des niveaux de gris

Les expériences de détection dans les zones uniformes ont été répétées pour des images de contenus dif-

férents. Les résultats concernant l'extraction du filigrane dans les zones uniformes de l'image 'fruit' pour les attaques citées précédemment et dans les mêmes conditions d'insertion (tatouage du troisième LSB), sont représentés dans la figure 4.8.

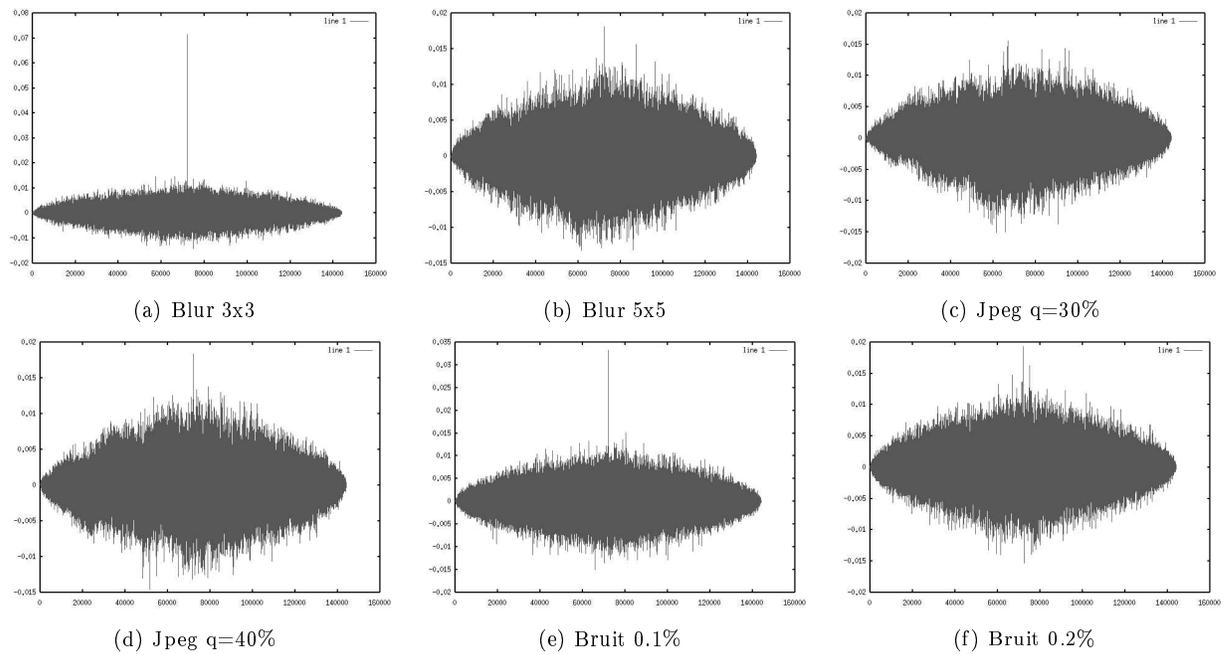


FIG. 4.8 – InterCorrélation pour l'image "fruit"

Si ces résultats sont assez satisfaisants pour des images telles que 'bateau' ou 'fruit', nous pouvons cependant constater sur la figure 4.9 que pour une image contenant beaucoup de textures, telle que l'image 'iba', les résultats de détection de la marque sont assez décevants.

Toutefois, pour cette même image 'iba', les résultats de la détection d'une marque insérée dans toute l'image, ne sont pas plus concluants, et concernant un schéma opérant une insertion de la marque dans les contours ou dans les textures, les résultats seraient désastreux.

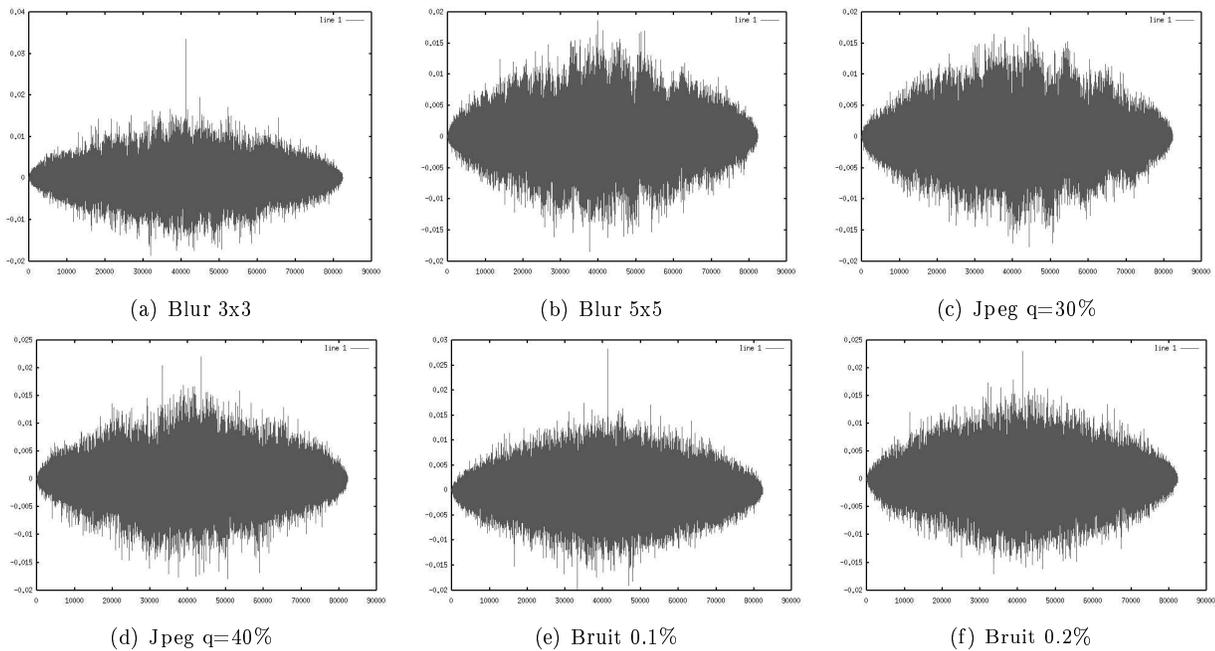


FIG. 4.9 – InterCorrélation pour l'image “iba”

Au vu des résultats obtenus dans les figures 4.4 à 4.9, la création du masque psychovisuel présenté section 4.2 a été amélioré.

Tout d’abord, la sous-bande $(III, 1)$ que nous avons considérée lors de la création du masque JND dans la section 4.2, n’est à présent plus prise en compte dans l’élaboration du masque, car la sélection des zones uniformes est considérée comme appartenant aux basses fréquences de l’image.

Les pixels du masque n’étant pas classifiés dans les zones uniformes, mais provenant de la sélection des sous-bandes, subissent la relation de l’équation 4.8.

De la même façon que les contours d’une image sont les zones les plus rapidement et les plus fortement dégradées, les bits de poids faibles (LSB pour Less Significant Bits), sont aussi les bits les moins robustes, comme nous l’avons vu à la section 3.6. L’insertion de la marque dans ces bits de poids faibles provoquerait une détection biaisée, c’est à dire un taux de fausse alarme important. Pour ces raisons, l’insertion du filigrane s’effectue non pas dans les LSB de l’image, mais seulement au bit de poids maximum autorisé par le masque. Un seul bit est donc inséré par pixel. Le schéma d’insertion est présenté figure 4.10 et le processus permettant tout aussi bien la détection que l’extraction, est présenté sur la figure 4.11.

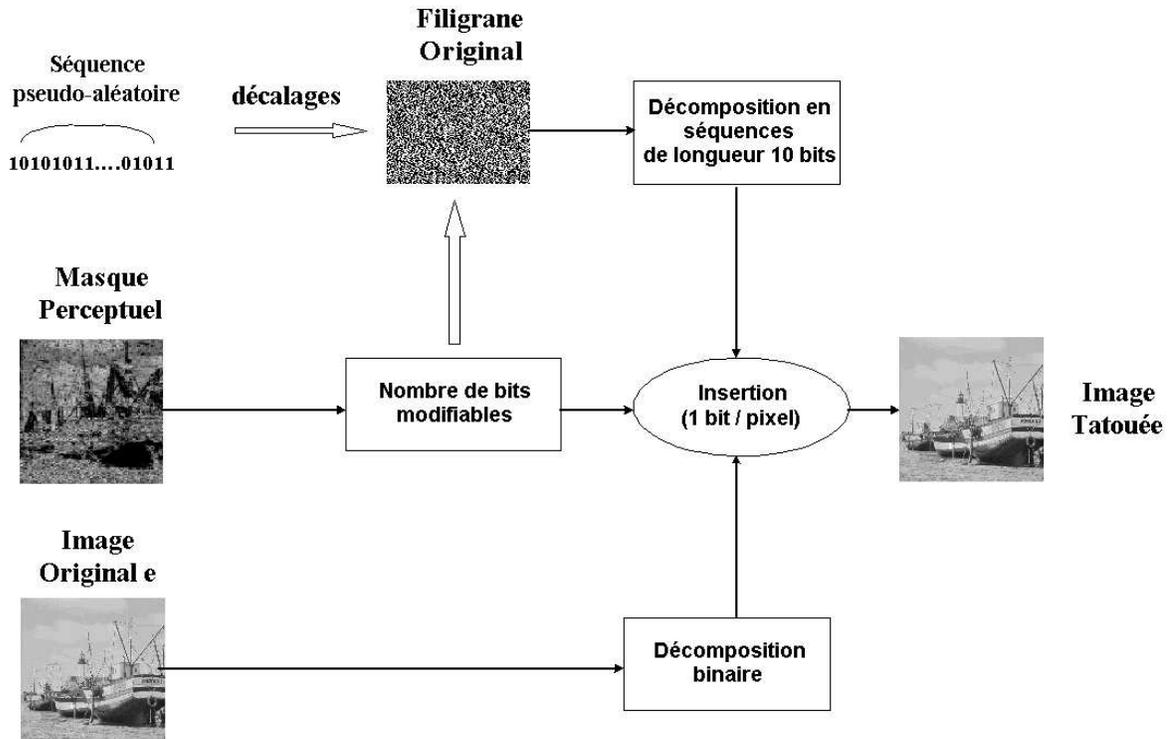


FIG. 4.10 – Schéma d'insertion de la marque dans les bits de l'image

Dans le but d'optimiser les résultats de détection et de résister à certaines déformations géométriques telles que la translation ou le cropping, l'insertion s'est effectuée sur tous les pixels disponibles du masque. La marque est alors étalée sur toute l'image et le processus d'insertion calcule, pour chaque pixel, un XOR entre chaque bit du filigrane et le bit tatouable de l'image originale fourni par le masque. Une marque de longueur 10 bits est créée par un générateur pseudo-aléatoire. Cette marque subit un nombre suffisant de décalages afin de combler tous les bits fournis par le masque perceptuel.

Notre première étude concernant la méthode d'extraction consistait à extraire l'ensemble des marques de 10 bits insérées, puis à calculer, pour chacune de ces séquences, la distance de Hamming entre les séquences binaires originales et extraites.

On définit la distance de Hamming sur \mathcal{F}^n entre $u = (u_1, u_2, \dots, u_n)$ et $v = (v_1, v_2, \dots, v_n)$ comme :

$$d(u, v) = |\{i = 1, \dots, n \mid u_i \neq v_i\}|. \quad (4.20)$$

On pourra considérer $d(u, v) = \underbrace{wt(xor(u, v))}_{u+v \in \mathcal{F}^n}$ où $wt(u^p) = \sum_{i=0}^{n-1} bit(p, i)$.

On aimerait aussi avoir une idée de la dispersion de C dans \mathcal{F}^n (plus les éléments seront éloignés les uns

des autres, meilleur sera notre code). On introduit donc :

$$d(C) = \min\{d(u, v) | u, v \in C, u \neq v\}. \quad (4.21)$$

Dans \mathcal{F}^n il est plus rapide de considérer :

$$d(C) = \min\{d(u + v, 0) | u, v \in C, u \neq v\} = \min\{wt(u) | u \in C, u \neq 0\} \quad (4.22)$$

Il suffit alors de $2^k - 1$ évaluations à la place de $(2^k - 1) 2^k - 1$.

Finalement, si l'on souhaite connaître la distance d'un élément $u \in \mathcal{F}^n$ à son plus proche voisin dans C :

$$d(u, C) = \min\{d(u, v) | v \in C\} \quad (4.23)$$

Cette approche permettait la détection des filigranes par portions de l'image. Une telle détection était surtout utile dans les cas d'attaques géométriques peu destructrices, (cropping, ou attaques localisées), mais ne s'avérait cependant pas très fiable face à des attaques dégradant globalement l'image (filtrages, codages JPEG, ...). De plus, comme nous l'avions introduit dans la section 1.1, certaines études [25] portant sur la capacité de détection ou d'extraction en fonction de la quantité d'information insérée, ont montré que des marques de longueurs importantes sont bien plus performantes que des marques courtes. C'est pourquoi, lors de l'utilisation de séquences de 10 bits de long, la faible longueur des marques fournissait un nombre important de fausses détections. C'est à dire que parfois le processus de détection pouvait détecter certaines marques sur des images non tatouées.

Dans le but d'améliorer le processus de détection, nous avons donc opté pour le calcul du coefficient d'intercorrélacion, que nous avons présenté équation 4.19. Les résultats de détection face à diverses attaques sont donnés dans le paragraphe 4.3.4.

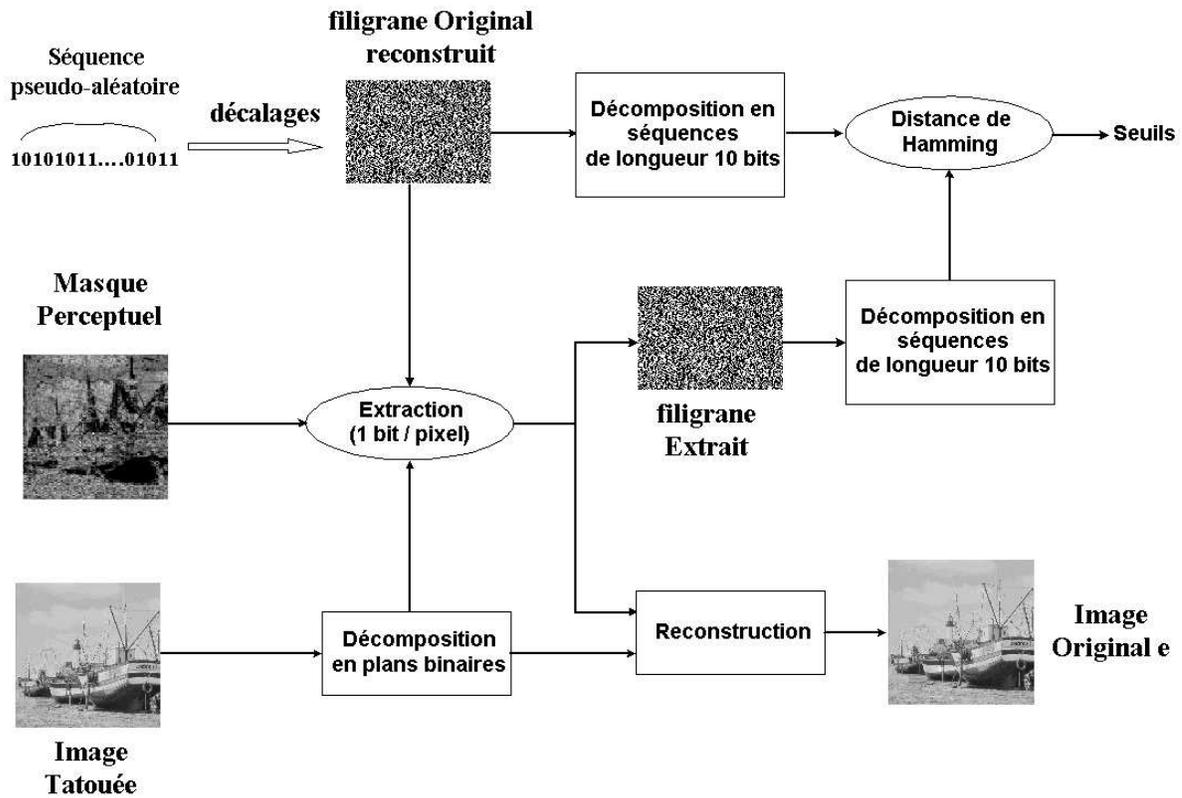


FIG. 4.11 – Schéma d'extraction

Lors de l'extraction, les données nécessaires à prouver l'authenticité des données sont : L'image tatouée et potentiellement attaquée (ou image potentiellement tatouée), le masque perceptuel, et la séquence pseudo-aléatoire du filigrane original, de longueur 10 bits.

Il est nécessaire de stocker le masque psychovisuel, car ce dernier ne peut pas être reconstruit. Effectivement, l'application du modèle visuel sur deux images différentes ne peut pas fournir deux masques identiques, donc l'application du modèle visuel sur l'image tatouée ne peut fournir un masque identique à celui obtenu sur l'image de référence. Cependant, si le stockage du masque est obligatoire, nous pouvons noter que ce dernier est beaucoup moins volumineux que l'image originale. Rappelons que le masque fournit un nombre de bits tatouables par pixel, il peut n'être codé que sur le nombre de bits maximum permis (en général entre 3 et 5 bits). Cette méthode fournit une solution intermédiaire en terme de capacité de stockage, entre les méthodes de tatouage nécessitant la conservation de l'image originale et les méthodes dites aveugles (sans image de référence).

Le masque nous permet de définir non seulement les pixels tatoués, mais aussi pour chaque pixel du masque,

le bit de l'image ayant été modifié. L'extraction de ces bits nous permet donc d'obtenir le filigrane extrait que nous comparons avec le filigrane original préalablement stocké par l'auteur.

4.3.4 Expérimentations

La figure 4.12 ci-dessous présente pour l'image bateau, le masque obtenu, ainsi que l'image tatouée par le schéma décrit précédemment (4.10). Nous constatons que la marque est adaptée au contenu de l'image, les régions uniformes sont incluses dans le masque et nous pouvons observer les directions des sous-bandes sélectionnées $\{(IV, 3) \text{ et } (V, 4)\}$ retranscrites dans le masque.

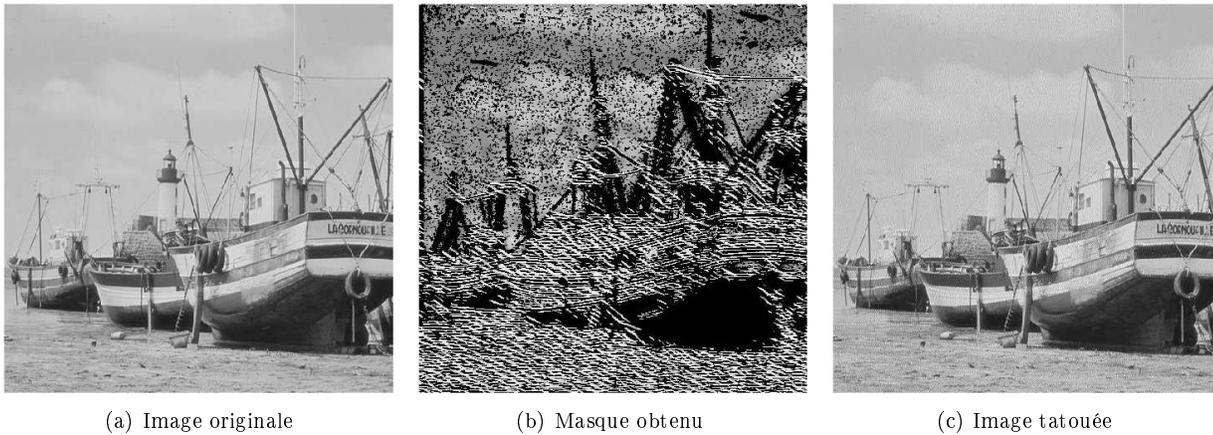
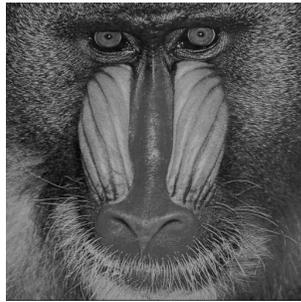


FIG. 4.12 – Images de l'algorithme de tatouage perceptuel

Dans l'image représentée figure 4.12(b), les pixels sombres représentent les zones non modifiables alors que les pixels de forte intensité représentent les zones permettant des modifications importantes.

Des tests de robustesse ont été effectués pour différentes images de référence. Les images tatouées ont subi plusieurs attaques, afin de déterminer une probabilité de récupération de la marque. La figure 4.13 présente quelques exemples des attaques que nous avons testées ici. Comme la plupart des méthodes d'insertion spatiales, cette méthode de tatouage par modification des bits de l'image ne permet pas de résultats significatifs face aux déformations géométriques. Effectivement, comme nous l'avons vu au chapitre 3, certaines déformations géométriques (rotations, décalages, suppression de lignes/colonnes, ...) opèrent une désynchronisation lors de la détection. Dans de tels cas, le filigrane n'est aucunement dégradé, mais il est simplement déplacé, la difficulté consiste donc à savoir où le chercher. De telles attaques ne sont donc pour l'instant pas prises en compte, nous pouvons envisager l'utilisation d'un schéma de 're-synchronisation' de la marque avant d'appliquer le processus de détection / extraction. Nous présenterons par la suite différents schémas de tatouage résistant à de telles

attaques.



(a) Modification de la palette



(b) Passe-bas (5×5)



(c) Passe-bas (5×5)



(d) Modification de la palette



(e) Rehaussement des niveaux de gris de 120%



(f) Addition de bruit 0.1%



(g) Addition de bruit 0.3%



(h) Suppression de 30 colonnes

FIG. 4.13 – Exemples d'attaques sur les images test

La figure 4.14 ci-dessous présente, pour un schéma d'insertion / détection dans le masque perceptuel, les résultats de la fonction d'intercorrélation entre la séquence originale et la séquence extraite face à diverses attaques.

Nous pouvons observer sur ces courbes, le pouvoir de détection

permis par le masque perceptuel. Précédemment, lors des tests de détection dans les zones uniformes, les contours ou toute l'image, chaque bit à tatouer était imposé et donc totalement indépendant du contenu de l'image. Une telle méthode pourrait entraîner une dégradation visible lors de l'insertion du filigrane. Ici, le masque garantit l'invisibilité en tout point et permet donc d'optimiser le tatouage autant dans l'aspect de l'invisibilité que dans l'aspect de la détection.

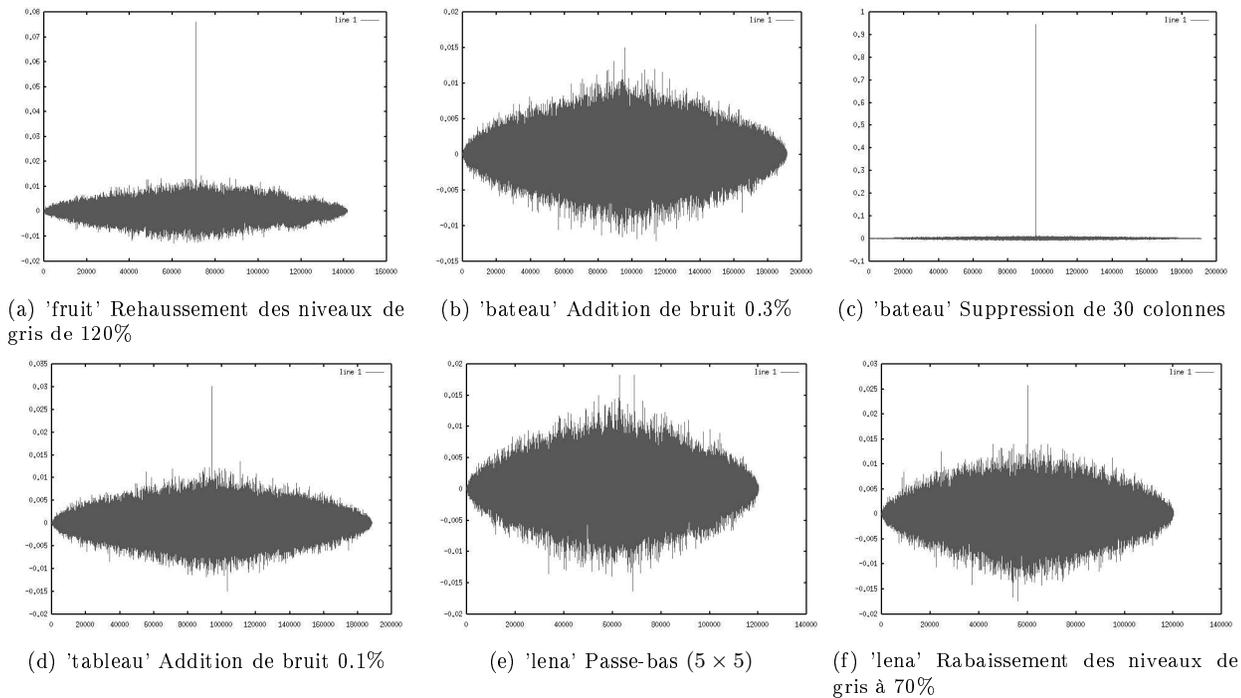


FIG. 4.14 – InterCorrélation entre la séquence originale et la séquence extraite pour différentes images

Nous pouvons aussi noter que le calcul du pourcentage total de similarités entre la séquence originale et la séquence extraite (tableau 4.5) permet aussi, mais dans une moindre mesure, d'anticiper sur le résultat de l'intercorrélation.

| % | Image tatouée | Passe-bas (3×3) | Passe-bas (5×5) | Rehaussement de contrastes | Jpeg q=40% | Jpeg q=50% |
|-------------------|---------------|----------------------------|----------------------------|----------------------------|------------|------------|
| Zones uniformes | 100 | 53 | 51 | 50 | 51 | 51 |
| Contours | 100 | 49 | 51 | 49 | 49 | 49 |
| Toute l'image | 100 | 52 | 50 | 50 | 50 | 50 |
| Masque perceptuel | 100 | 53 | 50 | 50 | 50 | 51 |

TAB. 4.5 – Pourcentage de similarités entre la séquence originale et la séquence extraite pour différentes zones

Nous avons pu constater les bons résultats fournis par cette utilisation simple du masque perceptuel. L'élaboration d'un schéma d'insertion dans les bits de poids faible de l'image doit passer par la sélection des bits à tatouer. Sans l'utilisation d'un masque perceptuel, la sélection des bits doit être faite de façon aléatoire. Une telle sélection des bits est très risquée car peut entraîner tout aussi bien des dégradations visibles, qu'un tatouage trop faible ne permettant pas de détection fiable.

Effectivement, une simple insertion dans les bits de poids faibles de l'image, sans l'exploitation d'un masque

perceptuel offrirait des résultats moins significatifs. Cependant, nous avons vu sur la figure 4.9 que pour une image comportant peu de zones uniformes, les résultats de détection sont encore à améliorer. Dans cette optique, nous présentons un schéma plus approfondi dans la section suivante, qui, comme nous le verrons, présente de meilleurs résultats pour les images ne comportant que peu de zones uniformes.

4.4 Le tatouage psychovisuel

Nous avons vu que la création d'un masque perceptuel permet d'envisager différentes méthodes d'insertion, et ceci dans différents domaines auxquels il suffit d'adapter le masque. Après avoir étudié la création ainsi que l'adaptation de masque J.N.D., nous avons choisi de créer une méthode complète de tatouage exploitant au mieux l'ensemble des critères psychovisuels détaillés précédemment. La figure 4.15 montre l'algorithme proposé. Ici, la sélection des sites, ainsi que le calcul des forces dans ces sites est similaire au travaux présentés ci-dessus (section 4.2).

4.4.1 Insertion de la marque

Rappelons que les valeurs de contraste seuil, $\Delta C_{i,j}$ ont été déterminés pour chaque sous-bande (i, j) , et dépendent d'une part de la sensibilité du SVH aux fréquences spatiales et d'autre part du contenu de l'image et des effets de masquage que ce contenu peut engendrer sur un contraste spatialement localisé et spectralement limité.

A présent, nous cherchons à insérer directement la marque dans les sous-bandes visuelles. Ici, nous ne faisons aucune distinction entre les zones uniformes et les contours, la marque est insérée sur toute l'étendue des sous-bandes traitées. Pour insérer la marque dans les sous-bandes, il nous faut définir la force maximale permise par le modèle psychovisuel. L'insertion d'un filigrane doit, pour chaque sous-bande, être majorée par la valeur du seuil de visibilité introduit précédemment $\Delta C_{i,j}$.

En supposant que le tatouage d'une sous-bande n'influe pas sur la luminance moyenne locale (les fréquences inférieures) et d'après l'équation 4.6, nous pouvons considérer qu'un tatouage par addition du seuil de contraste aurait la forme :

$$C_{i,j}^{t_M}(m, n) = C_{i,j}(m, n) + \Delta C_{i,j} = \frac{L_{i,j}^{t_M}(m, n)}{L_{i,j}(m, n)} \quad (4.24)$$

où $C_{i,j}^{t_M}(m, n)$ et $L_{i,j}^{t_M}(m, n)$ représentent respectivement le contraste tatoué et la luminance tatouée par la force maximale permise par le modèle.

Nous pouvons donc assurer que la luminance tatouée au maximum possible, en tout point (m, n) de la sous-bande (i, j) est donnée par l'équation :

$$L_{i,j}^{t_M}(m, n) = C_{i,j}^{t_M}(m, n) \times \overline{L_{i,j}}(m, n) \quad (4.25)$$

soit, en développant $C_{i,j}^{t_M}(m, n)$:

$$\left(\frac{L_{i,j}(m, n)}{\overline{L_{i,j}}(m, n)} + \Delta C_{i,j} \right) \times \overline{L_{i,j}}(m, n) = L_{i,j}(m, n) + \Delta C_{i,j} \times \overline{L_{i,j}}(m, n) \quad (4.26)$$

L'invisibilité est donc assurée par le modèle psychovisuel, si en tout point des sous-bandes, les valeurs des sous-bande tatouées restent inférieures à $L_{i,j}(m, n) + \Delta C_{i,j} \times \overline{L_{i,j}}(m, n)$.

L'insertion du filigrane, (figure 4.15) se fait alors spatialement sur des sous-bandes sélectionnées en fonction de leur activité. Effectivement, pour chaque couronne, la sous-bande d'écart type le plus important est sélectionnée en vue du tatouage. Le contraste de ces sous-bandes est alors calculé, puis tatoué comme décrit dans les équations 4.25 et 4.26. Une fois la force maximale allouée à la sous-bande, l'insertion se produit, de sorte à respecter le seuil supérieur fourni par le masque. La luminance de la sous-bande tatouée ($L_{i,j}^t(m, n)$) doit donc rester dans l'intervalle $[L_{i,j}(m, n) \leq L_{i,j}^t(m, n) \leq L_{i,j}^{t_M}(m, n)]$

La sous-bande tatouée peut alors être représentée spectralement, afin de permettre la reconstruction du spectre par simple addition des sous-bandes tatouées et de celles n'ayant subi aucune modification. Il est donc très important de noter que lors de ce processus, le contenu fréquentiel du filigrane doit rester confiné dans la sous-bande traitée, si ce filtrage n'est pas respecté, la marque rayonne sur toute l'étendue du spectre, ce qui affaiblira le pouvoir de détection.

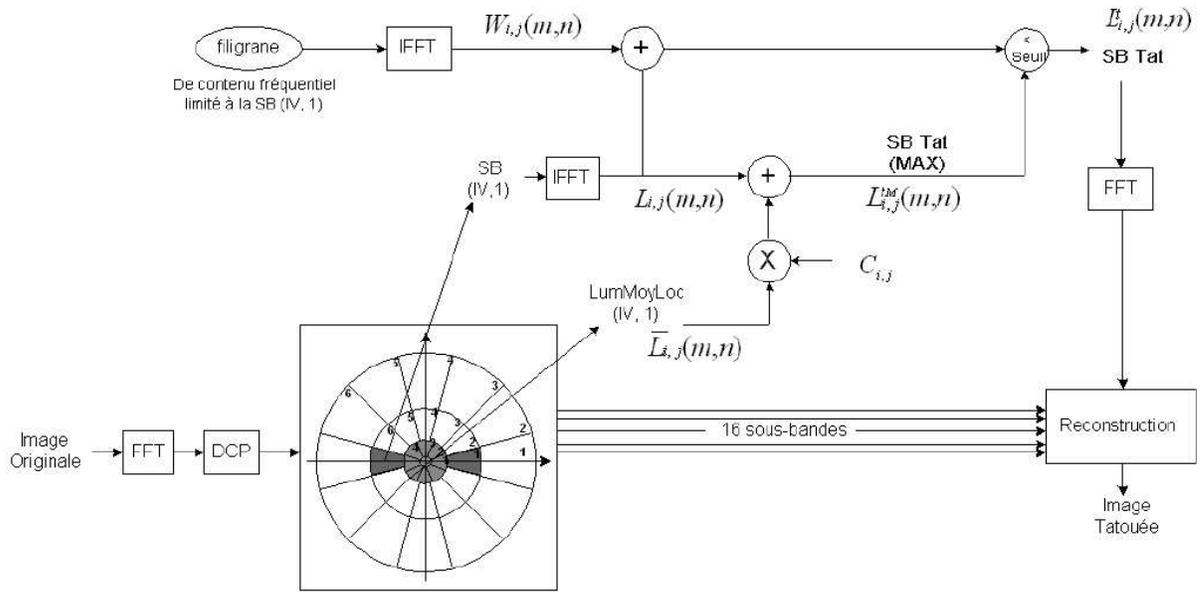


FIG. 4.15 – Schéma de tatouage psychovisuel, restreint à la sous-bande (II,1)

Nous pouvons observer sur la figure 4.15, les différentes étapes qui conduisent au tatouage de l'image. Dans un souci de simplification du schéma, nous présentons uniquement ici le traitement de la sous-bande (II,1). Nous constatons notamment deux étapes importantes, l'étape de tatouage maximal des contrastes, ainsi que le filtrage du filigrane avant l'insertion dans le masque. Concernant le traitement de la marque, comme nous le constatons sur le schéma, le filigrane est créé dans le domaine fréquentiel, il doit alors subir une transformation de Fourier inverse avant d'être inséré dans le contenu spectral des sous-bandes. Concernant l'insertion, comme nous l'avons vu précédemment, elle doit simplement respecter l'intervalle $[L_{i,j}(m,n) \leq L_{i,j}^t(m,n) \leq L_{i,j}^M(m,n)]$.

4.4.2 Processus de détection

Le processus de vérification de la marque utilise la théorie classique de la détection et ne requiert pas l'image originale. La fonction d'intercorrélation (equation : 4.19) entre le filigrane original $W_{i,j}(m,n)$ et la sous-bande potentiellement tatouée $L_{i,j}^{pt}(m,n)$ est calculée pour les sous-bandes d'énergie maximale, de la façon suivante. Pour simplifier cette équation, nous notons $W(i)$ le filigrane en une seule dimension, $L^T(i)$ est la luminance de la sous-bande tatouée, m_W et m_{L^T} représentent respectivement la moyenne de la marque et celle de la sous-bande tatouée.

$$r(d) = \frac{\sum_i [(W(i) - m_W) * (L^T(i - d) - m_{L^T})]}{\sqrt{\sum_i (W(i) - m_W)^2} \sqrt{\sum_i (L^T(i - d) - m_{L^T})^2}} \quad (4.27)$$

Le schéma ci-dessous présente le processus de détection aveugle employé. Le schéma de détection procède à l'extraction de la sous-bande potentiellement tatouée, puis calcule la fonction d'intercorrrelation entre la sous-bande spatiale et la représentation spatiale de la marque insérée.

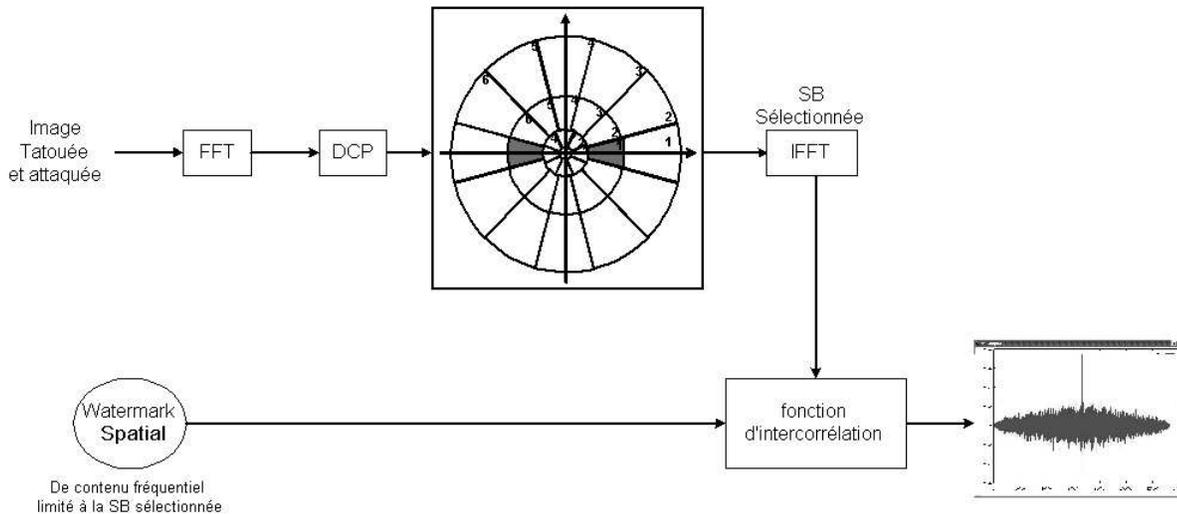


FIG. 4.16 – Schéma de détection de la marque

L'évaluation d'un coefficient de corrélation par sous-bande permet d'accroître la robustesse de l'algorithme. En effet, un filtrage passe-bas par exemple conservera les basses fréquences de l'image et supprimera les hautes fréquences conservant les coefficients de corrélation relatif aux basses fréquences et dégradant ceux des hautes fréquences. A l'inverse, une opération de rehaussement des contrastes permettra une meilleure détection de la marque dans les hautes fréquences.

Toutefois, il est important de noter que les tests de détection conduits sur les différentes fréquences radiales (couronnes), les sous-bandes de la couronne (IV) offraient de meilleurs résultats. Ceci s'explique par le manque de robustesse des hautes fréquences, d'une part et la faible capacité d'insertion des basses fréquences d'autre part. Le meilleur compromis semble donc être, comme nous l'avons vu, de tatouer les sous-bandes de moyennes fréquences.

Par exemple, pour l'image "lena", nous pouvons observer sur la figure 4.17 les résultats de détection de la marque sur les sous-bandes (III,1), (IV,2) et (V,I) face à un filtrage passe-bas de noyau 3×3 .

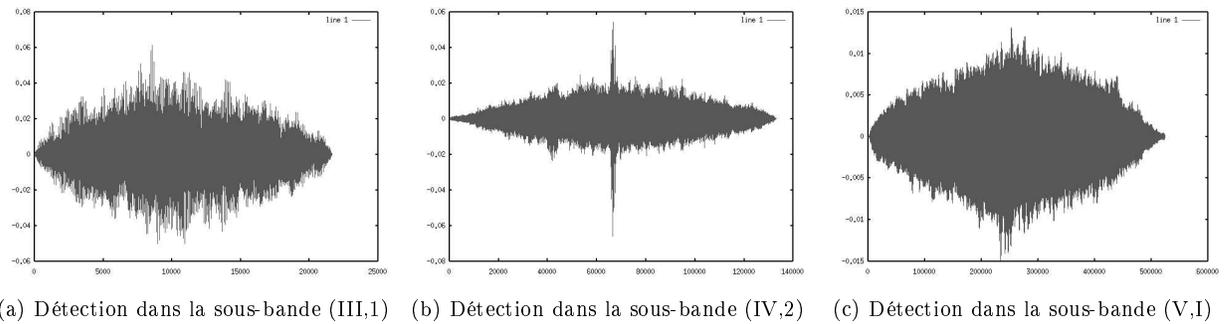


FIG. 4.17 – Exemples de détection en fonction de la fréquence radiale

4.4.3 Expérimentations

Si l'invisibilité de la marque était ici garantie par le modèle perceptuel, il a néanmoins fallu tester la robustesse du schéma vis-à-vis de plusieurs attaques. L'insertion du filigrane s'opérant ici dans le domaine spatial, ce schéma n'est pas adapté aux attaques géométriques. Nous étudierons par la suite une mise en œuvre adaptée fréquemment résistante à de telles attaques, que nous ne prenons pas en compte pour l'instant. Cependant, il est toujours possible comme nous l'avons vu dans le chapitre 3 d'envisager une compensation des déformations géométriques.

Nous présentons dans cette section, pour trois images ('bateau', 'singe' et 'tableau') de contenu fréquentiel différent, les expériences conduites en terme de robustesse de la marque vis-à-vis de 9 attaques différentes.

La figure 4.18 présente les résultats de l'intercorrélation pour l'image 'bateau', face à un filtrage passe-bas de noyau 3×3 , un découpage de l'image (cropping), un flou gaussien, et une modification de la palette. Ces attaques sont variées et présentent des résultats différents. Cependant, pour les quatre attaques citées, la fonction d'intercorrélation permet la détection de la marque.

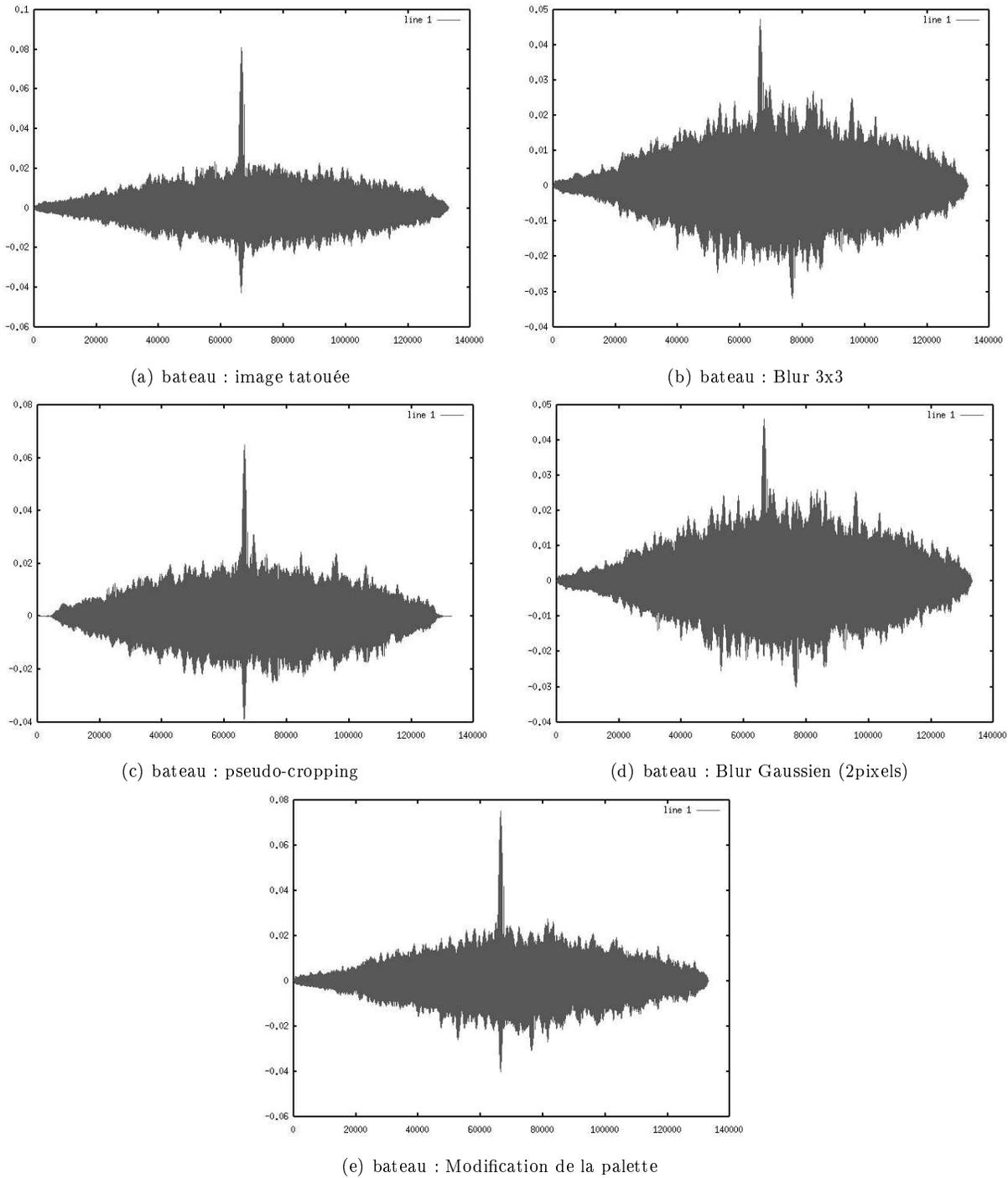


FIG. 4.18 – fonctions d'intercorrélation du filigrane avec les sous-bandes tatouées, image 'bateau'

La figure 4.19 présente, respectivement les résultats d'une addition de bruit, une quantification, un rehaussement de contrastes et une rotation de 5° . Comme nous l'avons prévu, la rotation ne permet pas la

synchronisation et donc ne permet pas non plus la détection de la marque. Cependant, les trois premières attaques, bien que parfois sévères, permettent toutefois la détection du filigrane.

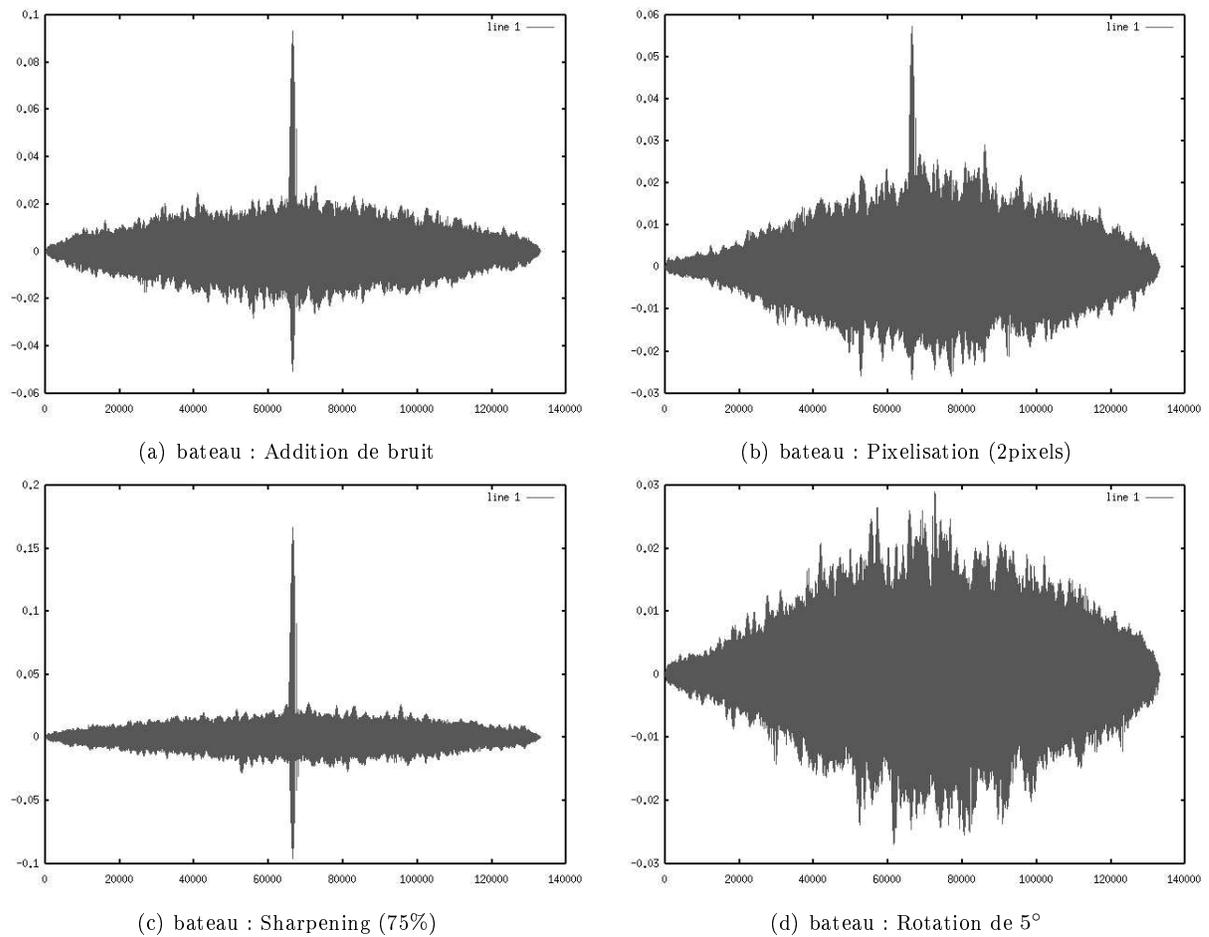


FIG. 4.19 – fonctions d'intercorrélation du filigrane avec les sous-bandes tatouées, image 'bateau'

Les figures 4.20 et 4.21 ci dessous, présentent les résultats de la fonction d'intercorrélation pour l'image singe. Nous pouvons observer que cette méthode de tatouage dans les sous-bandes psychovisuelles permet de très bons résultats de détection pour des images de contenus variés, contrairement au schéma présenté précédemment qui offrait essentiellement de bons résultats lorsque l'image comportait des zones uniformes (résultats présentés section 4.3.4).

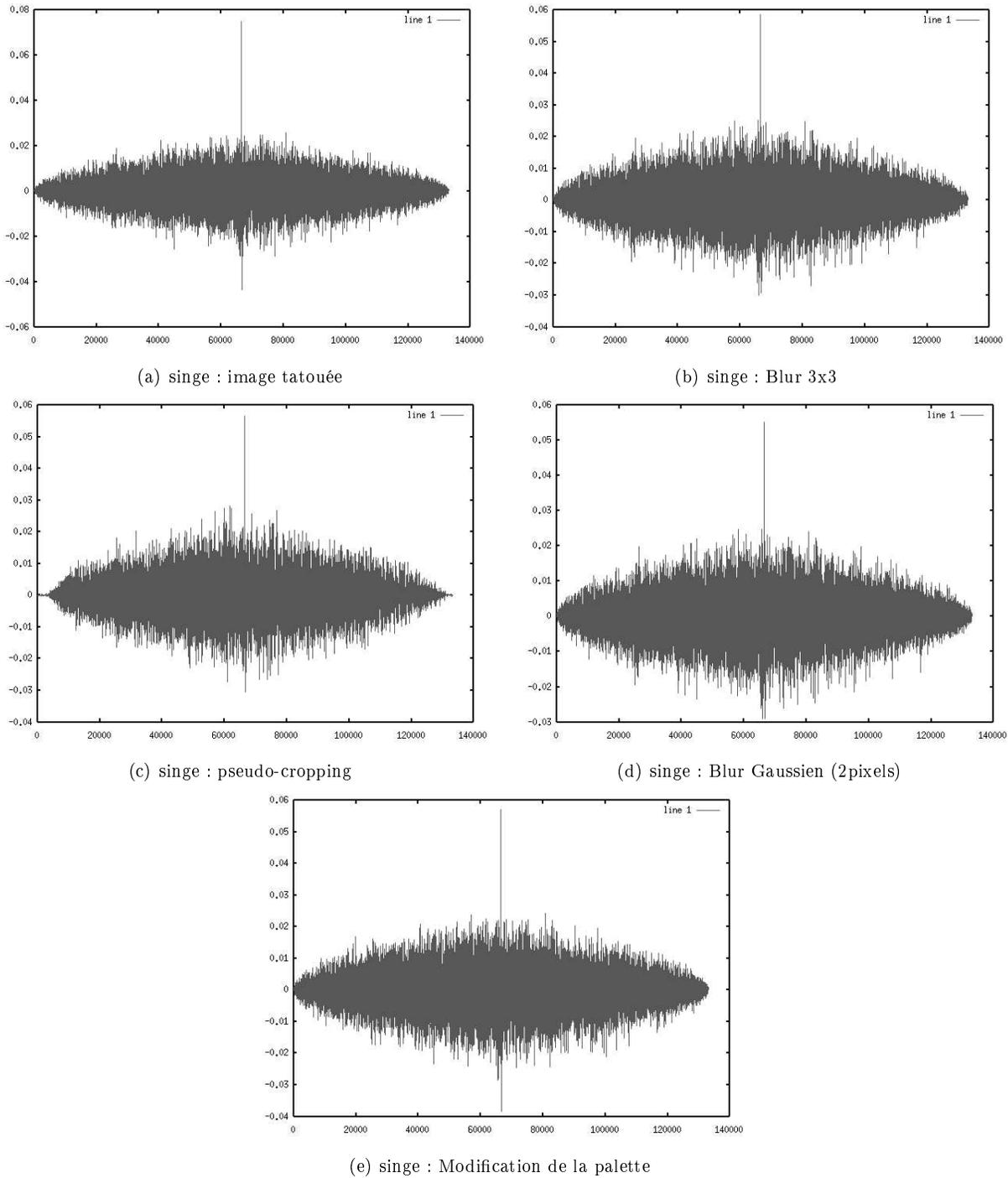
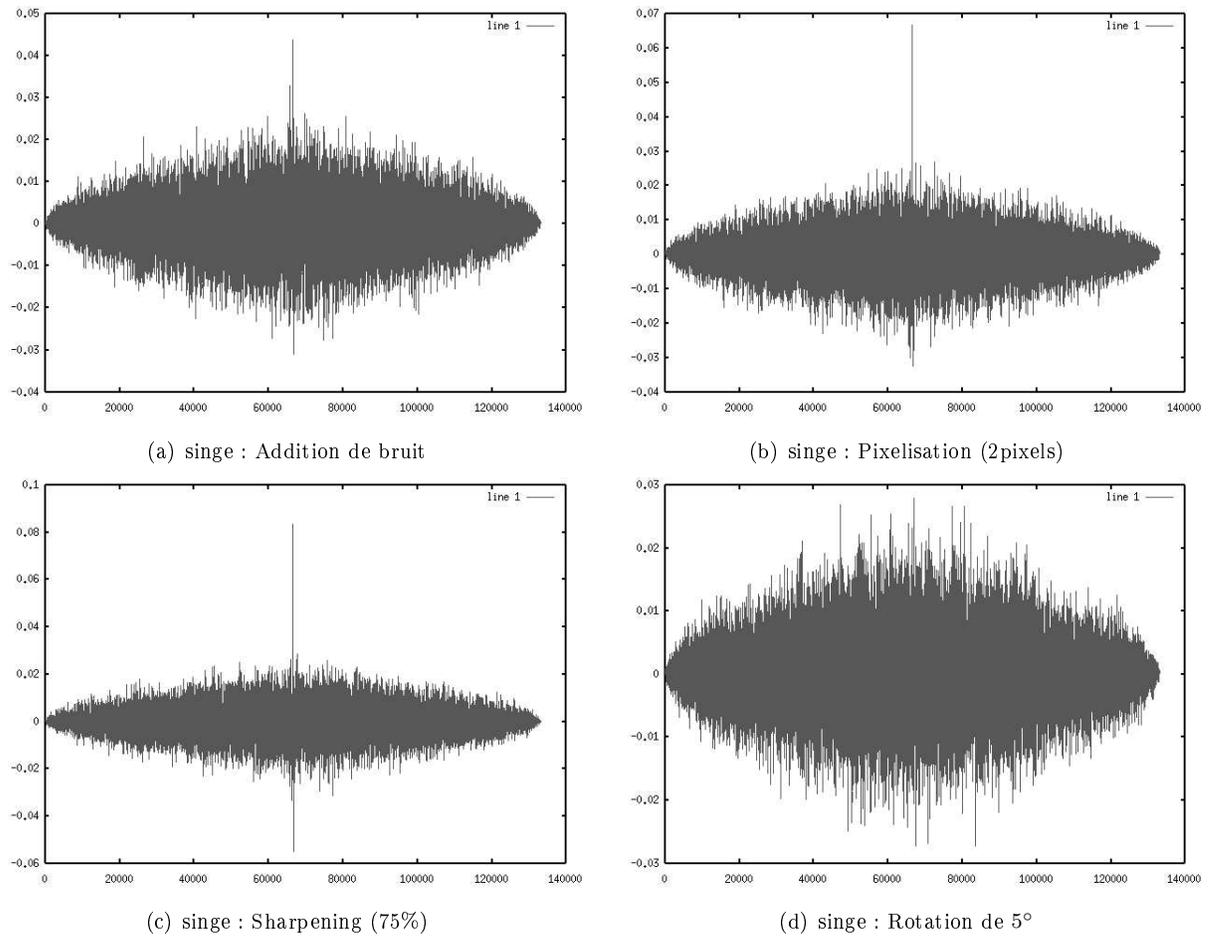
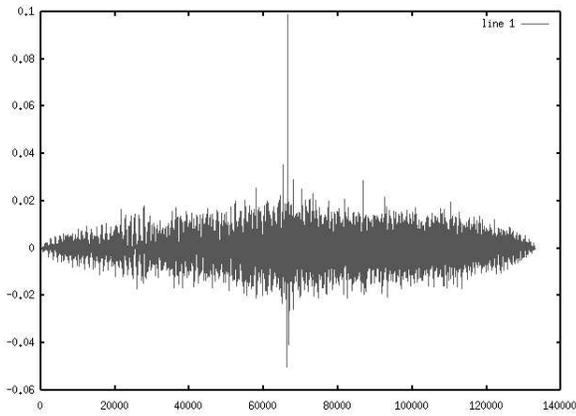


FIG. 4.20 – fonctions d'intercorrélation du filigrane avec les sous-bandes tatouées, image singe

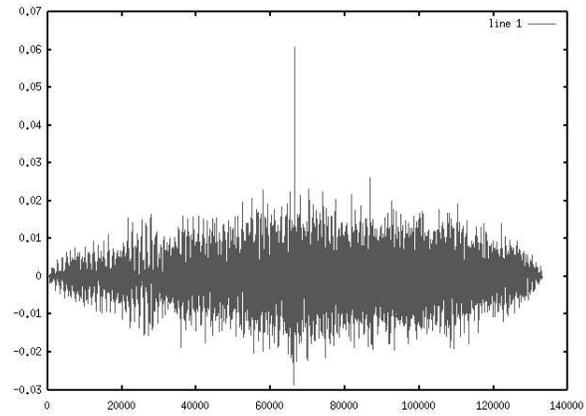
Encore pour l'image 'singe', nous présentons ci dessous, tout comme pour la figure 4.19, les résultats de détection d'une addition de bruit, d'un rehaussement de contrastes, et d'une rotation de 5° .

FIG. 4.21 – fonctions d'intercorrélation du filigrane avec les sous-bandes tatouées, image *singe*

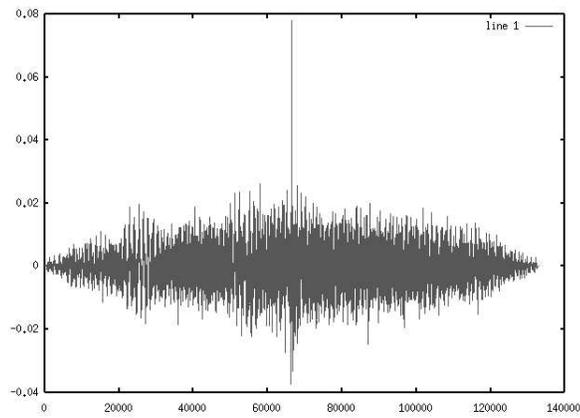
Observons à présent les résultats pour l'image *tableau* face aux mêmes attaques que précédemment :



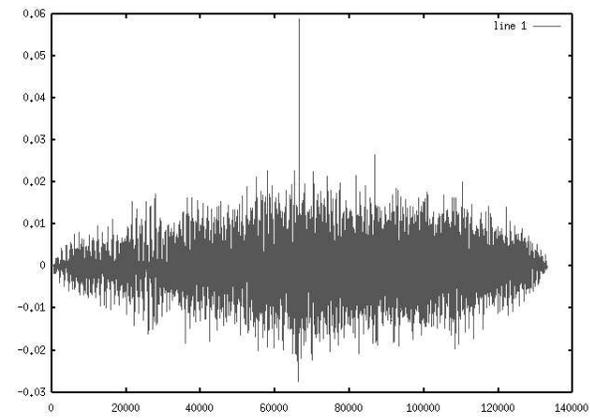
(a) tableau : image tatouée



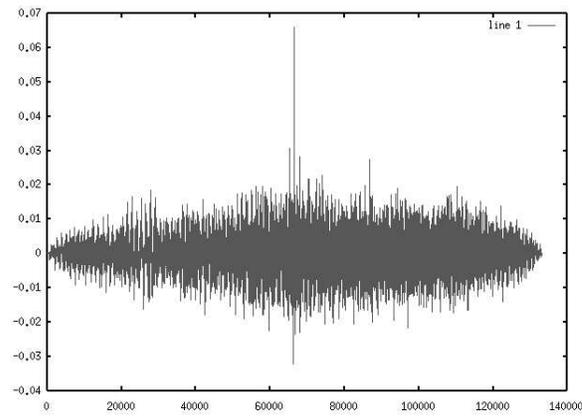
(b) tableau : Blur 3x3



(c) tableau : pseudo-cropping



(d) tableau : Blur Gaussien (2pixels)



(e) tableau : Modification de la palette

FIG. 4.22 – fonctions d'intercorrélation du filigrane avec les sous-bandes tatouées, image 'tableau'

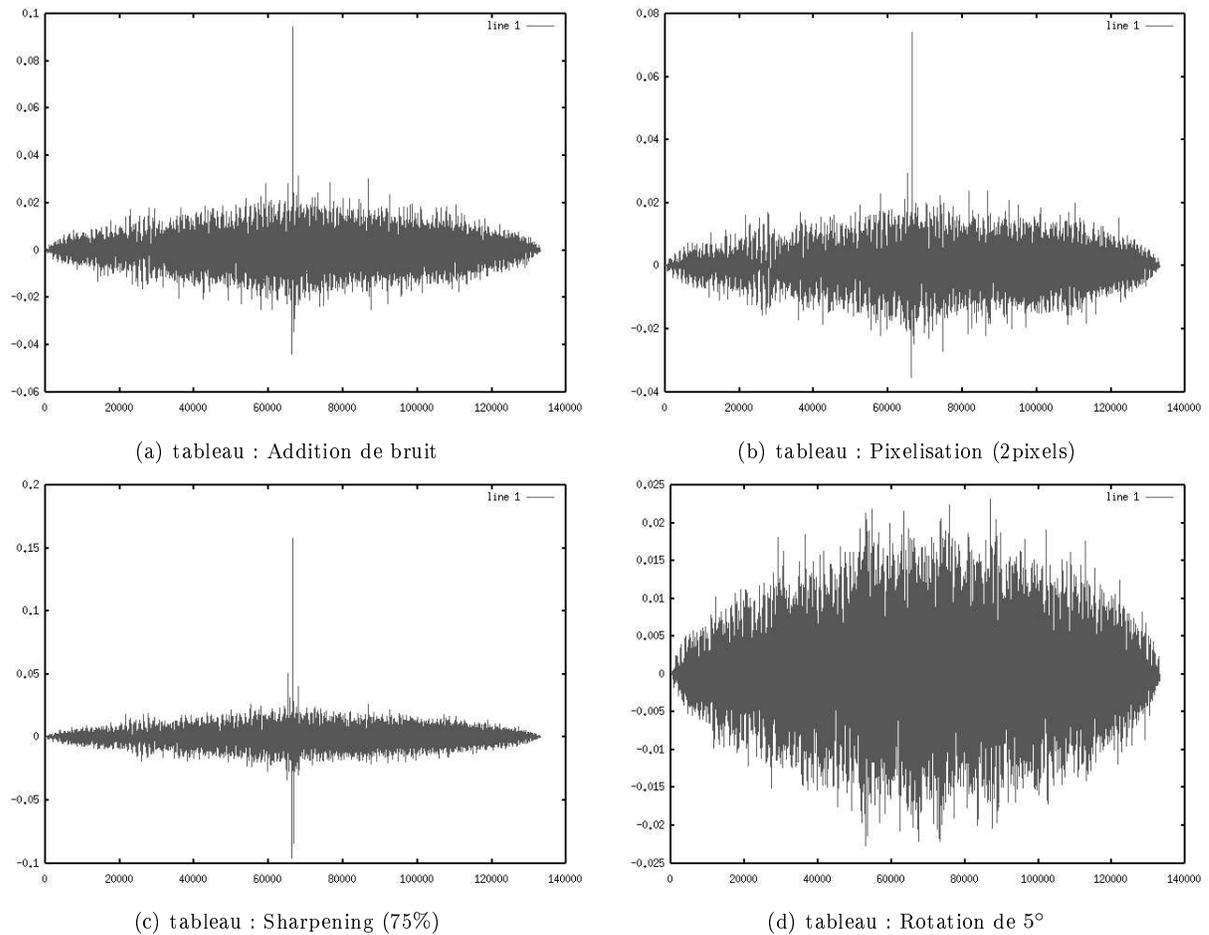


FIG. 4.23 – fonctions d’intercorrélacion du filigrane avec les sous-bandes tatouées, image tableau

Nous constatons au vu de ces résultats, que le comportement du schéma de détection ne varie pas de façon importante en fonction du contenu fréquentiel des images traitées.

Une étude de ces figures confirme le manque de robustesse de la marque pour des déformations géométriques (rotation de 5°), mais démontre cependant les bons résultats obtenus par l’utilisation du modèle perceptuel. Les lacunes des méthodes de tatouage spatial en termes de robustesse sont bien connues et proviennent d’un problème de synchronisation de la marque lors de la détection (cf. chapitre 3).

Nous avons testé le schéma de détection face à des attaques très variées, telles que des additions de bruit, du filtrage passe-bas, des rehaussements de contraste, et même une attaque géométrique n’opérant pas de désynchronisation. L’algorithme de détection s’est trouvé fructueux face à toutes ces attaques, le seul problème subsistant est celui des attaques géométriques. L’objectif des chapitres suivants est de résoudre ce problème.

4.5 Conclusion

Le tatouage d'image utilisant un masque psychovisuel (JND), comporte différents avantages et inconvénients. Le masque JND permet d'atteindre les forces maximales en tout point de l'image et donc, d'obtenir le meilleur compromis entre l'invisibilité et la robustesse.

L'avantage principal de l'utilisation des masques perceptuels est la simplicité d'adaptation aux différents domaines. Nous avons pu voir des résultats intéressants de conversion du masque psychovisuel spatial aux domaines de la DCT ou des ondelettes.

Cependant, de telles méthodes de tatouage ne sont pas résistantes face aux attaques géométriques. Effectivement, les attaques telles que des découpages d'images, des zoom ou des rotations, ne permettent pas de synchroniser la marque dans le masque psychovisuel et donc de la détecter.

Nous avons pu observer dans ce chapitre les bons résultats fournis par l'utilisation du masque perceptuel sur un schéma simple de tatouage des bits, puis nous avons pu constater l'efficacité d'un schéma de tatouage spatial des sous-bandes face à diverses attaques non géométriques.

Sans compter l'apport indéniable des aspects psychovisuels sur la robustesse des schémas de tatouage, l'utilisation des critères visuels permet de garantir l'invisibilité de la marque en toute circonstance. Certains travaux emploient des critères empiriques pour insérer la marque dans l'image, d'autres préfèrent assurer l'invisibilité en insérant une marque d'amplitude trop faible. Comme nous l'avons constaté, l'étude du système visuel humain permet de définir en tout point de l'image les forces maximales applicables au filigrane. Ceci permet donc d'insérer la marque au seuil de visibilité, garantissant ainsi le meilleur compromis entre l'invisibilité et la robustesse.

Chapitre 5

Tatouage Mojette

Sommaire

| | | |
|------------|--|------------|
| 5.1 | Introduction | 142 |
| 5.2 | Utilisation des fantômes dans une méthode de tatouage | 142 |
| 5.2.1 | Introduction | 142 |
| 5.2.2 | Schéma de tatouage | 143 |
| 5.2.3 | Gestion des clefs | 147 |
| 5.2.4 | Conclusion | 151 |
| 5.3 | Exploitation de la redondance : tatouage de flux Mojette | 152 |
| 5.3.1 | Introduction | 152 |
| 5.3.2 | Brouillage des plans binaires | 152 |
| 5.3.3 | Reconstruction des plans binaires | 157 |
| 5.3.4 | Conclusion | 161 |
| 5.4 | Une méthode de tatouage spatial adaptable à la stéganographie | 162 |
| 5.4.1 | Introduction | 162 |
| 5.4.2 | Algorithme de protection des droits d'auteurs | 162 |
| 5.4.3 | Algorithme de Stéganographie | 164 |
| 5.4.4 | Résultats | 166 |
| 5.4.5 | Conclusion | 177 |
| 5.5 | Conclusion | 177 |

5.1 Introduction

Nous proposons dans ce chapitre trois méthodes de tatouage exploitant les propriétés de la transformation Mojette. Ces schémas mettent à profit différents aspects de la transformation Mojette, que ce soit la notion de fantôme ou la redondance. Dans ce chapitre, par le biais de la Mojette, nous explorerons les liens qui peuvent exister entre le tatouage, la stéganographie, ou encore la cryptographie. Nous étudierons différents aspects connexes au tatouage parmi lesquels la gestion des clefs, le tatouage multiple ou encore la complexité opératoire.

La première méthode exploite les propriétés des fantômes Mojette pour une application de tatouage d'images médicales. Ce schéma étant adapté aux contraintes imposées par le traitement d'images médicales, nous proposons une méthode de tatouage invisible, réversible avec une extraction aveugle de la marque. Ce schéma permet un tatouage multiple concernant d'une part les informations du praticien et d'autre part, les informations relatives au patient. Du fait de cette application, la gestion des clefs est détaillée dans le paragraphe 5.2.3.

Le second schéma exploite essentiellement la redondance de la transformation Mojette. Ce schéma est entièrement paramétrable et peut s'adapter aux contraintes imposées. Ce schéma peut tout aussi bien être employé pour du tatouage invisible que pour de la cryptographie, là encore, nous proposons un schéma réversible, à extraction aveugle. Une des applications envisagée pour cette méthode est de proposer une accessibilité graduable aux données. Ici, l'insertion d'une très faible quantité d'information peut entraîner une très grande complexité d'extraction pour toute personne ne possédant pas la clef adéquate.

Le troisième schéma présenté ici, exploite les deux propriétés citées dans les schémas précédents : les fantômes et la redondance. Ici, nous proposons un schéma permettant aussi bien la protection des droits d'auteurs que la stéganographie. La gestion de la clef permet ainsi de basculer d'une application à l'autre. La robustesse du schéma sera évaluée dans la section 5.4.4. Tout comme la méthode citée précédemment, celle-ci est particulièrement adaptée aux problèmes de transmission de l'information.

5.2 Utilisation des fantômes dans une méthode de tatouage

5.2.1 Introduction

L'application du tatouage en imagerie médicale est très intéressante car ce mode de tatouage ne comporte pas les mêmes contraintes que les autres applications, telles que la protection des droits d'auteurs. Dans le domaine médical, un médecin ne peut se permettre de porter un diagnostic sur une image modifiée. Toute

méthode de tatouage appliquée sur des images médicales doit donc permettre une extraction totale de la marque. Étant donné que la marque doit être extraite avant tout diagnostic, l'invisibilité n'est donc pas la préoccupation principale lors de l'élaboration d'une méthode de tatouage de données médicales.

L'utilisation des images médicales numériques est régie par le standard DICOM [70] (Digital Imaging and Communication in Medicine). Ce standard international est très complet, il prend en compte les différents aspects du traitement de l'information numérique, tels que la transmission, les formats, la sécurité, la compression ou encore l'archivage.

Actuellement, la seule sécurité mise en œuvre dans le standard DICOM est une empreinte de chaque image stockée pour savoir si l'image a subi une dégradation. Le point faible du standard DICOM est la nécessité de stocker de façon sûre toutes les images. Il est important de noter que les images médicales sont généralement codées sur 12 bits/pixel ou 16 bits/pixel, et sont donc très encombrantes. Nous proposons donc ici une méthode de tatouage invisible, réversible avec une extraction aveugle, comportant la gestion des clefs par un tiers de confiance. De plus, la méthode proposée permet aussi le contrôle d'intégrité obligatoire dans une telle application.

5.2.2 Schéma de tatouage

Dans le paragraphe 5.2.2.1, on explicite comment le tatouage est mis en œuvre dans l'image médicale en construisant un fantôme à l'aide d'un ensemble de projections et comment on enlève ce tatouage en utilisant les paramètres Mojette (paragraphe 5.2.2.2) comme information disponible chez le tiers de confiance. En enrichissant un peu cette information détenue par le tiers, on montre comment rechercher et corriger des dégradations éventuelles effectuées dans l'image. Les aspects de gestion des différentes clefs sont discutés dans le paragraphe 5.2.3.

Une caractéristique originale de ce schéma est le double tatouage. Il est à noter que de multiples tatouages sont possibles. L'autre originalité réside dans le rôle de ce double tatouage. Pour une fois, le médecin et le patient sont considérés comme deux acteurs "de même poids" pour la gestion de l'image médicale. Ce fait est très important si on veut effectuer un stockage réparti des données médicales, où, par exemple chaque patient conserve sur CD-ROM les images de ses examens. Ce type de fonctionnement permet au médecin de continuer à confier les images au patient, comme c'était le cas avec les radiographies. Ceci est une façon commode et très efficace de gestion du dossier patient notamment dans le cas d'une urgence.

5.2.2.1 Insertion des marques

On supposera disposer d'une image rectangulaire $P \times Q$ et d'un ensemble de directions de projections discrètes $S_I = \{(p_i, q_i), i \in I\}$ tel que $P_I < P$ et $Q_I < Q$. On forme alors le fantôme correspondant à S_I par suite de dilatations, comme nous l'avons vu au chapitre 2. Sur un ensemble de B plans de bits les moins significatifs de l'image, on ajoute par une opération XOR, le fantôme localisé sur des points de l'image particuliers qui seront générés par le praticien et le patient et fournis au tiers de confiance sous forme chiffrée. Ainsi le patient utilisera son secret (partagé avec le tiers de confiance) pour chiffrer cette double suite de points de l'image $P_{IPat} = \{(k_i, l_i), i \in I_{Pat}\}$ et d'angles discrets $S_{IPat} = \{(p_i, q_i), i \in I_{Pat}\}$, et le médecin chiffrera sa propre suite de points de l'image P_{IMed} et d'angles discrets S_{IMed} avec sa clé privée afin de signer son tatouage.

Le schéma global d'insertion de la marque est présenté sur la figure 5.1. Il est à noter que ce schéma d'insertion doit être répété deux fois lors de la création de toute image médicale, de sorte à insérer dans l'image à la fois une information patient et une information médecin; chacune des deux parties ne devant pas connaître les données de l'autre.

L'insertion s'effectuant par une opération binaire, réversible et commutative, le processus d'extraction de la marque n'est donc pas contraint à respecter l'ordre employé lors de l'insertion.

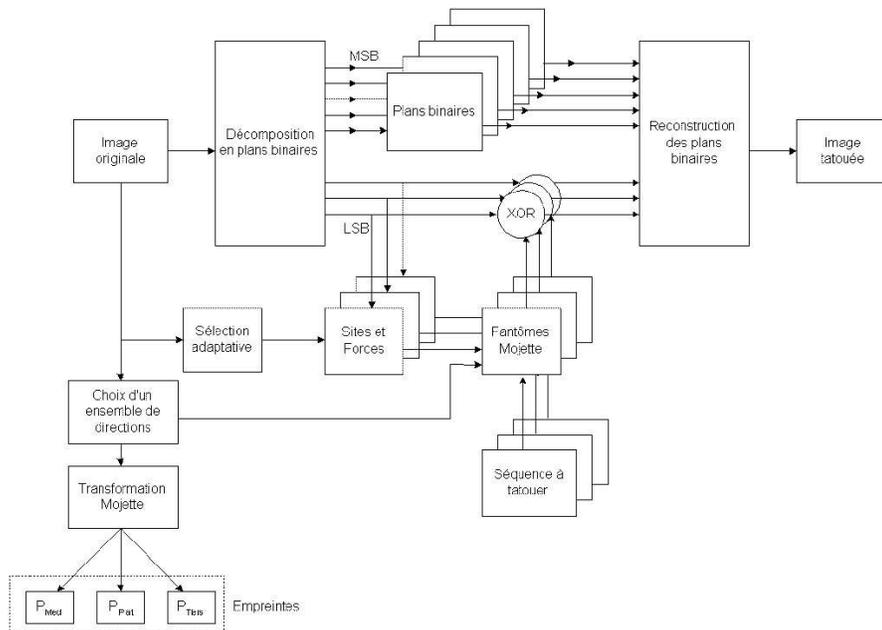


FIG. 5.1 – Schéma de tatouage par fantômes

Comme décrit sur la figure 5.1, l'image originale est décomposée en plans binaires (typiquement 12 ou 16 plans binaires). Les plans de poids forts (MSB) sont laissés intacts, alors que les plans binaires de poids faible (LSB) seront tatoués.

Lors du processus de tatouage une sélection de zones est effectuée directement sur l'image originale, cette dernière peut être basée sur des critères psychovisuels (luminances, fréquences spatiales) ou en fonction des régions d'intérêt (du point de vue médical mais non diagnostic) de l'image. Un fantôme Mojette est alors construit, comme explicité dans la section 2.3.2. Une séquence cryptographique binaire comportant des informations sur le patient (ou sur le médecin) est alors insérée dans les plans binaires par un simple OU exclusif binaire (XOR). Finalement, une reconstruction des plans binaires fournit l'image tatouée.

5.2.2.2 Extraction de la marque

L'extraction de la marque ne peut être exécuté qu'avec la connaissance du secret conjoint du patient et du médecin ou bien de celle détenue par le tiers. Ayant cette connaissance, il suffit de refaire un XOR du fantôme généré lors de la phase de tatouage pour obtenir l'image originale. Évidemment, la dégradation volontaire ou non de l'image ne modifie pas l'extraction du tatouage. Il est donc nécessaire de disposer d'une information supplémentaire comme par exemple une projection de l'image originale pour chacune des deux parties : lorsque cette projection appartient à l'ensemble ayant servi pour la construction des fantômes respectifs du praticien et du patient, la projection de l'image tatouée dans cette direction est égale à celle de l'image originale. C'est donc une information plus facile à utiliser qu'une empreinte qui donnerait dans un cas classique deux résultats différents sur l'image originale et sur l'image tatouée. Il est également important que la projection choisie pour l'empreinte d'une des deux parties n'appartiennent pas à l'ensemble choisi par l'autre partie, afin de permettre une vérification croisée.

Il y a deux cas de figures à considérer selon que l'image a subi ou non des dégradations.

Hypothèse 1 : Image non modifiée

La figure 5.2 présente le processus d'extraction de la marque. Le schéma est aveugle (l'image originale n'est pas requise lors de l'extraction) et inversible (il permet de récupérer l'image originale). Ce schéma doit être répété par le médecin et par le patient. L'image tatouée est décomposée en plans binaires, la clef du patient et celle du médecin sont utilisées conjointement dans le but de recréer le fantôme Mojette, la séquence cryptographique binaire est alors insérée dans ce dernier.

Le calcul du XOR entre les fantômes des LSB et les plans binaires tatoués fournit au système les plans binaires originaux de l'image, puis la reconstruction (avec les plans binaires non modifiés) permet d'obtenir

l'image originale. L'authenticité de l'image ainsi obtenue peut être vérifiée par un simple calcul des projections Mojette conservées par le patient et par le médecin (ces projections faisant office d'empreintes).

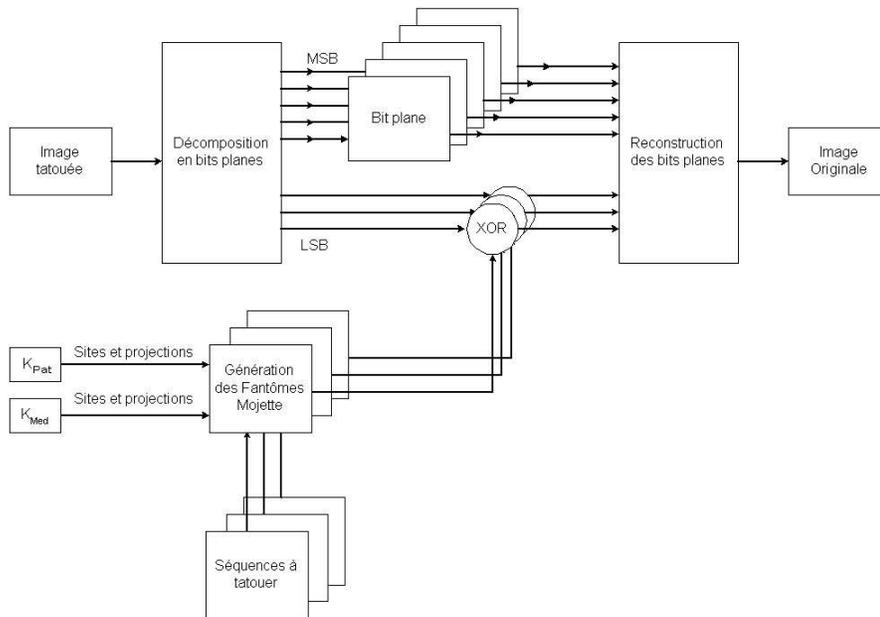


FIG. 5.2 – Schéma d'extraction des fantômes

Hypothèse 2 : Image modifiée

Dans l'hypothèse où l'image tatouée a été attaquée (empreintes ne coïncidant pas), dans le standard mis en place par DICOM, l'image est considérée comme non exploitable par le praticien. Dans le schéma proposé ici, en cas de modification de l'image, le tiers est invoqué. Celui-ci doit stocker une certaine quantité d'informations nécessaire à une récupération de données utiles.

Lors de l'examen par le couple patient - médecin d'une image, chacune des deux parties ayant donné son secret permet l'extraction de la marque.

Chacune des deux parties vérifie également que la projection de l'image détatouée est conforme à celle qu'il a enregistré. Si le tiers est invoqué, il doit disposer aussi d'informations complémentaires comme la projection qui sert d'empreinte pour le patient (resp. le praticien) pour la recherche de dégradations éventuelles de l'image tatouée. Une attaque malicieuse pourrait être mise en œuvre par le praticien et le patient réunis. Cette attaque consisterait à créer et insérer un fantôme à l'aide de l'ensemble de directions en leur possession. Ceci leur permettrait de dégrader l'image, tout en obtenant les mêmes empreintes. Pour faire face à une telle attaque, il suffit au tiers de confiance de stocker une troisième projection calculée pour une troisième direction et inconnue autant du patient que du médecin. Le tiers découvrant une anomalie sur cette projection pourra

alors découvrir la coalition mise en place par le patient et le médecin.

A partir de l'image obtenue après l'extraction, le tiers calcule l'ensemble disponible de projections et peut comparer le résultat avec les projections stockées. Sur les projections de différences DP , (la transformation Mojette étant linéaire) le tiers repère ainsi les bins en défaut et localise les pixels défaillants par une rétro-projection de ces bins dans une image D : les pixels défaillants sont à l'intersection d'une ligne de chacune des projections.

En supposant que le nombre de ces modifications est faible, il suffit d'essayer toutes les combinaisons possibles de valeurs des pixels de D donnant $DP = 0$. En ajoutant les images D ainsi déterminées à l'image obtenue après extraction, on forme les images corrigées.

Il peut ne pas être nécessaire de reconstruire l'image entière en cas de litige mais uniquement une certaine zone d'intérêt sur laquelle porte un diagnostic. Les données du patient et du médecin ayant été disposées dans l'image de façon à couvrir une région d'intérêt prédéfinie, lors du processus d'extraction expliqué ci-dessus, l'image récupérée n'est pas l'originale.

Il est important de retenir que dans un contexte de tatouage d'images médicales, les attaques n'ont pas pour objectif une détérioration de la marque insérée, mais plutôt l'addition ou la soustraction d'une pathologie. De ce fait, les attaques seront probablement localisées dans un espace restreint et donc plus aisément décelables par le dispositif proposé ici.

5.2.3 Gestion des clefs

Nous avons pu constater l'intérêt du tiers de confiance lors du processus d'extraction. Ce dernier doit conserver un certain volume d'informations. Outre les projections servant d'empreintes, le tiers doit conserver les différentes clefs d'insertion. Ces clefs, peu volumineuses sont indispensables lors de l'extraction. Le schéma 5.3 résume leur utilité.

Chaque clef est constituée de trois paramètres : un premier définissant la position des fantômes dans l'image, le second paramètre donne les directions des éléments structurants ayant établis ces fantômes. Finalement, la clef comporte une projection de l'image originale (ou tatouée) selon l'une des directions du fantôme. Cette projection aura une fonction d'empreinte.

Nous constatons donc sur la figure 5.3 que le médecin peut à l'aide de sa clef (K_{Med}), reconstituer son filigrane, l'extraire, puis vérifier l'authenticité de son image tatouée I_{TM} selon les directions qu'il détient.

Le patient peut indépendamment du praticien effectuer les mêmes opérations avec sa clef K_{Pat} . Il peut extraire son filigrane F_P pour obtenir et vérifier l'authenticité de son image tatouée I_{TP} .

Le patient et le praticien réunis peuvent alors extraire les deux marques et, ainsi obtenir l'image originale.

En cas de litige, le tiers de confiance peut être invoqué. Ce dernier peut reproduire à lui seul les opérations d'extraction du patient et du médecin. Comme nous l'avons vu précédemment, il peut également, à l'aide de projections supplémentaires, détecter les zones de l'image ayant éventuellement subi des dégradations.

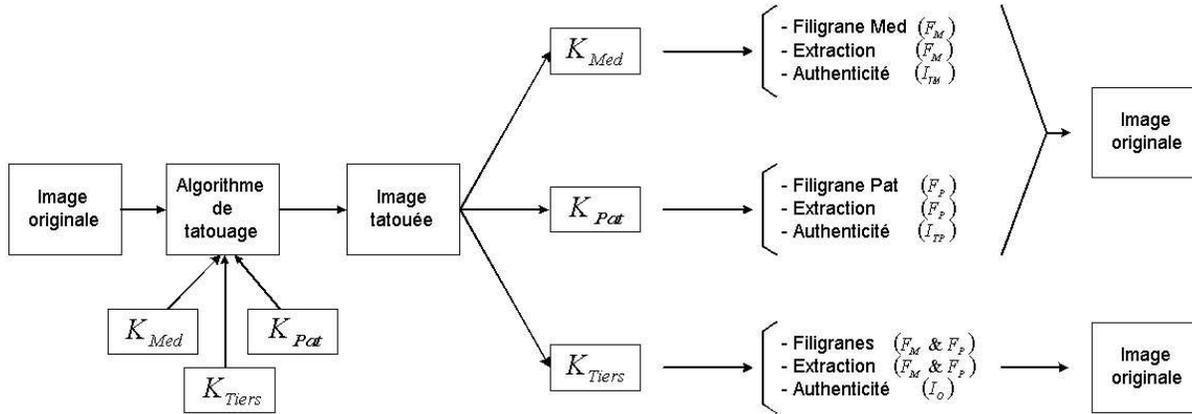
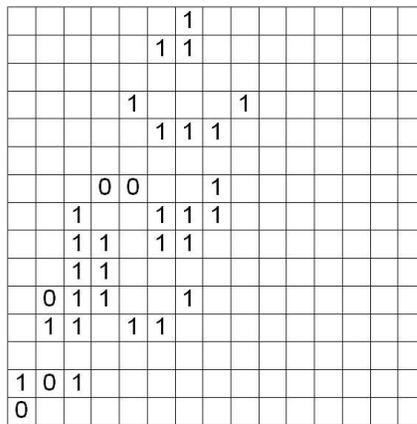
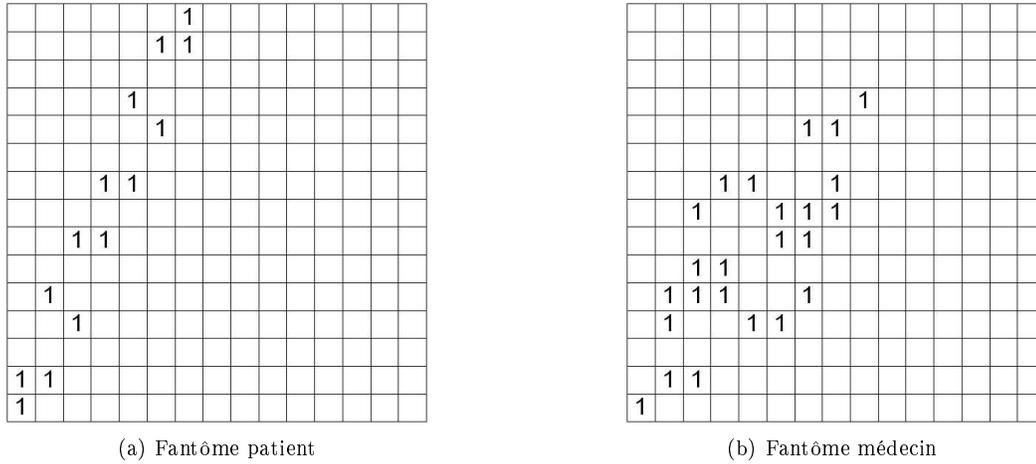


FIG. 5.3 – Gestion des clefs

La figure 5.4(a) représente un exemple de fantôme Mojette (appartenant au patient) construit à partir de l'ensemble de projections : $\{1,1\}$, $\{0,1\}$, $\{1,2\}$, $\{1,3\}$, $\{3,7\}$.

La figure 5.4(b) représente un fantôme Mojette (appartenant au médecin) construit à partir de l'ensemble de projections : $\{1,1\}$, $\{1,3\}$, $\{1,4\}$, $\{3,2\}$, $\{2,1\}$.

La figure 5.4(c) représente le filigrane final inséré dans l'image. Ce dernier est le produit d'un XOR entre le "fantôme patient" et le "fantôme médecin" (figures 5.4(a) et 5.4(b)).



(c) Résultat du XOR entre les deux fantômes

FIG. 5.4 – Exemples de fantômes

La figure 5.5(a) représente un agrandissement d’une image de fantôme générée à partir des six directions suivantes :

$$\{(10,21), (10,31), (10,41), (20,21), (20,31) \text{ et } (20,41)\}$$

Sur la figure 5.5(b), ce fantôme a été dupliqué sur les quatre sites suivants :

$$\{(250,90), (380,160), (220,350), (420,300)\}$$

avant d’être inséré par XOR dans les plans binaires LSB d’une image 512×512 .

A partir de ces informations, nous pouvons estimer la complexité du processus d’extraction, nous avons :

$$\sum |p| = 90 \text{ et } \sum |q| = 186$$

Considérons le cas où un seul fantôme est inséré dans l’image, selon l’équation 2.4 du paragraphe 2.3.2 :

$$(512 - 186) \times (512 - 90) = 326 \times 422 = 137572 \text{ positions.}$$

Donc, si nous considérons l’insertion de deux fantômes, composés du même ensemble d’éléments structurels

à deux pixels (ES2P), le nombre de positions possibles est majoré par l'équation :

$$C_n^k = \frac{n!}{k! \times (n-k)!}, \quad (5.1)$$

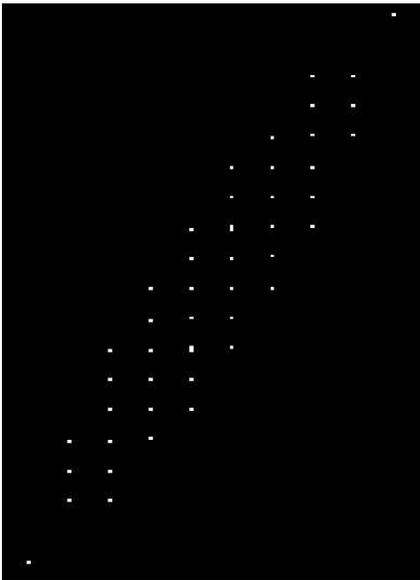
avec k =nombre de fantômes à insérer et n =nombre de positions possibles par fantôme.

soit pour deux fantômes, ici : $k = 2$ et $n = 137572$, $C_n^k = 9,4630.10^9$,

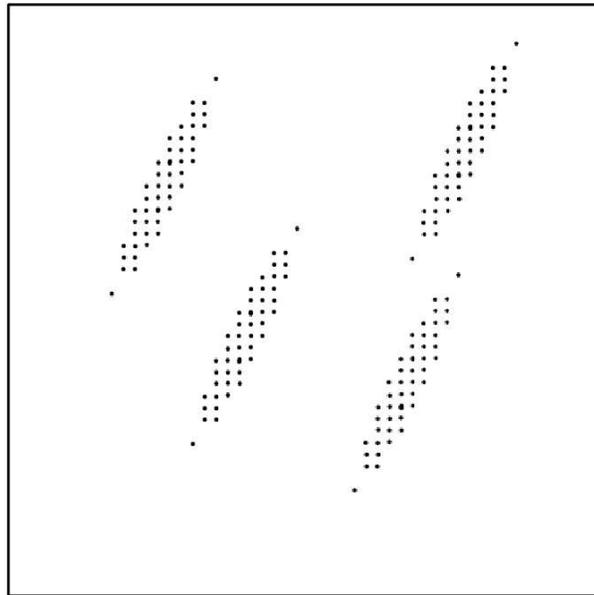
et pour trois fantômes : $C_n^k = 4,3394.10^{14}$

Il est évident que la complexité augmente lorsque le nombre de fantômes augmente, ce schéma de tatouage est donc paramétrable en fonction du degré de sécurité souhaité.

Nous pouvons noter que dans le cas cité ci-dessus, (4 sites d'insertion et six directions) la taille de la clef comportant les paramètres du fantôme est de 97 octets seulement. La taille des projections que devront conserver indépendamment le patient, le médecin et le tiers en vue de la vérification de l'intégrité dépend des directions choisies. La taille minimum des projections correspond aux directions (1,0) ou (0,1), ce qui représente pour une image de taille 512×512 , une projection de 512 bins codés chacun sur un seul bit (projection XOR), soit une clef de 512 bits.



(a) Zoom sur le fantôme Mojette



(b) Insertion multiple du fantôme

FIG. 5.5 – Fantômes Mojette

Pour ces conditions, le tableau 5.1 inventorie les données à conserver par le tiers, le praticien et le patient. Le volume d'information stockées est donc faible comparé au volume initial.

| | Informations conservées | Total |
|---------------------------|---|---------------------|
| Patient | Sites patient (776 bits), Projections du fantôme patient (0,1) : (512 bits), 1 projection (512 bits) | (1800 bits) |
| Médecin | Sites Médecin (776 bits), Projections du fantôme médecin (1,0) : (512 bits), 1 projection (512 bits) | (1800 bits) |
| Tiers de confiance | Sites (Patient + Médecin) (1552 bits), Projections des fantômes (Patient+Médecin) (1024 bits) 3 projections (1,2), (2,1), (1,3) : (1536+1536+5120 = 8192) | (10768 bits) |

TAB. 5.1 – Volume d'information stocké

La somme totale des quantités d'informations conservées par les trois parties est donc égale à 14368 bits soit moins de 2 kilo octets. Dans une architecture DICOM, l'image originale est intégralement conservée, soit, encore pour une image de taille 512×512 en supposant une profondeur de 16 bits : $512 \times 512 \times 16 = 4194304$ bits soit plus de 524 kilo octets.

5.2.4 Conclusion

Nous avons décrit dans cette section, un schéma de tatouage aveugle et réversible exploitant la notion de fantômes Mojette. Ici, l'invisibilité est uniquement assurée de façon empirique, par insertion dans les bits de poids faible de l'image. L'utilisation d'un masque perceptuel tel que celui décrit au chapitre précédent pourrait être envisagé, cependant, dans le contexte présenté ici, l'invisibilité n'est pas la préoccupation principale.

Le schéma proposé ici présente plusieurs avantages par rapport à la norme DICOM actuellement mise en place. L'atout principal est la faible quantité d'informations stockées. La notion de partage des clefs et de tatouage multiple est aussi très intéressante et peut permettre de faire face à différentes fraudes. Finalement, la possibilité de réellement insérer de l'information dans l'image (nom du praticien, nom du client, nom de l'hôpital ...) peut tout aussi bien être appréciée.

Pour une application autre que l'imagerie médicale, on pourrait souhaiter apporter plus de robustesse au schéma proposé ici, il est alors possible de l'adapter à un domaine fréquentiel. L'insertion des fantômes se produirait par exemple directement dans le spectre de Fourier de l'image, par simple multiplication par $+1$ ou -1 de certains coefficients, on obtiendrait le spectre tatoué. La détection du filigrane s'effectuerait alors par calcul de la fonction d'intercorrélation entre le spectre potentiellement tatoué et le fantôme potentiellement inséré.

5.3 Exploitation de la redondance : tatouage de flux Mojette

5.3.1 Introduction

Dans cette section, nous présentons un schéma de tatouage extensible à la cryptographie [6]. La redondance procurée par la transformation Mojette va être mise à profit dans ce schéma. La particularité de ce dernier, consiste à insérer la marque dans les bins non utilisés des projections Mojette et de rétroprojeter dans l'image des bins corrompus. Deux options sont alors possibles pour reconstruire l'image : utiliser les bins non modifiés (gardés secrets) ou utiliser n'importe quels bins. La première option conduit à l'image originale, la seconde conduit à une version dégradée.

Pour réaliser ce processus, une première étape consiste à effectuer une première transformation Mojette directe, suivie d'une transformation inverse, afin de déterminer les bins utilisés lors de la rétroprojection, les bins restants seront alors modifiés. La clef du tatouage permet de définir les bins n'ayant pas été corrompus par la marque, et de cette façon, permet de rétroprojeter les bins originaux dans l'image et d'obtenir l'image originale. Pour un utilisateur ne possédant pas cette clef, les bins corrompus seront immanquablement rétroprojetés, produisant une version dégradée de l'image. Ce schéma permettant de réaliser différents niveaux de visibilité du tatouage, il peut être employé dans des applications gérant différents niveaux de sécurité des données.

Une autre application potentielle d'un tel schéma de tatouage est la méthode dite de 'try and buy' qui permet au client une prévisualisation du produit (des données) avant de choisir de l'acheter.

Enfin, cette méthode peut être utilisée en complément d'une technique de tatouage quelconque, dans un but de sécurisation des données. Un pirate se trouverait face à deux difficultés, tout d'abord décoder l'image reçue, qui peut, comme nous le verrons être sécurisée avec différents niveaux de confidentialité. La seconde difficulté reste celle de la suppression de la marque elle même.

5.3.2 Brouillage des plans binaires

Ce schéma un peu particulier se place dans un contexte de transmission de l'information. L'algorithme présenté permet plusieurs applications, parmi lesquelles nous trouvons la protection des droits d'auteurs, le contrôle de l'intégrité, le contrôle d'accès aux données et la cryptographie. Les personnes autorisées (possédant la clef de décryptage) peuvent récupérer, après transmission des projections, l'image originale. Toutefois, un pirate ne pourra quant à lui reconstruire l'image originale, mais seulement une version dégradée. Cette version dégradée pourra être soit une image tatouée, soit une image cryptée. Le schéma global de tatouage de flux Mojette est présenté ci-dessous.

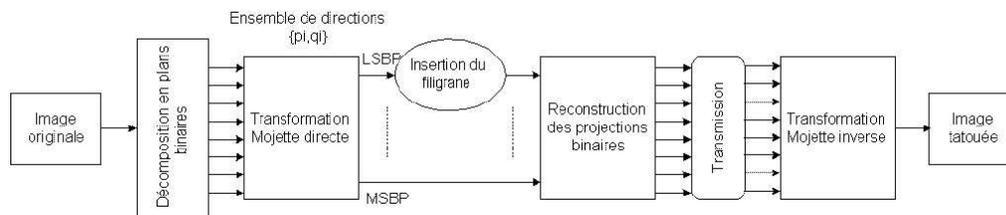


FIG. 5.6 – Schéma global de tatouage de flux

L'originalité d'un tel schéma réside dans le fait que dans l'hypothèse d'une récupération des données par une personne autorisée, l'image ne subit aucune modification lors de son parcours sur le réseau. Cependant, dans la seconde hypothèse, le pirate récupère des données, dans lesquelles se trouve l'image originale, mais ne sait pas comment reconstruire le « puzzle ».

L'image originale est décomposée en plans binaires, après une transformation Mojette directe sur chacun de ces plans, les 3 LSB sont soumis à la procédure de tatouage de projections, fournissant une clef privée, alors que les MSB sont laissés intacts. Une transmission de ces projections permettra alors au détenteur de la clef, de reconstituer les plans binaires corrompus et donc, de récupérer l'image originale.

Observons plus avant la procédure de tatouage des projections. Comme nous l'avons constaté précédemment lors de la présentation de la transformation Mojette, la transformation inverse s'effectue par rétroprojection des bins à correspondance 1-1.

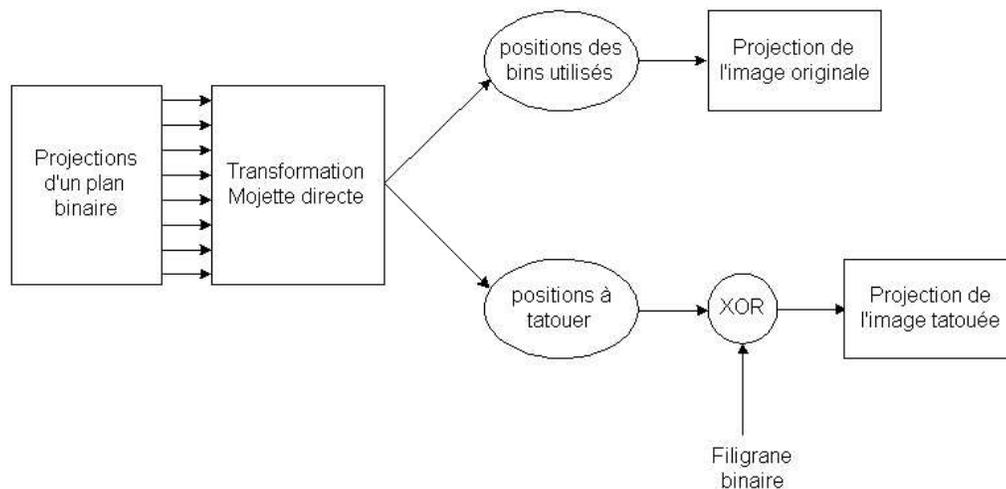


FIG. 5.7 – Schéma d'insertion du filigrane

Étant donné la redondance qu'apporte la transformation Mojette, il paraît évident que nous possédons plus

de bins lors de la réception de l'information qu'il n'en faudra lors de la transformation inverse. Il est cependant important de noter qu'il existe une multitude de combinaisons de ces bins pour reconstruire l'image originale.

Le principe de base consiste donc à tatouer (par addition d'une séquence pseudo-aléatoire) uniquement les bins que nous n'utiliserons pas pour rétroprojeter. La clef que nous devons alors transmettre sera constituée des positions des bins restés inchangés (ou inversement de ceux à éviter qui seront moins nombreux).

Nous pouvons observer figure 5.8 un exemple de Mojette inverse pour une image de taille 3×3 avec l'ensemble de directions $\{(1, 0), (1, 1), (-1, 1)\}$.

Les bins à correspondance univoque sont représentés dans les rectangles. Les bins qui seront rétroprojetés à l'étape suivante sont représentés en gras et soulignés. De plus, les pixels reconstruits ainsi que les positions des bins rétroprojetés sont représentés en grisés. Dans la figure 5.8, la valeur des bins importe peu, seule la position des bins est représentée.

Ici, les bins sont numérotés de 1 à 13, à l'issue de la reconstruction, les positions des bins rétroprojetés permettant la reconstruction de l'image originale sont les suivantes :

[6, 10, 1, 13, 5, 8, 11, 9 et 2]

Si nous utilisons ces positions de bins comme clef de tatouage, il est évident que le système n'est pas optimisé : nous avons une clef composée de neuf éléments pour une image de neuf pixels. Effectivement, la transformation Mojette inverse utilise obligatoirement un bin pour reconstruire un pixel.

La solution optimale serait d'utiliser un générateur pseudo-aléatoire fournissant la position des bins à rétroprojeter. Seule la graine du générateur pseudo-aléatoire sera nécessaire pour récupérer l'image originale.

Cette séquence aléatoire ne devra pas fournir les positions des bins tels que nous les avons numérotés dans la figure 5.8, mais ne doit donner que des positions de bins à correspondance univoque. Le générateur pseudo-aléatoire doit donc fournir les positions des bins dans le 'buffer' contenant les correspondances *pixel* \leftrightarrow *bin*.

Dans le cas présent, un générateur pseudo-aléatoire fournissant la séquence suivante :

[2, 1, 3, 3, 4, 2, 1, 3 et 3]

permettrait de rétroprojeter tout d'abord le bin correspondant à la seconde position indiquée dans la liste des correspondances univoques (rétroprojection du sixième bin), puis le bin correspondant à la première position dans la liste suivante (rétroprojection du dixième bin), etc...

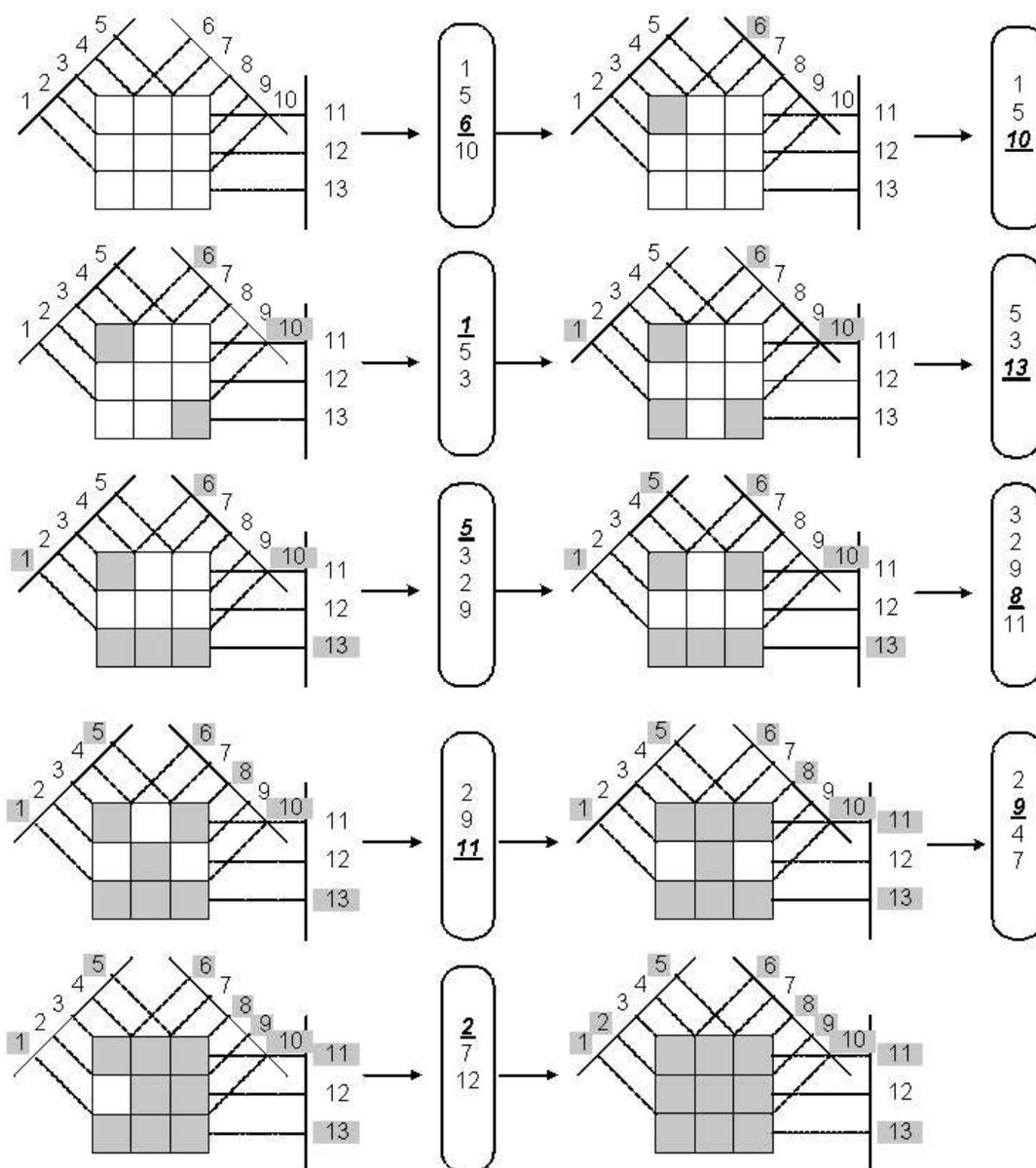


FIG. 5.8 – exemple d'ordre choisi de rétroprojections

Nous avons donc bien 9 bins rétrojetés dans une image de 9 pixels, sur un ensemble de 13 bins. Les 4 bins restants (de positions : 3, 4, 7 et 12) sont alors tatoués et la séquence utilisée pour définir l'ordre des rétroprojections est conservée comme clef privée du tatouage.

Selon un tel exemple, la complexité (nombre de combinaisons possibles) semble restreinte, cependant, comme nous allons le voir dans le tableau 5.2, une faible redondance peut être suffisante pour obtenir une complexité très importante.

Les chances d'un pirate de rétroprojeter les bins non modifiés sont majorées par l'équation 5.1 que nous rappelons ici :

$$C_n^k = \frac{n!}{k! \times (n-k)!} \quad (5.2)$$

Le tableau 5.2 ci-dessous, montre le comportement des C_n^k pour une image de taille 4×4 en fonction de différentes directions de projections, soit différentes redondances.

| Nb Tot bins | 17 | 19 | 20 | 22 | 23 | 24 | 25 | 28 | 30 | 34 |
|-------------|----------------|-------|----------------|-------|----------------|-------------------------|-------------|-----------------------------------|-------------------------|----------------------------------|
| C_n^k | 17 | 969 | 4845 | 74613 | 245157 | $7,3.10^5$ | $2,04.10^6$ | $3,04.10^7$ | $1,45.10^8$ | $2,20.10^9$ |
| Directions | {0,1} {1,3} | {1,5} | {1,1} {1,3} | {1,6} | {1,2} {1,3} | {1,0} {1,1} {1,3} | {1,7} | {1,1} {1,0} {-1,1} {1,2} | {1,1} {1,2} {1,3} | {1,0} {1,1} {1,2} {1,3} |
| Redondance | 0.062 | 0.187 | 0.250 | 0.375 | 0.437 | 0.500 | 0.562 | 0.750 | 0.875 | 1.125 |

TAB. 5.2 – C_n^k et Redondance en fonction des directions de projection

La Figure 5.9 montre respectivement l'image originale ainsi que l'image obtenue après une reconstruction illicite. Ici, le tatouage s'est effectué sur les 3 LSB de l'image.



(a) Image originale



(b) Image tatouée, 3LSBP brouillés

FIG. 5.9 – tatouage d'image par rétroprojection corrompue

Dans un tel cas, l'image modifiée ne présente que de très faibles dégradations, une application en tatouage d'images peut alors être envisagée. Nous verrons par la suite que la prise en compte des plans binaires de poids supérieur permet l'adaptation du schéma au cryptage des données.

5.3.3 Reconstruction des plans binaires

Lors de l'extraction, il suffit au possesseur de la clef, de ne rétroprojeter que les bins non modifiés pour obtenir l'image originale. Toutefois, toute personne non autorisée se verra obligée de rétroprojeter des bins erronés et n'obtiendra donc qu'une version dégradée de l'image, comme nous l'avons vu sur la Figure 5.9(b).

Ce schéma de tatouage est modulable, tant par l'aspect de l'invisibilité (tatouage de 2 LSB, ne pas tatouer tous les bins non utilisés...) que par l'aspect robustesse (ajout de projections tatouées lors de la transmission).

Les images représentées ci-dessous présentent les résultats du brouillage des différents plans binaires de l'image.

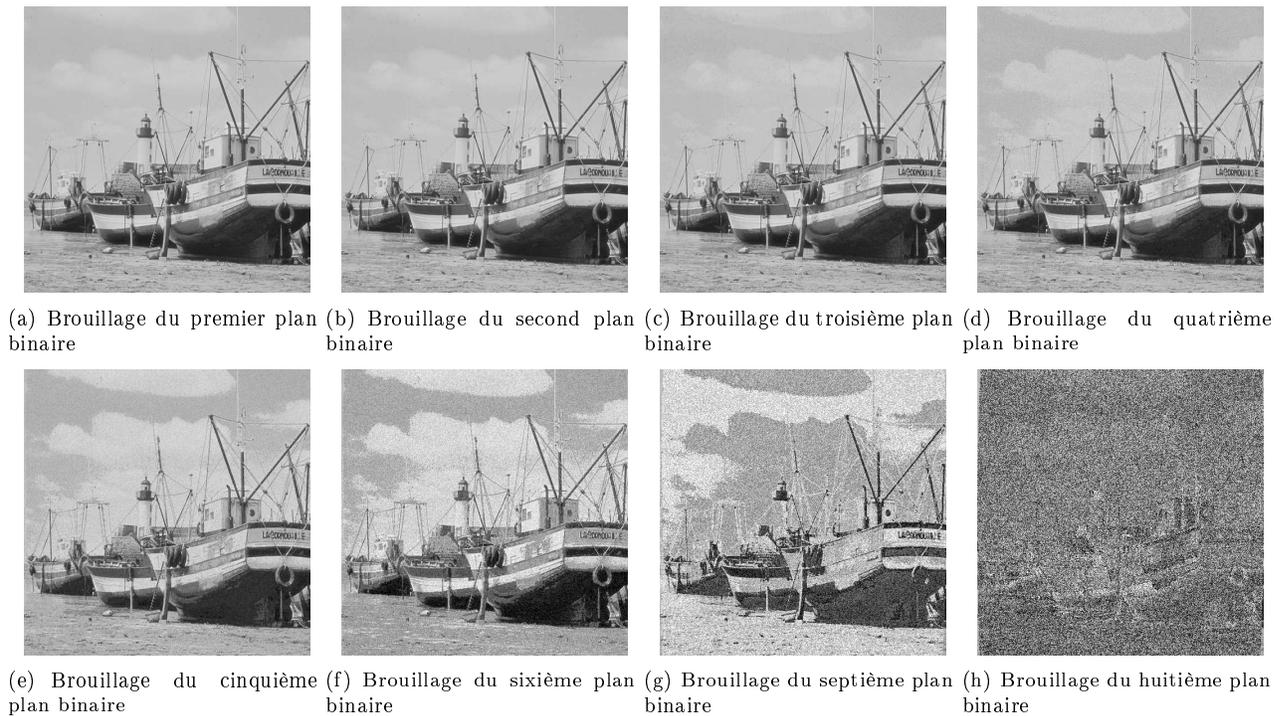


FIG. 5.10 – cryptage des différents plans binaires

Nous pouvons constater sur la figure 5.10 que le tatouage des trois premiers plans binaires ne dégrade pas perceptuellement l'image, et peut être employé dans des applications nécessitant une marque invisible tel que la stéganographie. Le tatouage des plans binaires supérieurs, est plus adapté à la cryptographie.

Le schéma présenté figure 5.11 décrit le processus d'extraction de la marque. Nous constatons que la clef permet d'extraire les bins originaux de l'ensemble de projections disponibles. La possession de ces bins originaux permet donc la reconstruction des plans binaires originaux, et par ce biais, la reconstruction de l'image originale.

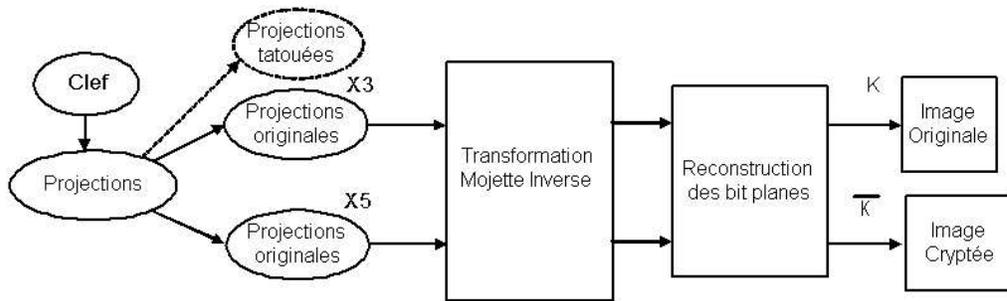


FIG. 5.11 – récupération de l'image originale

Le schéma d'insertion présenté sur la figure 5.6, peut subir de légères modifications pour crypter des images à l'aide du même procédé. Comme nous l'avons vu précédemment, la modification peut s'opérer directement sur les plans binaires de poids fort. Pour une meilleure sécurité, cet algorithme peut être appliqué directement sur les pixels de l'image en niveaux de gris. Ceci permet l'obtention d'un algorithme de cryptage d'images. L'image 5.12 présente un exemple de cryptage de l'image 'bateau'.

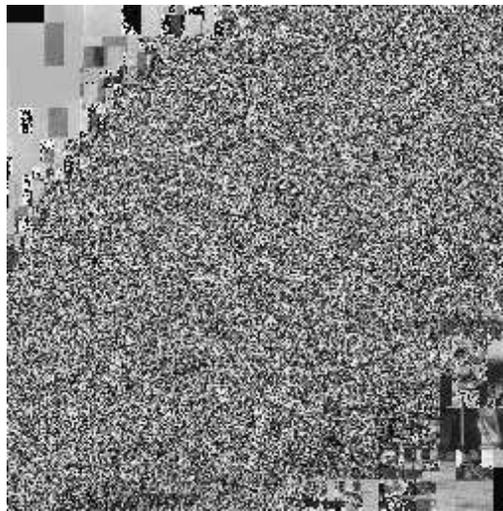


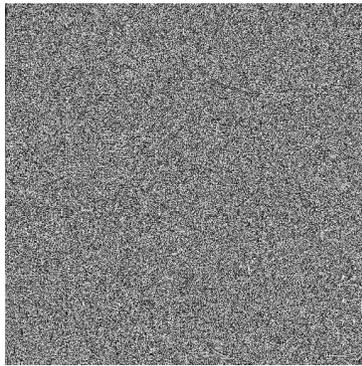
FIG. 5.12 – Image Cryptée

La figure 5.13 présente pour l'image 'singe' (voir Annexe C) différentes méthodes de cryptage. La transformation Mojette Directe a été appliquée sur l'image selon les directions suivantes :

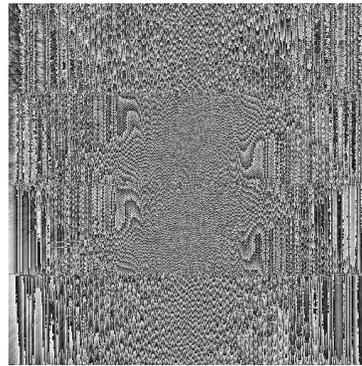
$$\{1,126\}, \{1,128\}, \{1,129\}, \{1,-128\} \text{ et } \{0,1\}$$

Nous avons choisi de ne tatouer (ici par une simple addition unitaire) qu'une seule projection, la première pour les figures 5.13(a) et 5.13(c) et la troisième pour les figures 5.13(b) et 5.13(d). Puis le calcul de la transformation inverse sur l'ensemble de projections fournit les images cryptées de la figure 5.13. Les figures

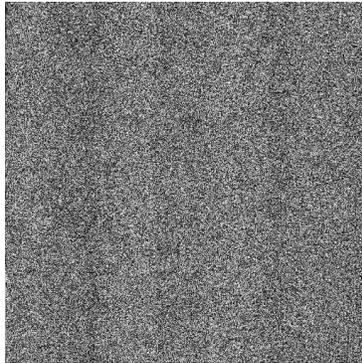
5.13(a) et 5.13(b) découlent d'une transformation Mojette Modulo 256 alors que les figures 5.13(c) et 5.13(d) découlent d'une transformation Mojette en Niveaux de Gris.



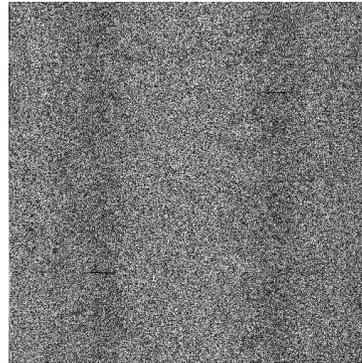
(a) Rétroprojection Modulo 256 avec tatouage de la direction No 0



(b) Rétroprojection Modulo 256 avec tatouage de la direction No 2



(c) Rétroprojection en NG avec tatouage de la direction No 0



(d) Rétroprojection en NG avec tatouage de la direction No 2

FIG. 5.13 – cryptage d'image par rétroprojection corrompue

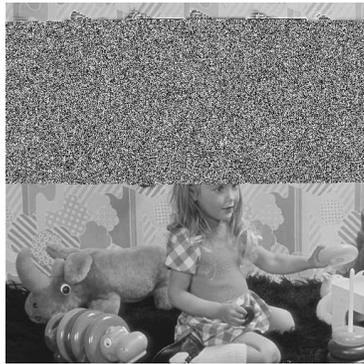
L'aspect de propagation des erreurs dues à la rétroprojection est bien représenté dans la figure 5.14 ci-dessous, ou toutes les image n'ont subi la modification que d'un seul bin et d'un unique niveau de gris. Si le bin modifié est judicieusement sélectionné, la dégradation est très rapidement diffusée sur toute l'image. Ces images représentent clairement l'imprévisibilité du processus de transformation Mojette inverse.



(a) cryptage de l'image 'lena'



(b) cryptage de l'image 'fruit'



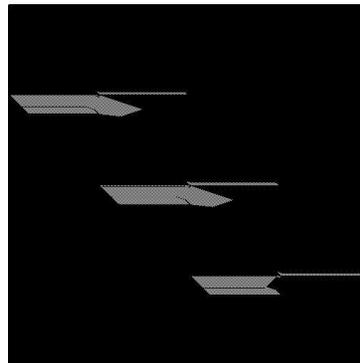
(c) cryptage de l'image 'enfant'



(d) cryptage de l'image 'clown'



(e) cryptage de l'image 'iba'



(f) modification spatiale engendrée pour l'image 'iba'

FIG. 5.14 – cryptages d'images par substitution d'un seul bin

les directions de projections utilisées pour ces 5 images sont respectivement :

$\{1,127\}$, $\{1,-128\}$, $\{1,129\}$ et $\{1,128\}$ pour l'image 'lena'.

$\{1,101\}$, $\{1,-102\}$, $\{1,102\}$, $\{1,-103\}$ et $\{1,104\}$ pour l'image 'fruit'.

$\{83,1\}$, $\{84,1\}$, $\{85,1\}$, $\{-86,1\}$, $\{87,1\}$ et $\{-87,1\}$ pour l'image 'enfant'.

$\{71,1\}$, $\{-72,1\}$, $\{73,1\}$, $\{-73,1\}$, $\{74,1\}$, $\{-74,1\}$ et $\{75,1\}$ pour l'image 'clown'.

et enfin,

$\{126,127\}$, $\{129,-128\}$, $\{-130,129\}$ et $\{-127,128\}$ pour l'image 'iba'.

Dans un tel cas de figure, la clef ne contient donc que la position du bin "défectueux".

Nous avons choisi de n'ajouter qu'un seul niveau de gris au premier bin de la première projection, afin d'être sûr que le bin modifié soit rétroprojeté.

Nous pouvons constater sur ces images que les directions de projection comprenant des valeurs de 'p' égales à '1' provoquent des zones de dégradation verticales sur l'image. à l'opposé, les directions de projection dont les valeurs des 'q' sont égales à '1' provoquent des dégradations horizontales. La largeur de ces zones est totalement dépendante des valeurs des vecteurs de direction (p, q) . Les directions de projection plus complexes employées pour l'image 'iba' entraînent un résultat surprenant. Dans ce dernier cas, les modifications apportées à l'image sont parfaitement imperceptibles. La figure 5.14(f) représente le résultat d'une égalisation d'histogramme sur l'image de différence entre l'image de référence 'iba' et sa version dégradée. La différence maximale entre les deux images est de treize NG.

5.3.4 Conclusion

Nous avons présenté ici une méthode de tatouage par brouillage des plans binaires d'une image, extensible directement aux niveaux de gris des pixels afin d'obtenir une méthode de cryptage. La clef appropriée permet le "décryptage" de l'image originale, toute personne non autorisée ne pouvant obtenir qu'une version dégradée. Il est possible de tatouer différents plans binaires pour obtenir une image tatouée plus ou moins sévèrement, ce qui permet de paramétrer la sécurité du protocole. De plus, nous avons vu que les différentes applications d'un tel schéma sont vastes : complément d'une méthode de tatouage, cryptographie, stéganographie, application de 'try and buy', etc ...

Cependant, le comportement de la transformation Mojette inverse sur des bins modifiés reste encore très difficile à anticiper. Concernant l'aspect cryptographique du schéma, cette caractéristique en est justement le point fort, car elle permet une très grande complexité pour tout pirate souhaitant accéder aux données originales et, ce, pour une très grande simplicité de codage.

5.4 Une méthode de tatouage spatial adaptable à la stéganographie

5.4.1 Introduction

Dans [5], nous avons présenté une méthode de tatouage par insertion de fantômes Mojette dans le domaine spatial, employant un algorithme de détection original basé sur les projections Mojette. Cette méthode s'est basée sur les deux méthodes précédentes et exploite donc autant la notion de fantômes que la redondance fournie par la transformation Mojette. Ici, le principe est d'insérer un fantôme dans l'image, puis de transmettre différentes projections, parmi lesquelles se trouveront non seulement des projections originales, mais aussi certaines projections tatouées.

Tout comme pour le schéma présenté précédemment, la clef adéquate permet de connaître les projections originales, et donc permet aussi l'obtention de l'image originale.

Comme nous le verrons dans la section 5.4.3, une légère modification de la gestion des clefs, permet d'adapter ce schéma au domaine de la stéganographie. Détaillons à présent ces deux algorithmes.

5.4.2 Algorithme de protection des droits d'auteurs

Le premier algorithme a pour objectif la protection des droits d'auteurs. A partir d'un ensemble de I projections, nous produisons un fantôme selon $(I - 1)$ directions, insérons alors ce fantôme dans les pixels de l'image par une simple addition de niveaux de gris. Il est important de noter que ce schéma peut tout aussi bien être envisagé avec l'insertion de plusieurs fantômes de même clef ou de clefs différentes.

Selon le lemme de Katz, l'ensemble I de projections, permet de reconstruire l'image si la somme des $|p_i|$ (respectivement, la somme des $|q_i|$) est supérieure à la dimension horizontale (respectivement verticale) de l'image.

Ici, l'ensemble $(I - 1)$ est choisi volontairement de façon à ne pouvoir reconstruire l'image.

Dans le schéma de protection du copyright, l'image tatouée est obtenue par addition du (des) fantôme(s) Mojette dans les niveaux de gris de l'image originale. Le fantôme étant de moyenne nulle, les modifications de chaque pixel sont compensées par celles de son voisinage. La force du fantôme peut aussi être modulée afin d'assurer le meilleur compromis entre l'invisibilité et la robustesse. Il est aussi envisageable d'employer un masque psychovisuel lors de l'insertion.

La transformation Mojette est à nouveau appliquée sur l'image, selon la direction modifiée par l'insertion du fantôme (soit, la I^{eme} direction). La I^{eme} projection est donc conservée par l'utilisateur, comme la clef du schéma de tatouage. Cette projection erronée va donc remplacer sa version originale lors de la transmission de

l'information.

La figure 5.15 présente l'algorithme d'insertion présenté ci-dessus.

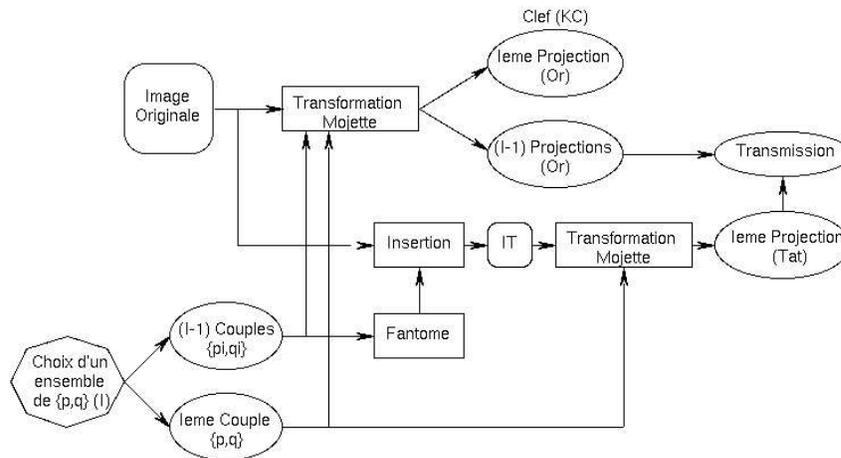


FIG. 5.15 – Schéma d'insertion du copyright

Un des avantages que procure ce schéma réside dans la double utilité de la clef. Effectivement, cette dernière peut aussi être utilisée comme empreinte de l'image.

Lors de l'extraction du copyright, le possesseur de la clef peut remplacer la projection corrompue par l'originale, puis procéder à la transformation Mojette inverse, afin d'obtenir l'image originale. Toute attaque malveillante opérant directement sur les bins Mojette provoquerait immédiatement des dégradations sévères et irréversibles, comme nous l'avons vu dans la section 5.3.

La détection de la marque consiste à calculer la I^{eme} projection sur l'image tatouée, un pic représentant le filigrane doit être introduit dans la projection, voir fig 5.20.

L'obtention de la projection du filigrane extrait s'effectue tout simplement par une différence entre la I^{eme} projection tatouée et la clef.

La fonction d'inter-corrélation, (définie dans la section 4.3.3 et rappelée dans l'équation 4.19) entre la projection du filigrane original f_O et la projection du filigrane extrait f_E , fournit un pic de corrélation si la marque est détectée dans l'image.

$$r(d) = \frac{\sum_i [(f_O(i) - m_{f_O}) * (f_E(i - d) - m_{f_E})]}{\sqrt{\sum_i (f_O(i) - m_{f_O})^2} \sqrt{\sum_i (f_E(i - d) - m_{f_E})^2}} \quad (5.3)$$

où m_{f_o} et m_{f_e} sont les valeurs moyennes des filigranes f_o et f_e .

Si la détection du filigrane s'effectue par le calcul de l'inter-corrélation, son extraction est simplement réalisée par une soustraction du fantôme dans l'image. La figure 5.16 présente le processus d'extraction. Dans ce schéma 'IT' représente l'image tatouée par addition du fantôme.

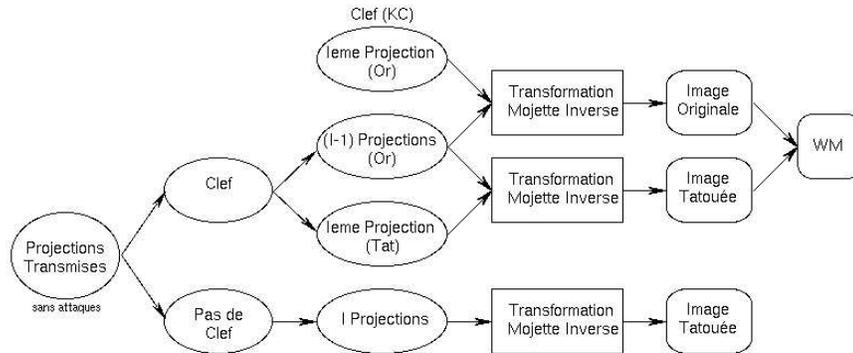


FIG. 5.16 – Schéma d'extraction du copyright

Cette figure montre que l'utilisation de la clef permet de séparer la projection tatouée des projections originales. Une fois cette sélection effectuée, il suffit d'ajouter la I^{eme} projection afin de reconstruire l'image originale.

En fonction du degré de sécurité souhaité, il est possible d'ajouter des projections à l'ensemble des projections transmises. L'ajout de projections originales permet d'accroître la robustesse du schéma. Ceci permet une reconstruction plus fiable de l'image. Cependant, l'ajout de projections tatouées augmente la complexité d'extraction pour toute personne n'étant pas en possession de la clef.

5.4.3 Algorithme de Stéganographie

Une adaptation très simple du schéma précédent nous conduit à un schéma de stéganographie. Le processus d'insertion suit exactement les mêmes étapes que le schéma de protection de copyright, la seule différence réside dans la clef du tatouage. La clef du schéma de stéganographie devient non plus la projection originale mais la projection tatouée. De cette façon, toute personne ne possédant pas cette clef ne pourra que reconstruire l'image originale, alors que le destinataire, pourra reconstruire aussi l'image tatouée, afin d'extraire le message par une simple différence des images. Le schéma d'insertion présenté figure 5.17 est très semblable au schéma de protection du copyright, mais nous constatons que seules les projections originales sont transmises. Sur cette figure, 'IS' représente l'image de stéganographie, comportant le message secret (fantôme Mojette).

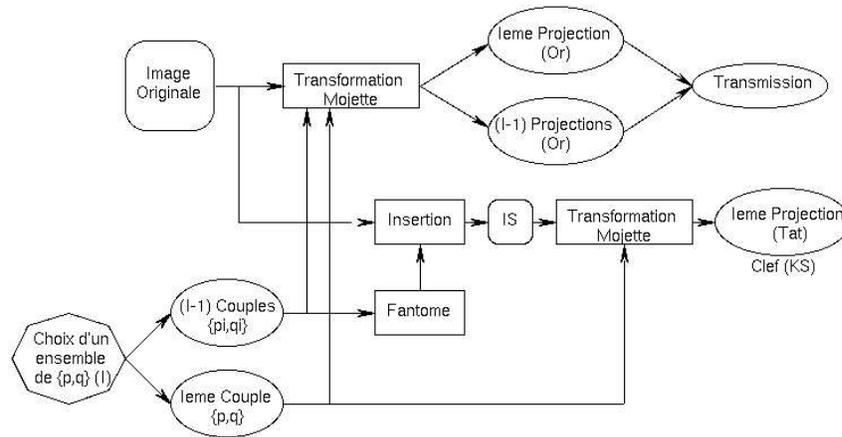


FIG. 5.17 – Schéma de stéganographie, insertion de la marque

Les étapes d'extraction du message secret sont évidemment les mêmes que pour l'extraction du schéma précédent, mais ici, la projection originale est remplacée par celle qui contient le message.

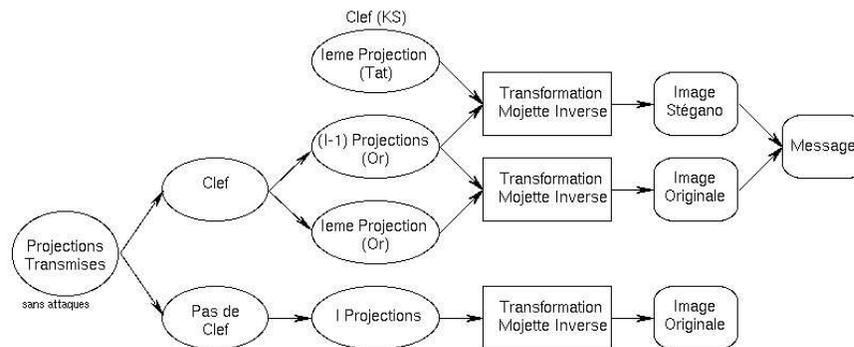


FIG. 5.18 – Schéma de stéganographie, insertion de la marque

Habituellement, les schémas de stéganographie, effectuent la transmission du média tatoué avec le message secret enfoui dedans. L'image stéganographiée est donc accessible à quiconque ne possédant pas la clef. Dans le schéma présenté ici, les données transmises comportent l'image, le message secret et une certaine redondance de l'information de l'image. Le possesseur de la clef pourra reconstituer l'image tatouée ainsi que sa version originale, et donc, par la même occasion, le message secret. Une personne non autorisée ne pourra que reconstruire l'image originale.

5.4.4 Résultats

Le premier algorithme présenté, a pour objectif la protection des droits d'auteurs, il doit donc satisfaire une contrainte de robustesse. La figure 5.19 ci dessous, présente les images du processus de tatouage. L'image originale est représentée figure 5.19(a), la figure 5.19(b) représente l'image tatouée et le filigrane employé est donné en 5.19(c).

Le fantôme de la figure 5.19(c) est créé selon l'ensemble de directions (p, q) suivant :

$$\{(-1,68) (-1,59) (-1,57) (-1,50) (-1,35) (-1,32) (-1,26) (1,16) (1,18) (1,31) (1,33) (1,34) (1,52)\}$$

et ce fantôme a été dupliqué dans chacun des sites $[m; n]$ suivants, où m représente le numéro de la ligne et n le numéro de la colonne, l'origine étant placée en haut à gauche de l'image :

$$\{[511;40], [511;80], [511;120], [511;160], [511;200], [511;240], [511;280], [511;320], [511;360], [511;400]\}$$

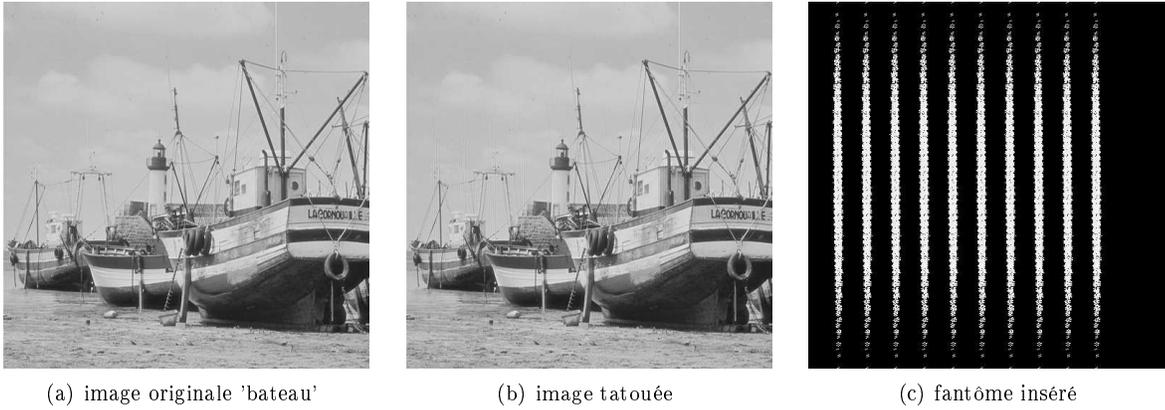


FIG. 5.19 – Images du processus de tatouage

Les figures 5.20(a) et 5.20(b) sont les projections respectives de l'image originale (5.19(a)) et de l'image tatouée (5.19(b)) selon la direction $(0,1)$.

La figure 5.20(c) représente la différence entre 5.20(a) et 5.20(b), la projection du fantôme selon la direction $(0,1)$ est donnée en 5.20(d) et finalement, le résultat de la fonction d'intercorrélation entre la projection du filigrane original et celle du filigrane extrait est donné figure 5.20(e).

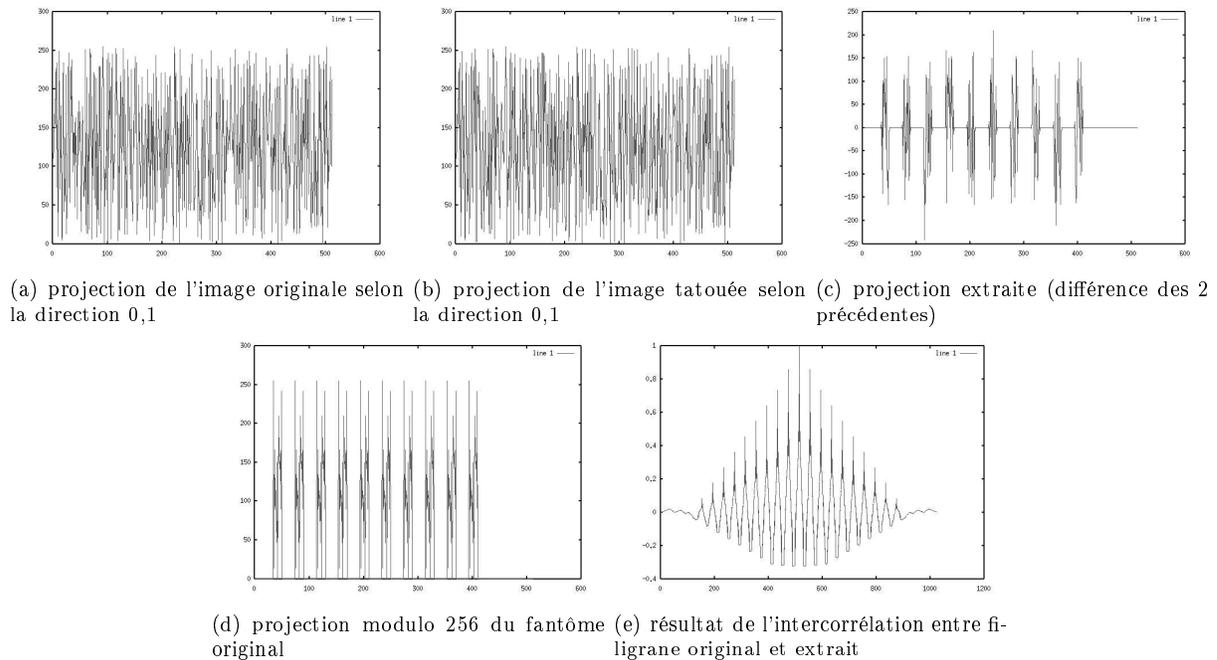


FIG. 5.20 – exemple de tatouage par fantômes

Différentes attaques de Stirmark ont été appliquées sur un ensemble de 8 images tatouées. Certains résultats sont présentés ci-dessous, figure 5.21 à 5.29 sur des images de contenu fréquentiel varié. Ces résultats représentent l'application du schéma de détection sur des images ayant subi le procédé de tatouage, suivi d'une attaque quelconque. Ici, nous n'étudions pas les aspects de fausse alarme (tentative de détection sur des images ne comprenant pas la marque).

La figure 5.21(a) présente le résultat de la fonction d'intercorrélation normalisée sur l'image tatouée, la figure 5.21(b) présente le résultat de l'intercorrélation après avoir fait subir un filtrage moyenné (3×3) à l'image tatouée. La figure 5.21(c) présente la même attaque mais avec un noyau de taille (5×5). Le résultat d'un filtrage médian de taille 3 est représenté figure 5.21(d), puis les courbes représentées de la figure 5.22(a) à 5.22(d) représentent respectivement le processus de détection pour des codages JPEG de facteurs de qualité $q=10\%$, 20% , 30% et 40% .

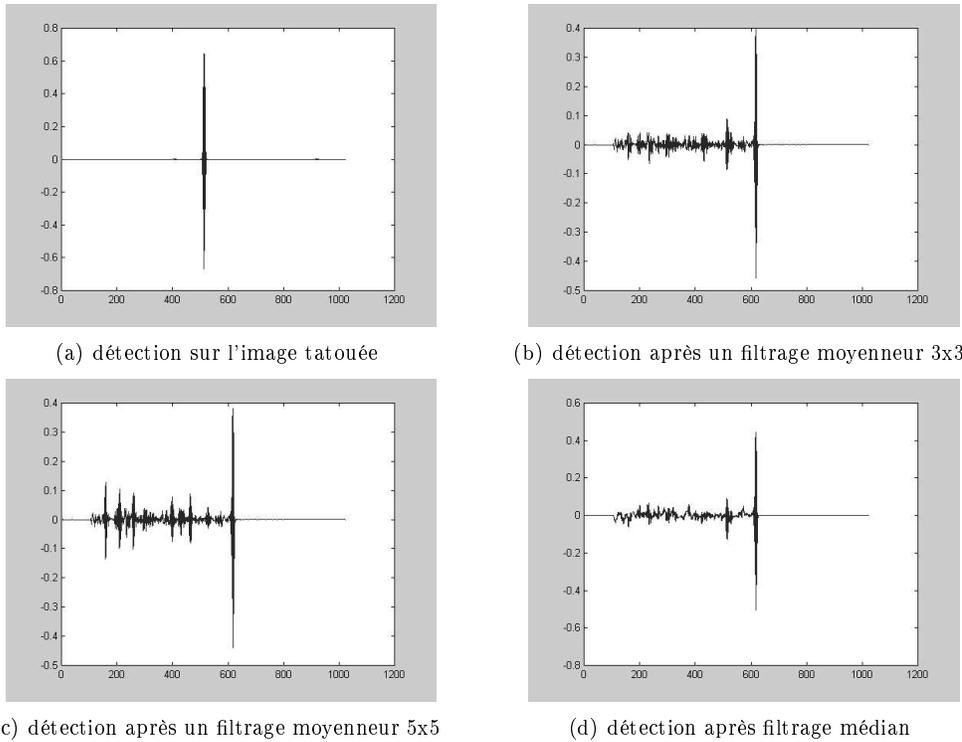


FIG. 5.21 – Résultats de détection pour l'image 'bateau' (1/3)

Nous pouvons constater sur la figure 5.21 que le schéma de tatouage par fantômes permet une détection fructueuse face à divers filtres passe bas ou médian. Nous verrons par la suite que ce comportement varie en fonction du contenu fréquentiel de l'image. Observons à présent les résultats de la détection du filigrane face à différents taux de compression JPEG (figure 5.22) :

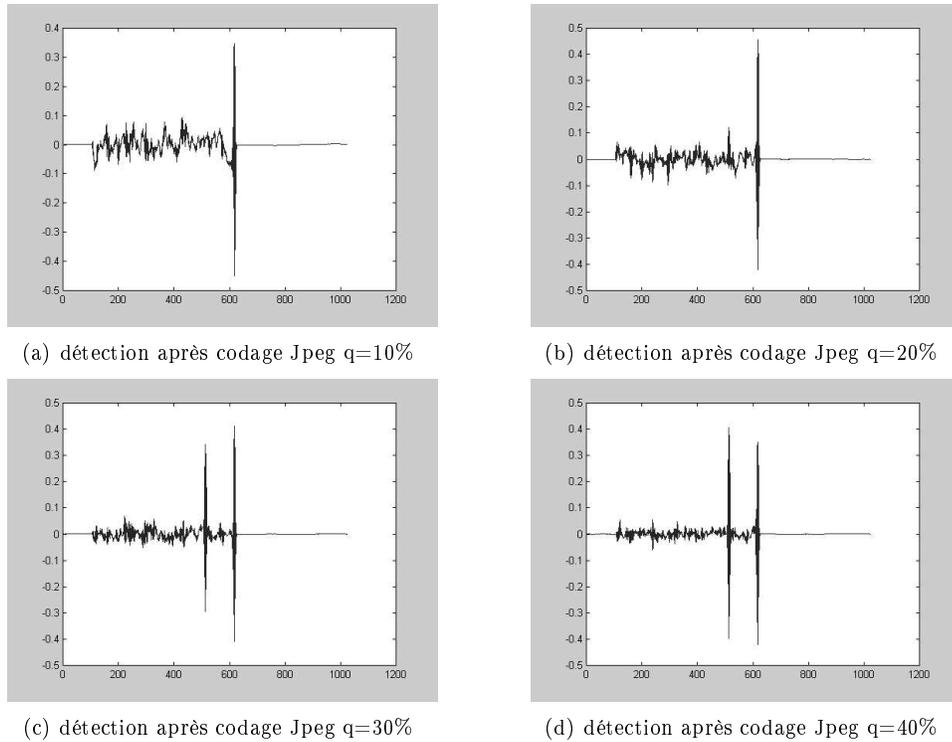


FIG. 5.22 – Résultats de détection pour l'image 'bateau' (2/3)

Encore une fois, les résultats de détection face à la compression JPEG sont positifs, la marque à été détectée pour les quatre facteurs de compression testés. Nous présentons, encore pour l'image 'bateau', quelques résultats face à d'autres attaques dans la figure 5.23 ci dessous.

La courbe représentée en 5.23(a) représente la détection sur l'image ayant subi un échantillonnage. Les figures 5.23(b), 5.23(c) et 5.23(d), présentent des rotations de respectivement 2, 5 et 10 degrés. Le résultat d'un rehaussement du contraste de facteur 75% est présenté figure 5.23(e) et, finalement, un brouillage local des pixels sur un voisinage de deux pixels fournit le résultat présenté sur la courbe 5.23(f).

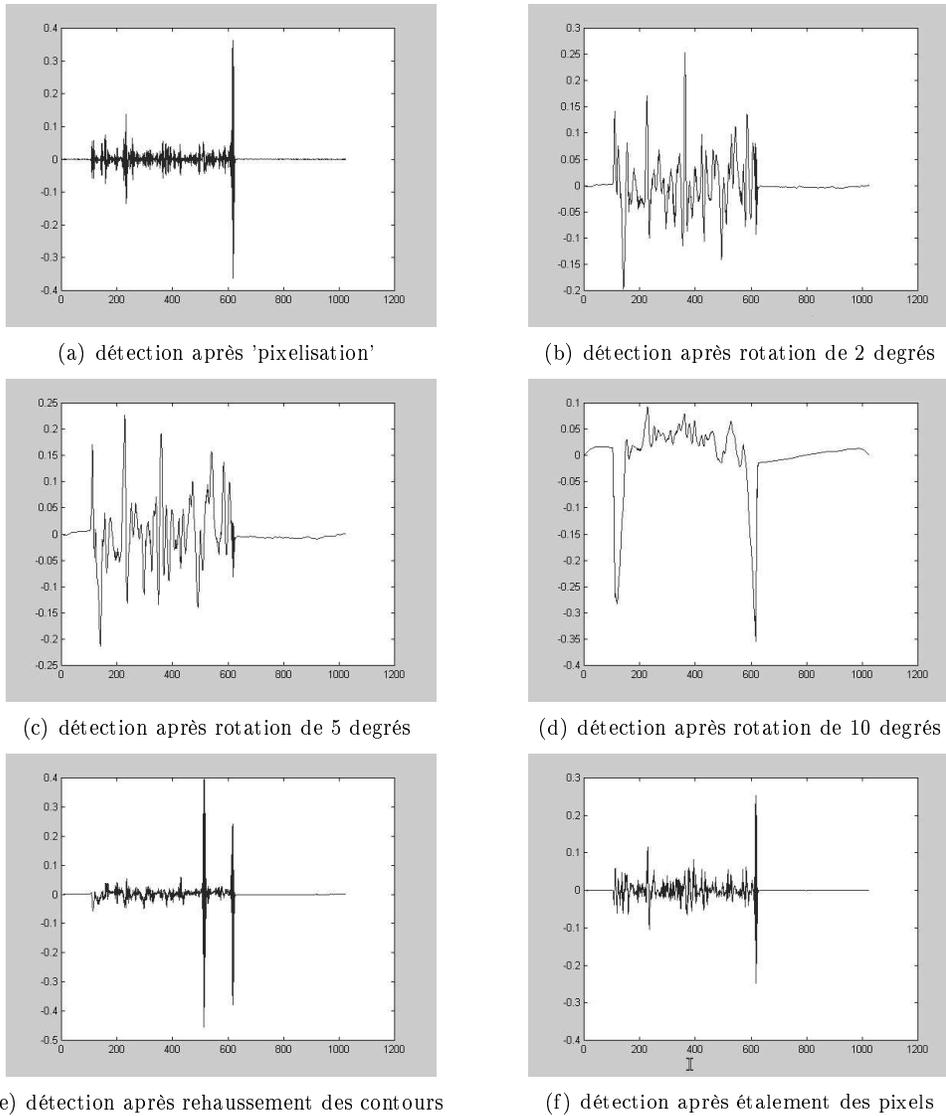


FIG. 5.23 – Résultats de détection pour l'image 'bateau' (3/3)

Nous constatons ici que le schéma proposé fournit des résultats très encourageants pour la plupart des attaques testées, mais reste cependant peu performant face à des attaques géométriques opérant une désynchronisation, telles que les rotations présentées ici. Un tel procédé pourrait permettre l'utilisation d'un schéma de re-synchronisation approximatif de l'image (compensation des déformations géométriques). La plupart des schémas de tatouage nécessitent une compensation exacte des déformations géométriques, ici, une compensation approximative peut être suffisante car une légère rotation permet encore la détection de la marque.

Concernant l'image 'singe', les résultats sont fournis sur les figures 5.24, à 5.26. L'image 5.24(a) présente le

résultat de la détection sur l'image tatouée. Tout comme pour la figure 5.21, nous présentons ici respectivement un filtrage passe bas de noyau (3×3) (fig. 5.24(b)), de noyau (5×5) (fig. 5.24(c)) et un filtrage médian (fig. 5.24(d)).

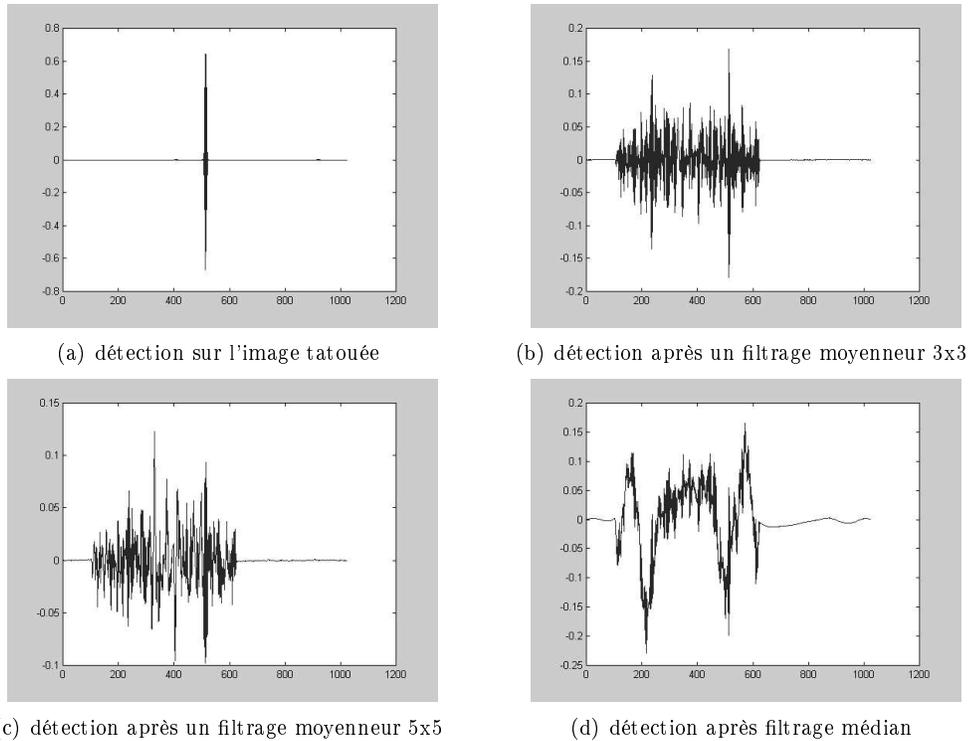


FIG. 5.24 – Résultats de détection pour l'image 'singe' (1/3)

La figure 5.25 présente les courbes de détection face à des compressions JPEG de facteurs de qualité de valeurs respectives : $q=10\%$, 20% , 30% et 40% dans les figures 5.25(a) à 5.25(d).

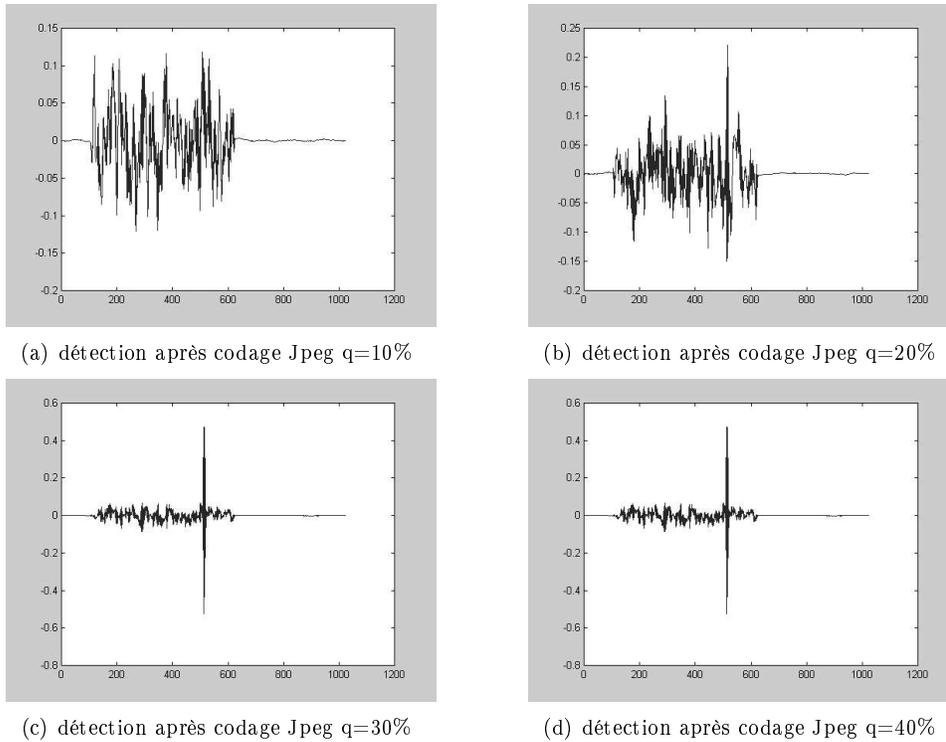


FIG. 5.25 – Résultats de détection pour l'image 'singe' (2/3)

Nous observons pour l'image 'singe', que le filtrage médian ou le codage JPEG avec $q=10\%$ sont des attaques auxquelles le schéma ne parvient pas à résister. Ceci provient probablement du fait que l'image contient plus de textures que de contours et que celles-ci sont moyennées sur les projections.

Concernant toujours l'image 'singe', la suite des résultats est donnée ci-dessous pour un échantillonnage, des rotations, un rehaussement de contraste et un étalement des pixels, tout comme pour la figure 5.23.

Ces résultats confirment la faiblesse du schéma face à certaines attaques géométriques.

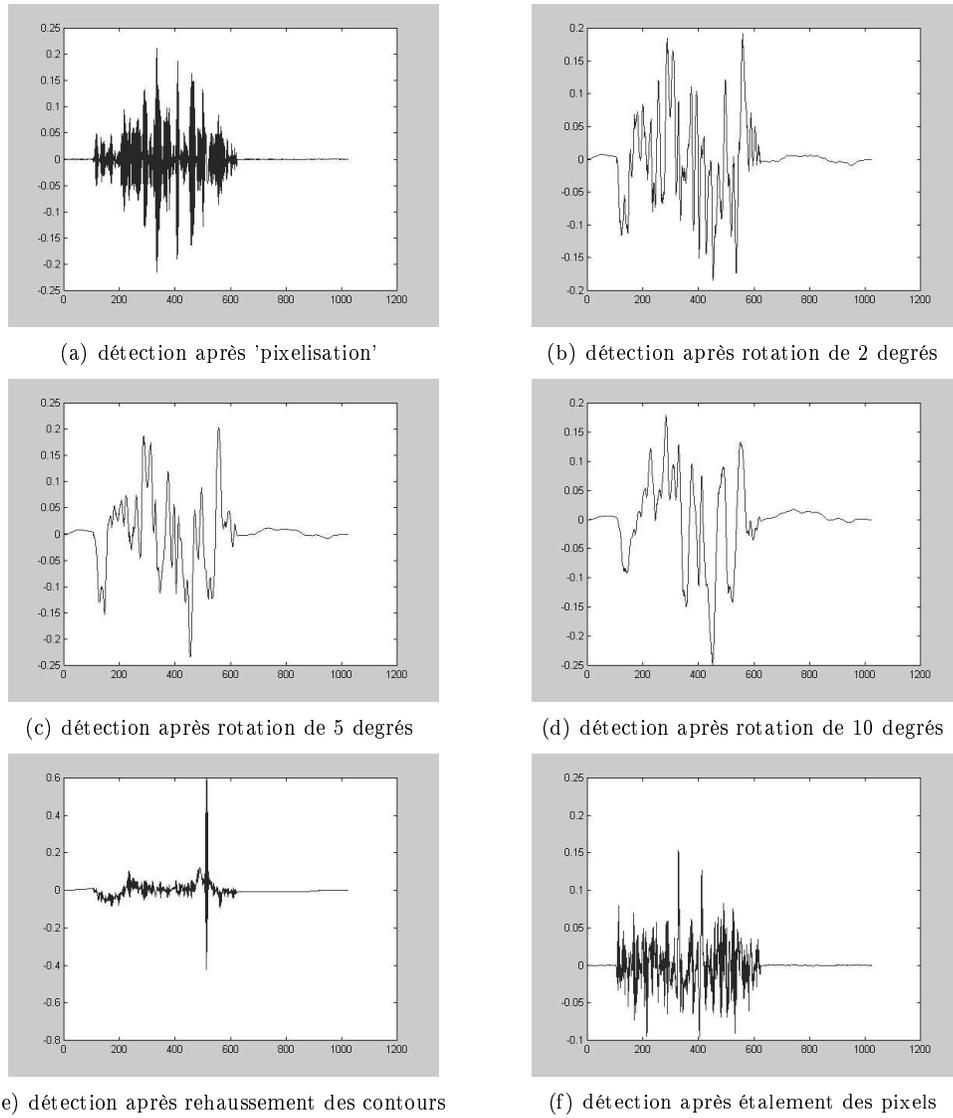
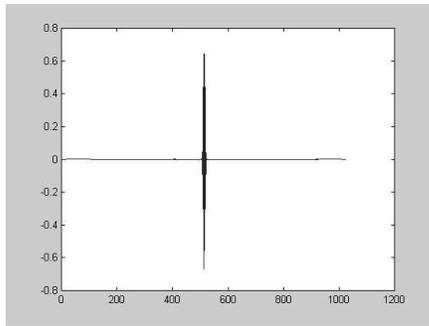
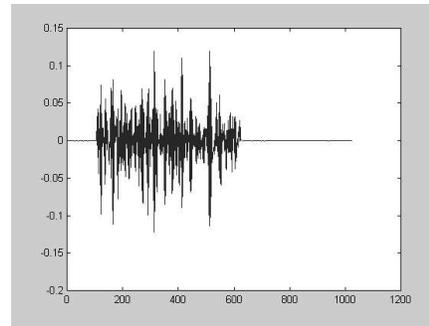


FIG. 5.26 – Résultats de détection pour l'image 'singe' (3/3)

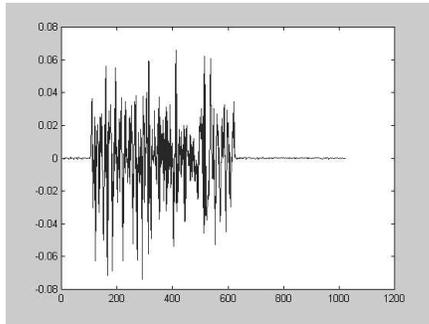
Les mêmes résultats concernant l'image 'Iba' sont donnés sur les figures 5.27 à 5.29.



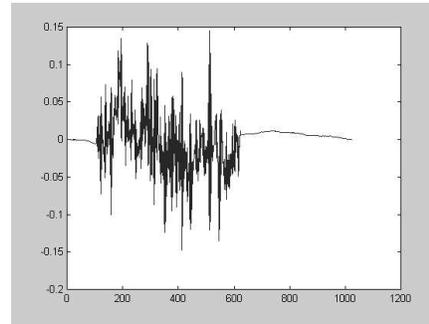
(a) détection sur l'image tatouée



(b) détection après un filtrage moyenneur 3x3



(c) détection après un filtrage moyenneur 5x5



(d) détection après filtrage médian

FIG. 5.27 – Résultats de détection pour l'image 'iba' (1/3)

Tout comme pour l'image 'singe', l'image 'iba' comporte de nombreuses zones texturées, ceci se traduit encore une fois par une diminution du pouvoir de détection du processus. Cette tendance se confirme dans les résultats suivants mettant en œuvre la compression JPEG.

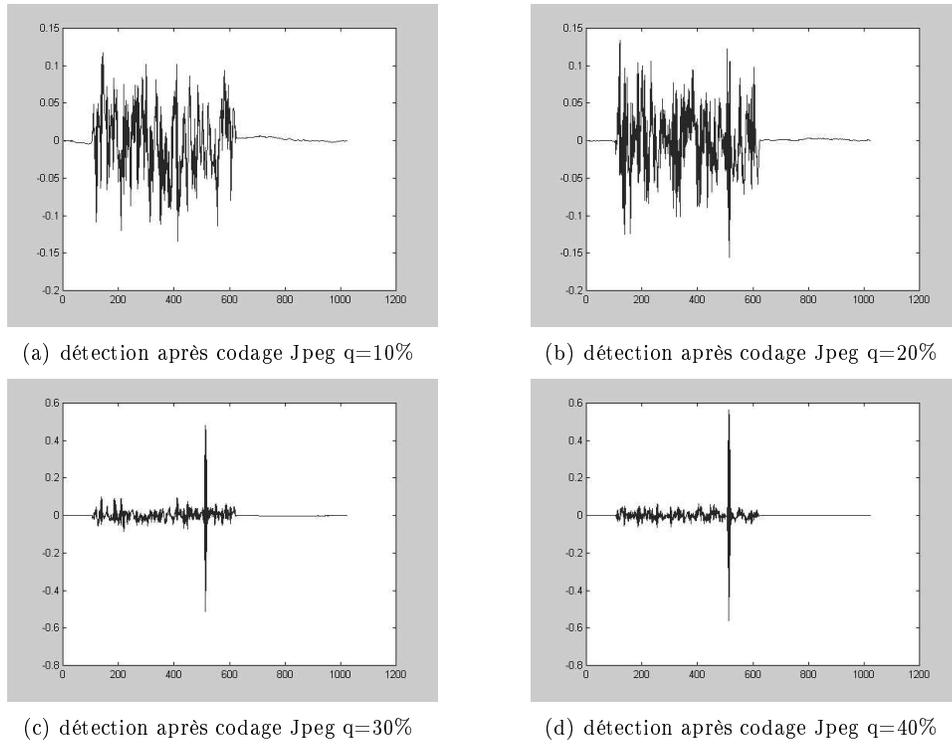


FIG. 5.28 – Résultats de détection pour l'image 'iba' (2/3)

Dans la figure suivante 5.29, nous constatons qu'une déformation géométrique faible (rotation de 2° ou 5° , fig. 5.29(b) et 5.29(c)) peut permettre la détection de la marque.

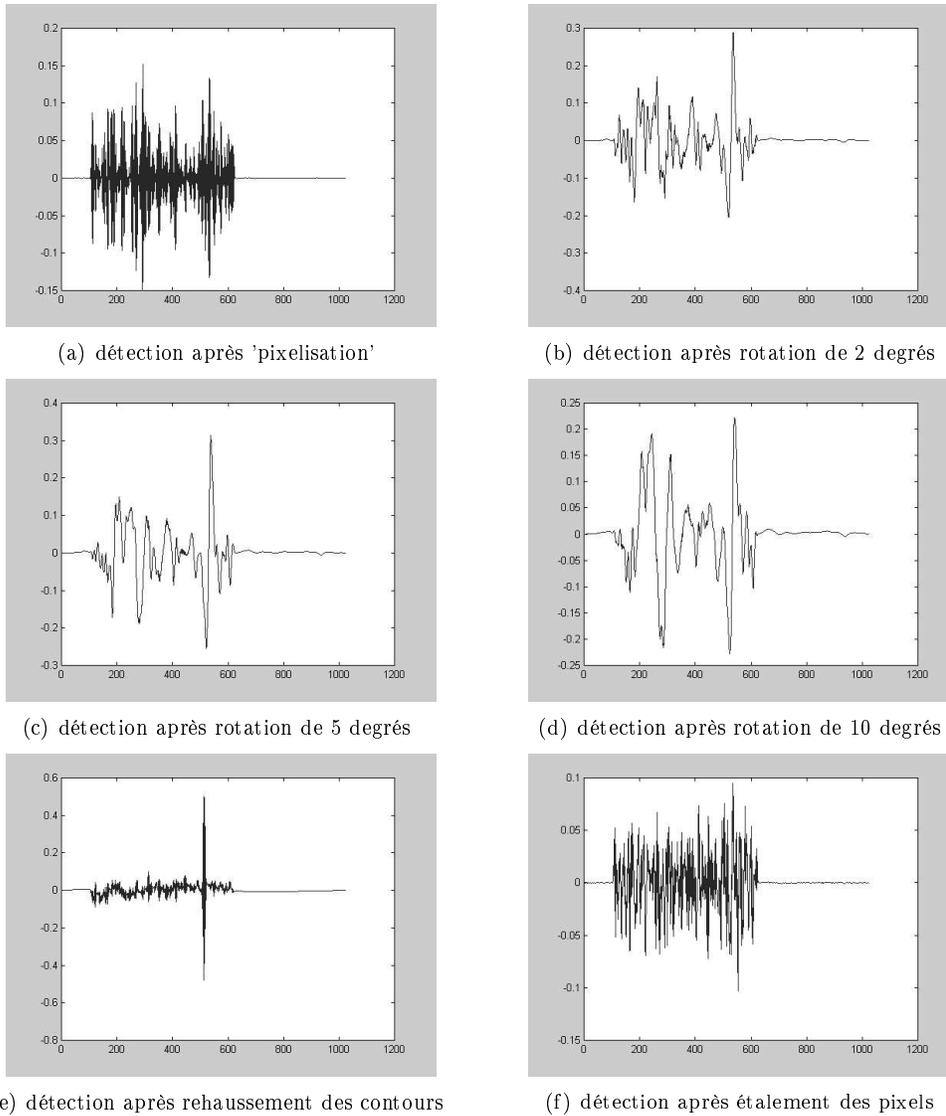


FIG. 5.29 – Résultats de détection pour l'image 'iba' (3/3)

Nous pouvons donc constater que cet algorithme de tatouage résiste à diverses modifications de l'image et peut donner des résultats encourageants pour de faibles attaques géométriques. Face aux autres attaques, telles que les filtrages, la compression JPEG ou encore le rehaussement de contraste, le schéma proposé ici fournit d'assez bons résultats.

Concernant certaines attaques géométriques assez importantes opérant une dé-synchronisation de la marque, ce schéma ne permet pas en l'état, la détection de la marque. Cependant, du fait de sa résistance à de faibles déformations, une compensation des attaques géométriques, même assez grossière pourrait être suffisante pour retrouver la marque.

5.4.5 Conclusion

Nous avons présenté dans cette section, un procédé de tatouage d'images basé sur les propriétés étudiés dans les sections 5.2 et 5.3. Ce schéma exploite conjointement les aspects de redondance et de fantômes Mojette. Par une modification très simple de la gestion des clefs, ce schéma est adaptable au domaine de la stéganographie, comme à celui de la protection de copyright. Ce processus permet donc d'assurer autant l'aspect de l'invisibilité que celui de la robustesse de la marque.

Le schéma de protection de copyright a été testé face à différentes attaques, dont certaines déformations géométriques faibles et a montré globalement de très bons résultats.

5.5 Conclusion

Ce chapitre était dévolu aux méthodes de tatouage basées sur la transformation Mojette, nous avons présenté trois schémas différents, chacun ayant une utilisation particulière.

Le premier schéma proposé est adapté aux contraintes du tatouage des images médicales. Il permet un tatouage invisible, réversible et une extraction aveugle, tout en traitant le problème des empreintes nécessaire dans de telles applications. Un schéma complet de gestion des clefs y est détaillé et le recours à un tiers de confiance est proposé.

Le second schéma proposé ici est plutôt adapté au domaine de la transmission. Les paquets d'information transmis sont bruités, dans le but de ne permettre qu'aux personnes autorisées de reconstruire l'image originale. Dans un tel procédé, les personnes ne possédant pas la clef appropriée, ne pourront reconstruire qu'une version dégradée, voire cryptée de l'image. Ce schéma est totalement paramétrable en fonction de la sécurité souhaitée.

Le troisième schéma propose un lien entre le domaine du tatouage de protection de droits d'auteurs et le domaine de la stéganographie. En effet, ce schéma exploite les propriétés des deux précédents, mais permet par une simple substitution des clefs, de s'adapter à ces deux domaines pourtant bien distincts. Ce dernier schéma fournit des résultats assez satisfaisants autant en terme d'invisibilité que de robustesse, face à diverses attaques.

La transformation Mojette, ainsi que le concept de fantômes, permettent donc de proposer plusieurs schémas de tatouage d'images dans des applications très variées. Un aspect très important de la transformation Mojette, correspondant aux attentes des schémas de tatouage est la notion de redondance. Effectivement, l'objectif d'un schéma de tatouage est d'insérer de l'information dans un support sans le dégrader. L'exploitation de la redondance apportée par la transformation Mojette nous semble donc être une voie à explorer, nous avons

proposé ici quelques idées dans ce sens.

Le seul point faible commun aux algorithmes présentés jusqu'ici est le manque de robustesse vis-à-vis de certaines attaques géométriques. L'objectif du chapitre suivant est double :

- l'exploration d'un lien entre le domaine psychovisuel et la transformation Mojette.
- la proposition d'une approche originale permettant une meilleure robustesse face aux attaques de désynchronisation.

Chapitre 6

Une méthode alliant masques

psychovisuels et transformation Mojette

Sommaire

| | | |
|-----|---|-----|
| 6.1 | Introduction | 180 |
| 6.2 | Choix de l'espace de tatouage | 180 |
| 6.3 | Création du masque | 185 |
| 6.4 | Schéma d'insertion | 187 |
| 6.5 | Schéma de détection | 190 |
| 6.6 | Expérimentations | 195 |
| 6.7 | Conclusion | 203 |

6.1 Introduction

Nous avons présenté dans le chapitre 4, les différentes étapes de la création d'un masque perceptuel, ce masque a ensuite été validé respectivement dans le domaine DCT, des études ont été réalisées dans le domaine DWT et enfin le masque a été validé dans le domaine spatial. Nous avons ensuite présenté un schéma de tatouage des sous-bandes visuelles.

Dans le chapitre 5, nous avons conduit des études dans le domaine Mojette. La première étude portait sur l'insertion spatiale de fantômes dans l'optique du tatouage d'images médicales. Le second schéma opérait directement dans les bins Mojette en vue d'une application de transmission. Ce schéma était extensible au domaine de la cryptographie. Finalement, une combinaison de ces deux méthodes, a abouti à un schéma de tatouage adaptable à la stéganographie.

Nous avons pu constater que chacune de ces méthodes comportait certains avantages, mais aussi parfois quelques inconvénients. La faiblesse principale des méthodes de tatouage prenant en compte des critères visuels est le manque de robustesse vis-à-vis des attaques géométriques. Les méthodes de tatouage exploitant la transformée Mojette (chapitre 5) n'utilisaient jusqu'à présent que des critères empiriques pour assurer l'invisibilité de marque.

Comme nous l'avons vu, la décomposition en canaux perceptuels modélisant le comportement du SVH, s'effectue dans l'espace de Fourier. Nous allons donc étudier le comportement de ce domaine transformé face aux déformations géométriques. Nous constaterons que ce domaine présente des caractéristiques très intéressantes.

Nous avons donc opté pour la mise en œuvre d'un schéma de tatouage exploitant le modèle psychovisuel décrit précédemment. Concernant la détection du filigrane, la technique de détection Mojette détaillée dans la section 5.4 sera encore une fois mise à profit, mais à présent dans le domaine de Fourier.

6.2 Choix de l'espace de tatouage

Une étude préalable du domaine de Fourier est importante et permet de mieux appréhender les différentes attaques. Nous allons à présent tester le comportement du spectre d'une image vis-à-vis de différentes déformations géométriques.

Supposons dans le domaine continu, le signal $f(x, y)$ de spectre :

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(u.x+v.y)} dx.dy \quad (6.1)$$

Une translation de (x_0, y_0) dans les deux directions du signal $f(x, y)$ peut s'écrire $g(x - x_0, y - y_0)$ et a

pour transformée :

$$G(u, v) = F(u, v)e^{-j(ux_0+vy_0)} \quad (6.2)$$

Pour le cas d'une homothétie dans une direction, soit par exemple la dilatation de $f(x, y)$ dans la direction Ox donne :

$$g(x, y) = f(kx, y) \quad (6.3)$$

de transformée

$$G(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(kx, y)e^{-j(u.x+v.y)} dx.dy \quad (6.4)$$

Le changement de variable $z = kx$, permet d'obtenir

$$G(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(z, y)e^{-j(\frac{u.z}{k}+v.y)} dz.dy \quad (6.5)$$

Soit

$$G(u, v) = \frac{1}{k}F\left(\frac{u}{k}, v\right) \quad (6.6)$$

La dilatation du signal se traduit donc par une homothétie inverse de la variable correspondante, (ici, une réduction) dans le plan des fréquences. Si cette même dilatation d'un facteur k est effectuée dans les 2 directions $g(x, y) = f(kx, ky)$ il y aura une contraction de $\frac{1}{k}$ de la transformée de Fourier.

Supposons à présent une rotation de l'image :

$$g_{\theta}(x, y) = f(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta) \quad (6.7)$$

Par transformée de Fourier, nous obtenons :

$$G_{\theta}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta) e^{-j(u.x+v.y)} dx.dy \quad (6.8)$$

En effectuant les changements de variables suivants :

$$s = x \cos \theta + y \sin \theta \qquad x = s \cos \theta - t \sin \theta$$

soit

$$t = -x \sin \theta + y \cos \theta \qquad y = s \sin \theta + t \cos \theta$$

nous obtenons :

$$G_{\theta}(u, v) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j(us \cos \theta - ut \sin \theta + us \sin \theta + vy \cos \theta)} f(s, t) ds dt \quad (6.9)$$

$$G_{\theta}(u, v) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j[(u \cos \theta + v \sin \theta)s + (-u \sin \theta + v \cos \theta)t]} f(s, t) ds dt \quad (6.10)$$

Nous constatons que la rotation se traduit par une rotation de même angle dans le plan des fréquences.

Considérons à présent l'inclinaison,

$$g(x, y) = f(x - my, y) \quad (6.11)$$

La transformée de Fourier donne alors

$$G(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - my, y) e^{-j(ux + vy)} dx dy \quad (6.12)$$

effectuons d'abord l'intégration par rapport à x :

$$G(u, v) = \int_{-\infty}^{\infty} e^{-jvy} \left[\int_{-\infty}^{\infty} f(x - my, y) e^{-jux} dx \right] dy \quad (6.13)$$

nous pouvons alors effectuer le changement de variable $x - my = z$

$$G(u, v) = \int_{-\infty}^{\infty} e^{-jvy} \left[\int_{-\infty}^{\infty} f(z, y) e^{-juz + jummy} dz \right] dy \quad (6.14)$$

L'intégrale en fonction de la variable z est la transformée de Fourier d'une ligne horizontale de l'image $f(z, y)$. On lui fait subir un déphasage e^{jummy} correspondant à une translation de longueur m .

$$G(u, v) = \int_{-\infty}^{\infty} e^{-jvy + jummy} \left[\int_{-\infty}^{\infty} f(z, y) e^{-juz} dz \right] dy \quad (6.15)$$

$G(u, v)$ est la transformée de Fourier de $f(z, y)$ calculée pour u et $(v - um)$

$$G(u, v) = F(u, v - um) \quad (6.16)$$

La transformée d'un signal incliné par rapport à l'axe vertical subit une inclinaison identique mais par rapport à l'axe horizontal. Si la fonction est inclinée selon une pente m , sa transformée de Fourier est inclinée selon une pente $\frac{-1}{m}$.

Cette étude du comportement de la transformation de Fourier face à différents types d'attaques géométriques, permet de dresser le tableau suivant :

| Attaque Spatiale | Attaque fréquentielle |
|------------------------------------|--|
| Rotation de θ° | Rotation de θ° |
| Translation de (m,n) | Translation de (m,n) et multiplication par une exponentielle |
| Inclinaison de (m) selon l'axe x | Inclinaison de $(-1/m)$ selon l'axe v |
| Homothétie de (m,n) | contraction de $(1/m, 1/n)$ |
| Changement d'échelle | Changement d'échelle |
| Miroir | Miroir |
| Suppression de lignes | Suppression de lignes |

TAB. 6.1 – Comportement sur le spectre des attaques

Ces déformations ont été définies pour la transformation de Fourier dans le domaine analogique. Les phénomènes induits dans le domaine numérique sont légèrement différents. Supposons à présent une image numérique $s(x, y)$, dont la transformation de Fourier discrète est donnée par la formule 6.17

$$S(u, v) = \frac{1}{M \cdot N} \sum_{m=0}^M \sum_{n=0}^N s(m, n) e^{-j2\pi \left(\frac{u \cdot m}{M} + \frac{v \cdot n}{N} \right)}. \quad (6.17)$$

De façon similaire, le spectre de Fourier peut être retranscrit dans le domaine spatial par l'équation 6.18

$$s(m, n) = \frac{1}{M \cdot N} \sum_{u=0}^M \sum_{v=0}^N S(u, v) e^{+j2\pi \left(\frac{u \cdot m}{M} + \frac{v \cdot n}{N} \right)}. \quad (6.18)$$

Dans le domaine numérique, les déformations induites sur le spectre sont un peu plus complexes à cause de la numérisation. Une déformation géométrique, appliquée sur l'image numérique, va le plus souvent entraîner un repliement du spectre (aliasing en anglais). Lors du traitement d'une image numérique, l'étude porte sur un support borné, d'où un nombre d'échantillons limités, induisant un problème d'ondulation.

La figure 6.1 présente des tests effectués sur l'image 'lena' illustrant les résultats donnés par ce tableau. Pour une meilleure visualisation, les spectres des images ont subi une égalisation d'histogramme.

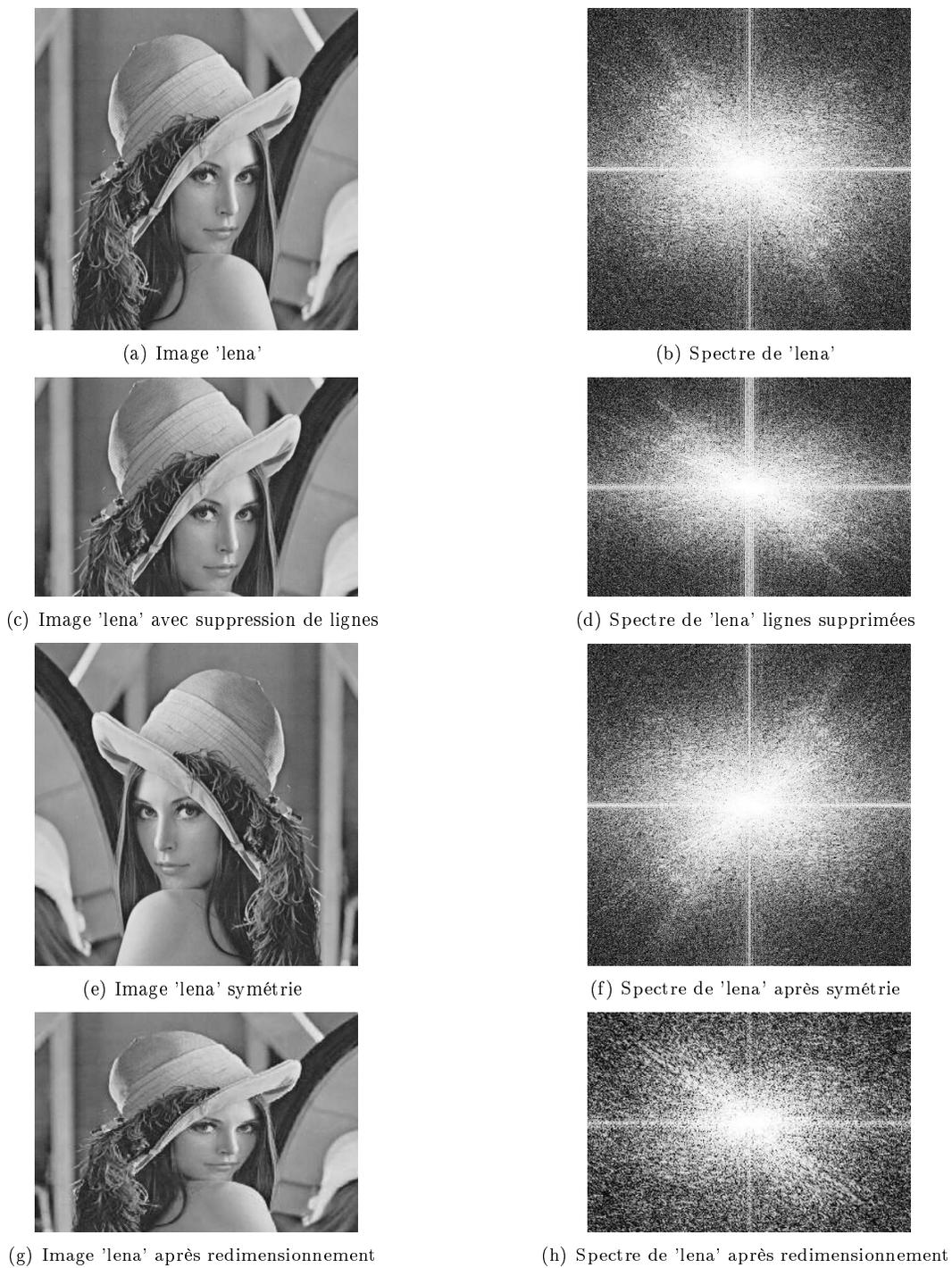


FIG. 6.1 – Comportement du spectre en fonction de différentes attaques géométriques

La méthode présentée maintenant supporte différentes attaques géométriques et évite les problèmes de dé-synchronisation de la marque lors de la détection.

6.3 Création du masque

Au vu des comportements décrits dans la section précédente, nous avons opté pour une méthode de tatouage dans le domaine de Fourier. Les études psychovisuelles détaillées précédemment vont à présent être appliquées non plus spatialement mais directement sur les coefficients du spectre de Fourier. Le masque est donc identique à celui de la section 4.2, mais à présent, l'insertion, doit s'effectuer dans le domaine fréquentiel, tout en vérifiant que la représentation spatiale de la marque reste contenue dans le masque JND. Le schéma permettant de maintenir les dégradations inférieures au masque perceptuel est donné dans la section 6.4.

Comme nous le constaterons par la suite, outre l'application spectrale du masque perceptuel, une des étapes les plus importantes du processus d'insertion est l'étape de sélection des sous-bandes à tatouer. Comme nous l'avons vu dans la section 4.4, la détection est optimale dans les sous-bandes de la couronne IV de la DCP. Il est toujours envisageable de sélectionner une sous-bande par couronne. Cependant, dans un souci d'optimisation et de simplification, nous nous limiterons (section 6.6) à l'insertion d'un filigrane dans la sous-bande d'activité maximale de la couronne IV uniquement.

Les tableaux 6.2 à 6.8 représentent pour sept images de test, les valeurs des écarts types pour chacune des dix sept sous-bandes de la DCP.

| arbre (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|-------------|------------|-----------|-----------|-----------|----------|----------|
| BF (I+II) | 341.809204 | X | X | X | X | X |
| III | 12.336727 | 12.030947 | 15.339738 | 15.488437 | X | X |
| IV | 3.200491 | 2.712337 | 2.879495 | 3.380801 | 3.493072 | 3.514742 |
| V | 1.653143 | 1.368634 | 1.128130 | 1.114534 | 1.371187 | 1.783859 |

TAB. 6.2 – Ecart types des sous-bandes pour l'image 'arbre'

| bateau (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|--------------|------------|----------|-----------|----------|----------|----------|
| BF (I+II) | 588.983948 | X | X | X | X | X |
| III | 9.711377 | 6.254829 | 12.073222 | 5.545864 | X | X |
| IV | 2.932935 | 1.978387 | 1.649386 | 2.793088 | 1.341148 | 2.008909 |
| V | 1.256965 | 0.795067 | 0.597858 | 0.872892 | 0.485526 | 0.766162 |

TAB. 6.3 – Ecart types des sous-bandes pour l'image 'bateau'

| clown (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|-------------|------------|----------|----------|----------|----------|----------|
| BF (I+II) | 243.524536 | X | X | X | X | X |
| III | 10.555946 | 6.543955 | 9.367599 | 7.047615 | X | X |
| IV | 2.543078 | 1.389775 | 1.179486 | 1.640078 | 1.252609 | 1.382022 |
| V | 1.259909 | 0.430199 | 0.330565 | 0.374363 | 0.322601 | 0.499565 |

TAB. 6.4 – Ecart types des sous-bandes pour l'image 'clown'

| Enfant (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|--------------|------------|----------|----------|----------|----------|----------|
| BF (I+II) | 513.730530 | X | X | X | X | X |
| III | 7.841820 | 6.854107 | 7.146344 | 6.892491 | X | X |
| IV | 1.759055 | 1.899534 | 1.392024 | 1.688545 | 2.356708 | 1.737225 |
| V | 0.687658 | 0.508869 | 0.372738 | 0.610564 | 0.485695 | 0.507844 |

TAB. 6.5 – Ecart types des sous-bandes pour l'image 'enfant'

| Fruit (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|-------------|------------|----------|----------|----------|----------|----------|
| BF (I+II) | 538.623657 | X | X | X | X | X |
| III | 10.782708 | 6.755310 | 8.025631 | 6.738794 | X | X |
| IV | 1.936553 | 1.409636 | 1.122137 | 1.297739 | 1.043040 | 1.153632 |
| V | 0.492676 | 0.214220 | 0.194130 | 0.351562 | 0.202181 | 0.228981 |

TAB. 6.6 – Ecart types des sous-bandes pour l'image 'fruit'

| iba (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|-----------|------------|----------|-----------|----------|----------|----------|
| BF (I+II) | 295.611633 | X | X | X | X | X |
| III | 11.546186 | 9.407474 | 22.695704 | 9.634499 | X | X |
| IV | 2.980126 | 1.496651 | 2.704341 | 4.926414 | 1.739292 | 1.557589 |
| V | 1.426783 | 0.742161 | 0.841185 | 1.338308 | 0.620061 | 0.813069 |

TAB. 6.7 – Ecart types des sous-bandes pour l'image 'iba'

| tableau (ECT) | Dir. 1 | Dir. 2 | Dir. 3 | Dir. 4 | Dir. 5 | Dir. 6 |
|---------------|------------|----------|-----------|----------|----------|----------|
| BF (I+II) | 428.410614 | X | X | X | X | X |
| III | 9.572443 | 6.126048 | 12.069341 | 5.693418 | X | X |
| IV | 2.128053 | 1.858001 | 1.192279 | 2.786710 | 1.074091 | 1.186200 |
| V | 0.895006 | 0.670664 | 0.374387 | 0.889688 | 0.343620 | 0.484845 |

TAB. 6.8 – Ecart types des sous-bandes pour l'image 'tableau'

Pour toute image traitée, le calcul de ces valeurs d'écart types, permet de sélectionner la sous-bande à tatouer pour chaque couronne de la DCP. Compte tenu de la procédure d'insertion détaillée par la suite, l'insertion de la marque peut s'effectuer soit dans la sous-bande d'écart type maximal, soit dans la sous-bande

d'écart type minimal. En nous basant sur l'idée qu'un pic fréquentiel sera moins aisément décelable si il est "noyé" dans une zone fortement texturée (d'écart type important), nous avons opté pour l'insertion de la marque dans les sous-bandes d'écart types maximums.

Nous avons vu précédemment que les attaques les plus dévastatrices sont fréquemment les déformations géométriques opérant une dé-synchronisation de la marque. Dans ce schéma, nous proposons d'exploiter la sélectivité angulaire de la DCP. Ce schéma est établi dans le but de résister au plus grand nombre possible d'attaques géométriques. Cependant, comme nous le verrons par la suite, ce schéma peut aussi permettre de très bons résultats de robustesse face à des attaques non géométriques (filtrages, compression, quantification, ...). Nous constaterons que ce schéma de détection débute par une sélection de la sous-bande potentiellement tatouée. Une fois cette sous-bande sélectionnée, le processus d'extraction, semblable à celui présenté dans la section 5.4 est mis en œuvre.

6.4 Schéma d'insertion

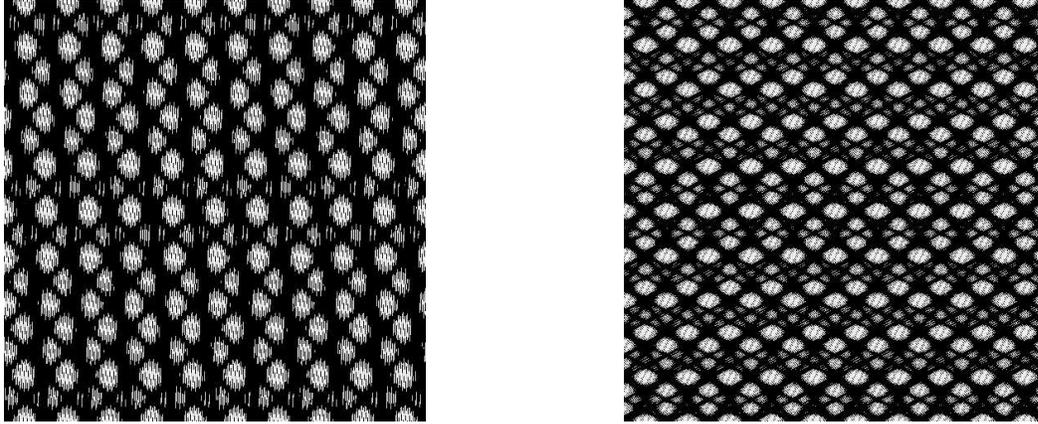
Une fois les sous-bandes sélectionnées, l'insertion d'un fantôme Mojette dans le spectre doit se produire tout en respectant le masque perceptuel fourni par la sous-bande. Nous avons vu dans la section précédente que l'aspect psychovisuel est garanti ici par l'utilisation du masque perceptuel présenté dans la section 4.2. Quant au processus d'insertion, il est similaire à celui que nous avons présenté dans la section 5.2, mais adapté au domaine fréquentiel.

Des exemples de représentation spatiale de fantôme Mojette sont représentés sur la figure 6.2 ci dessous. L'image 6.2(a) présente un fantôme positionné au pixel (250, 140) et construit avec les directions suivantes :

$$\{(10,9) (10,-9) (10,1)\}$$

L'image 6.2(b) représente un fantôme positionné au pixel de coordonnées (80, 140) et construit à partir des directions suivantes :

$$\{(10,9) (10,19) (10,-9) (10,-19)\}$$



(a) Représentation spatiale d'un fantôme confiné à la sous-bande (III, 1) (b) Représentation spatiale d'un fantôme confiné à la sous-bande (IV, 4)

FIG. 6.2 – Représentation spatiale de fantômes Mojette

La figure 6.3 montre le processus d'adaptation fréquentielle du masque spatial. Le filigrane fréquentiel, qui est un fantôme Mojette, subit une transformation de Fourier inverse. Sa représentation spatiale, notée W_s , est alors comparée avec le masque psychovisuel spatial (M_s).

Si la force du filigrane est inférieure à celle fournie par le masque, celle-ci est alors incrémentée, puis une nouvelle transformation de Fourier inverse permet à nouveau une comparaison du filigrane avec le masque. Ce procédé est alors répété jusqu'à ce que la force de la représentation spatiale du filigrane se trouve au seuil de visibilité défini par le masque psychovisuel. Le dépassement en un seul point de l'image, du masque par la marque, entraîne la fin du processus d'incrémentation de la force. Ce schéma itératif est représenté ci-dessous :

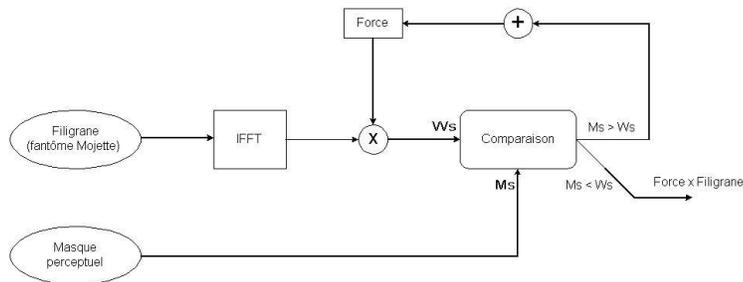


FIG. 6.3 – Adaptation du masque spatial au domaine de Fourier

Une fois cette force maximale calculée, le filigrane est donc pondéré par ce coefficient, puis inséré dans la sous-bande correspondante par une simple multiplication du spectre par le fantôme. Les seuils $\Delta C_{i,j}$ représentant les plus petits pas de quantification constants par sous-bande, un masque psychovisuel doit être mis en

œuvre pour chaque sous-bande de la DCP. Le processus d'insertion permet donc l'obtention des sous-bandes tatouées. Afin de respecter les masques perceptuels fournis pour chaque sous-bande visuelle, la position du fantôme doit être étudiée de façon à maintenir la marque confinée dans l'étendue spectrale de la sous-bande traitée.

Un ensemble de fantômes à été créé afin de pouvoir insérer différentes marques dans chaque sous-bande.

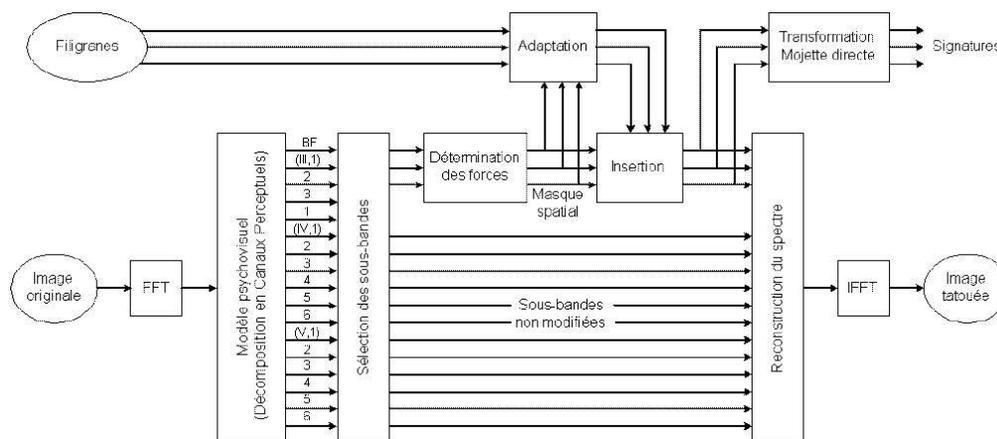


FIG. 6.4 – Insertion dans les coefficients de Fourier

L'insertion, dont le schéma global est présenté figure 6.4, effectue une multiplication des coefficients de Fourier par le fantôme pondéré.

Nous observons sur la figure 6.4 les différents traitements à appliquer sur les images à tatouer. La transformée de Fourier doit être appliquée sur l'image originale, afin de procéder, comme nous l'avons vu dans le premier chapitre, à la Décomposition en Canaux Perceptuels.

Une fois la DCP appliquée sur le spectre, le schéma présenté ci-dessus montre que l'insertion peut s'effectuer sur plusieurs sous-bandes. Une fois ces sous-bandes sélectionnées, le processus de détermination des forces maximales est mis en œuvre. L'insertion du filigrane s'effectue alors dans le spectre, tout en respectant le masque perceptuel. Ensuite, la reconstruction totale du spectre est accomplie avec les autres sous-bandes n'ayant pas subi le procédé de tatouage. Une transformation Mojette est effectuée sur les sous-bandes tatouées afin de conserver une empreinte des données. Finalement, une transformation de Fourier inverse du spectre modifié permet l'obtention de l'image tatouée.

Les images données dans la figure 6.5 représentent respectivement l'image originale en 6.5(a), l'image tatouée en 6.5(b), le module du spectre de l'image originale est représenté figure 6.5(d), et finalement, le module du spectre de l'image tatouée est donné en 6.5(e).

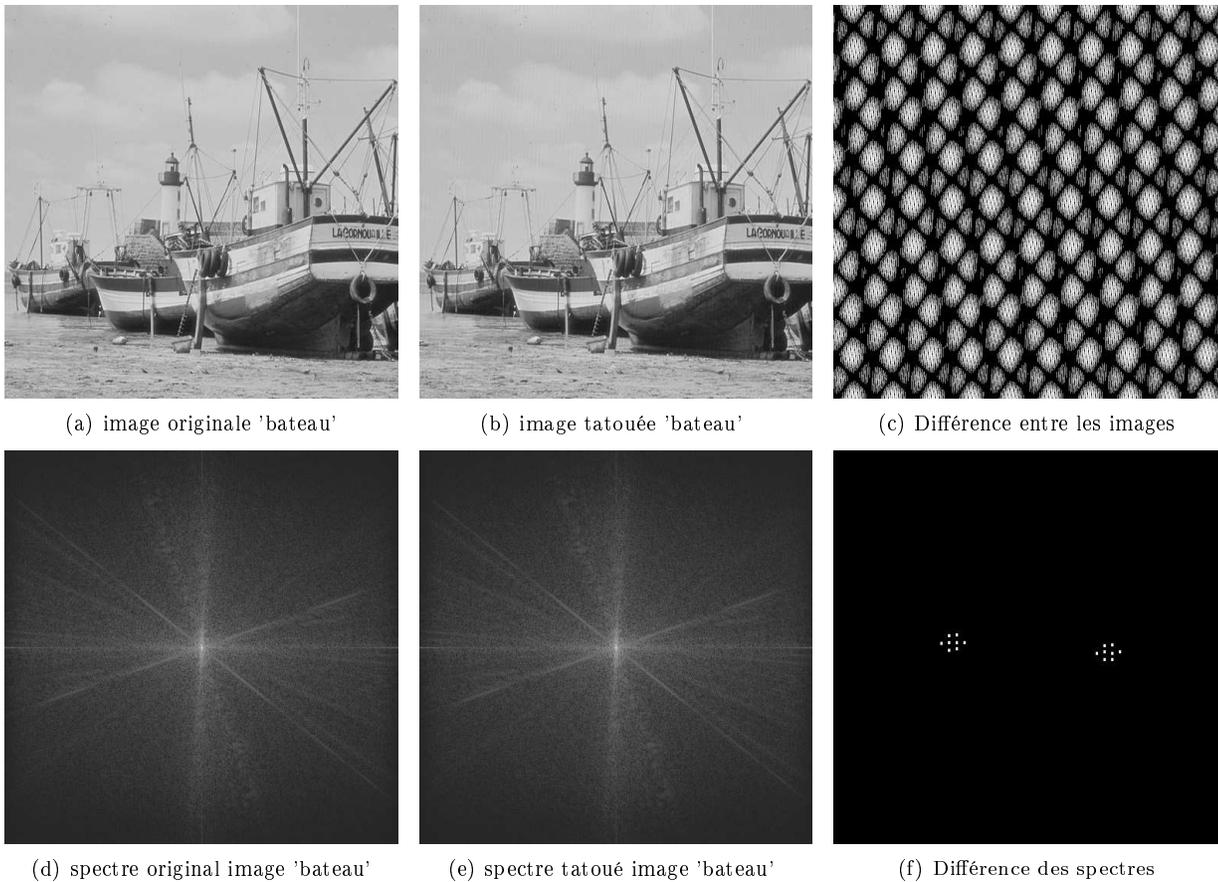


FIG. 6.5 – Images et spectres de l'algorithme de tatouage

Nous pouvons observer sur ces images qu'aucune différence n'est détectable, autant sur les images que sur leurs spectres respectifs.

6.5 Schéma de détection

Le processus de détection de la marque est similaire à celui que nous avons découvert dans la section 5.4. Rappelons que dans cette section, le principe du processus de détection était de calculer la fonction d'intercorrélation entre la projection Mojette du filigrane et la projection de l'image potentiellement tatouée selon le même angle Mojette. Ce principe ayant fourni de très bons résultats dans le domaine spatial, il est repris ici, mais doit être adapté au domaine fréquentiel.

Les différentes étapes du processus de détection sont présentées dans le schéma 6.6 ci-dessous :

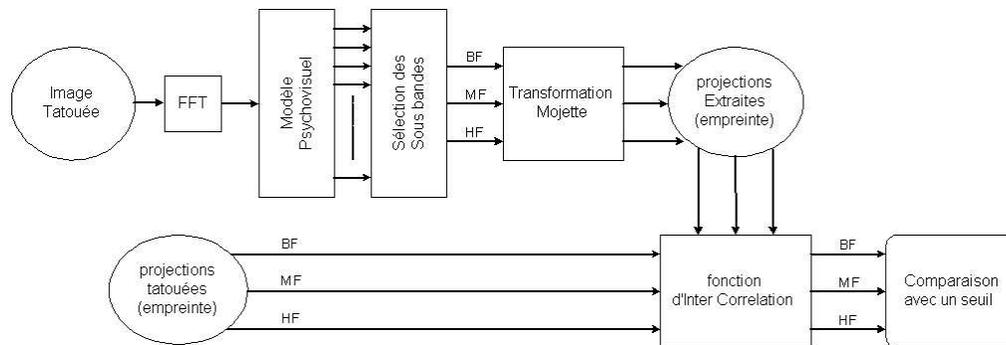


FIG. 6.6 – Extraction Mojette du filigrane fréquentiel

Nous observons sur cette figure que l'extraction des sous-bandes suit les mêmes étapes que précédemment. Une transformation de Fourier permet l'obtention du spectre, sur lequel la DCP est appliquée afin d'extraire les sous-bandes d'énergie maximale. Ces sous-bandes d'énergie maximale subissent une transformation de Fourier selon l'angle Mojette défini lors du processus d'insertion. La fonction d'intercorrélacion détaillée précédemment est ensuite appliquée entre les projections Mojette obtenues sur l'image potentiellement tatouée et les projections Mojette du fantôme inséré.

Nous pouvons observer sur la figure 6.7 différentes images du processus de tatouage. Les images 6.7(a) et 6.7(b) représentent respectivement le spectre original de l'image 'lena' et le spectre tatoué. Finalement, l'image 6.7(c) représente la sous-bande sélectionnée par l'algorithme.

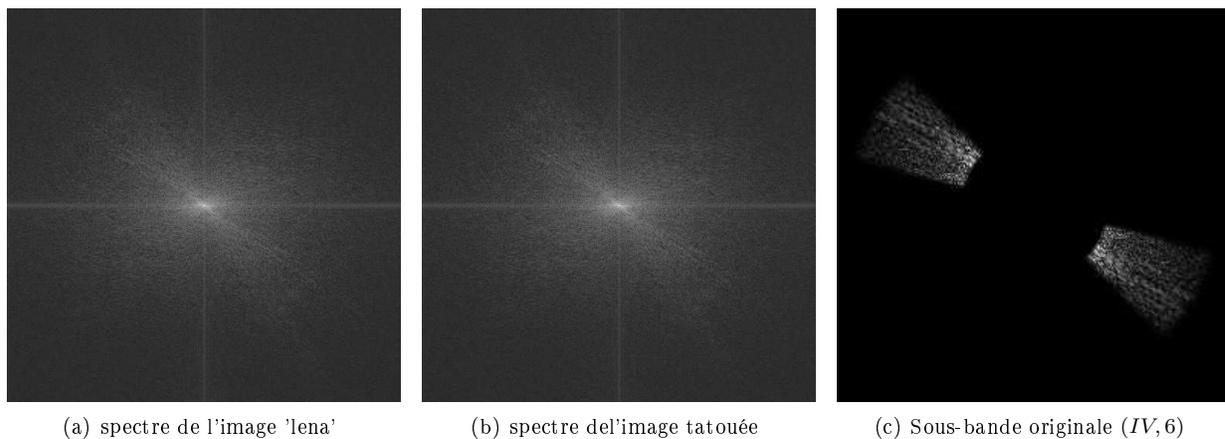


FIG. 6.7 – Images et spectres de l'algorithme de tatouage (1/4)

Nous pouvons constater sur ces images que le spectre tatoué est perceptuellement identique au spectre original. Un pirate ne peut déceler la présence du filigrane ni dans la représentation spatiale de l'image tatouée,

ni dans sa représentation fréquentielle.

Les figures 6.8 et 6.10 représentent différentes projections Mojette du schéma de tatouage. L'image 6.8(a) représente la projection Mojette de la partie réelle du spectre original selon la direction (1,1). La projection Mojette de la partie réelle du spectre tatoué est donnée en 6.8(b).

Rappelons que dans les courbes présentant les projections Mojette, l'axe des abscisses correspond à la position des bins dans la projection, l'ordonnée correspond aux valeurs de bins.

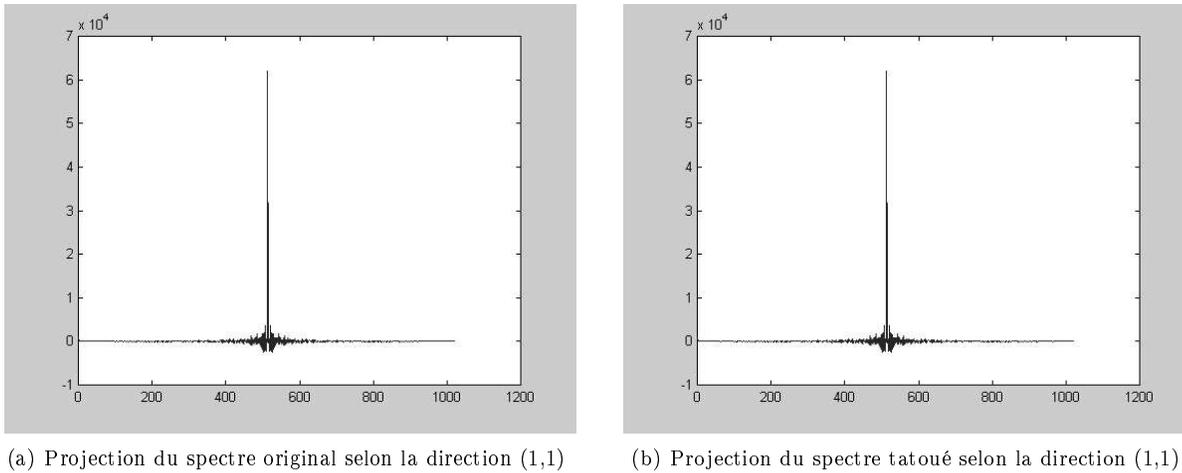


FIG. 6.8 – Images et spectres de l’algorithme de tatouage (2/4)

Selon ces images, la présence du filigrane dans le spectre est indécélable. Un pirate ne peut pas se baser sur ces projections pour déceler la présence d’une marque éventuelle.

Cependant, la connaissance du filigrane nous permet d’effectuer un zoom sur la zone tatouée de la sous-bande, ceci dévoile la présence du fantôme. Les figures 6.9 ci-dessous présentent ces zooms et montrent la présence du fantôme dans le spectre.

Ici, nous présentons (figure 6.9) différents zooms sur les projections Mojette, afin de découvrir la présence du fantôme dans le spectre. Un zoom sur la projection de la sous-bande originale est présenté figure 6.9(a). Un second zoom sur la projection du fantôme est donné sur l’image 6.9(b). Un troisième zoom sur la sous-bande tatouée est présenté sur l’image 6.9(c). Et enfin, la courbe présentée en 6.9(d) indique un zoom sur la projection de la sous-bande tatouée et attaquée par un filtrage passe-bas de noyau 3×3 .

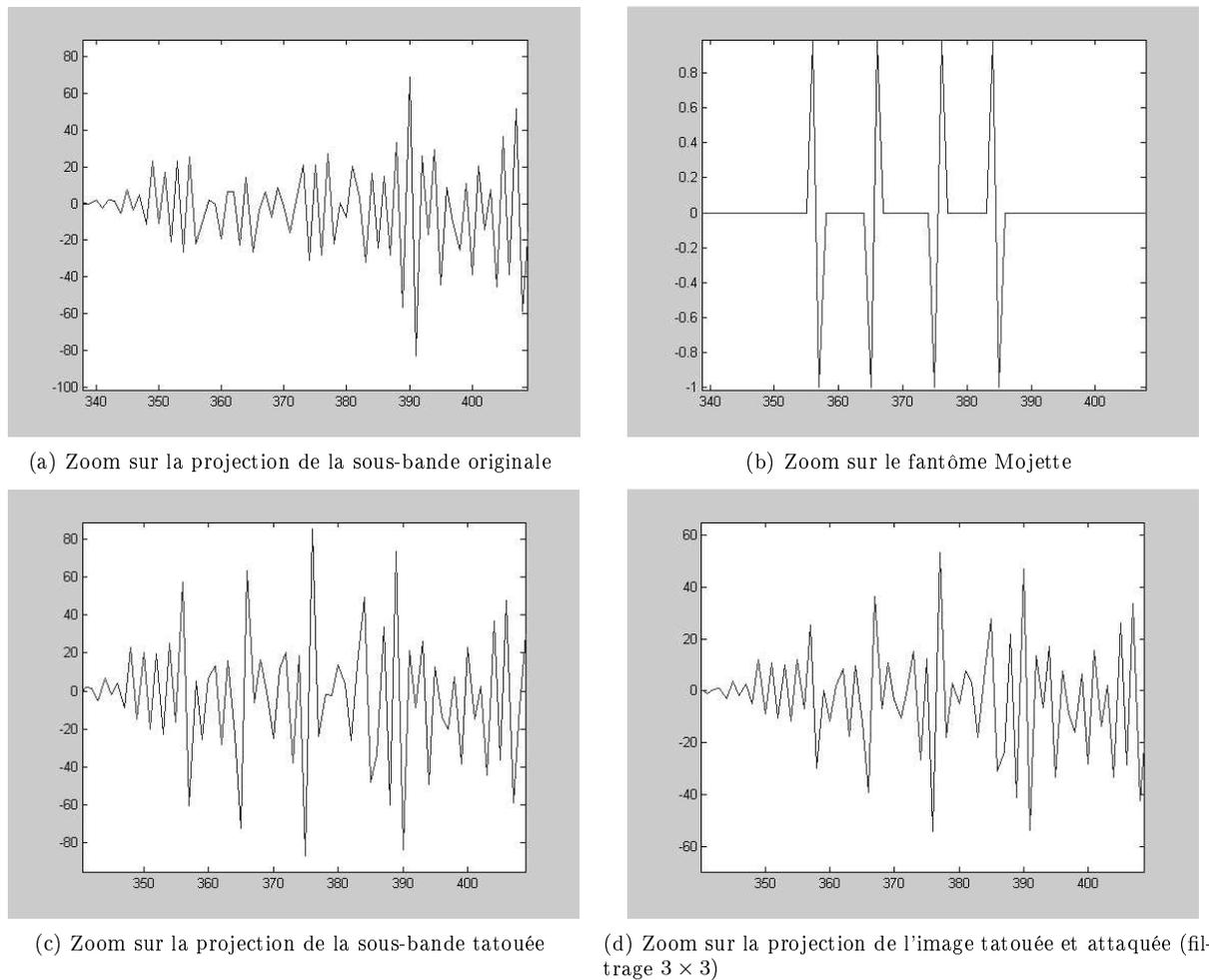
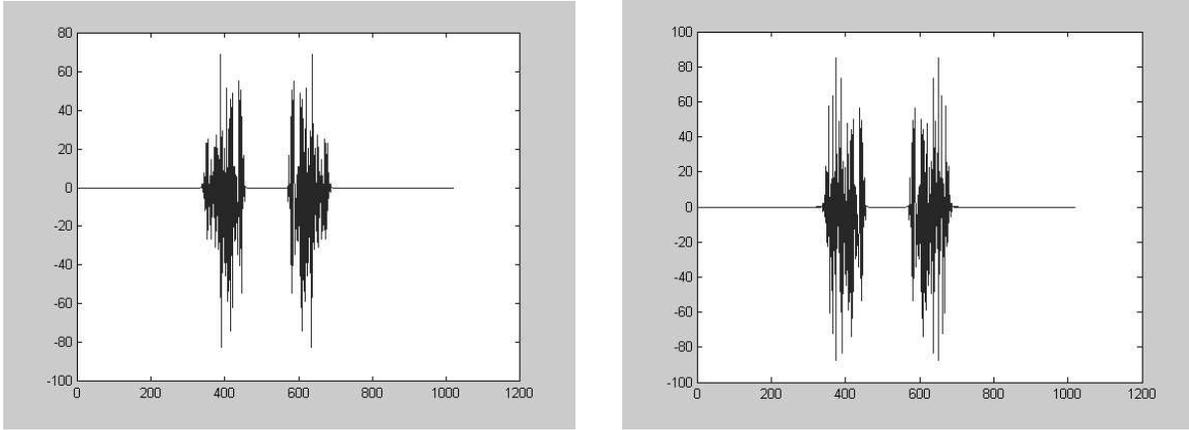


FIG. 6.9 – Images et spectres de l’algorithme de tatouage (3/4)

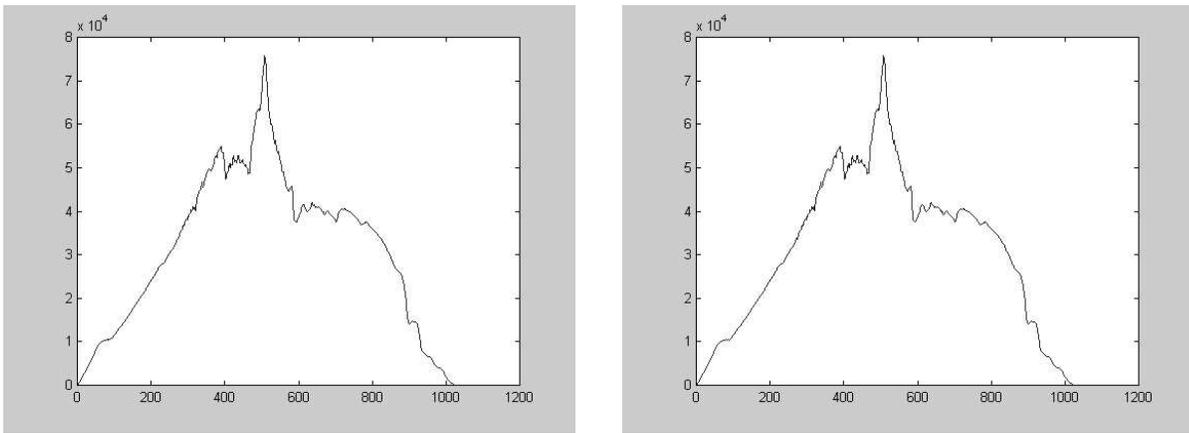
Ces images confirment clairement la présence du fantôme non seulement dans le spectre tatoué, mais aussi dans la projection Mojette des sous-bandes. Nous observons qu’après une attaque de l’image telle qu’un filtrage passe-bas, la marque est toujours présente dans le spectre.

Les projections des représentations fréquentielle des sous-bandes originale et tatouée sont présentées sur les images 6.10(a) et 6.10(b).

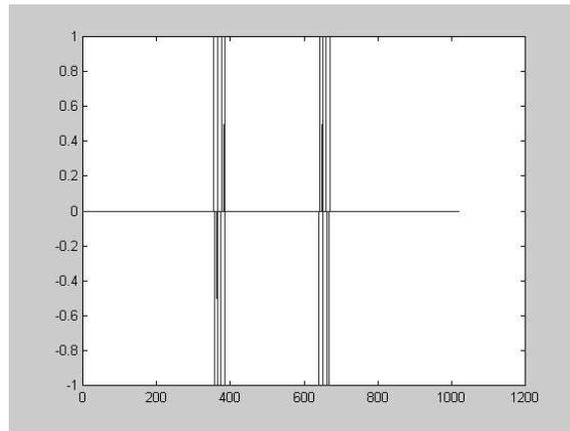
Enfin, les images, 6.10(c) et 6.10(d) représentent respectivement la projection de l’image originale et celle de l’image tatouée selon la direction $(1, 1)$. L’image 6.10(e) représente la projection du fantôme confiné à la sous-bande $(IV, 6)$ toujours selon la direction $(1, 1)$.



(a) Projection de la sous-bande originale selon la direction (1,1) (b) Projection de la sous-bande tatouée selon la direction (1,1)



(c) Projection de l'image 'lena' selon la direction (1,1) (d) Projection de l'image tatouée selon la direction (1,1)



(e) Projection du fantôme selon la direction (1,1)

FIG. 6.10 – Images et spectres de l’algorithme de tatouage (4/4)

6.6 Expérimentations

Comme nous l'avons vu, l'utilisation de critères perceptuels, permet ici de garantir l'invisibilité de la marque. La robustesse du schéma est étudiée ci-dessous par des tests de détection face à différentes attaques.

Nous pouvons observer l'allure générale des courbes d'intercorrélation sur la figure 6.11. Ici, l'insertion du filigrane s'est effectuée dans les moyennes fréquences horizontales du spectre (sous-bande $(IV, 1)$).

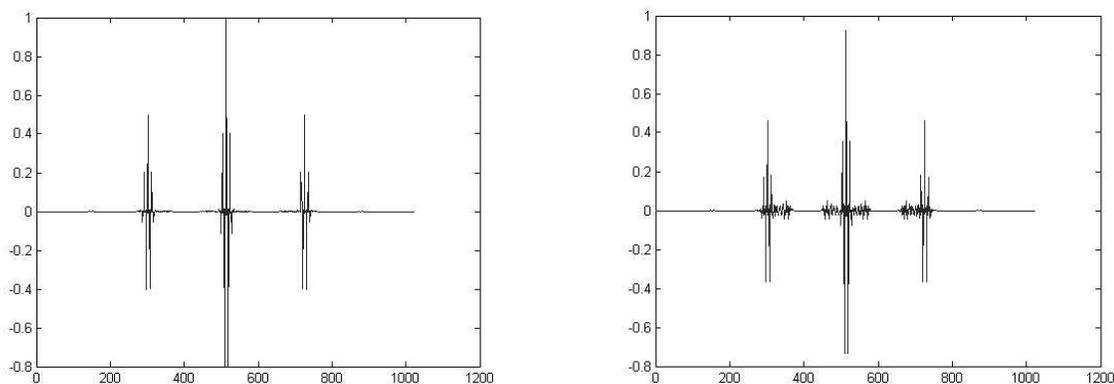
Deux possibilités de détection s'offrent à nous ici, la première option (figure 6.11(a)) consiste à calculer la fonction d'intercorrélation entre la projection du fantôme et le résultats de la différence entre la projection du spectre tatoué et la projection du spectre original. Cette méthode consiste à calculer la corrélation entre le filigrane original et le filigrane extrait.

La seconde option (figure 6.11(b)) consiste à calculer la fonction d'intercorrélation entre la projection du fantôme Mojette et la projection de la sous-bande tatouée.

Bien que la première méthode offre généralement de meilleurs résultats, l'inconvénient majeur de cette dernière est la nécessité de conserver (en plus de la projection du fantôme) la projection de l'image originale afin de calculer la projection de différence.

Pour la seconde option, les seules données à conserver sont les valeurs de la projection du fantôme Mojette. Cette solution offre tout de même de très bons résultats et du fait de la faible quantité d'informations stockées, elle est plus facilement applicable dans la pratique.

Les résultats donnés dans cette section proviennent donc de la seconde option, la détection s'effectue par fonction d'intercorrélation entre la projection du fantôme et la projection de la sous-bande tatouée.



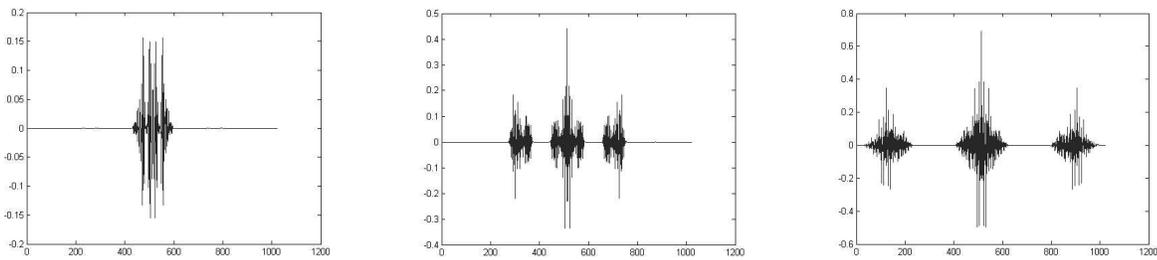
(a) Intercorrélation entre la marque et la différence des projections (b) Intercorrélation entre la marque et la projection tatouée

FIG. 6.11 – Résultats d'intercorrélations pour l'image 'bateau' tatouée

Nous constatons sur cette courbe la présence de trois pics de corrélation dus à la symétrie imposée par l'espace de Fourier. Le pic de corrélation situé au centre de la courbe correspond à la corrélation sans déphasage entre la projection du fantôme et celle de la sous-bande tatouée. Le pic de corrélation situé sur la gauche de la courbe (environ à l'abscisse 300) correspond au calcul de la corrélation de la moitié droite du fantôme avec la moitié gauche de la sous-bande. Et enfin, le pic de corrélation situé à droite (abscisse de 750) correspond à la corrélation entre la partie gauche du fantôme et la partie droite du spectre.

Les figures 6.12 et 6.13 présentées ci-dessous, présentent les courbes de l'algorithme de détection pour l'image 'iba' en fonction des fréquences radiales du spectre.

La détection de la marque pour une image tatouée (sans attaques) dans les sous-bandes $(III, 3)$, $(IV, 4)$ et $(V, 4)$ sont respectivement illustrées par les courbes 6.12(a), 6.12(b) et 6.12(c).

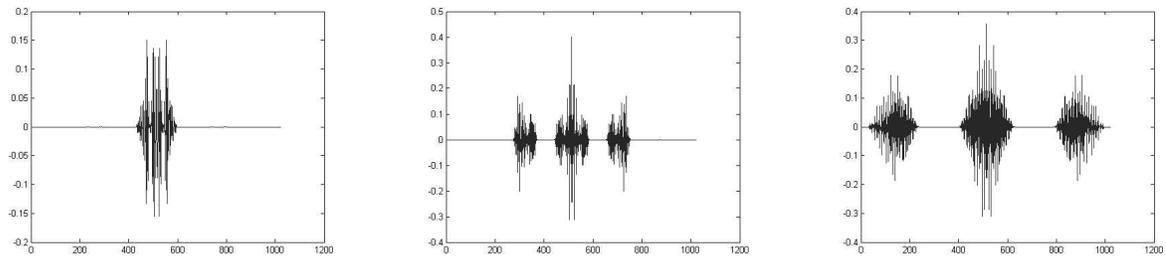


(a) Détection dans la sous-bande $(III, 3)$ (b) Détection dans la sous-bande $(IV, 4)$ (c) Détection dans la sous-bande $(V, 4)$

FIG. 6.12 – Résultats de détection dans les sous-bandes pour l'image 'iba'

Nous constatons sur ces courbes que le tatouage de la sous-bande $(III, 3)$ (Basses fréquences) n'autorise pas l'application d'une force élevée, ceci se traduit donc par un faible pouvoir de détection. Comme nous l'avons vu précédemment, il est reconnu que l'insertion d'une marque dans les basses fréquences de l'image entraîne des modifications aisément perceptibles, mais cependant plus robustes. A l'opposé, l'insertion d'une marque dans les hautes fréquences reste plus facilement invisible, mais est toutefois moins robuste.

Ce phénomène se traduit ici dans la figure 6.13 par les courbes de détection dans les sous-bandes précédentes, mais à présent, une attaque est appliquée sur l'image tatouée. Une compression JPEG de facteur de qualité $q=20\%$ est appliquée sur chacune des trois images tatouées. Le processus de détection fournit la courbe 6.13(a) pour l'image tatouée dans la sous-bande $(III, 3)$, l'image 6.13(b) concerne la sous-bande $(IV, 4)$ et enfin, la courbe 6.13(c) représente la détection dans la sous-bande $(V, 4)$.



(a) Détection dans la sous-bande $(III, 3)$ après codage JPEG ($q=20\%$) (b) Détection dans la sous-bande $(IV, 4)$ après codage JPEG ($q=20\%$) (c) Détection dans la sous-bande $(V, 4)$ après codage JPEG ($q=20\%$)

FIG. 6.13 – Résultats de détection dans les sous-bandes pour l'image 'iba'

Les résultats présentés ici confirment une seconde fois le choix que nous avons porté dans la section 4.4.2 sur les sous-bandes de la couronne IV . Effectivement, nous observons que le processus de détection présente de meilleurs résultats dans la sous-bande $(IV, 4)$ (figure 6.13(b)) que dans les couronnes III et V .

Les figures 6.14 à 6.16 présentent les résultats de l'algorithme en terme de robustesse pour l'image 'bateau'. Les courbes données en figure 6.14 concernent la détection après filtrage (6.14(a)) et respectivement les taux de compression JPEG suivants : 5% pour l'image 6.14(b), 10% pour l'image 6.14(c) et 15% pour l'image 6.14(d).

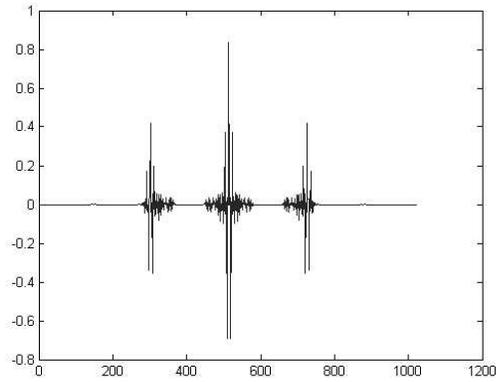
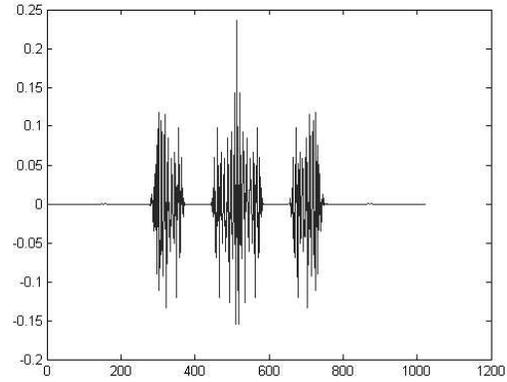
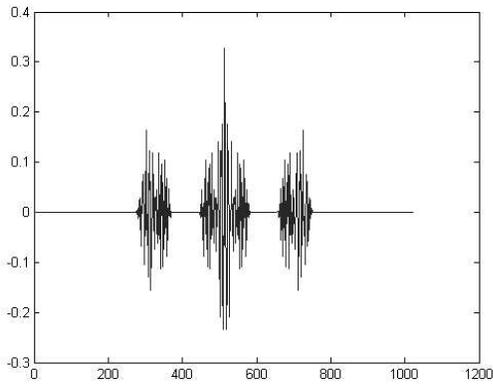
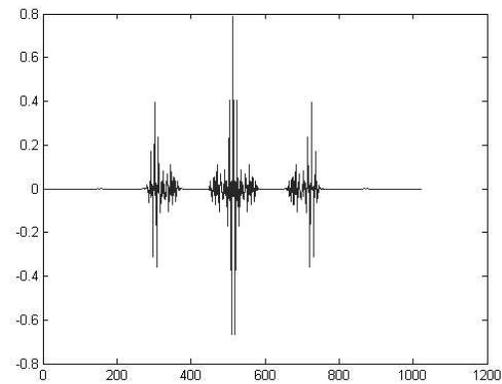
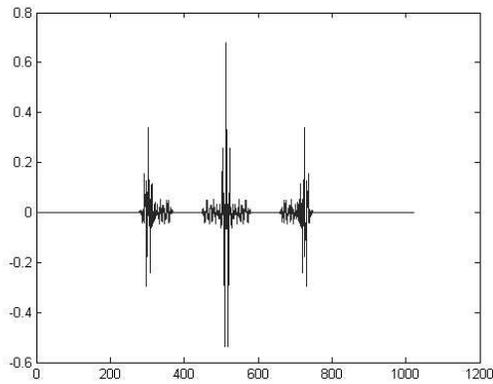
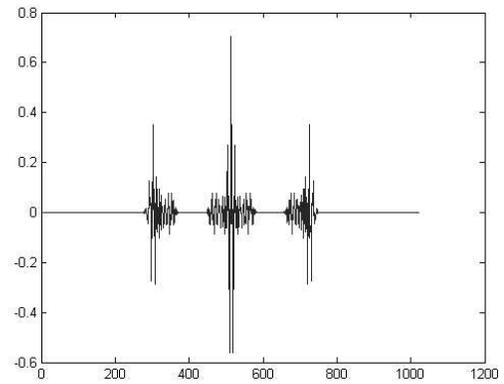
(a) Filtrage Médian (3×3)(b) Codage JPEG ($q = 5\%$)(c) Codage JPEG ($q = 10\%$)(d) Codage JPEG ($q = 15\%$)

FIG. 6.14 – Résultats d'intercorrélation entre le fantôme original et l'image tatouée

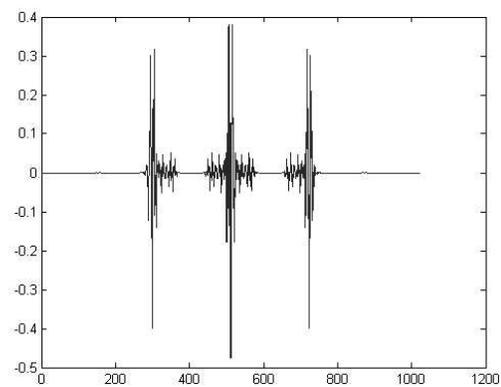
Cet algorithme de tatouage n'ayant pas pour objectif principal la résistance face aux attaques propres au contenu de l'image (filtrages, compression), nous observons tout de même dans ce contexte de très bons résultats (figure 6.14).



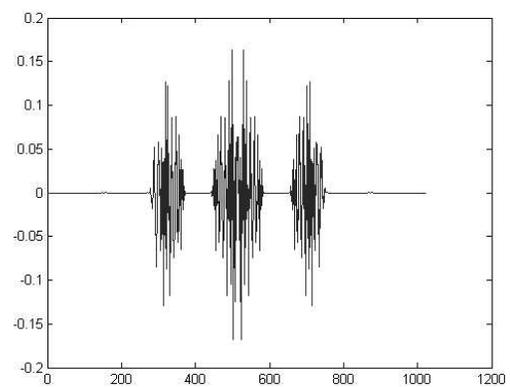
(a) Suppression d'une ligne et d'une colonne



(b) Suppression de 17 lignes et de 5 colonnes



(c) Cropping (1%)



(d) Cropping (10%)

FIG. 6.15 – Résultats d'intercorrélation entre le fantôme original et l'image tatouée

Les résultats de la détection concernant la rotation sont donnés sur la figure 6.16, nous constatons que le schéma proposé résiste à des rotations de 5° , figure 6.16(a), ou de 0.75° , figure 6.16(b). Les trois pics de corrélations apparaissent clairement sur ces courbes.

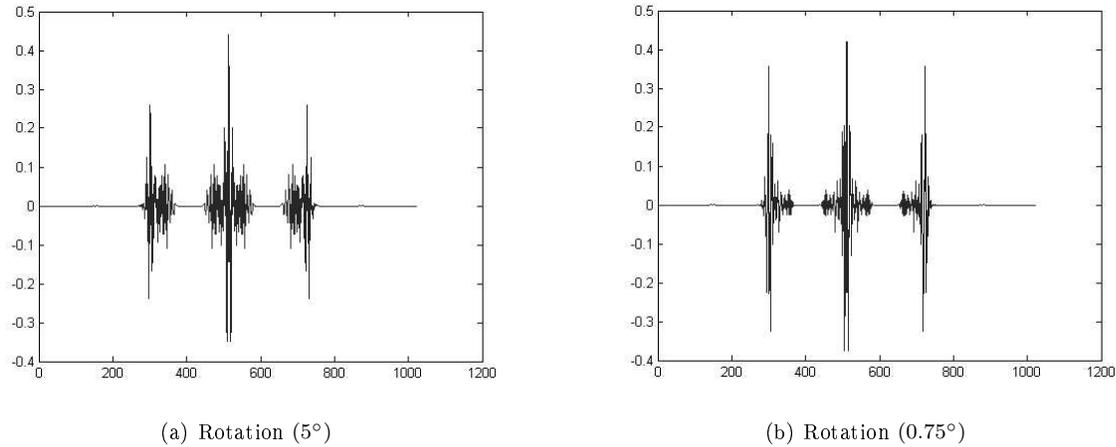


FIG. 6.16 – Résultats d'intercorrélations entre le fantôme original et l'image tatouée

Observons à présent les résultats de détection pour une seconde image. Les résultats de détection obtenus pour l'image 'fruit' sont présentés ci-dessous.

Les deux possibilités de détection détaillées précédemment sont présentées figure 6.17, la première option (corrélations entre la marque originale et la marque extraite) est donnée sur la figure 6.17(a) alors que la seconde option (corrélations entre la marque et la sous-bande tatouée) est donnée par la figure 6.17(b). Comme nous l'avons prévu précédemment, nous observons que la première option fournit de meilleurs résultats, mais cependant les deux possibilités permettent la détection de la marque.

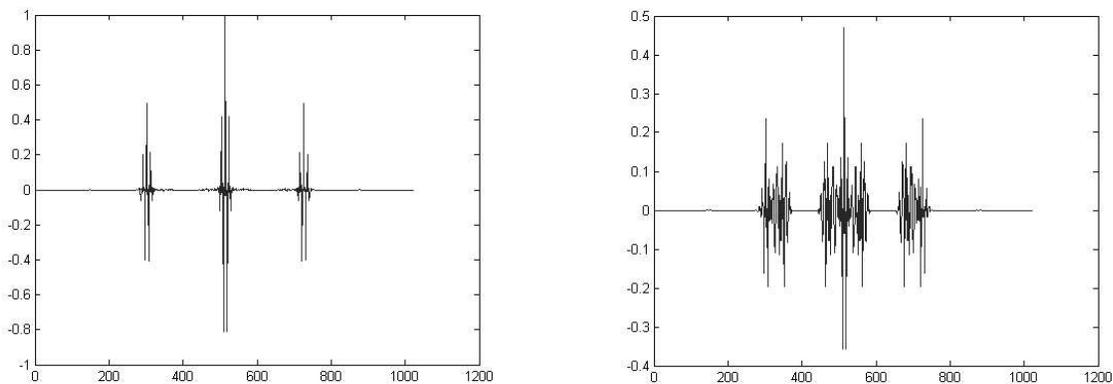
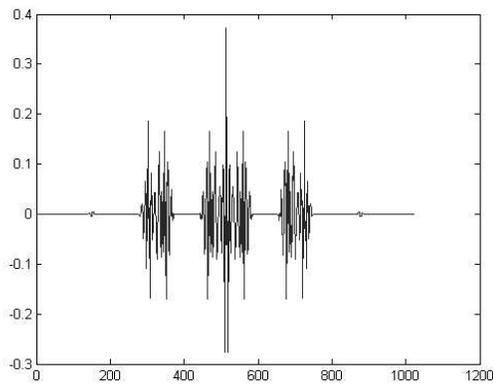


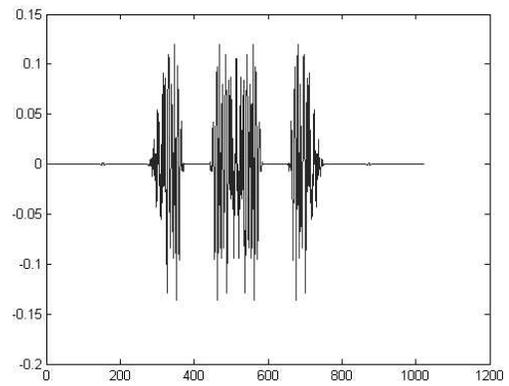
FIG. 6.17 – Résultats d'intercorrélations pour l'image tatouée

Les tests de détection pour des filtrages et taux de compression JPEG étudiés précédemment pour l'image

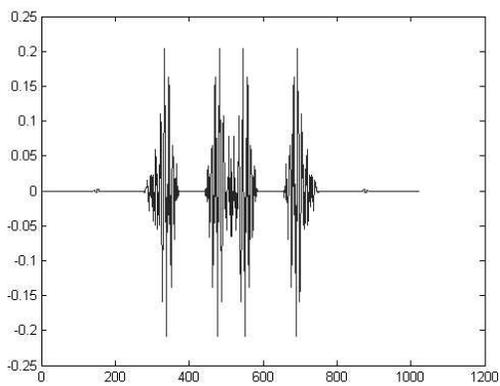
'bateau' sont reconduits ici pour l'image 'fruit' (figure 6.18).



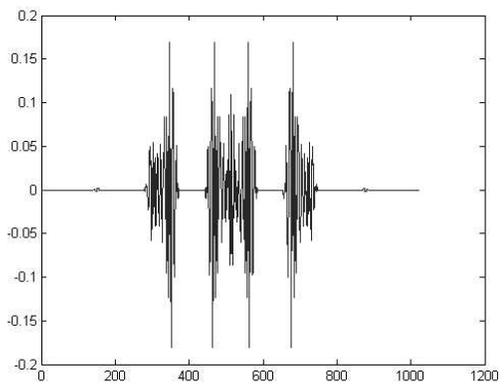
(a) Filtrage Médian (3×3)



(b) Codage JPEG ($q = 5\%$)



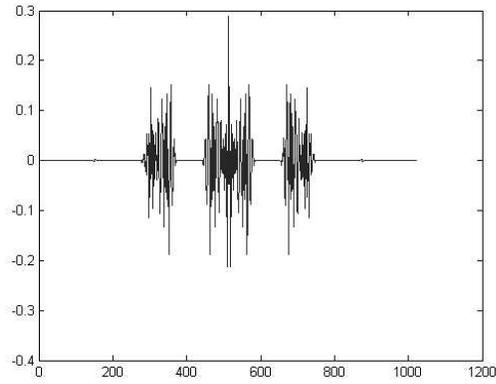
(c) Codage JPEG ($q = 10\%$)



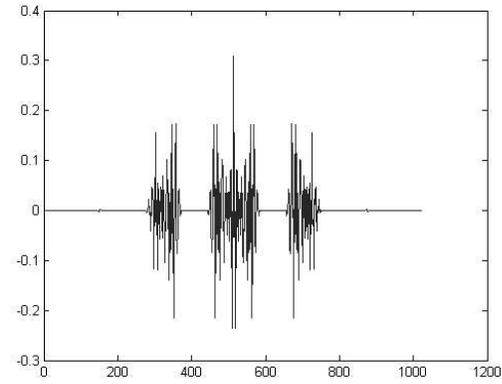
(d) Codage JPEG ($q = 15\%$)

FIG. 6.18 – Résultats d'intercorrélation entre le fantôme original et l'image tatouée

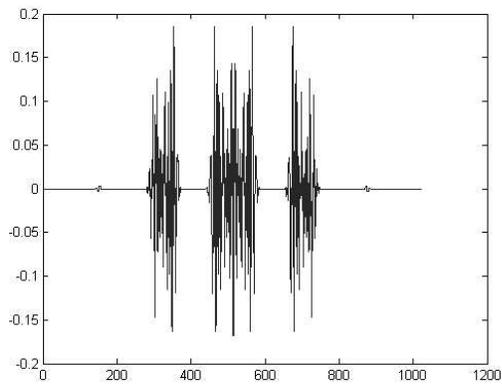
Les résultats du processus de détection pour certaines attaques géométriques sont présentés ci-dessous dans la figure 6.19. Des suppressions de lignes et de colonnes sont présentées sur les courbes 6.19(a) et 6.19(b).



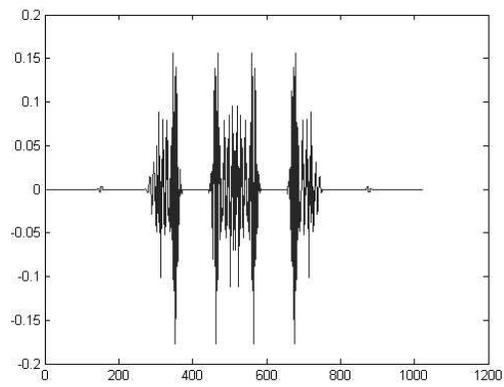
(a) Suppression d'une ligne et d'une colonne



(b) Suppression de 17 lignes et de 5 colonnes



(c) Cropping (1%)



(d) Cropping (10%)

FIG. 6.19 – Résultats d'intercorrélation entre le fantôme original et l'image tatouée

Une rotation de l'image tatouée de respectivement 5° et 0.75° fournit les résultats donnés sur les courbes 6.20(a) et 6.20(b).

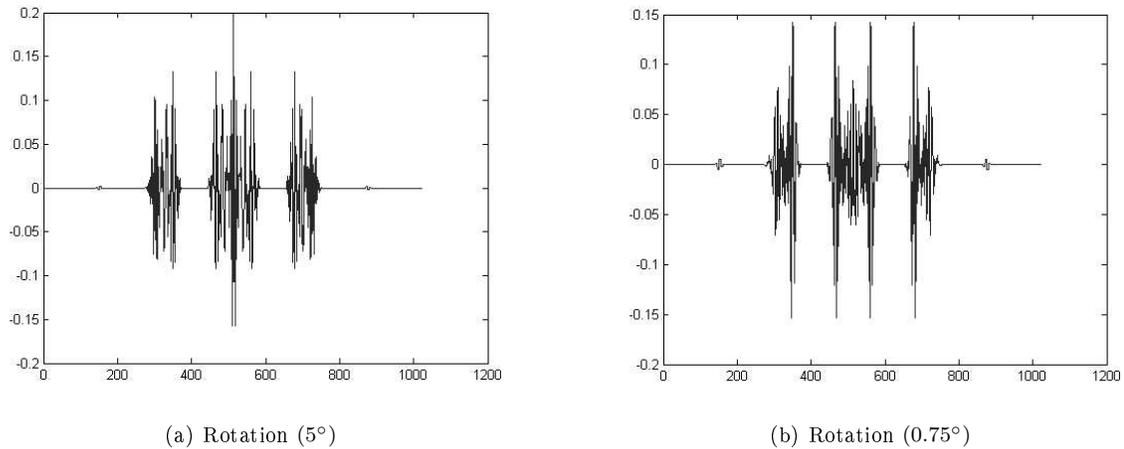


FIG. 6.20 – Résultats d'intercorrélation entre le fantôme original et l'image tatouée

6.7 Conclusion

Nous avons présenté dans ce chapitre un schéma de tatouage exploitant conjointement les propriétés de la transformation Mojette, respectivement détaillées et mise en œuvre dans les chapitre 2 et 5 et la modélisation du S.V.H. détaillée et utilisée dans les chapitres 1 et 4. Le schéma de tatouage proposé ici, opère directement sur les coefficients du spectre de Fourier. Un fantôme Mojette pondéré à l'aide des critères psychovisuels est inséré dans une sous-bande moyennes fréquences sélectionnée en fonction de son activité.

Le processus de détection exploite les projections Mojette. La présence du filigrane dans l'image est détectée par le calcul de la fonction d'intercorrélation entre la projection potentiellement tatouée et la projection du fantôme selon la même direction Mojette.

Nous constatons que le schéma proposé est très résistant face à diverses attaques dont plusieurs attaques géométriques opérant une dé-synchronisation de la marque.

Chapitre 7

Contribution au projet Aquamars

Sommaire

| | | |
|------------|-----------------------------------|------------|
| 7.1 | Introduction | 206 |
| 7.2 | Contexte général du projet | 206 |
| 7.3 | Description des travaux | 207 |
| 7.3.1 | Détail des sous-projets | 207 |
| 7.3.2 | Contribution de l'équipe IVC | 208 |
| 7.4 | Conclusion | 209 |

7.1 Introduction

Une partie importante des travaux effectués au cours de cette thèse, notamment sur l’aspect psychovisuel, et plus particulièrement sur la création de masques perceptuels, se sont inscrits dans le contexte du projet “AQUAMARquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en toute Sécurité sur les réseaux hétérogènes de télécommunication” (AQUAMARS). Aquamars était un projet précompétitif labellisé par le Réseau National de Recherche en Télécommunications (RNRT).

Ce projet a officiellement débuté le 1/03/1999 et s’est déroulé sur une période de 24 mois. Différents partenaires, industriels et universitaires, ont collaboré sur la mise en oeuvre d’un schéma complet de tatouage d’images.

Les équipes ayant participé au projet sont :

- Heudiasyc de l’Université de Compiègne (UTC)
- TIMC-IMAG de l’Université Joseph Fourier de Grenoble, actuellement LIS
- SEI-IVC de l’Université de Nantes (IRESTE), actuellement IRCCyN
- l’Institut Eurecom
- l’Ecole Nationale Supérieure des Télécommunications ENST
- Thomson CSF Communications (TCC), actuellement THALES
- Netimage

Nous verrons dans ce chapitre, le contexte général du projet, ainsi qu’une description des travaux réalisés par chaque partenaire.

7.2 Contexte général du projet

Nous avons pu voir précédemment qu’il existe différentes applications au tatouage d’image, le projet RNRT Aquamars se positionne dans le contexte de la préservation et de la protection des droits de propriété de matériel audiovisuel. Cet aspect de protection des droits d’auteurs, peut toutefois englober différentes approches de tatouage, tels l’identification de l’oeuvre, le suivi de l’oeuvre, par le biais de fingerprints, ou même le contrôle de l’intégrité et de l’authenticité. Cependant, il n’était pas exclus dans le projet Aquamars d’utiliser le watermarking comme indicateur de contenu, pour de l’indexation ou du contrôle parental par exemple.

Le projet Aquamars avait donc pour but, la création d’une chaîne complète de tatouage concernant les images fixes, une méthode de tatouage de flux MPEG-2 compressé, ainsi qu’une préparation a l’introduction du tatouage dans les techniques de compression, tel que du tatouage d’objets MPEG 4. Il était aussi envisagé

de développer un schéma complet de protection de flux 4 :2 :2.

Le projet s'est organisé autour des six sous-projets suivants :

- Approche système de tatouage
- Approche psychovisuelle
- Développement de tatoueurs efficaces
- Réalisation de démonstrateurs image fixe et MPEG 4
- Marquage de flux MPEG 2
- Application au contrôle d'intégrité

La liaison de ces différents sous-projets a permis de nombreuses interactions entre les différents partenaires.

Nous allons à présent détailler ces différents sous-projets, ainsi que les implications des partenaires.

7.3 Description des travaux

7.3.1 Détail des sous-projets

Le premier sous-projet concernant l'“Approche système du tatouage”, a impliqué tous les partenaires. Son enjeu était de proposer au concepteur d'une chaîne de communication d'imagerie, une technique de tatouage ainsi que ses paramètres. Ce premier sous-projet intégrait aussi le développement du système d'enregistrement et de certification, partie mise à la charge de Netimage et TCC. La responsabilité de ce sous-projet était confiée à l'ENST en termes d'aspects théoriques ainsi qu'à Netimage pour les aspects d'intégration et de développement.

Le second sous-projet, concernant l'approche psychovisuelle du tatouage et impliquant l'IRESTE, sera détaillé plus avant dans la section suivante.

Le troisième sous-projet nommé “développement de techniques efficaces et adaptées au tatouage” était coordonné par TIMC-IMAG. Dans ce sous-projet, une étude des différents espaces de tatouage, ainsi qu'une évaluation de ces méthodes en termes de robustesse était envisagée. Parmi les différentes approches, une méthode basée sur la démarche utilisée en compression fractale, a été retenue. L'objectif est donc ici, de chercher à résister aux attaques géométriques.

Le quatrième sous-projet avait pour objectif, la création de démonstrateurs, l'un pour MPEG 4 et l'autre pour des images de musées. Ici, le coordonnateur était TCC. Concernant MPEG 4, une définition des attaques que peut subir le flot de données, et en particulier, les objets audiovisuels était envisagée. Dans le cadre des applications aux oeuvres d'art, l'effort était placé sur la compatibilité entre l'aspect invisibilité de la marque

et les contraintes de copyright de documents de très forte valeur.

Le cinquième sous-projet, organisé par TCC, avait pour objectif l'adaptation d'une technique existante de tatouage de flux MPEG-2 d'une part et la validation des travaux des différents partenaires par l'intégration dans ce même procédé d'autre part. Plusieurs coopérations ont eu lieu, parmi lesquelles, par exemple l'adaptation de la méthode proposée par TIMC-IMAG au tatouage de flux MPEG, l'évaluation de la qualité par l'IRESTE, sur les images tatouées, ou encore l'analyse de la robustesse par l'ENST.

Et enfin, le sixième sous-projet, dirigé par Eurécom visait à étudier des applications de tatouage autres que la protection de propriété. La vérification de l'intégrité des images faisait partie de ces applications.

| Organisme | Sous-projet |
|----------------------------|-------------|
| TCC | 1-2-3-4-5-6 |
| ENST, Paris | 1-2-3-4-5-6 |
| TIMC - IMAG - UJF Grenoble | 1-3-4-5 |
| IRESTE - Nantes | 1-2-4 |
| Heudiasyc - UTC Compiègne | 1-3-4-5 |
| Eurécom - Sophia Antipolis | 1-4-6 |
| Netimage | 1-2-4 |

TAB. 7.1 – Participation des organismes aux sous-projets

Les contributions des partenaires prévues dans les différents sous-projets sont représentées dans le tableau 7.1 ci-dessus.

7.3.2 Contribution de l'équipe IVC

Comme nous l'avons vu, le rôle de l'IRESTE, plus précisément de l'équipe Image et Vidéocommunications (IVC), était de développer et valider l'approche psychovisuelle du tatouage. L'étude psychovisuelle du tatouage s'est effectuée selon deux axes principaux.

L'utilisation d'un critère objectif de qualité des images tatouées était le premier axe des travaux de l'équipe IVC dans le projet AQUAMARS. Ces travaux étant déjà réalisés au sein de l'équipe et ayant fait l'objet de plusieurs thèses [15], [22], nous ne les développons pas ici.

Le second axe, en rapport plus direct avec le domaine du tatouage d'images était la création d'un masque psychovisuel ayant pour but l'optimisation de la méthode de tatouage, avec une étude préalable des espaces de tatouages possibles. Cette étude des espaces s'est effectuée principalement sur deux espaces transformés : le domaine de Fourier ainsi que le domaine DCT. Les résultats obtenus à l'IRESTE sur les aspects psychovisuels ont été utilisés dans les sous-projets 1, 3, 4 et 5.

Notre participation au projet a consisté tout d'abord à dresser d'une part, un état de l'art exhaustif de modèles psychovisuels, ainsi que le modèle proposé par l'équipe IVC [106], et d'autre part, un état de l'art concernant les techniques de tatouage [105].

Concernant la mise en oeuvre d'un masque psychovisuel spatial, les travaux présentés dans les chapitres 4.2 et 4.3.3 ont fait partie intégrante de la participation de l'équipe IVC dans le projet AQUAMARS. Le masque JND présenté a donc pu être adapté par les différents partenaires pour l'adapter à leurs méthodes de tatouage ainsi qu'à leur espace d'insertion. De cette façon, l'ENST a pu adapter sa méthode de tatouage par blocs DCT au masque spatial [128].

7.4 Conclusion

Nous avons vu que le projet Aquamars avait pour but de proposer des solutions à l'un des maillons de la chaîne de sécurité des réseaux de télécommunications : le tatouage d'images ou de documents vidéo, appelé watermarking par les anglo-saxons ou aquamarquage pour sa traduction littérale. Le projet Aquamars a donc réuni sept partenaires industriels ou universitaires, afin de bénéficier des compétences de chacun dans les différents aspects du tatouage. Deux démonstrateurs ont été produits à l'issue du projet, l'un pour du tatouage de flux MPEG-2 et l'autre pour du tatouage de banque d'images fixes et de lecture des marques. La participation de l'IRESTE dans le projet Aquamars s'est effectuée sur les aspects psychovisuels du tatouage.

Dans cette étude, nous avons détaillé la création d'un masque perceptuel, ce dernier a été adapté dans différents domaines transformés (DCT, DWT), puis amélioré lors de sa mise en œuvre dans une technique complète de tatouage opérant dans le domaine spatial. L'utilisation de masques perceptuels permet de définir pour tout pixel de l'image la puissance maximale permise à la marque pour optimiser le compromis entre l'invisibilité et la robustesse. Cependant, l'utilisation d'un masque contraint le tatoueur à conserver ce dernier en vue de procéder à la détection ou à l'extraction de la marque. Une méthode de tatouage opérant directement dans la représentation spatiales des sous-bandes psychovisuelles a ensuite été établie. Les techniques présentées de tatouage perceptuel présentent de nombreux avantages : garantie de l'invisibilité, masques perceptuels adaptables dans différents domaines. Nous avons pu constater que de telles méthodes comportent néanmoins une faiblesse importante en terme de robustesse face à certaines attaques géométriques dé-synchronisantes.

En vue de résister à de telles attaques, une étude portant sur la création de schémas de tatouage Mojette a été conduite. Là encore, différentes applications du tatouage ont été abordées. Nous avons présenté une technique de tatouage adapté au domaine spécifique du tatouage des images médicales. Ce schéma respecte les exigences spécifiques à un tel domaine : nécessité de fournir une méthode de tatouage réversible à extraction aveugle, contrôle de l'intégrité. Dans cette seconde partie, nous avons présenté une méthode de tatouage paramétrable et adaptable au chiffrement des données opérant une modification des projections Mojette. Enfin, dans un schéma de tatouage par fantômes Mojette aisément adaptable par une simple permutation des clefs, nous avons pu explorer certains liens entre tatouage et stéganographie. Ce schéma a permis la mise en œuvre d'une technique originale et efficace de détection de la marque utilisant les projections Mojette. Il permet de plus, de profiter de la redondance fournie par la transformation Mojette pour insérer des messages secrets de longueur importante.

Cependant les schémas de tatouage par transformation Mojette sont fondés uniquement sur des considérations heuristiques en terme d'invisibilité du filigrane. De ce fait aucune optimisation du compromis invisibilité-robustesse n'est réalisée dans de tels schémas.

Cette optimisation était l'objectif du dernier schéma proposé, dans lequel nous avons détaillé une étude portant sur l'adaptation des critères visuels et de la transformation Mojette. L'insertion de fantômes Mojette

directement dans les sous-bandes fréquentielles du spectre de Fourier permettent une optimisation conjointe des avantages fournis par la modélisation psychovisuelle d'une part et par ceux de la transformation Mojette d'autre part. Ici, le compromis requis entre l'invisibilité et la robustesse de la marque est optimal. La technique de détection est emprunté au schéma de tatouage-stéganographie détaillé précédemment.

Nous avons découvert dans cette seconde partie que l'étude des critères perceptuels nous à permis d'assurer l'invisibilité de tout filigrane. Cependant, nous avons pu constater l'impuissance des méthodes de "tatouage perceptuel" vis-à-vis des attaques géométriques dé-synchronisantes. Outre l'exploitation de l'explosion combinatoire apportée par la transformation Mojette dans une méthode de cryptographie et de stéganographie, l'objectif principal de l'étude de la transformation Mojette était de proposer une solution à cette lacune. Conformément à nos attentes, l'alliance des aspects perceptuels et Mojette a permis de résister à de telles attaques.

Conclusion Générale

Comme nous l'avons vu au cours de ce mémoire, les applications du tatouage sont multiples : protection des droits d'auteurs, vérification de l'authenticité, indexation ou encore stéganographie. L'élaboration d'une méthode de tatouage peut passer par l'étude de différents aspects, parmi lesquels nous pouvons citer la gestion des clefs, le cryptage des données, la recherche de l'invisibilité de la marque, ou encore l'étude des attaques envisageables. Nos travaux ont essentiellement porté sur le tatouage de copyright. Les deux principales caractéristiques attendues des schémas de tatouage d'image ayant pour objectif la protection des droits d'auteurs, sont l'invisibilité et la robustesse. Ces deux caractéristiques antinomiques en imposent une troisième qui est la capacité d'information insérée. Une lacune importante commune à la plupart des méthodes de tatouage concerne les aspects perceptuels. L'optimisation conjointe de l'invisibilité et de la robustesse peut se faire soit par le biais d'une étude détaillée du système visuel humain soit, plus simplement par l'utilisation de critères empiriques. La plupart des méthodes de tatouage présentes dans la littérature proposent un procédé d'insertion basé sur des critères empiriques.

La contribution que nous avons souhaitée apporter au domaine du tatouage des images concerne justement l'exploitation des critères visuels. L'étude du système visuel humain permet l'élaboration de masques perceptuels définissant les composantes invisibles offrant la meilleure résistance aux attaques. Une telle étude est parfaitement adaptée à la problématique du tatouage d'images. Par ce biais, nous avons ainsi pu établir différentes méthodes de tatouage offrant de très bons résultats tant en termes d'invisibilité qu'en termes de robustesse.

Nous avons tout d'abord mis en œuvre trois méthodes de tatouage exploitant les critères visuels étudiés précédemment. La première est consacrée à l'élaboration ainsi qu'à la validation d'un masque perceptuel, appelé masque J.N.D. Dans ces travaux, un masque est établi en prenant en compte le découpage fréquentiel opéré par le SVH, puis, différents tests de robustesse dans ce masque nous ont conduit à y apporter quelques modifications. Ce masque perceptuel a ensuite été adapté et testé dans différents domaines : DCT par Byron

Yu de l'ENST Paris, ondelettes par Anne Manoury de l'IRCCyN de Nantes. De plus, nous avons mis en œuvre une méthode d'adaptation du masque dans le domaine spatial par modification des bits de poids faible. Ces différentes méthodes de tatouage ont été exposées. Nous avons ensuite élaboré une méthode de tatouage des sous-bandes visuelles. L'avantage principal offert par ces différents schémas est la garantie de l'invisibilité quelle que soit la marque ou le domaine de tatouage. Cependant, ces schémas comportent certains inconvénients. Les deux inconvénients majeurs des schémas exploitant le masque psychovisuel sont d'une part la nécessité de conserver le masque et d'autre part l'incapacité à détecter la marque lors de déformations géométriques sévères. Pour résister à de telles attaques, un procédé de compensation des attaques géométriques doit être envisagé le plus souvent.

Nous avons alors souhaité apporter une amélioration de la robustesse vis-à-vis des attaques géométriques, par l'étude d'une transformation présentant des aspects intéressants : la transformation Mojette. Cette dernière comporte certaines propriétés intéressantes telles que la notion de fantôme Mojette ou encore la redondance. L'étude de la transformation Mojette nous a permis l'exploration de domaines connexes au tatouage d'images tels que la stéganographie, la cryptographie, le tatouage multiple, la gestion des clefs ou encore la vérification de l'authenticité.

Nous avons présenté trois méthodes de tatouage basées sur les propriétés de la transformation Mojette. Le premier schéma a pour objectif le tatouage d'images médicales. Celui-ci opère l'insertion d'un fantôme Mojette dans les plans binaires de poids faible de l'image. Dans cette méthode, des études portant sur la gestion des clefs, ainsi que la complexité du schéma ont été nécessaires. La seconde méthode concerne une application de tatouage de flux Mojette. Ici, l'insertion s'opère par addition d'information sur les éléments des projections Mojette. La transformation Mojette inverse des projections tatouées induit naturellement un brouillage de l'image reconstruite. Ce procédé a été appliqué sur différents plans binaires de l'image voire même directement sur les pixels en niveaux de gris. Le but de ce schéma de tatouage est de pouvoir s'adapter à diverses contraintes telles que le tatouage imperceptible, la cryptographie ou encore l'accessibilité progressive aux données. La complexité de détection (de décryptage), aspect très important de cette méthode, a aussi été détaillée. Nous avons enfin présenté un schéma de protection de copyright adaptable à la stéganographie. Ce schéma est basé sur les deux précédents, et se place dans un contexte de transmission. L'insertion d'un fantôme Mojette dans les plans binaires de poids faible permet la diffusion des projections originales et de la projection tatouée sur un réseau de transmission. Comme nous l'avons vu, le possesseur de la clef pourra remplacer la projection tatouée par sa version originale et ainsi, reconstruire l'image originale. Toute personne ne possédant pas la clef se verra contrainte de rétroprojeter la projection erronée et n'obtiendra ainsi qu'une version tatouée

de l'image. Comme nous avons pu le constater, l'échange de la projection tatouée par sa version originale permet la constitution d'un schéma de stéganographie.

L'inconvénient majeur rencontré dans ces méthodes de tatouage basées sur la transformation Mojette est l'absence de critères visuels. Effectivement lors des méthodes de tatouage Mojette, l'invisibilité est garantie par le tatouage des plans binaires de poids faible. De plus, l'exploitation de la transformation Mojette pour des techniques de tatouage dans le domaine spatial permet la résistance du filigrane face à de faibles déformations géométriques, mais reste tout de même impuissante vis-à-vis des attaques de plus grande amplitude opérant une désynchronisation sévère de la marque, comme les rotations.

Nous avons alors souhaité réaliser une technique de tatouage mettant à profit les différents avantages fournis tant par l'utilisation du modèle perceptuel dans l'optimisation conjointe de l'invisibilité et de la robustesse, que par les propriétés spécifiques de la transformation Mojette en termes de robustesse.

Deux caractéristiques principales du domaine de Fourier nous ont conduits à exploiter cet espace transformé lors de l'élaboration de la méthode de tatouage. Tout d'abord, le domaine de Fourier comporte des caractéristiques intéressantes face aux attaques géométriques. Ensuite, la modélisation du système visuel exploité ici opère également dans l'espace de Fourier. Nous avons donc envisagé l'élaboration d'une technique de tatouage du spectre de l'image. Le filigrane se présente ici sous la forme d'un fantôme Mojette, ce dernier est pondéré par le modèle visuel utilisé précédemment puis inséré dans la représentation spectrale d'une sous-bande visuelle. Les résultats encourageants d'une technique de détection de la marque employée précédemment lors d'un procédé de tatouage dans l'espace Mojette est encore exploité ici avec succès face à diverses attaques. Une telle méthode permet la détection de la marque pour une image ayant subi différentes attaques géométriques, sans avoir besoin d'effectuer au préalable de compensation des attaques.

Perspectives

Les aspects perceptuels n'ont pas été suffisamment étudiés dans les applications de tatouage d'images. De nombreuses études doivent encore être effectuées afin de permettre une optimisation conjointe de l'insertion et de la détection de la marque. Une étude des dégradations causées par les différentes attaques permettrait d'anticiper le comportement de l'image lors de l'attaque et rendrait ainsi possible le tatouage des zones de l'image les moins affectées par la dégradation.

Parallèlement à de telles études, une évaluation des dégradations permettrait de classer les attaques fournissant des images exploitables, des attaques trop sévères ne permettant plus la commercialisation ou une quelconque exploitation de l'image.

Les méthodes utilisant des masques perceptuels souffrent actuellement d'un manque certain de robustesse vis-à-vis des attaques géométriques. De ce fait, différents travaux portant sur l'adaptation des masques perceptuels aux domaines fréquentiels seraient souhaitables. Une étude approfondie sur la prise en compte des attaques, ainsi que sur le contenu des images doit absolument être envisagée lors de l'élaboration des masques perceptuels. En résumé, pour ces deux types d'études futures à conduire, il faudrait un schéma d'optimisation de la marque en boucle fermée avec des dégradations et l'analyse psychovisuelle. On peut dire que l'analyse en boucle ouverte réalisée dans le cadre de cette thèse fournit un premier pas prometteur. La seconde étape sera d'incorporer la phase d'optimisation globale comportant les blocs de création de masque psychovisuel, d'insertion du tatouage, de dégradations et de récupération de cette marque. Le but de cette optimisation sera à la fois d'obtenir un filigrane de longueur maximale pour tatouer tout en restant perceptuellement invisible et décodable.

Une extension des travaux portant sur l'utilisation de la transformation Mojette serait souhaitable. Cette transformation présentant des caractéristiques très intéressantes vis-à-vis des attaques géométriques ou encore en termes de capacité d'insertion de la marque.

Les liens existant entre le tatouage d'images, la stéganographie ou encore la cryptographie peuvent être

explorés par le biais de la transformation Mojette. Certains aspects des méthodes de tatouage situées dans un contexte de transmission, telles que certaines méthodes présentées dans ce document peuvent être améliorés. Par exemple, le couplage d'une méthode de tatouage avec un processus de codage à description multiple, ou encore un codage conjoint source canal peuvent être étudiés. Une étude des dégradations apportées par la transmission des données sur un réseau permettrait une optimisation de la résistance de l'algorithme aux erreurs binaires de transmission.

Nos travaux portant sur l'insertion d'un fantôme Mojette dans les sous-bandes visuelles du spectre permettent la prise en compte du SVH dans un procédé de tatouage résistant aux déformations géométriques. Cependant, une optimisation de l'ajustement de la représentation spatiale du fantôme dans le masque permettrait d'accroître la robustesse de la méthode tout en conservant la certitude de l'invisibilité de la marque insérée.

Annexe A

Psychovisuel

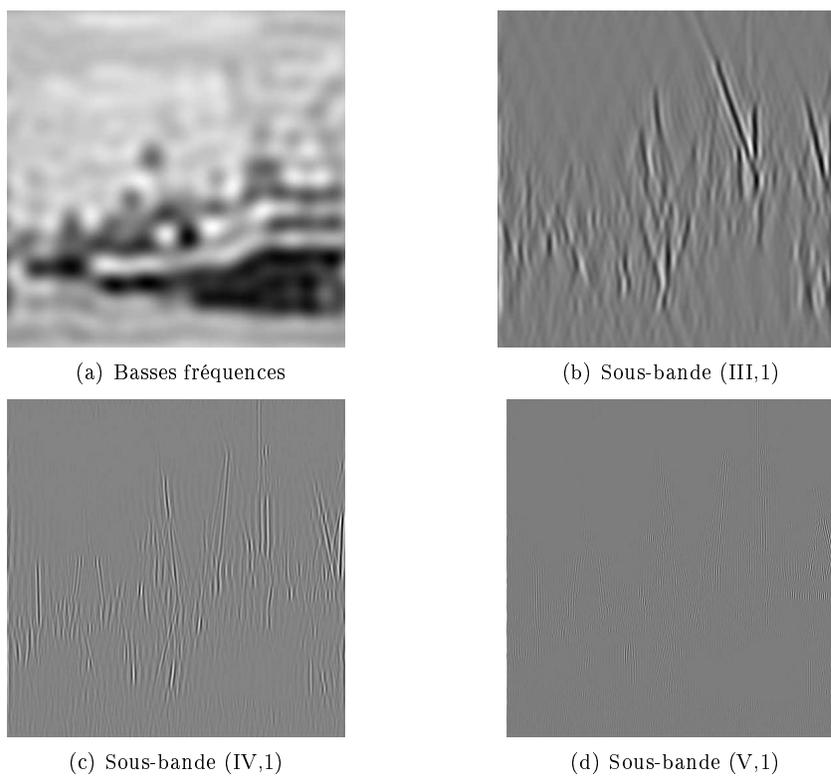
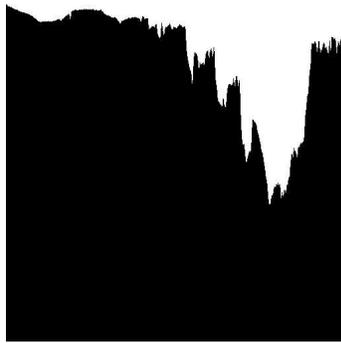


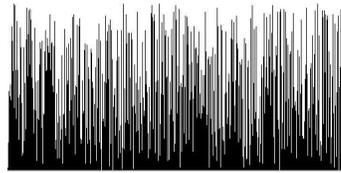
FIG. A.1 – Représentation spatiale de sous-bandes visuelles

Annexe B

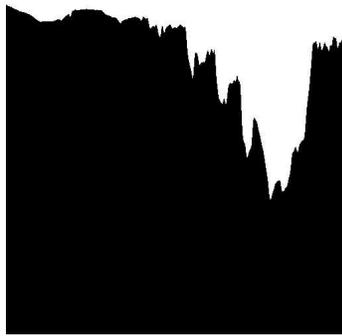
Mojette



(a) projection selon la direction (1,0)



(b) projection Modulo selon la direction (1,0)



(c) projection selon la direction (1,0) de l'image filtrée



(d) projection Modulo selon la direction (1,0) de l'image filtrée

FIG. B.1 – (a) : Projections Mojette de l'image 'bateau' selon la direction (1,0), (b) projection Modulo, (c) projection de l'image filtrée, (d) projection modulo de l'image filtrée

Annexe C

Tatouage



(a) Image 'arbre'



(b) Image 'bateau'



(c) Image 'clown'



(d) Image 'enfant'

FIG. C.1 – Les images de tests utilisées (1/3)



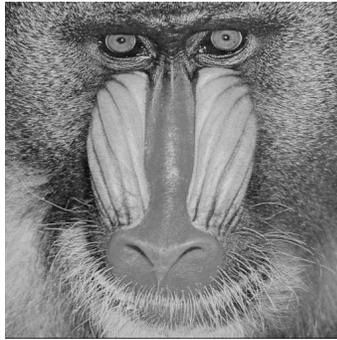
(a) Image 'fruit'



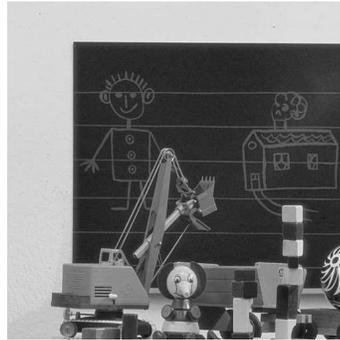
(b) Image 'iba'



(c) Image 'lena'



(d) Image 'singe'



(e) Image 'tableau'

FIG. C.2 – Les images de tests utilisées (2/3)



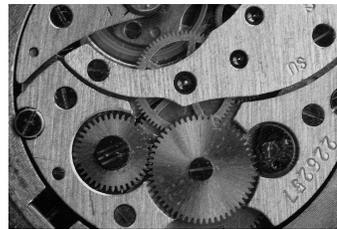
(a) Image 'eglise'



(b) Image 'bureau'



(c) Image 'chambre'



(d) Image 'engrenage'



(e) Image 'water'



(f) Image 'franck'

FIG. C.3 – Les images de tests utilisées (3/3)

Annexe D

Attaques



(a) Suppression de 17 lignes et de 5 colonnes



(b) Suppression de 1 ligne et de 1 colonne



(c) Filtrage médian de noyau 2×2



(d) Filtrage médian de noyau 4×4

FIG. D.1 – Attaques Stirmark (1/5)



(a) Cropping de 10%



(b) Cropping de 25%



(c) Cropping de 50%



(d) Cropping de 75 %



(e) Attaque FMLR

(f) Filtrage Gaussien 3×3

FIG. D.2 – Attaques Stirmark (2/5)



(a) Codage JPEG 10% de qualité



(b) Codage JPEG 20% de qualité



(c) Codage JPEG 30% de qualité



(d) Ratio 80% horizontalement, 100% verticalement



(e) Ratio 120% horizontalement, 100% verticalement



(f) Rotation 10°

FIG. D.3 – Attaque Stirmark (3/5)



(a) Rotation 45°



(b) Rotation 90°



(c) Rotation 0.25° et remise à l'échelle



(d) Rotation 15° et remise à l'échelle



(e) Rotation 30° et remise à l'échelle



(f) Redimensionnement 90%

FIG. D.4 – Attaque Stirmark (4/5)



(a) Redimensionnement 2%



(b) Sharpening 3×3



(c) Shearing 1 horizontalement, 1 verticalement



(d) Shearing 5 horizontalement, 5 verticalement

FIG. D.5 – Attaques Stirmark (5/5)

Bibliographie

- [1] S.J. Anderson, D.C. Burr, and M. Morrone. Two dimensional spatial and spatial frequency selectivity of motion sensitive mechanisms in human vision. *J.O.S.A.*, 8, pages : 1340–1351, 1991.
- [2] J. Brassil and S. Low, N. Maxemchuk, and L. O’Gorman. Electronic marking and identification techniques to discourage document copying. *Proceedings of Infocom’94*, 1994.
- [3] Daniel Augot, Jean-Marc Boucqueau, Jean-François Delaigle, Caroline Fontaine, and Eddy Goray. Secure delivery of images over open networks. *Proceedings of the IEEE*, 87(7), pages : 1251–1267, July 1999.
- [4] Florent Autrusseau. Modélisation psychovisuelle pour le tatouage des images. *Rapport de stage de DEA*, page 56, Septembre 1999.
- [5] Florent Autrusseau and Jeanpierre Guédon. Image watermarking for copyright protection and data hiding via the mojette transform. *SPIE Electronic Imaging*, 4675, 2002.
- [6] Florent Autrusseau, Jeanpierre Guédon, and Yves Bizais. Watermarking and cryptographic schemes for medical imaging. *Accepted for publication in SPIE Medical Imaging, Image processing*, 15-20 February 2003.
- [7] Florent Autrusseau, Abdelhakim Saadane, and Dominique Barba. Psychovisual selection of auspicious sites for watermarking. *SPIE PICS*, pages 326–329, 2000.
- [8] Florent Autrusseau, Abdelhakim Saadane, and Dominique Barba. A psychovisual approach for watermarking. *SPIE Electronic Imaging*, 4314(41), pages : 378–386, 2001.
- [9] Mauro Barni, Franco Bartolini, and Vito Cappellini and Alessandro Piva. A DCT-domain system for robust image watermarking. *Signal Processing*, 66, pages : 357–372, 1998.
- [10] P. Barten. Evaluation of subjective image quality with the square-root integral method. *Journal of Optical Society of America*, 7, pages : 2024–2031, 1990.

- [11] P.G.J. Barten. Contrast sensitivity of the human eye and its effects on image quality. *SPIE, Bellingham*, 1999.
- [12] F. Bartolini, M. Barni, V. Cappellini, and A. Piva. Mask building for perceptual hiding frequency embedded watermarks. *Proceedings of the International Conference on image Processing*, 1, pages : 450–454, 1998.
- [13] F. Bartolini, G. Bini, V. Cappellini, A. Fringuelli, G. Meucci, and A. Piva. Enforcement of copyright laws for multimedia through blind, detectable, reversible watermarking. *Proceedings of IEEE International Conference on Multimedia Computing and Systems*, Juin 1999.
- [14] Laurent Bedat. *Aspects psychovisuels de la perception des couleurs. Application au codage d'images couleurs fixes avec compression de l'information*. PhD thesis, Université de Nantes, 1998.
- [15] Nachida Bekkat. *Critère objectif de qualité subjective d'images monochromes. Conception du modèle et validation expérimentale*. PhD thesis, Université de Nantes, 1999.
- [16] W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for data hiding. *IBM systems journal*, 35(3/4), pages : 313–336, 1996.
- [17] A.P. Bradley. A wavelet visible difference predictor. *IEEE Transactions on Image Processing*, 8(5), pages : 717–730, 1999.
- [18] Gordon W. Braudaway. Protecting publicly-available images with an invisible image watermark. *Proceedings of the International Conference on Image Processing*, 1, pages : 524–527, 1997.
- [19] D.C. Burr and S. A. Wijesundra. Orientation discrimination depends on spatial frequency. *Vis. Research*, 31(7/8), pages : 1449–1452, 1991.
- [20] C. Busch, P. Nesi, M. Schmucker, and M. Spinu. Evolution of music score watermarking algorithm. *Proceedings of SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents IV*, 4675(21), 2002.
- [21] C. Busch and A. Pant. Digital watermarking for the protection of music scores. *Proceedings of SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents IV*, 4314(11), 2001.
- [22] Patrick Le Callet. *Critères objectifs avec référence de qualité visuelle des images couleur*. PhD thesis, Université de Nantes, 2001.
- [23] F.W. Campbell and J.G. Robson. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, pages : 551–566, 1968.

- [24] CCIR. Projet de révision de la recommandation 500-4 : Méthode d'évaluation subjective de la qualité des images de télévision. *Document commission d'études du CCIR*, 11/BL/51-F, 1992.
- [25] Yoonki Choi and Kiyoharu Aizawa. Universal watermark estimation scheme based on error probabilities. *SPIE, Electronic Imaging, Security and Watermarking of multimedia contents*, January 2001.
- [26] S. Comes and B. Macq. Human visual quality criterion. *Proceedings of the SPIE, Visual Communications and Image Processing*, 1360, pages : 2–13, 1990.
- [27] I. J. Cox, J. Kilian, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6(12), pages : 1673–1687, 1997.
- [28] Ingemar J. Cox, Joe Kilian, Tom Leighton, and Talal Shamoon. Secure spread spectrum watermarking for images, audio and video. *Proceedings of the International Conference on Image Processing*, pages 243–246, 1996.
- [29] S. Craver, N. Memon, B. Yeo, and M. Yeung. Can invisible watermarks resolve rightful ownerships? *Technical report RC 20509, IBM Research division*, july 1996.
- [30] Scott Craver. Zero knowledge watermark detection. *Workshop on Information Hiding*, Sept. 29 - Oct. 1, 1999.
- [31] Scott Daly. The visible difference predictor : An algorithm for the assessment of image fidelity. *Proc. of the SPIE, Human Vision, Visual Processing and Digital Display*, III, pages : 2–15, 1992.
- [32] Scott Daly. A visual model for optimizing the design of image processing algorithms. *International Conference on Image Processing, ICIP*, 3, pages : 16–20, 1994.
- [33] J.F. Delaigle, C. De Vleeschouwer, and B. Macq. Watermarking algorithm based on a human visual model. *Signal Processing*, 66, pages : 319–335, 1998.
- [34] O.D. Faugeras. Digital color image processing within the framework of a human visual model. *IEEE Trans. ASSP*, 27(4), pages : 380–393, 1979.
- [35] J. M. Foley. Human luminance pattern mechanisms : Masking experiments require a new model. *Journal of the Optical Soc. of America*, 11(6), pages : 1710–1719, 1994.
- [36] Caroline Fontaine. Contribution à la recherche de fonctions booléennes hautement non linéaires, et au marquage d'images en vue de la protection des droits d'auteur. *Thèse de Doctorat de l'université Paris 6*, November 1998.
- [37] J. Fridrich. Applications of data hiding in digital images. In *Tutorial for the ISSPA '99 Conference*, pages 22–25, Brisbane, Australia, August 1999.

- [38] Jiri Fridrich. Visual hash for oblivious watermarking. *Proc. SPIE, Electronic Imaging, Security and watermarking of multimedia contents*, Jan 24-26 2000.
- [39] Jiri Fridrich, Arnold C. Baldoza, and Richard J. Simard. Robust digital watermarking based on key-dependent basis functions. *Information Hiding, Second International Workshop*, pages 143–157, 1998.
- [40] Jiri Fridrich and Miroslav Goljan. Robust hash functions for digital watermarking. *Proc. of the The Int.l Conf. on Information Technology : Coding and Computing, (ITCC 2000).*, March 2000.
- [41] Jiri Fridrich, Miroslav Goljan, and Rui Du. Invertible authentication. *Proc. SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents*, January 2001.
- [42] Miroslav Goljan, Jiri Fridrich, and Rui Du. Distortion-free data embedding for images. *4th Information Hiding Workshop*, April 2001.
- [43] J.N. Graham. Detection of grating patterns containing two spatial frequencies : A comparison of a single channel and multiple channel models. *Vis. Research*, 11, pages : 251–259, 1971.
- [44] Jeanpierre Guédon. *Un tomographe discret : L'opérateur Mojette - Applications en image et réseaux*. Habilitation à diriger des recherches de l'université de nantes, Université de Nantes, 18 Décembre 2001.
- [45] Jeanpierre Guédon and Dominique Barba. Décomposition psychovisuelle et transmission rapide d'images. *Journées sur la compression et transmission de signaux audiovisuels (CORESA), CNET, Rennes CCETT, (6 pages)*, pages 562–572, Janvier 1995.
- [46] Jeanpierre Guédon, Dominique Barba, and N. Burger. psychovisual image coding via an exact discrete radon transform algorithm. *Proc. SPIE, VCIP'95, Tai Pai*, 2501, pages : 562–572, May 1995.
- [47] Jeanpierre Guédon and Nicolas Normand. Spline mojettettransform. applications in tomography and communications. *IEEE EUSIPCO'02*, 3-6 Septembre 2002.
- [48] Jeanpierre Guédon, Nicolas Normand, Pierre Verbert, Benoit Parrein, and Florent Autrusseau. Load-balancing and scalable multimedia distribution using the mojette transform. *SPIE, Internet Multimedia Management Systems II, ITCOM*, 4519(25), pages : 226–234, August 2001.
- [49] Jeanpierre Guédon, Benoît Parrein, and Nicolas Normand. Internet distributed image information system. *Integrated Computer Aided Engineering*, 89, pages : 205–214, 2001.
- [50] C. F. Hall and E. L. Hall. A nonlinear model for the spatial characteristics of the human visual system. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7, pages : 161–170, 1977.
- [51] Franck Hartung and Bernd Girod. Watermarking of MPEG-2 encoded video without decoding and re-encoding. *Multimedia Computing and Networking SPIE*, 3020, pages : 264–273, 1997.

- [52] Frank Hartung and Bernd Girod. Fast public key watermarking of compressed video. *Proceedings of the International Conference on Image Processing*, pages 528–531, 1997.
- [53] D. J. Heeger. Normalisation of cells responses in cat striates cortex. *Visual Neuroscience*, 9, pages : 181–198, 1992.
- [54] A. Herrigel, H. Petersen, J.J.K. Ó Ruanaidh, T. Pun, and P. Shelby. Copyright techniques for digital images based on asymmetric cryptographic techniques. *Workshop on Information Hiding*, Apr 1998.
- [55] A. Herrigel, J.J.K. Ó Ruanaidh, H. Petersen, S. Pereira, and T. Pun. Secure copyright protection techniques for digital images. *Information Hiding (Lecture Notes in Computer Science)*, 1525, pages : 169–190, 1998.
- [56] R. F. Hess, A. Bradley, and L. Piotrowski. Contrast-coding in amblyopia difference in the neural basis of the human amblyopia. *Proceeding R. Society of London*, B217, pages : 309–330, 1983.
- [57] Chiou-Ting Hsu and Ja-Ling Wu. Hidden signatures in images. *Proceedings of the International Conference on Image Processing*, 3, pages : 223–226, 1996.
- [58] Chiou-Ting Hsu and Ja-Ling Wu. Hidden digital watermarks in images. *IEEE Transactions on image processing*, 8(1), 1999.
- [59] Hisashi Inoue, Akio Miyazaki, Akihiro Yamamoto, and Takashi Katsura. A digital watermark based on the wavelet transform and its robustness on image compression. *Proceedings of the International Conference on Image Processing*, 2, pages : 391–395, October 1998.
- [60] Mohan S. Kankanhalli, Rajmohan, and K. R. Ramakrishnan. Content based watermarking of images. *ACM Multimedia - Electronic Proceedings*, 1998.
- [61] S. Karunasekera and N. Kingsbury. A distortion measure for coding artifacts in images : Implementation aspects. *International report, signal Processing group, Department of Engineering*, 1995.
- [62] S.A. Karunasekera and N.G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensivity. *Proceedings of SPIE, Visual Communications and Image Processing*, 2094, pages : 474–486, 1993.
- [63] S.A. Karunasekera and N.G. Kingsbury. A distortion measure for image artifacts based on human visual sensivity. *Proceedings of ICASSP-IEEE*, 5, pages : 117–120, 1994.
- [64] M. Katz. Questions of uniqueness and resolution in reconstruction from projections. *Lecture Notes in Biomathematics, Springer Verlag*, 1997.

- [65] D. Kirovski and H. S. Malvar. Robust covert communication over a public audio channel using spread spectrum. *Information Hiding Workshop*, April 2001.
- [66] Deepa Kundur and Dimitrios Hatzinakos. A robust digital image watermarking method using wavelet-based fusion. *Proceedings of the International Conference on Image Processing*, 1, pages : 544–547, 1997.
- [67] Deepa Kundur and Dimitrios Hatzinakos. Digital watermarking using multiresolution wavelet decomposition. *Proceedings of the ICASSP*, 5, pages : 2659–2662, May 1998.
- [68] Martin Kutter and Fabien A. P. Petitcolas. A fair benchmark for image watermarking systems. *Electronic Imaging'99, Security and Watermarking of Multimedia Contents*, 3657, January 1999.
- [69] G. E. Legge and J. M. Foley. Contrast masking in human vision. *Journal of the Optical Soc. of America*, 70, pages : 1458–1471, 1980.
- [70] Andrei Leontiev. Dicom 2001 : the old, the new and the future, February 2002. <http://medical.nema.org/>.
- [71] J.O. Limb. Distorsion criteria of the human viewer. *IEEE Transaction on system, Man and cybernetics*, SMC-9(12), pages : 778–793, 1979.
- [72] J. Lubin. The use of psychophysical data and models in the analysis of display system performance. *Digital Images and Human Vision*, pages 162–178, 1993.
- [73] F.X.J. Lukas and Z.L. Budrikis. Picture quality prediction based on a visual model. *IEEE Trans. Comm*, 30(7), pages : 1679–1692, 1982.
- [74] J. L. Mannos and D. J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions on information theory*, 20(4), pages : 535–536, 1974.
- [75] Anne Manoury. *Tatouage d'images numériques par paquets d'ondelettes*. PhD thesis, Université de Nantes, Novembre 2001.
- [76] André Marion. Acquisition et visualisation des images. *Eyrolles*, page 615, 1997.
- [77] Peter Meerwald and Shelby Pereira. Attacks, applications and evaluation of known watermarking algorithms with Checkmark. In *Proceedings of SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents IV*, volume 4675, San Jose, CA, USA, January 2002. SPIE.
- [78] Bijan G. Mobasserri. Direct sequence watermarking of digital video using m-frames. *Proceedings of the International Conference on Image Processing*, 2, pages : 399–403, October 1998.

- [79] P. Moon and D. E. Spencer. Visual data applied to lighting design. *JOSA*, 34, pages : 605–617, Oct 1944.
- [80] Marcus J. Nadenau, Stefan Winkler, David Alleysson, and Murat Kunt. Human vision models for perceptually optimized image processing - a review. *submitted to Proceedings of the IEEE*, Septembre 2000.
- [81] P. Nguyen and S. Baudry. Tatouage temps réel de flux vidéo CCIR-656 et MPEG-2. *CORESA*, June 1999.
- [82] Nicolas Normand. *Représentation d'images et distances discrètes basées sur les éléments structurants à deux pixels*. PhD thesis, Université de Nantes, Janvier 1997.
- [83] Nicolas Normand and Jeanpierre Guédon. La transformation mojette : une représentation redondante pour l'image. *Comptes-rendus de l'Académie des sciences de Paris, Informatique Théorique*, pages 123–126, 1998.
- [84] Joseph J. K. O'Ruanaidh and Thierry Pun. Rotation, scale and translation invariant digital image watermarking. *Proc. IEEE Int. Conf. Image Processing*, 1, pages : 536–539, Oct 1997.
- [85] Benoît Parrein, Nicolas Normand, and Jeanpierre Guédon. Multiple description coding using exact discrete radon transform. *IEEE Data Compression Conference*, Mars 2001.
- [86] Benoît Parrein, Pierre Verbert, Nicolas Normand, and Jeanpierre Guédon. Scalable multiple description on packets network via the n-dimensional mojette transform. *SPIE, Quality of Service over next-Generation Data Networks*, 4524(36), pages : 243–252, August 2001.
- [87] M. Pavel, G. Sperling, T. Riedl, and A. Vanderbeek. Limits of visual communication : the effect of signal-to-noise ratio on the intelligibility of american sign language. *J.O.S.A.*, A4, pages : 2355–2365, 1987.
- [88] E. Peli. Contrast in complex images. *J.O.S.A.*, 7(10), pages : 2032–2040, 1990.
- [89] Shelby Pereira, Joseph J. K. Ó Ruanaidh, Frédéric Deguillaume, Gabriela Csurka, and Thierry Pun. Template based recovery of fourier-based watermarks using log-polar and log-log maps. *IEEE Int. Conf. on Multimedia Computing and Systems (ICMCS'99)*, pages 870–874, June 1999.
- [90] H. A. Peterson, A. J. Ahumada Jr, and A. B. Watson. Improved detection model for dct coefficient quantization. *Proc. of the SPIE Conference on Human Vision, Visual Processing and Digital Display IV*, 1913, pages : 191–201, February 1993.
- [91] Fabien A. P. Petitcolas. Attaques et évaluation des filigranes numériques. *CORESA*, June 1999.

- [92] Fabien A.P. Petitcolas, Ross J. Anderson, and Markus G. Kuhn. Attacks on copyright marking systems. *Second Workshop on information hiding*, 1525, pages : 218–238, 1998.
- [93] Olivier Philippé. *Représentation d'image pour le codage conjoint source-canal pour les réseaux à qualité de service*. PhD thesis, Université de Nantes, 1998.
- [94] G.C. Philips and H. Wilson. Orientation bandwidths of spatial mechanisms measured by masking. *J.O.S.A.*, 1(2), pages : 226–232, 1984.
- [95] R. Pickholtz, D. Schilling, and L. Milstein. Theory of spread-spectrum communications - a tutorial. *IEEE Transactions on Communications*, COM-30, pages : 855–884, 1982.
- [96] I. Pitas. A method for signature casting on digital images. *Proceedings of ICIP, IEEE press*, vol 3, pages : 215–218, 1996.
- [97] A. Piva, M. Barni, F. Bartolini, and V. Cappellini. DCT-based watermark recovering without resorting to the uncorrupted original image. *Proceedings of the International Conference on Image Processing*, 1, pages : 520–523, 1997.
- [98] Christine I. Podilchuk and Wenjun Zeng. Digital image watermarking using visual models. *IEEE Journal on selected Areas in Communications*, 16(4), pages : 525–539, 1998.
- [99] Christine I. Podilchuk and Wenjun Zeng. Image-adaptive watermarking using visual models. *IEEE Journal on Selected Areas in Communications*, 16(4), pages : 525–539, May 1998.
- [100] William Puech and Michel Dumas. Transfert sécurisé d'images par combinaison de techniques de cryptographie et de tatouage. *Proc. Colloque CORESA'2001*, Novembre 2001.
- [101] J. Radon. Über die bestimmung von functionen durch ihre integralwerte langs gewisser mannigfaltigkeiten. *Berichte Sachsische Academie der Wissenschaften*, 69, pages : 262–267, 1917.
- [102] J.J.K Ó Ruanaidh, W.J. Dowling, and F.M. Boland. Phase watermarking of digital images. *Proc. IEEE International Conference on Image Processing*, III, pages : 239–242, 1996.
- [103] Joseph O Ruanaidh, Holger Petersen, Alexander Herrigel, Shelby Pereira, and Thierry Pun. Cryptographic copyright protection for digital images based on watermarking techniques. *Theoretical Computer Science*, 226, pages : 117–142, Sept 1999.
- [104] G.S. Rubin and K. Siegel. Recognition of low-pass filtered faces and letters. *Invest Ophthalmol. Vis. Sci.*, 25, 1984.

- [105] A. Saadane and F. Atrousseau. Analyse et spécifications structurelles du modèle de vision et de qualité image. Technical report, Rapport du Contrat RNRT AQUAMARS, Septembre 1999, Rapport D 2.1., SEI-IRESTE, Septembre 1999.
- [106] A. Saadane and F. Atrousseau. Modélisation psychovisuelle pour le tatouage des images. Technical report, Rapport du Contrat RNRT AQUAMARS, Septembre 1999, Rapport D 2.2.1., SEI-IRESTE, Septembre 1999.
- [107] M.B. Sachs, J. Nachmias, and J.G. Robson. Spatial frequency channels in human vision. *Journal of the optical society of America*, 61(9), pages : 1176–1186, 1971.
- [108] Robert Safranek. A comparison of the coding efficiency of perceptual models. *Proceedings of the IEEE Conf. on Human Vision, Visual Processing and Digital Display VI*, 1995.
- [109] B. Sakmann and O. D. Creutzfeldt. Scotopic and mesopic light adaptation in the cat's retina. *Pflügers archiv*, 313, pages : 168–185, 1969.
- [110] D.J. Sakrison. On the role of the observer and a distortion measure in image transmission. *IEEE Trans. on Com.*, 25(11), pages : 1251–1267, 1977.
- [111] O.H. Schade. Optical and photoelectric analog of the eye. *J. Opt. Soc. Am.*, 46(9), pages : 721–739, 1956.
- [112] B. Schneier. Applied cryptography, second edition. *John Wiley and Sons Ed.*, 1996.
- [113] R.G. Van Schyndel, A.Z. Tirkel, and C.F. Osborne. A digital watermark. *Proceedings of the International Conference on Image Processing*, 2, pages : 86–90, 1994.
- [114] J. M. Shapiro. Embedded image coding using zerotrees of wavelets coefficients. *IEEE Transactions on Signal Processing*, 41, pages : 3445–3462, 1993.
- [115] Hakim Sénane. *Représentation d'images en sous-bandes visuelles. Application au codage d'images de télévision sans défauts visibles*. PhD thesis, Université de Nantes, 1996.
- [116] Mitchell D. Swanson, Bin Zhu, and Ahmed H. Tewfik. Transparent robust image watermarking. *Proceedings of the International Conference on Image Processing*, 3, pages : 211–214, October 1996.
- [117] P.C. Teo and D. J. Heeger. Perceptual images distortion. *Proc. of SPIE*, 2179, pages : 127–141, 1994.
- [118] R. Valois, K, and Valois. Spatial vision. *Oxford Univ. Press.*, 1988.
- [119] Pierre Verbert and Jeanpierre Guédon. An exact discrete backprojection operator. *IEEE EUSIPCO'02*, 3-6 Septembre 2002.

- [120] Pierre Verbert and Jeanpierre Guédon. N-dimensional mojette transform. application to multiple description. *IEEE Int Conf on Digital Signal Processing (DSP)*, II, pages : 1211–1214, 1-3 Juillet 2002.
- [121] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor. Visual thresholds for wavelet quantization error. *SPIE Proceedings, Human Vision and Electronic Imaging*, 1996.
- [122] Andrew B. Watson. The cortex transform : Rapid computation of simulated neural images. *Computer Vision, Graphics, And Image Processing*, 39, pages : 311–327, 1987.
- [123] Andrew B. Watson. Efficiency of a model human image code. *J. O. S. A.*, 4(12), pages : 2401–2417, 1987.
- [124] Andrew B. Watson. Dct quantization matrices visually optimized for individual images. *Proc. of the SPIE Conference on Human Vision, Visual Processing and Digital Display IV*, 1913, pages : 202–216, February 1993.
- [125] S.J.P. Western, R.L. Lagendijk, and J. Biemond. Perceptual image quality based on a multiple channel HVS model. *Proceedings of ICASSP*, 4, pages : 2351–2354, 1995.
- [126] Raymond B. Wolfgang, Christine I. Podilchuk, and Edward J. Delp. Perceptual watermarks for digital images and video. *Proceedings of the IEEE*, 87(7), pages : 1108–1127, July 1999.
- [127] Min Wu and Bede Liu. Watermarking for image authentication. *Proceedings of the International Conference on Image Processing*, 2, pages : 437–441, October 1998.
- [128] Byron Yu. Estimating the watermarking capacity of images. Technical report, Documentation of Summer Internship, Ecole Nationale Supérieure des Télécommunications, Paris, 2000.
- [129] C. Zetsche and G. Hauske. Multiple channel model for the prediction of subjective image quality. *Proceedings of the SPIE, Human Vision, Visual Processing and Digital Display*, pages 209–216, 1989.

Production scientifique

PUBLICATIONS EN REVUES :

- Abdelhakim Saadane et Florent Atrousseau, “Watermarking perceptuel et adaptatif d’images fixes”, Accepté pour publication dans la revue *Traitement du Signal*, Numéro spécial *Watermarking*, 2002 (32 pages).

PUBLICATIONS EN CONGRES INTERNATIONAUX AVEC COMITE DE LECTURE ET ACTES :

- Florent Atrousseau, Abdelhakim Saadane et Dominique Barba, “Psychovisual selection of auspicious sites for watermarking”, *SPIE, PICS, Portland, OR*, 2000, pages : 326-329.
- Florent Atrousseau, Abdelhakim Saadane et Dominique Barba, “A Psychovisual approach for watermarking”, *SPIE Electronic Imaging, Security and Watermarking of Multimedia Contents III, San Jose, CA*, vol. 4314, pages : 495-504, January 2001.
- Florent Atrousseau, Jeanpierre Guédon, “Image watermarking for copyright protection and data hiding via the Mojette transform”, *Security and Watermarking of Multimedia Contents IV, San Jose, CA, SPIE Electronic Imaging*, vol. 4675, pages : 378-386, January 2002.
- Florent Atrousseau, Jeanpierre Guédon, “Image Watermarking in the Fourier Domain Using the Mojette Transform”, *Santorini Greece, 14th IEEE International Conference on Digital Signal Processing (DSP2002)*, vol. II, pages : 725-728, July 2002.
- Jeanpierre Guédon, Nicolas Normand, Pierre Verbert, Benoit Parrein et Florent Atrousseau, “Load-balancing and scalable multimedia distribution using the Mojette transform”, *Internet Multimedia Management Systems II, ITCOM, Denver*, vol. 4519, pages : 226-234, August 2001.
- Florent Atrousseau, Jeanpierre Guédon, Yves Bizais, "Watermarking and cryptographic schemes for medical imaging", Accepté pour publication dans *SPIE Medical Imaging, Image processing*, 15-20 February

2003, San Diego California, USA.

RAPPORTS DE CONTRATS :

- A. Saadane et F. Autrusseau, “Aquamarquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en toute sécurité sur les réseaux hétérogènes de télécommunication. Analyse et spécifications”, Contrat RNRT AQUAMARS, Septembre 1999, Rapport D 2.1 (47 pages).
- A. Saadane et F. Autrusseau, “Aquamarquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en toute sécurité sur les réseaux hétérogènes de télécommunication. Modélisation psycho-visuelle pour le tatouage des images”, Contrat RNRT AQUAMARS, Octobre 1999, Rapport D 2.2.1 (23 pages).