

MA-004-13-1

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saâd Dahlab de Blida



**Faculté des sciences, Département d'informatique**

**En vue d'obtention du diplôme de Master**

**En Sécurité informatique.**

**Spécialité : Informatique/ Option : Génie des logiciels.**

*Sécurisation des documents imprimés par tatouage  
Numérique*

**Présenté par :** Melle : MAHAMED I Imene

Melle : HEMINNA Amel

**Promotrice :** Mme K. Hamoudi Ait Saadi

**Co-promoteur :** Mr S. FERFERA

**Organisme d'accueil :** Centre de Développement des Technologies Avancées (CDTA),

Division : Architecteur Des Système Multimédia



**Soutenu le :** 13.07.2010, devant le jury composé de :

Nom. Président du jury, *B. Babouche*

**Président**

Nom examinateur 1, *H. Hamoudi*

**Rapporteurs**

Nom examinateur 2, *E. Elghes*

**Examineurs**

MA-004-13-1

## Résumé

L'objectif de ce mémoire est de protéger les droits de propriété des droits d'auteurs des documents imprimés par tatouage numérique, en insérant d'une manière imperceptible et robuste l'information du propriétaire. Pour y parvenir, deux solutions d'insertion sont proposées et implémentées, l'insertion peut être effectuée directement dans le domaine spatial du document imprimé, ou dans le domaine fréquentiel après avoir appliqué une Transformation en Cosinus Discrète (DCT) sur les données numériques du document.

Le système a été implémenté en utilisant le langage visuel C#.

**Mots clés :** Protection du contenu, tatouage numérique, droits d'auteur.

## Abstract

The aim of this is to protect the copyright of the printed documents by digital watermarking. This is done by inserting imperceptible and robust watermark in the document. To achieve this, two solutions are proposed and implemented. The insertion can be performed directly in spatial domain of the printed document or in the frequency domain after applying a Discrete Cosine Transform (DCT) on digital data of the document.

The system was implemented using Visual C # Language.

**Keywords:** Content protection, digital watermarking, copyright.

# Remerciement

*Nous remercions ALLAH de nous avoir donné le courage, la patience, la santé et la motivation d'entamer et de finir ce mémoire de fins d'études dans les meilleures conditions.*

*Nous tenons à remercier très chaleureusement les personnes qui ont contribué de près ou de loin à la réalisation de ce travail et particulièrement « nos parents ».*

*Notre promotrice Mm. Ait saadi Karima de nous avoir proposé ce sujet, et pour son soutien et ses conseils tout au long du projet. Ce projet qui a développé en nous une capacité de recherche et d'adaptation.*

*Un grand merci est adressé au notre promoteur au niveau d'université Saad Dahleb Mr. FERFERA Soufiane pour son soutien et conseils et pour sa disponibilité tout au long de l'année.*

*Nos sentiments de profonde gratitude vont à nos professeur qui tout au long de notre cursus nous ont transmis leur savoir sans réserve.*

*Tout notre respect et nos remerciements les plus sincères vont vers les membres de jury qui vont pleinement consacrer leur temps et leur attention afin d'évaluer notre travail, qui espérons le sera à la hauteur de leur attente.*

*Que toutes et tous ceux qui ont fait que ce travail aboutisse trouvent ici l'expression de nos remerciements les plus sincères.*

*Heminna Amel et Mahamedi Imene.*

# Dédicace

*Ce n'est qu'à grâce à l'aide d'ALLAH que j'ai finalisé  
cet humble travail que je dédie à toutes les personnes qui  
me sont chères et à qui je tiens vraiment en  
reconnaissance de tout le soutiens qu'elles m'ont apporté  
durant les moments difficiles*

*A mes très chers PARENTS que j'aime  
énormément, et qui se sont tant sacrifiés pour moi, pour  
m'offrir le bonheur.*

*A mes frères ABDELHAK, ABDELHAMID et  
OMAR EL FAROUK et à ma petite sœur ASMA.*

*A toute ma famille et particulièrement ma grande  
mère « jedà »*

*A mon binôme AMEL, et à toute sa famille.*

*A tout mes copains et copines sans exception.*

*A tout mes collègues et amis de la promotion un par  
un.*

*A tout ceux qui m'aiment et qui me connaissent, à  
tous ceux qui je compte.*

# Dédicace

*Ce n'est qu'à grâce à l'aide d'ALLAH que j'ai finalisé  
cet humble travail que je dédie à toutes les personnes qui  
me sont chères et à qui je tiens vraiment en  
reconnaissance de tout le soutiens qu'elles m'ont apporté  
durant les moments difficiles*

*A mes très chers PARENTS que j'aime  
énormément, et qui se sont tant sacrifiés pour moi, pour  
m'offrir le bonheur.*

*A mes sœurs YASMINA et AMINA et à mon  
unique frère OUSSAMA.*

*A toute ma famille.*

*A mon binôme IMENE et à toute sa famille.*

*A tout mes copains et copines sans exception.*

*A tout mes collègues et amis de la promotion un par  
un.*

*A tout ceux qui m'aiment et qui me connaissent, à  
tous ceux qui je compte.*

# Sommaire

<b>Introduction générale</b> .....	1
<b>CHAPITRE I : Généralités et Définitions</b>	
I. Introduction.....	4
II. Historique.....	4
III. Notions de base du tatouage.....	5
III.1. Définition.....	5
III.2. Application de tatouage numérique.....	6
III.3. Caractéristiques du tatouage numérique.....	8
III.4. Classification des systèmes du tatouage numérique.....	10
III.4.1. Classification selon la manière d'insertion.....	10
III.4.2. Classification selon le domaine d'insertion.....	11
III.4.3. Classification selon le type de la marque insérée.....	12
III.4.4. Classification selon le type du tatouage.....	12
III.5. Les attaques.....	13
III.6. La notion d'embrouillage.....	13
III.7. Quelques techniques existant de tatouage numérique des documents imprimés.....	13
IV. Conclusion.....	16
<b>CHAPITRE II : Système de tatouage proposé</b>	
I. Introduction.....	17
II. Conception du système de tatouage proposé.....	17
II.1. Schéma générale du système de tatouage proposé.....	17
II.2. Diagramme de cas d'utilisation.....	18
II.2.1. Cas d'utilisation général du système.....	18
II.2.2. Cas d'utilisation « Processus d'insertion ».....	19
II.2.3. Cas d'utilisation « Insertion dans le domaine spatial ».....	20
II.2.4. Cas d'utilisation « Insertion dans le domaine fréquentiel ».....	21
II.2.5. Cas d'utilisation « Extraction ».....	22
II.2.6. Cas d'utilisation « Extraction dans le domaine spatial ».....	23
II.2.7. Cas d'utilisation « Extraction dans le domaine fréquentiel ».....	24
II.3. Diagramme de séquence.....	25

II.3.1. Diagramme de séquence du cas d'utilisation « Insertion dans le domaine spatial ».....	25
II.3.2. Diagramme de séquence du cas d'utilisation « extraction dans le domaine spatial ».....	27
III.Fonctionnement du système selon le type du document à tatoué.....	28
IV.Description du processus d'insertion et d'extraction.....	29
IV.1.Processus d'insertion dans le domaine spatial.....	29
IV.1.1.La binarisation.....	30
IV.1.1.1.La méthode <i>Ordred</i> .....	31
IV.1.1.2.La méthode <i>Floyd-Steinberg</i> .....	33
IV.1.2.Duplication de la marque.....	34
IV.1.3.Embrouillage de la marque.....	35
IV.1.4.Classification de blocs.....	36
IV.1.5.Insertion de la marque.....	39
IV.2.Processus d'extraction dans le domaine spatial.....	40
V.2.1. Classification des blocs du document tatoué et reconstruction de la clé publique.....	41
V.2.2. Extraction de la marque insérée.....	41
V.2.3. Amélioration de la qualité de la marque.....	42
V.2.4. Analyse expérimentale des résultats.....	44
V. Insertion dans le domaine fréquentiel.....	46
V.1. Processus d'insertion.....	47
V.1.1.La segmentation du document en blocs de taille 8x8.....	48
V.1.2.Application de la DCT 2D.....	48
V.1.3.L'insertion de la marque.....	49
V.2. Processus d'extraction de la marque.....	51
V.3. Analyses expérimentales des résultats .....	51
V.4. Processus d'insertion pour la DCT 2D 4x4.....	55
V.5. Analyses expérimentales des résultats pour la DCT 2D 4x4.....	55
VI.Conclusion.....	57
<b>CHAPITRE III. Tests et attaques</b>	
I. Introduction.....	58
II. Robustesse de la méthode développée dans le domaine spatial.....	60

III. Robustesse de la méthode développée dans le domaine fréquentiel.....	62
III.1. Méthode basée sur la DCT 8x8.....	62
III.2. Méthode basée sur la DCT 4x4.....	63
IV. Conclusion .....	66

#### **CHAPITRE IV. Implémentation**

I. Introduction .....	67
II. Environnement d'application .....	67
III. Interface de l'application .....	67
IV. Conclusion.....	75
<b>Conclusion générale.....</b>	<b>76</b>
Bibliographie .....	78



## Liste des figures

### CHAPITRE I. Généralités et Définitions

<b>Figure I.1.</b> Schéma générale d'un système de tatouage numérique.....	6
<b>Figure I.2.</b> Illustration graphique du compromis entre les caractéristiques du tatouage numérique.....	10

### CHAPITRE II. Système de tatouage proposé

<b>Figure II.1.</b> Synoptique du système de tatouage proposé des documents imprimés.....	18
<b>Figure II.2.</b> Diagramme de cas d'utilisation général.....	19
<b>Figure II.3.</b> Diagramme de cas d'utilisation « processus d'insertion ».....	20
<b>Figure II.4.</b> Diagramme de cas d'utilisation « insertion dans le domaine spatial ».....	21
<b>Figure II.5.</b> Diagramme de cas d'utilisation « insertion dans le domaine fréquentiel ».....	22
<b>Figure II.6.</b> Diagramme de cas d'utilisation « processus d'extraction ».....	23
<b>Figure II.7.</b> Diagramme de cas d'utilisation « extraction dans le domaine spatial ».....	24
<b>Figure II.8.</b> Diagramme de cas d'utilisation « extraction dans le domaine fréquentiel ».....	25
<b>Figure II.9.</b> Diagramme de séquence du cas d'utilisation « insertion dans le domaine spatial ».....	26
<b>Figure II.10.</b> Diagramme de séquence du cas d'utilisation « extraction dans le domaine spatial ».....	27
<b>Figure II.11.</b> Fonctionnement du système selon le type du document à tatoué.....	28
<b>Figure II.12.</b> Illustration de la transformation de RVB vers le plan YCrCb et la transformation inverse.....	29
<b>Figure II.13.</b> Processus d'insertion dans le domaine spatial.....	30
<b>Figure II.14.</b> Résultat de la duplication de la marque.....	35
<b>Figure II.15.</b> Résultat de l'embrouillage de la marque dupliquée.....	36
<b>Figure II.16.</b> Organigramme de la classification et la génération des clés.....	38
<b>Figure II.17.</b> Résultat de la génération des clés publique et secrète.....	38
<b>Figure II.18.</b> Organigramme d'insertion de la clé publique.....	39
<b>Figure II.19.</b> Résultat du processus d'insertion.....	39
<b>Figure II.20.</b> Processus d'extraction dans le domaine spatial.....	40

<b>Figure II.21.</b> Organigramme des différentes étapes pour la reconstruction de la clé publique.....	41
<b>Figure II.22.</b> Organigramme d'extraction de la marque.....	42
<b>Figure II.23.</b> Résultat des opérations de correction et réduction.....	43
<b>Figure II.24.</b> Synoptique générale du tatouage numérique dans le domaine fréquentiel....	47
<b>Figure II.25.</b> Processus d'insertion dans le domaine fréquentiel.....	48
<b>Figure II.26.</b> Implémentation de la DCT 8x8.....	49
<b>Figure II.27.</b> Organigramme de l'opération d'insertion dans le domaine fréquentiel.....	50
<b>Figure II.28.</b> Opération d'extraction dans le domaine fréquentiel.....	51
<b>Figure II.29.</b> Courbe de variation du PSNR en fonction de $\alpha$ pour différentes k (DCT 2D).....	54
<b>Figure II.30.</b> Variation du PSNR en fonction de $\alpha$ (DCT 4x4).....	56

### CHAPITRE III. Attaques

<b>Figure III.1.</b> Histogramme de variation du PSNR et du taux de corrélation après l'attaque d'impression pour les différentes positions k (DCT 8x8).....	63
--	----

### CHAPITRE IV. Implémentation

<b>Figure IV.1.</b> Schéma générale de l'application.....	68
<b>Figure IV.2.</b> Menu principale.....	68
<b>Figure IV.3.</b> Calcul du taux de corrélation de l'extraction.....	68
<b>Figure IV.4.</b> Chargement du document en choisissant son type et la tâche à effectuée....	69
<b>Figure IV.5.</b> Le cas du chargement du document en niveaux de gris.....	70
<b>Figure IV.6.</b> Le cas où le document chargé est en couleurs.....	70
<b>Figure IV.7.</b> Insertion dans le domaine spatial (méthode de binarisation Ordred).....	71
<b>Figure IV.8.</b> Résultats d'insertion dans le domaine spatial (méthode de binarisation Ordred).....	71
<b>Figure IV.9.</b> La différence entre le document d'origine et le document tatoué.....	72
<b>Figure IV.10.</b> Insertion dans le domaine fréquentiel (transformée DCT 4x4).....	72
<b>Figure IV.11.</b> Résultats d'insertion dans le domaine fréquentiel (transformée DCT 4x4).	73
<b>Figure IV.12.</b> Extraction dans le domaine spatial (méthode de binarisation Ordred).....	73
<b>Figure IV.13.</b> Extraction dans le domaine fréquentiel (transformée DCT 4x4).....	74
<b>Figure IV.14.</b> L'attaque de recadrage.....	74

## Liste des tables

### CHAPITRE II. Système de tatouage proposé

<b>Table II.1.</b> Génération de la matrice de binarisation $B_8$ .....	32
<b>Table II.2.</b> Filtre d'erreur de Floyd et Steinberg.....	33
<b>Table II.3.</b> Les quatre paramètres utilisés pour le calcul du GCL.....	35
<b>Table II.4.</b> Table de classification de blocs.....	37
<b>Table II.5.</b> Table de référence pour les opérations de la correction et la rédaction.....	43
<b>Table II.6.</b> Résultats d'insertion et d'extraction dans le domaine spatial (Barbara 256x256).....	45
<b>Table II.7.</b> Résultats d'insertion et d'extraction dans le domaine spatial (Femme 256x256).....	46
<b>Table II.8.</b> Variation du PSNR selon $\alpha$ pour la position $k=1$ (DCT 2D 8x8).....	52
<b>Table II.9.</b> Variation du PSNR selon $\alpha$ pour la position $k=2$ (DCT 2D 8x8).....	52
<b>Table II.10.</b> Variation du PSNR selon $\alpha$ pour la position $k=3$ (DCT 2D 8x8).....	52
<b>Table II.11.</b> Variation du PSNR selon $\alpha$ pour la position $k=4$ (DCT 2D 8x8).....	53
<b>Table II.12.</b> Variation du PSNR selon $\alpha$ pour la position $k=5$ (DCT 2D 8x8).....	53
<b>Table II.13.</b> Variation du PSNR selon $\alpha$ pour la position $k=6$ (DCT 2D 8x8).....	53
<b>Table II.14.</b> Variation du PSNR selon $\alpha$ pour la position $k=7$ (DCT 2D 8x8).....	53
<b>Table II.15.</b> Les documents tatoués ainsi les différences pour $k$ varie de 4 à 7 (DCT 2D 8x8).....	54
<b>Table II.16.</b> Variation du PSNR selon de $\alpha$ pour la position $k=3$ (DCT 2D 4x4).....	56
<b>Table II.17.</b> Variation du PSNR selon de $\alpha$ pour la position $k=4$ (DCT 2D 4x4).....	56
<b>Table II.18.</b> Variation du PSNR selon de $\alpha$ pour la position $k=5$ (DCT 2D 4x4).....	56
<b>Table II.19.</b> Les documents tatoués ainsi les différences pour $k=3$ (DCT 2D 4x4).....	57

### CHAPITRE III. Attaques

<b>Table III.1.</b> Résultats d'extraction après attaques dans le domaine spatial	60
<b>Table III.2.</b> Résultats d'extraction après l'attaque d'impression-scan dans le domaine fréquentiel pour $k$ varie de 4 à 7 (DCT 2D 8x8).....	63
<b>Table III.3.</b> Résultats d'extraction après attaques dans le domaine fréquentiel (DCT 2D 8x8, $k=5$ ).....	64
<b>Table III.4.</b> Résultats d'extraction après attaques dans le domaine fréquentiel (DCT 4x4, $k=3$ ).....	65

## Liste des équations

### CHAPITRE II. Système de tatouage proposé

II.1. Calcule de a composante Y de l'espace YCrCb à partir de l'espace RVB.....	28
II.2. Calcule de a composante Cr de l'espace YCrCb à partir de l'espace RVB.....	28
II.3. Calcule de a composante Cb de l'espace YCrCb à partir de l'espace RVB.....	28
II.4. Calcule de a composante R de l'espace RVB à partir de l'espace YCrCb.....	29
II.5. Calcule de a composante V de l'espace RVB à partir de l'espace YCrCb.....	29
II.6. Calcule de a composante B de l'espace RVB à partir de l'espace YCrCb.....	29
II.7. Equation générale d'une méthode de binarisation (méthode Ordred).....	31
II.8. Matrice de primitive B_2 de binarisation.....	31
II.9. Matrice de primitive B_3 de binarisation.....	32
II.10. Génération d'une cellule en demi-teinte.....	33
II.11. Formule de génération de la séquence aléatoire pour le procédé d'embrouillage.....	35
II.12. Formule de calcul du MSE (l'erreur quadratique moyenne).....	44
II.13. Formule de calcul de PSNR (le rapport signal sur bruit).....	44
II.14. Formule de calcul du taux de corrélation de l'extraction.....	44
II.15. Formule générale de calcule de la DCT 8x8.....	49
II.16. Méthode de calcule simple pour la DCT 8x8.....	49
II.17. Formule générale de calcul de DCT inverse (IDCT 8x8).....	50
II.18. Méthode de calcule simple pour la DCT 4x4.....	55
II.19. Les coefficients de la DCT 4x4.....	55
II.20. Formule générale de calcul de DCT inverse (IDCT 4x4).....	55

<b>CHAPITRE III. Système de tatouage proposé</b>	
III.1. Matrice de coefficients pour le filtrage numérique.....	59
III.2. Formule pour calculer le filtre gaussien.....	59

## Introduction générale

L'analyse et la reconnaissance de documents ont pour but de convertir un document sous format papier en un format électronique compréhensible et réutilisable. Le document papier, une fois converti sous forme électronique, permet une recherche par le contenu, un transfert très rapide, un archivage et une gestion beaucoup plus aisée. De plus, une fois imprimé, le document électronique offre tous les avantages du format papier, donc une plus grande rapidité et efficacité de consultation et une lecture plus confortable.

Mais cette technologie numérique pose un certain nombre de problèmes concernant la copie numérique. Les moyens informatiques matériels, tels que les photocopieurs couleur, les scanners et les imprimantes, de même que les logiciels de retouches d'images de plus en plus sophistiqués, disponibles à très peu de frais dans la grande distribution, rendent la copie un « clone » du document d'origine. Ceci permet aux fraudeurs de tous bords de produire des falsifications de documents de plus en plus sophistiquées.

C'est la raison pour laquelle il est devenu indispensable de sécuriser avec des technologies d'avant-garde tous types de media : les documents numériques et papiers particulièrement sensibles tels que les Billets de banque, les chèques postaux, les vignettes d'affranchissement et les documents officiels tels que Les cartes ID, les permis de conduire, etc .... Les mesures techniques destinées à empêcher la copie numérique et sa diffusion sont le tatouage numérique, l'embrouillage et le cryptage.

La facilité avec laquelle les documents peuvent être copiés a conduit beaucoup de chercheurs à s'interroger sur de nouvelles méthodes pour protéger la propriété intellectuelle dans le monde numérique.

Comment protéger efficacement les documents de valeur ou confidentiels?

Plusieurs aspects doivent être considérés pour protéger un document :

- une sécurisation visible ou invisible;
- une solution résistante à tous types d'attaques tels que les attaques usuelles de traitements numériques d'images et particulièrement, le système de sécurisation doit être robuste aux attaques d'impression et ré-numérisation pour les documents papiers ;

- une résistance à l'élimination ou à l'altération de l'élément de sécurité;
- la distinction entre un original et une copie;
- une détection sur un fragment de document;
- etc.

Ces quelques considérations montrent qu'il n'y a pas de solution universelle et unique pour la protection de documents multimédia. Vu les progrès constants des équipements potentiellement aux mains des fraudeurs, il est nécessaire que les fournisseurs de solutions continuent à développer des parades. Une combinaison de diverses techniques est souvent une solution qui permet de répondre à des demandes de protection variées pour un même document.

Les tatouages numériques – par analogie aux filigranes apparaissant sur les billets de banque – ont été proposés comme un outil miracle pour contrôler l'usage de ces documents.

À cet égard il est nécessaire d'établir des procédés anti-copie afin d'empêcher la copie illégale. Les mesures techniques destinées à empêcher la copie numérique et sa diffusion sont le tatouage numérique et le cryptage.

On peut ainsi insérer aux documents numériques des éléments de sécurité tels que la signature du propriétaire cryptée (copyright) ou un logo qui permettent d'identifier son contenu ainsi de détecter et/ou faire le suivi de toute modification et de vérifier son authenticité.

Le problème est de cacher une marque dans la représentation analogique du document de telle façon que la qualité perçue de ce document ne soit pas réduite mais qu'il soit impossible d'effacer ou de rendre inutilisable cette marque sans altérer la qualité du document. Par exemple, le tatouage numérique doit être invisible dans le document et des distorsions évidentes doivent apparaître si une attaque est tentée.

Notre objectif dans ce mémoire est de rassembler les techniques de tatouage numérique et d'embrouillage afin d'aboutir à un système efficace de protection des droits d'auteurs des documents imprimés sensibles et officiels. Ainsi, nous voulons éviter et lutter contre la fraude et le piratage et d'assurer la protection des droits de propriété intellectuelle des documents qu'ils soient numériques ou imprimés.

La structure de ce mémoire reflète la logique de notre objectif et comprend quatre (04) chapitres qui sont présentés comme suit :

Le premier chapitre est scindé en deux sections. Dans la première section, nous présentons des définitions et quelques concepts de base. Nous décrivons les principales caractéristiques à travers les différentes définitions rencontrées dans la littérature. Aussi nous exposons les différents domaines d'application où le tatouage numérique est appliqué.

La seconde section comporte quelques méthodes de protection des documents imprimés par tatouage numérique. Toute en évoquant leurs avantages et inconvénients.

Une présentation détaillée du système de sécurisation des documents imprimés par tatouage numériques fait l'objet du second chapitre. Le système proposé opère dans deux différents domaines à savoir, le domaine spatial où les documents sont transformés en demi-teinte et le domaine fréquentiel qui prend en compte la transformation DCT 2D pour la taille des blocs 8x8 pixels et 4x4 pixels. Ce chapitre comporte également les différents tests et expérimentations réalisées afin de valider les résultats d'insertion et d'extraction de la marque.

Dans le chapitre trois, nous présentons les différents résultats d'évaluation face à la principale attaque d'impression/numérisation destinée pour les documents imprimés et aussi les résultats obtenus face aux attaques de traitements d'image usuels.

Le chapitre 4 présente l'environnement de programmation ainsi que l'interface de l'application et les différentes fonctionnalités offertes par le système.

Enfin, nous terminerons par une conclusion générale en énonçant quelques perspectives.

Il est important de porter à la connaissance du lecteur que notre travail rentre dans le cadre du projet « *Développement d'un système de sécurisation des documents numériques et imprimés par tatouage numérique* » initié par le CDTA dans le cadre des projets menées par le Réseau de Recherche Nationale dans le domaine des sciences et technologies de l'information et de la communication (2RSTIC) lancés par le Ministère de la poste et des technologies de l'information et de la communication au cours de l'année 2009.

# CHAPITRE I :

# GENERALITES ET DEFENITIONS

*Le savoir que l'on ne complète pas chaque  
jour diminue tous les jours.*  
Proverbe Chinois



**I. Introduction**

Différentes techniques ont été proposées pour la protection des droits d'auteurs des documents numériques, mais peu de travaux ont été effectués pour les documents imprimés ou papiers.

Avant de donner un aperçu sur les méthodes développées dans ce sens, il est nécessaire de donner du tatouage numérique, les définitions de base et les principales caractéristiques et applications de cette thématique.

**II. Historique**

L'idée de cacher une information existait depuis bien longtemps, il existe quelques événements historiques qui illustrent ce concept, un exemple a été cité dans le livre « *The Histories of Herodotus* », où « *histiaeus* » a rasé le crane d'un de ses esclaves, et a tatoué un message sur son crane, il a attendu que ses cheveux grandissent, ainsi le message a été caché, et cela dans le but de lancer une guerre contre l'empire perse [1].

Le travail du tatouage a été publié vers 1282 en Italie où les marques étaient constituées par l'addition d'un modèle de fil fin au module du papier. Le papier était devenu légèrement fin où le fil a été rajouté, et donc plus transparent. Néanmoins, le tatouage n'avait pas les mêmes objectifs que ceux de nos jours. Le but derrière l'insertion de la marque était incertain, la marque a été utilisée probablement pour identifier le moule avec lequel le papier a été fabriqué, ou pour représenter de signes de religion, ou simplement pour l'utiliser comme un outil de décoration [2].

Au début de 21<sup>ème</sup> siècle, l'utilisation du tatouage de papier est devenue de plus en plus claire en Europe et en Amérique et les marques servaient comme une marque de fabrication pour noter la date de fabrication du papier, et indiquer la taille originale de la feuille. Dans cette même période, les marques ont été utilisées aussi comme moyen contre la falsification des billets monétaires et d'autres documents confidentiels.

Le terme marque « watermark » a été utilisé et pour la première fois à la fin du 21<sup>ème</sup> siècle, et paraît comme étant une dérivée de la langue allemande et précisément du terme mark veut dire la marque. Cette appellation de la marque ne veut pas dire que

l'eau est nécessaire pour la création de la marque, mais probablement que la marque ressemble à l'effet de l'eau sur le papier [2].

En 1954, Emil Hembrooke de la Muzak Corporation a rangé un tatouage dans un travail musical. Un code d'identification a été inséré dans la musique par l'application d'un filtre faible à une fréquence de 1KHz.

En 1979, Szrpanski a décrit une machine qui détecte des modèles placés dans des documents pour le contrefaire. Neuf ans plus tard, Holt et Al ont décrit une méthode pour insérer un code d'identification dans un signal audio. Et c'était en 1988 que Komatsu et Tominga ont utilisé le terme tatouage numérique pour la première fois.

Dés 1995, l'utilisation de la technique du tatouage numérique est devenue trop large.

### **III. Notions de base du tatouage**

#### **III.1. Définitions**

Le tatouage numérique, digital watermarking en anglais, consiste à insérer un tatouage dans un document numérique (image, son, vidéo. . .). La modification s'effectue dans les composantes perceptibles (comme la luminance des pixels d'une image), et non dans l'en-tête d'un fichier par exemple. Ce tatouage doit pouvoir être détecté et décodé, mais doit être imperceptible, c'est-à-dire que la déformation doit être suffisamment faible pour que l'utilisateur ne puisse pas différencier le document tatoué de d'origine. Cette notion d'imperceptibilité et d'insertion dans la trame même du document rejoint la traduction littérale du terme digital watermark : "filigrane électronique" [3].

On peut trouver les premiers filigranes sur des papiers du treizième siècle, dans le but de garantir leur qualité. Sur un billet de banque, les fibres sont marquées au moment de la sortie du bain d'eau, ce qui est à l'origine du terme anglais water mark. De la même manière que sur un billet de banque, le filigrane électronique est d'abord invisible et n'est révélé que par une transformation spécifique. L'intérêt d'une telle opération est que le tatouage est indépendant du format de stockage des données, puisqu'il est intrinsèque au document [3].

Le schéma du tatouage numérique est résumé sur la figure I.1. Un message  $m$  contenant  $L$  bits d'information est transformé selon une clé  $k$  en un tatouage  $w$  qui est ensuite inséré dans le document d'origine  $x$  (aussi appelé "hôte") pour donner un document tatoué  $y$ . C'est la phase d'insertion. Ici,  $w$  est exprimé sous la forme d'un bruit qui est ajouté au document, la déformation dépendant de la puissance du bruit.  $k$  est secrète et spécifique au tatoueur.  $y$  est ensuite copié et attaqué, ce qui est modélisé par la transmission dans un canal soumis à du bruit. Le document reçu est appelé  $z$ . La réception d'un document consiste en deux parties : d'une part, la détection du tatouage et d'autre part, s'il est présent, son décodage.

La phase de détection consiste à prouver la présence d'un tatouage dans  $z$  grâce à  $k$ . La phase de décodage consiste à calculer une estimation  $m'$  de  $m$ . Si la taille du message inséré  $L$  est suffisamment grande et contient une information intelligible (par exemple, des caractères ASCII), certains auteurs considèrent que la détection devient inutile puisqu'on peut appliquer un simple décodage. Si la chaîne décodée est inintelligible (par exemple, non ASCII), on considère qu'il n'y a pas de tatouage.

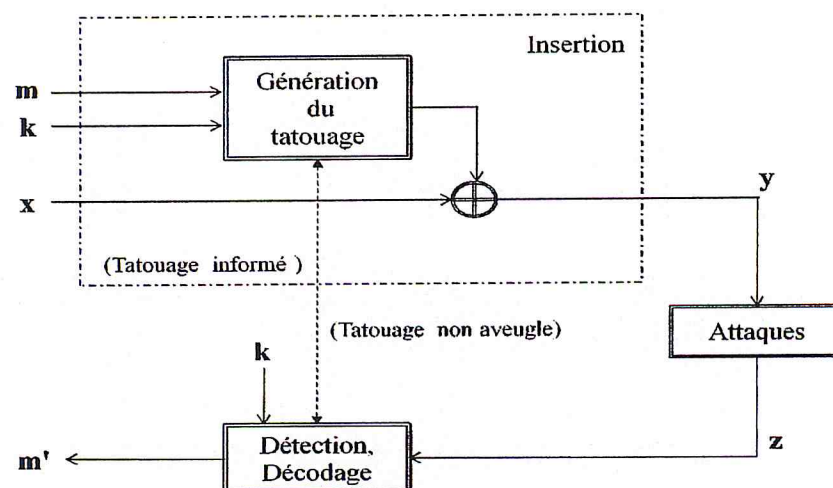


Figure I.1. Schéma générale d'un système de tatouage numérique.

### III.2. Applications de tatouage numérique

Le tatouage est adapté à un large champ d'applications. Les applications les plus courantes sont décrites ci-dessous :

- **Protection des droits d'auteur (copyright)**

C'est l'une des applications les plus porteuses : le tatouage numérique offre une alternative intéressante à la cryptographie, car il permet de protéger le document, même lorsque celui-ci est diffusé.

L'objectif est d'incruster une information dans la donnée source, typiquement le copyright du propriétaire, afin de prévenir toute revendication frauduleuse de propriété. Cette signature ne doit être connue que de la personne ou de l'organisme de tatouage. Elle dépend donc d'une clé secrète, qui permet son insertion et sa détection. Cette application nécessite la mise en place d'un algorithme très élevé. En effet, celui-ci ne doit pas être très ambigu et doit toujours déterminer l'appartenance du medium, même si d'autres insèrent également une marque [4].

- **Traçabilité (fingerprinting)**

Le but de cette application est de pouvoir contrôler et faire le suivi des copies de document et tracer les copies illégales de ce dernier (suivi des pirates). Ce type d'application engendre un marquage unique pour chaque document distribué. Ces marques doivent être très robustes, pour pouvoir résister à des attaques ayant pour but de détruire la marque [4].

- **Protection contre les copies**

Un souhait des distributeurs des documents numériques est l'existence d'un moyen de protection contre la copie, afin d'interdire une circulation de ces derniers illégaux. Cette application consiste à intégrer au document une marque "intelligente". Cela nécessite l'utilisation de matériel particulier. En effet, les appareils doivent pouvoir détecter la marque et agir en conséquence, c'est-à-dire en permettant ou non la lecture ou la copie du document [5].

- **Authentification**

Dans le cadre d'application telle que l'authentification, l'objectif est de détecter toutes modifications éventuelles des données, afin de pouvoir certifier si celles-ci ont été modifiées ou non.

Parmi l'ensemble des applications en tatouage, le tatouage pour l'authentification est donc celui qui use du niveau le plus faible de robustesse. Il doit être résistant à des attaques classiques mais doit être détruit en cas de modification de la donnée. Dans ce cas, la marque peut être intégrée sur les objets principaux de la donnée. Si un de ces objets est modifié ou supprimé, la marque est alors détruite [4].

- **Indexation**

Le domaine d'indexation des images consiste à classer de manière automatique des images selon leur contenu. Il permet de faciliter une recherche dans une base de données. Les techniques classiques utilisées consistent à effectuer un traitement automatique de l'image, de manière à dégager les composantes essentielles du contenu. Le tatouage d'un document permet ainsi d'insérer une information (contenant peu de bits) décrivant le contenu de l'image. Cela permet de qualifier sommairement l'image, ou d'insérer un pointeur vers une description plus complète [4].

- **Sécurité médicale**

Insertion d'un " identifiant" confidentiel assurant la correspondance entre le patient et la radio, afin d'éviter toute confusions [4].

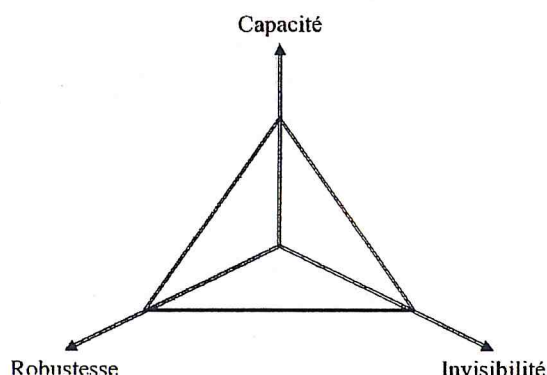
### **III.3. Caractéristique du tatouage numérique**

Il n'existe pas de caractéristiques générales pour tous les systèmes, chaque système sécurisé par le tatouage numérique a ses propres propriétés et cela selon les besoins de l'application, néanmoins, il existe quelques caractéristiques communes entre ces systèmes :

- **Imperceptibilité**

Désigne la similarité perceptuelle entre la donnée originale et celle tatouée. La procédure d'insertion doit assurer que la marque est imperceptible pour toutes les personnes. La marque est vraiment imperceptible si les humains ne peuvent pas faire la différence entre la donnée originale et la copie tatouée.

Certaines approches, dites perceptuelles, ont été proposées afin d'assurer l'imperceptibilité en exploitant les caractéristiques du Système Visuel Humain (SVH)



**Figure I.2.** Illustration graphique du compromis entre les caractéristiques du tatouage numérique.

### III.4. Classification des systèmes du tatouage numérique

On peut distinguer plusieurs critères de classification des algorithmes de tatouage : Les algorithmes du tatouage numérique sont classifiés selon plusieurs critères à savoir, le mode d'insertion de la marque (Schéma additif ou Schéma substitutif), la façon dont est insérée la marque : directement dans le document (domaine spatial), dans une transformé du document (domaine fréquentiel), selon le type du tatouage (aveugle ou non aveugle, perceptible ou imperceptible et symétrique ou asymétrique) et aussi selon le type de la marque insérée (fragile ou robuste). Pour mieux étudier les algorithmes de tatouage numérique, ces critères vont séparément traités.

#### III.4.1. Classification selon la manière d'insertion

- **Schéma additif** : Ce schéma se résume dans l'extraction des coefficients à modifier du document d'origine puis le tatouage de ce dernier s'effectue par l'ajout de la marque à ces coefficients. L'insertion peut s'effectuer soit directement sur le document, dans le domaine spatial, soit dans un domaine transformé. De ce fait, adapter la marque au document d'origine est une contrainte essentielle à respecter pour que le signal qu'elle représente ne soit ni trop faible (problème de robustesse) ni trop fort (dégradation du signal original) [8].

- **Schéma substitutif** : Dans les modes substitutifs, l'information à insérer est substituée à des caractéristiques du document. Par exemple, P.Bas et J. M. Chassery proposent dans [9] une méthode basée sur l'insertion de similarités. L'idée de base consiste donc à insérer une signature en modifiant le contenu structurel du document. Ainsi, l'étape d'insertion consiste d'une part à détecter les points d'intérêt et d'autre part

à insérer des similarités autour de ces points. A la suite, la détection de marque s'effectue par recherche de ces similarités.

### III.4.2. Classification selon le domaine d'insertion

- **Le domaine spatial** : Les méthodes qui viennent en premier à l'esprit sont celles du domaine spatial, où elles modifient et agissent directement sur la luminance des pixels [10]. Comme aucun traitement initial n'est requis, ces algorithmes sont très rapides et permettent de travailler en temps réel [11]. La plupart d'algorithmes utilisent le LSB bit (Least Significant Bit) pour l'insertion. Il y a beaucoup de variantes de cette technique. Elle implique essentiellement d'insérer la marque en remplaçant le moindre peu significatif des bits de document avec un bit de données de la marque [12].

- **Le domaine fréquentiel** : Domaine transformé ou domaine fréquentiel obtenu du domaine spatial par une transformation en une dimension ou deux dimensions. La transformation peut se réaliser sur tout le document ou sur des blocs obtenus par une subdivision de celui-ci [SW1]. L'avantage principal de ce domaine par rapport au domaine spatial est que l'insertion de la marque se fait dans les coefficients de la transformée, et ainsi, elle assure que les modifications appliquées sur un sous-ensemble de ces coefficients seront propagées à tout les pixels dans le domaine spatial. Ce qui rend ces modifications imperceptibles [13].

Les transformées les plus utilisées dans le domaine de tatouage numérique sont : DCT (*Discrete Cosine Transform* ou TCD *Transformée en Cosinus Discrète* en français), DFT (*Discrete Fourier Transform* ou TFD *Transformé en Fourier discrète*) et DWT (*Discret Wavelet Transform* ou TOD *Transformée en ondelettes*). Bien que l'insertion dans le domaine spatial soit simple et facile à appliquer, l'insertion dans le domaine fréquentiel offre plus d'avantages, particulièrement, en termes d'imperceptibilité et de robustesse.

### III.4.3. Classification selon le type de la marque insérée

- **Tatouage robuste** : Il a pour objectif de transmettre une information malgré la modification du document. Lors de la lecture de la marque, certains algorithmes permettent d'extraire un message complet (une suite de symboles), tandis que d'autres indiquent simplement si le document a été marqué ou pas (on parle de détection de marque). Le tatouage robuste est particulièrement adapté au suivi et à la gestion de

droits. Même si un fraudeur modifie le document, il est possible de retrouver l'auteur initial en insérant un numéro d'identification par tatouage robuste [14].

- **Tatouage fragile** : Il permet de vérifier l'intégrité du document marqué. Il est très fragile aux modifications, et permet de vérifier que le document n'a pas été retouché et donc de l'authentifier. Néanmoins, certains systèmes de tatouage fragile sont tout de même résistants aux traitements les plus usuels (compression avec perte notamment) afin de ne détecter que les modifications les plus préjudiciables vis-à-vis de l'interprétation du document. Ce type de schéma de tatouage est dit **semi-fragile**.

### III.4.4. Classification selon le type du tatouage

- **Tatouage imperceptible** : Dans ce type, on n'observe pas l'existence de la marque. En conséquence, elle n'affecte pas la qualité du document et ce dernier garde sa qualité commerciale [15].

- **Tatouage perceptible** : Par contre, dans ce type de tatouage, la marque est bien visible dans le document. Il est utilisé plus dans l'application non commerciale [15].

Aussi, les schémas de tatouages peuvent être classés suivant les éléments nécessaires pour l'extraction (lecture du message depuis le document) de la marque. Un schéma **aveugle** n'a pas besoin du document d'origine pour extraire la marque. Au contraire, un schéma **non aveugle** nécessite le media d'origine pour pouvoir lire correctement le message. Ces types de schémas sont de moins en moins étudiés, les applications concrètes étant assez rares [14].

Un dernier point discriminant est l'utilisation des clefs. La marque insérée est issue du codage du message à transmettre. Il est dépendant d'une clef. Si cette même clef est nécessaire au décodage (c'est-à-dire à l'extraction du message), le schéma est **symétrique** et dans le cas contraire, il est **asymétrique** (systèmes à clef privée et clef publique). On retrouve cette classification dans les algorithmes de cryptographie.

### III.5. Les attaques

Les attaques sont, le plus souvent, des traitements classiques qu'une personne effectue sur le support qu'elle utilise. Elles peuvent être des traitements visant soit à brouiller soit à enlever la marque de protection dans le document. On peut distinguer deux grandes familles d'attaques, les bienveillantes et les malveillantes [14] :



- **Attaques bienveillantes** : Il s'agit de traitement qui n'a pas initialement pour objectif d'empêcher la détection de la marque. Il peut s'agir des dégradations dues à une compression, à un changement de type de compression, à des filtrages (réduction de bruit), à un changement de résolution, etc.
- **Attaques malveillantes** : vise explicitement à rendre le tatouage inopérant. Ces attaques, comme souvent dans le domaine numérique, sont difficile à prouver d'un point de vue juridique. Toutefois, une attaque malveillante qui a réussi devra produire un contenu à la fois lavé de son tatouage et encore exploitable.

### **III.6. La notion d'embrouillage**

L'embrouillage est une opération destinée à transformer un signal numérique en un signal numérique aléatoire ou pseudo-aléatoire, de même signification et de même débit binaire, en vue d'en faciliter la transmission ou l'enregistrement.

### **III.7. Quelques techniques existantes du tatouage numérique des documents imprimés [16]**

Avec la prolifération des supports numériques tels que des images, audio et vidéo, les techniques robustes de tatouage numérique sont nécessaires pour la protection des droits d'auteur, de contrôle de copie et d'authentification.

Une variété de tatouage numérique et de techniques de dissimuler des données ont été proposées à ces fins. Cependant, la plupart des méthodes développées aujourd'hui sont pour des documents numériques en niveaux de gris et en couleurs. Ces techniques ne peuvent pas être appliquées directement à des documents imprimés ou papiers. L'inconvénient c'est que l'évolution des valeurs des pixels dans un document imprimés pose des détériorations qui sont visuellement perceptibles.

Dans la littérature, peut de travaux sont publier sur le tatouage numérique des documents imprimés.

Dans cet état de l'art, les principaux travaux publiés des différentes techniques traitant ce sujet important sont présentés, ainsi les discussions actuelles sur des questions importantes telles que la robustesse et la capacité de cacher des données.

Les tatouages numériques – par analogie aux filigranes apparaissant sur les billets de banque – ont été proposés comme un outil miracle pour contrôler l'usage de ces documents.

A cet égard il est nécessaire d'établir des procédés anti-copie afin d'empêcher la copie illégale. Les mesures techniques destinées à empêcher la copie numérique et sa diffusion sont le tatouage numérique et le cryptage.

On peut ainsi insérer aux documents numériques des éléments de sécurité tels que la signature du propriétaire cryptée (copyright) ou un logo qui permettent d'identifier son contenu ainsi que de détecter et/ou faire le suivi de toute modification et de vérifier son authenticité.

Le problème est de cacher une marque dans la représentation analogique du document de telle façon que la qualité perçue de ce document ne soit pas réduite mais qu'il soit impossible d'effacer ou de rendre inutilisable cette marque sans altérer la qualité du document. Par exemple, le tatouage numérique doit être invisible dans le document et des distorsions évidentes doivent apparaître si une attaque est tentée.

Pour Les documents imprimés, la technologie du tatouage inclue des méthodes qui cachent les informations (la marque) dans le texte du document papier après numérisation, et les méthodes basées sur l'insertion d'informations dans l'image du document imprimé en demi-teinte. Les méthodes de cette dernière catégorie sont conçues pour être robuste au processus d'impression et habituellement ces méthodes nécessitent la capture du document imprimé sous forme électronique, à partir de laquelle les données insérées peuvent être extraites ou une marque visible (logo) est révélée par un traitement approprié.

Pour les autres documents numériques (en niveaux de gris ou en couleurs), les techniques actuelles pour l'insertion d'information dans les images (documents) sont inspirées des méthodes de codage d'images et de compression. L'information est insérée en utilisant la transformée en cosinus discrète (TCD), la transformée de Fourier discrète (TFD) en amplitude et en phase, la transformation en ondelettes(TOD), le codage prédictive adaptative. Les méthodes proposées doivent résister à une large gamme de traitements d'images tels que le filtrage, les opérations de conversion numérique /

analogique / numérique, la compression JPEG, l'ajout de bruits, etc. et particulièrement la ré-numérisation et l'impression.

Dans ce paragraphe nous nous intéressons uniquement aux méthodes de tatouage des documents en demi-teinte.

Elles peuvent être classées en deux catégories. Une classe de techniques insère des données invisible dans l'image en demi-teinte de sorte que les données peuvent être lues à partir des images en demi-teinte et en appliquant des algorithmes d'extraction sur les images scannées. Dans d'autres travaux, les auteurs ont utilisé deux différentes matrices de dither (au lieu d'une seule matrice) pour la génération du document en demi-teinte de telle sorte que les différentes propriétés statistiques en raison d'utilisation des deux matrices de dither peuvent être détectées.

*Note* : La matrice de dither est la partie la plus importante de l'approximation des images tramées. En choisissant différentes matrices de dither, nous pouvons obtenir plusieurs niveaux différents d'intensité à partir de l'image initiale. C'est à dire, les configurations de dither ("dither patterns") affectent la qualité de l'image numérique. Afin d'obtenir une meilleure image, nous devons définir les matrices de dither qui sont les meilleures à appliquer.

Trois méthodes d'insertion des données dans des endroits pseudo-aléatoires dans les images en demi-teinte sans en tenir compte de l'image d'origine en couleurs et de la méthode de binarisation. Ces dernières, nommées DHST, DHPT et DHSPT, s'appliquent sur un pixel en demi-teinte pour insérer un bit de données. Dans DHST, N bits de données sont insérés dans N endroits en utilisant un basculement forcé, et cela quand le pixel dans l'endroit pseudo aléatoire diffère de la valeur à insérer, il est forcément basculé. Dans le processus d'extraction, les données sont lues directement à partir des N endroits pseudo-aléatoires. Cette méthode génère des parties indésirables de pixels noirs et blancs.

Dans DHPT, une paire de pixels noir et blanc (au lieu d'un seul dans DHST) est choisi pour basculer dans les endroits pseudo-aléatoires. Le DHSPT consiste à choisir des paires de pixels noir et blanc qui sont au maximum liés avec les pixels voisins avant le basculement. Ces derniers devenus les pixels les moins liés après l'insertion et les

parties résultantes devenues plus petites, ce qui prouve la qualité visuelle de cette méthode.

Un algorithme appelé « Sélection d'intensité » (SI) est proposé pour une meilleure sélection des endroits d'insertion. Il assure une meilleure qualité visuelle des images tatouées sans sacrifier la capacité d'insertion des données.

Généralement, cet algorithme sélectionne des endroits de pixels qui sont très claires ou très sombres. Il représente un bit de données en tant que parité de la somme des pixels en demi-teinte dans  $M$  endroits pseudo-aléatoires et sélectionne le meilleur des  $M$  endroits. Cependant, cet algorithme a besoin de l'image d'origine en niveaux de gris ou la binarisation inverse de l'image.

#### **IV. Conclusion**

Dans ce chapitre, une introduction générale au tatouage numérique a été présentée, avec les définitions de base et les principaux concepts, en décrivant les transformées les plus utilisés dans le tatouage numérique, ses principales caractéristiques et les domaines de son application.

Dans le prochain chapitre, deux techniques de tatouage numériques, notamment dans le domaine spatial et dans le domaine fréquentiel seront détaillées avec des tests et des résultats qui évalueront les performances de ces dernières.

## **I. Introduction**

Différentes techniques ont été proposées pour la protection des droits d'auteurs des documents numériques, mais peu de travaux ont été effectués pour les documents imprimés ou papiers.

Avant de donner un aperçu sur les méthodes développées dans ce sens, il est nécessaire de donner du tatouage numérique, les définitions de base et les principaux caractéristiques et applications de cette thématique.

## **II. Historique**

L'idée de cacher une information existait depuis bien longtemps, il existe quelques événements historiques qui illustrent ce concept, un exemple a été cité dans le livre « *The Histories of Herodotus* », où « *histiaeus* » a rasé le crâne d'un de ses esclaves, et a tatoué un message sur son crâne, il a attendu que ses cheveux grandissent, ainsi le message a été caché, et cela dans le but de lancer une guerre contre l'empire perse [1].

Le travail du tatouage a été publié vers 1282 en Italie où les marques étaient constituées par l'addition d'un modèle de fil fin au module du papier. Le papier était devenu légèrement fin où le fil a été rajouté, et donc plus transparent. Néanmoins, le tatouage n'avait pas les mêmes objectifs que ceux de nos jours. Le but derrière l'insertion de la marque était incertain, la marque a été utilisée probablement pour identifier le moule avec lequel le papier a été fabriqué, ou pour représenter de signes de religion, ou simplement pour l'utiliser comme un outil de décoration [2].

Au début de 21<sup>ème</sup> siècle, l'utilisation du tatouage de papier est devenue de plus en plus claire en Europe et en Amérique et les marques servaient comme une marque de fabrication pour noter la date de fabrication du papier, et indiquer la taille originale de la feuille. Dans cette même période, les marques ont été utilisées aussi comme moyen contre la falsification des billets monétaires et d'autres documents confidentiels.

Le terme marque « watermark » a été utilisé et pour la première fois à la fin du 21<sup>ème</sup> siècle, et paraît comme étant une dérivée de la langue allemande et précisément du terme mark veut dire la marque. Cette appellation de la marque ne veut pas dire que

l'eau est nécessaire pour la création de la marque, mais probablement que la marque ressemble à l'effet de l'eau sur le papier [2].

En 1954, Emil Hembrooke de la Muzak Corporation a rangé un tatouage dans un travail musical. Un code d'identification a été inséré dans la musique par l'application d'un filtre faible à une fréquence de 1KHz.

En 1979, Szrpanski a décrit une machine qui détecte des modèles placées dans des documents pour le contrefaire. Neuf ans plus tard, Holt et Al ont décrit une méthode pour insérer un code d'identification dans un signal audio. Et c'était en 1988 que Komatsu et Tominga ont utilisé le terme tatouage numérique pour la première fois.

Dés 1995, l'utilisation de la technique du tatouage numérique est devenue trop large.

### **III. Notions de base du tatouage**

#### **III.1. Définitions**

Le tatouage numérique, digital watermarking en anglais, consiste à insérer un tatouage dans un document numérique (image, son, vidéo. . .). La modification s'effectue dans les composantes perceptibles (comme la luminance des pixels d'une image), et non dans l'en-tête d'un fichier par exemple. Ce tatouage doit pouvoir être détecté et décodé, mais doit être imperceptible, c'est-à-dire que la déformation doit être suffisamment faible pour que l'utilisateur ne puisse pas différencier le document tatoué de d'origine. Cette notion d'imperceptibilité et d'insertion dans la trame même du document rejoint la traduction littérale du terme digital watermark : "filigrane électronique" [3].

On peut trouver les premiers filigranes sur des papiers du treizième siècle, dans le but de garantir leur qualité. Sur un billet de banque, les fibres sont marquées au moment de la sortie du bain d'eau, ce qui est à l'origine du terme anglais water mark. De la même manière que sur un billet de banque, le filigrane électronique est d'abord invisible et n'est révélé que par une transformation spécifique. L'intérêt d'une telle opération est que le tatouage est indépendant du format de stockage des données, puisqu'il est intrinsèque au document [3].

Le schéma du tatouage numérique est résumé sur la figure I.1. Un message  $m$  contenant  $L$  bits d'information est transformé selon une clé  $k$  en un tatouage  $w$  qui est ensuite inséré dans le document d'origine  $x$  (aussi appelé "hôte") pour donner un document tatoué  $y$ . C'est la phase d'insertion. Ici,  $w$  est exprimé sous la forme d'un bruit qui est ajouté au document, la déformation dépendant de la puissance du bruit.  $k$  est secrète et spécifique au tatoueur.  $y$  est ensuite copié et attaqué, ce qui est modélisé par la transmission dans un canal soumis à du bruit. Le document reçu est appelé  $z$ . La réception d'un document consiste en deux parties : d'une part, la détection du tatouage et d'autre part, s'il est présent, son décodage.

La phase de détection consiste à prouver la présence d'un tatouage dans  $z$  grâce à  $k$ . La phase de décodage consiste à calculer une estimation  $m'$  de  $m$ . Si la taille du message inséré  $L$  est suffisamment grande et contient une information intelligible (par exemple, des caractères ASCII), certains auteurs considèrent que la détection devient inutile puisqu'on peut appliquer un simple décodage. Si la chaîne décodée est inintelligible (par exemple, non ASCII), on considère qu'il n'y a pas de tatouage.

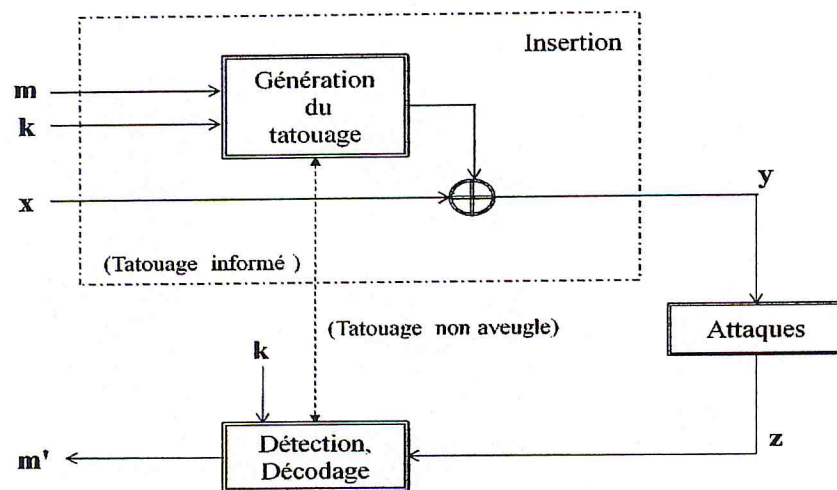


Figure I.1. Schéma générale d'un système de tatouage numérique.

### III.2. Applications de tatouage numérique

Le tatouage est adapté à un large champ d'applications. Les applications les plus courantes sont décrites ci-dessous :

- **Protection des droits d'auteur (copyright)**

C'est l'une des applications les plus porteuses : le tatouage numérique offre une alternative intéressante à la cryptographie, car il permet de protéger le document, même lorsque celui-ci est diffusé.

L'objectif est d'incruster une information dans la donnée source, typiquement le copyright du propriétaire, afin de prévenir toute revendication frauduleuse de propriété. Cette signature ne doit être connue que de la personne ou de l'organisme de tatouage. Elle dépend donc d'une clé secrète, qui permet son insertion et sa détection. Cette application nécessite la mise en place d'un algorithme très élevé. En effet, celui-ci ne doit pas être très ambigu et doit toujours déterminer l'appartenance du médium, même si d'autres insèrent également une marque [4].

- **Traçabilité (fingerprinting)**

Le but de cette application est de pouvoir contrôler et faire le suivi des copies de document et tracer les copies illégales de ce dernier (suivi des pirates). Ce type d'application engendre un marquage unique pour chaque document distribué. Ces marques doivent être très robustes, pour pouvoir résister à des attaques ayant pour but de détruire la marque [4].

- **Protection contre les copies**

Un souhait des distributeurs des documents numériques est l'existence d'un moyen de protection contre la copie, afin d'interdire une circulation de ces derniers illégaux. Cette application consiste à intégrer au document une marque "intelligente". Cela nécessite l'utilisation de matériel particulier. En effet, les appareils doivent pouvoir détecter la marque et agir en conséquence, c'est-à-dire en permettant ou non la lecture ou la copie du document [5].

- **Authentification**

Dans le cadre d'application telle que l'authentification, l'objectif est de détecter toutes modifications éventuelles des données, afin de pouvoir certifier si celles-ci ont été modifiées ou non.



Parmi l'ensemble des applications en tatouage, le tatouage pour l'authentification est donc celui qui use du niveau le plus faible de robustesse. Il doit être résistant à des attaques classiques mais doit être détruit en cas de modification de la donnée. Dans ce cas, la marque peut être intégrée sur les objets principaux de la donnée. Si un de ces objets est modifié ou supprimé, la marque est alors détruite [4].

- **Indexation**

Le domaine d'indexation des images consiste à classer de manière automatique des images selon leur contenu. Il permet de faciliter une recherche dans une base de données. Les techniques classiques utilisées consistent à effectuer un traitement automatique de l'image, de manière à dégager les composantes essentielles du contenu. Le tatouage d'un document permet ainsi d'insérer une information (contenant peu de bits) décrivant le contenu de l'image. Cela permet de qualifier sommairement l'image, ou d'insérer un pointeur vers une description plus complète [4].

- **Sécurité médicale**

Insertion d'un " identifiant " confidentiel assurant la correspondance entre le patient et la radio, afin d'éviter toute confusions [4].

### **III.3. Caractéristique du tatouage numérique**

Il n'existe pas de caractéristiques générales pour tous les systèmes, chaque système sécurisé par le tatouage numérique a ses propres propriétés et cela selon les besoins de l'application, néanmoins, il existe quelques caractéristiques communes entre ces systèmes :

- **Imperceptibilité**

Désigne la similarité perceptuelle entre la donnée originale et celle tatouée. La procédure d'insertion doit assurer que la marque est imperceptible pour toutes les personnes. La marque est vraiment imperceptible si les humains ne peuvent pas faire la différence entre la donnée originale et la copie tatouée.

Certaines approches, dites perceptuelles, ont été proposées afin d'assurer l'imperceptibilité en exploitant les caractéristiques du Système Visuel Humain (SVH)

dans le cas des images et celles du Système Audible Humain (SAH) dans le cas des signaux audio [6].

- **Robustesse**

Détermine l'habilité de détecter et/ ou d'extraire la marque même si la qualité de la données tatouée est dégradée des éventuelles transformations. Ces transformations peuvent être des manipulations usuelles ou des attaques malveillantes. Dans le premier cas, la marque est affecté sans avoir l'intension de l'enlever ou le détruire (par exemple : compression avec perte, impression, etc.). Dans le deuxième cas la donnée tatouée est manipulée afin d'extraire ou détruire la marque insérée (par exemple : rotation, cropping, translation, etc.). A l'exception de l'application de protection d'intégrité, la marque doit survivre et résister à toutes les attaques susceptibles de se produire entre le moment de l'insertion de la marque et celui de sa détection/ extraction [6].

- **Capacité**

Désigne le nombre de bits qui peuvent être insérés dans le document d'origine tout en respectant les exigences des autres caractéristiques. Bien qu'une grande capacité soit désirée, son insertion cause des dégradations en termes de robustesse et d'imperceptibilité. Plus la capacité est grande, plus la qualité perceptuelle du document tatoué sera dégradée et moins robuste sera la marque insérée [6].

- **Sécurité**

Représente la sécurité de l'information insérée. Une technique du tatouage est vraiment sécurisée si la connaissance des algorithmes exacts d'insertion et d'extraction de la marque ne sert pas à enlever ou détruire la marque par une personne non autorisée. Dans la plupart des systèmes, cette sécurité est assurée en utilisant des clés secrètes et pseudo-aléatoires. Ces clés servent à générer la séquence représentant la marque ou à déterminer les emplacements où la marque sera insérée [7].

Les trois premières caractéristiques sont fortement liées les unes aux autres et un schéma du tatouage numérique s'inscrit donc dans un compromis entre l'imperceptibilité, la robustesse et la capacité. Ce dernier est illustré sur la figure I.2.

à insérer des similarités autour de ces points. A la suite, la détection de marque s'effectue par recherche de ces similarités.

### **III.4.2. Classification selon le domaine d'insertion**

- **Le domaine spatial** : Les méthodes qui viennent en premier à l'esprit sont celles du domaine spatial, où elles modifient et agissent directement sur la luminance des pixels [10]. Comme aucun traitement initial n'est requis, ces algorithmes sont très rapides et permettent de travailler en temps réel [11]. La plupart d'algorithmes utilisent le LSB bit (Least Significant Bit) pour l'insertion. Il y a beaucoup de variantes de cette technique. Elle implique essentiellement d'insérer la marque en remplaçant le moindre peu significatif des bits de document avec un bit de données de la marque [12].

- **Le domaine fréquentiel** : Domaine transformé ou domaine fréquentiel obtenu du domaine spatial par une transformation en une dimension ou deux dimensions. La transformation peut se réaliser sur tout le document ou sur des blocs obtenus par une subdivision de celui ci [SW1]. L'avantage principal de ce domaine par rapport au domaine spatial et que l'insertion de la marque se fait dans les coefficients de la transformée, et ainsi, elle assure que les modifications appliquées sur un sous ensemble de ces coefficients seront propagées à tout les pixels dans le domaine spatial. Ce qui rend ces modifications imperceptibles [13].

Les transformées les plus utilisées dans le domaine de tatouage numérique sont : DCT (*Discrete Cosine Transform* ou TCD *Transformée en Cosinus Discrète* en français), DFT (*Discrete Fourier Transform* ou TFD *Transformé en Fourier discrète*) et DWT (*Discret Wavelet Transform* ou TOD *Transformée en ondelettes*). Bien que l'insertion dans le domaine spatial soit simple et facile à appliquer, l'insertion dans le domaine fréquentiel offre plus d'avantages, particulièrement, en termes d'imperceptibilité et de robustesse.

### **III.4.3. Classification selon le type de la marque insérée**

- **Tatouage robuste** : Il a pour objectif de transmettre une information malgré la modification du document. Lors de la lecture de la marque, certains algorithmes permettent d'extraire un message complet (une suite de symbole), tandis que d'autres indiquent simplement si le document a été marqué ou pas (on parle de détection de marque). Le tatouage robuste est particulièrement adapté au suivi et à la gestion de

droits. Même si un fraudeur modifie le document, il est possible de retrouver l'auteur initial en insérant un numéro d'identification par tatouage robuste [14].

- **Tatouage fragile** : Il permet de vérifier l'intégrité du document marqué. Il est très fragile aux modifications, et permet de vérifier que le document n'a pas été retouché et donc de l'authentifier. Néanmoins, certains systèmes de tatouage fragile sont tout de même résistants aux traitements les plus usuels (compression avec perte notamment) afin de ne détecter que les modifications les plus préjudiciables vis-à-vis de l'interprétation du document. Ce type de schéma de tatouage est dit **semi-fragile**.

#### **III.4.4. Classification selon le type du tatouage**

- **Tatouage imperceptible** : Dans ce type, on n'observe pas l'existence de la marque. En conséquence, elle n'affecte pas la qualité du document et ce dernier garde sa qualité commerciale [15].

- **Tatouage perceptible** : Par contre, dans ce type de tatouage, la marque est bien visible dans le document. Il est utilisé plus dans l'application non commerciale [15].

Aussi, les schémas de tatouages peuvent être classés suivant les éléments nécessaires pour l'extraction (lecture du message depuis le document) de la marque. Un schéma **aveugle** n'a pas besoin du document d'origine pour extraire la marque. Au contraire, un schéma **non aveugle** nécessite le media d'origine pour pouvoir lire correctement le message. Ces types de schémas sont de moins en moins étudiés, les applications concrètes étant assez rares [14].

Un dernier point discriminant est l'utilisation des clefs. La marque insérée est issue du codage du message à transmettre. Il est dépendant d'une clef. Si cette même clef est nécessaire au décodage (c'est-à-dire à l'extraction du message), le schéma est **symétrique** et dans le cas contraire, il est **asymétrique** (systèmes à clef privée et clef publique). On retrouve cette classification dans les algorithmes de cryptographie.

#### **III.5. Les attaques**

Les attaques sont, le plus souvent, des traitements classiques qu'une personne effectue sur le support qu'elle utilise. Elles peuvent être des traitements visant soit à brouiller soit à enlever la marque de protection dans le document. On peut distinguer deux grandes familles d'attaques, les bienveillantes et les malveillantes [14] :

- **Attaques bienveillantes** : Il s'agit de traitement qui n'a pas initialement pour objectif d'empêcher la détection de la marque. Il peut s'agir des dégradations dues à une compression, à un changement de type de compression, à des filtrages (réduction de bruit), à un changement de résolution, etc.
- **Attaques malveillantes** : vise explicitement à rendre le tatouage inopérant. Ces attaques, comme souvent dans le domaine numérique, sont difficile à prouver d'un point de vue juridique. Toutefois, une attaque malveillante qui a réussi devra produire un contenu à la fois lavé de son tatouage et encore exploitable.

### **III.6. La notion d'embrouillage**

L'embrouillage est une opération destinée à transformer un signal numérique en un signal numérique aléatoire ou pseudo-aléatoire, de même signification et de même débit binaire, en vue d'en faciliter la transmission ou l'enregistrement.

### **III.7. Quelques techniques existantes du tatouage numérique des documents imprimés [16]**

Avec la prolifération des supports numériques tels que des images, audio et vidéo, les techniques robustes de tatouage numérique sont nécessaires pour la protection des droits d'auteur, de contrôle de copie et d'authentification.

Une variété de tatouage numérique et de techniques de dissimuler des données ont été proposées à ces fins. Cependant, la plupart des méthodes développées aujourd'hui sont pour des documents numériques en niveaux de gris et en couleurs. Ces techniques ne peuvent pas être appliquées directement à des documents imprimés ou papiers. L'inconvénient c'est que l'évolution des valeurs des pixels dans un document imprimés pose des détériorations qui sont visuellement perceptibles.

Dans la littérature, peut de travaux sont publier sur le tatouage numérique des documents imprimés.

Dans cet état de l'art, les principaux travaux publiés des différentes techniques traitant ce sujet important sont présentés, ainsi les discussions actuelles sur des questions importantes telles que la robustesse et la capacité de cacher des données.

Les tatouages numériques – par analogie aux filigranes apparaissant sur les billets de banque – ont été proposés comme un outil miracle pour contrôler l'usage de ces documents.

A cet égard il est nécessaire d'établir des procédés anti-copie afin d'empêcher la copie illégale. Les mesures techniques destinées à empêcher la copie numérique et sa diffusion sont le tatouage numérique et le cryptage.

On peut ainsi insérer aux documents numériques des éléments de sécurité tels que la signature du propriétaire cryptée (copyright) ou un logo qui permettent d'identifier son contenu ainsi que de détecter et/ou faire le suivi de toute modification et de vérifier son authenticité.

Le problème est de cacher une marque dans la représentation analogique du document de telle façon que la qualité perçue de ce document ne soit pas réduite mais qu'il soit impossible d'effacer ou de rendre inutilisable cette marque sans altérer la qualité du document. Par exemple, le tatouage numérique doit être invisible dans le document et des distorsions évidentes doivent apparaître si une attaque est tentée.

Pour Les documents imprimés, la technologie du tatouage inclue des méthodes qui cachent les informations (la marque) dans le texte du document papier après numérisation, et les méthodes basées sur l'insertion d'informations dans l'image du document imprimé en demi-teinte. Les méthodes de cette dernière catégorie sont conçues pour être robuste au processus d'impression et habituellement ces méthodes nécessitent la capture du document imprimé sous forme électronique, à partir de laquelle les données insérées peuvent être extraites ou une marque visible (logo) est révélée par un traitement approprié.

Pour les autres documents numériques (en niveaux de gris ou en couleurs), les techniques actuelles pour l'insertion d'information dans les images (documents) sont inspirées des méthodes de codage d'images et de compression. L'information est insérée en utilisant la transformée en cosinus discrète (TCD), la transformée de Fourier discrète (TFD) en amplitude et en phase, la transformation en ondelettes(TOD), le codage prédictive adaptative. Les méthodes proposées doivent résister à une large gamme de traitements d'images tels que le filtrage, les opérations de conversion numérique /

analogique / numérique, la compression JPEG, l'ajout de bruits, etc. et particulièrement la ré-numérisation et l'impression.

Dans ce paragraphe nous nous intéressons uniquement aux méthodes de tatouage des documents en demi-teinte.

Elles peuvent être classées en deux catégories. Une classe de techniques insère des données invisible dans l'image en demi-teinte de sorte que les données peuvent être lues à partir des images en demi-teinte et en appliquant des algorithmes d'extraction sur les images scannées. Dans d'autres travaux, les auteurs ont utilisé deux différentes matrices de dither (au lieu d'une seule matrice) pour la génération du document en demi-teinte de telle sorte que les différentes propriétés statistiques en raison d'utilisation des deux matrices de dither peuvent être détectées.

*Note :* La matrice de dither est la partie la plus importante de l'approximation des images tramées. En choisissant différentes matrices de dither, nous pouvons obtenir plusieurs niveaux différents d'intensité à partir de l'image initiale. C'est à dire, les configurations de dither ("dither patterns") affectent la qualité de l'image numérique. Afin d'obtenir une meilleure image, nous devons définir les matrices de dither qui sont les meilleures à appliquer.

Trois méthodes d'insertion des données dans des endroits pseudo-aléatoires dans les images en demi-teinte sans en tenir compte de l'image d'origine en couleurs et de la méthode de binarisation. Ces dernières, nommées DHST, DHPT et DHSPT, s'appliquent sur un pixel en demi-teinte pour insérer un bit de données. Dans DHST, N bits de données sont insérés dans N endroits en utilisant un basculement forcé, et cela quand le pixel dans l'endroit pseudo aléatoire diffère de la valeur à insérer, il est forcément basculé. Dans le processus d'extraction, les données sont lues directement à partir des N endroits pseudo-aléatoires. Cette méthode génère des parties indésirables de pixels noirs et blancs.

Dans DHPT, une paire de pixels noir et blanc (au lieu d'un seul dans DHST) est choisi pour basculer dans les endroits pseudo-aléatoires. Le DHSPT consiste à choisir des paires de pixels noir et blanc qui sont au maximum liés avec les pixels voisins avant le basculement. Ces derniers devenus les pixels les moins liés après l'insertion et les

parties résultantes devenues plus petites, ce qui prouve la qualité visuelle de cette méthode.

Un algorithme appelé « Sélection d'intensité » (SI) est proposé pour une meilleure sélection des endroits d'insertion. Il assure une meilleure qualité visuelle des images tatouées sans sacrifier la capacité d'insertion des données.

Généralement, cet algorithme sélectionne des endroits de pixels qui sont très claires ou très sombres. Il représente un bit de données en tant que parité de la somme des pixels en demi-teinte dans  $M$  endroits pseudo-aléatoires et sélectionne le meilleur des  $M$  endroits. Cependant, cet algorithme a besoin de l'image d'origine en niveaux de gris ou la binarisation inverse de l'image.

#### **IV. Conclusion**

Dans ce chapitre, une introduction générale au tatouage numérique a été présentée, avec les définitions de base et les principaux concepts, en décrivant les transformées les plus utilisées dans le tatouage numérique, ses principales caractéristiques et les domaines de son application.

Dans le prochain chapitre, deux techniques de tatouage numériques, notamment dans le domaine spatial et dans le domaine fréquentiel seront détaillées avec des tests et des résultats qui évalueront les performances de ces dernières.



# CHAPITRE II

## SYSTEME DE TATOUAGE PROPOSE

Il est encore plus facile de juger de l'esprit d'un  
homme par ses questions que par ses réponses.

G. de Lévis

**I. Introduction**

Dans ce chapitre nous aborderons les différents concepts utilisés pour la réalisation du système de sécurisation des documents imprimés proposé basé sur le tatouage numérique. Le système ainsi proposé assurera la protection des droits d'auteur des documents imprimés ou papiers par insertion d'une marque invisible qui peut être un logo ou une signature de propriétaire. Le processus d'insertion de la marque opère dans deux domaines différents à savoir, le domaine spatial où la marque est ajoutée aux données d'origines du document, et le domaine fréquentiel où la marque est ajoutée aux coefficients transformés par la Transformation en Cosinus Discrète (TCD).

**II. Conception du système de tatouage proposé**

Dans cette section, nous présenterons les différentes étapes suivies pour la conception et la réalisation de notre projet. Nous décrirons, en premier lieu, l'architecture globale du système. En suite, nous expliquerons les différentes étapes de la modélisation, qui est illustrée par quelques différents diagrammes que propose le langage UML (Unified Modeling Language). Vu sa simplicité, clarté et efficacité il a été utilisé pour modéliser notre système. En effet, UML définit un langage souple et adaptable, offrant une notation graphique standard pour la création de modèles des systèmes à concevoir. Il offre un ensemble de diagrammes, qui peuvent être utilisés dans le processus de modélisation. Dans notre conception, nous avons utilisé deux diagrammes qui sont : diagramme de cas d'utilisation et diagramme de séquence.

**II.1. Schéma générale du système de tatouage proposé**

L'architecture générale du système de tatouage proposé est divisée en cinq (05) modules à savoir le module d'acquisition du document imprimé afin de le numériser (rendre sous format numérique), le module d'embrouillage qui permet d'embrouiller la marque afin de la perturber en utilisant une clé, le module d'insertion de la marque dans le document, le module d'extraction de cette dernière et le module de désembrouillage en utilisant la même clé d'embrouillage afin d'obtenir la marque d'origine. Ils sont présentés dans la figure II.1.

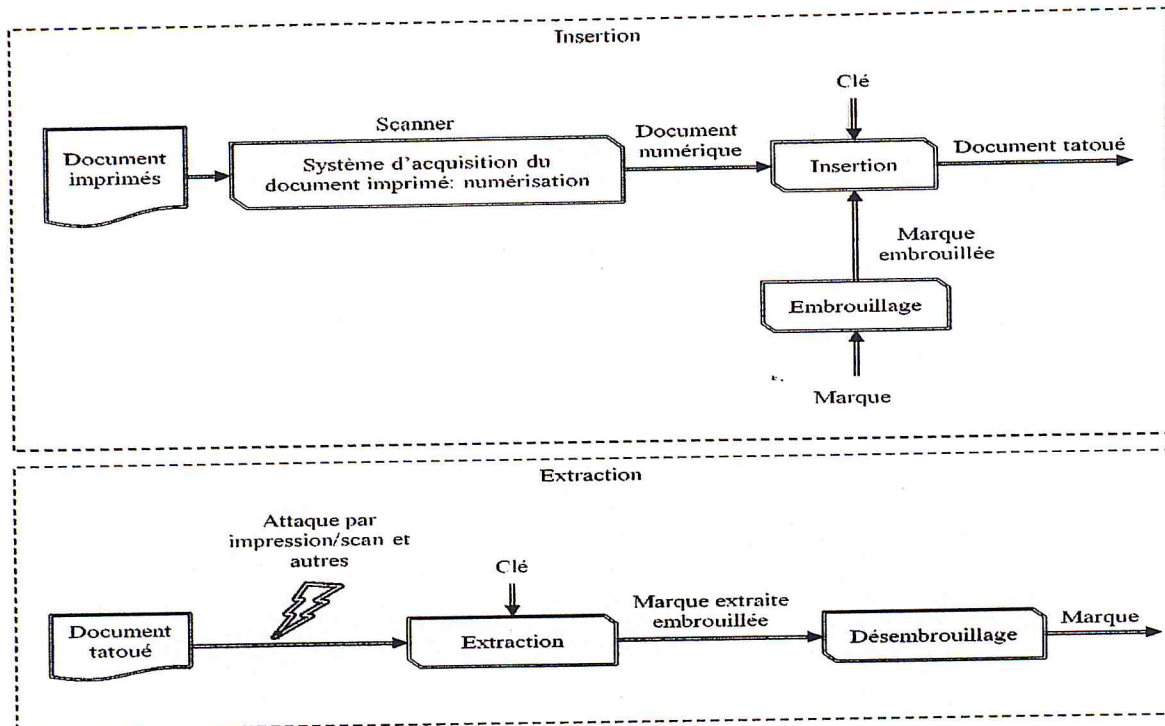


Figure II.1. Synoptique du système de tatouage proposé des documents imprimés.

## II.2. Diagramme de cas d'utilisation

Un diagramme de cas d'utilisation permet d'effectuer une bonne délimitation du système et aussi d'améliorer la compréhension de son fonctionnement. Il existe deux concepts fondamentaux dans ce type de diagramme :

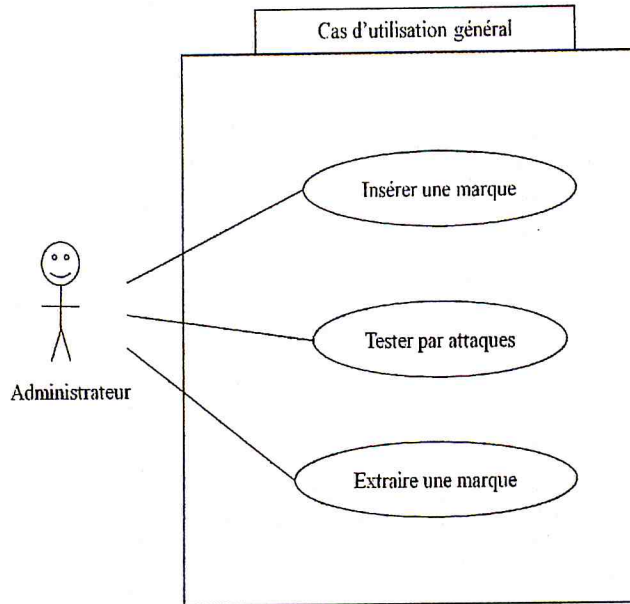
- Les acteurs : ce sont les utilisateurs du système (humain, matériels, logiciels).
- Les cas d'utilisation : représentent les actions des acteurs dans le système.

Dans le système proposé, nous avons un seul acteur qui est l'administrateur.

Nous déterminons dans ce qui suit les différents cas d'utilisation qui représentent le comportement du système. Dans ce cas, plusieurs cas d'utilisations sont énumérés, nous commençons par le cas d'utilisation générale.

### II.2.1. Cas d'utilisation général du système

Dans ce cas, l'administrateur exécute plusieurs opérations : l'insertion de la marque dans le document à protéger, tester la robustesse de cette dernière aux différentes attaques, et extraire la marque qui a été insérée pour vérifier les droits d'auteurs. La figure II.2 illustre le diagramme de cas d'utilisation générale du système.



**Figure II.2.** Diagramme de cas d'utilisation Général

**II.2.2. Cas d'utilisation « Processus d'insertion »**

Dans ce cas, l'administrateur insère la marque dans deux différents domaines, à savoir le domaine spatial, où le document est en niveau de gris, le domaine fréquentiel où le document est soit en couleur, soit en niveau de gris. Pour chaque domaine il ya deux choix de passages, dans le premier, deux méthodes de binarisation sont disponibles (Ordred et Floyd). Et pour passer au fréquentiel deux transformées sont disponibles, Transformée en Cosinus Discrète DCT 8x8, DCT 4x4. La figure II.3 illustre le diagramme de cas d'utilisation du processus d'insertion.

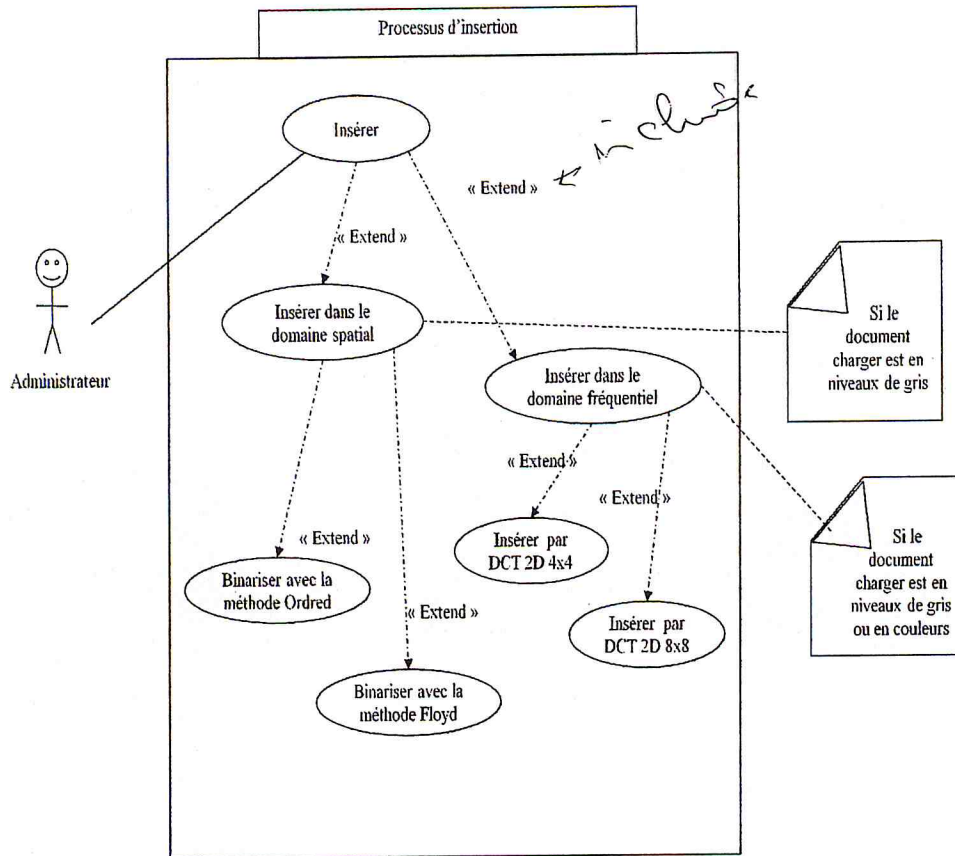


Figure II.3. Diagramme de cas d'utilisation « processus d'insertion »

II.2.3. Cas d'utilisation « Insertion dans le domaine spatial »

Dans ce cas d'utilisation, l'administrateur charge la marque à insérer, un processus d'embrouillage de cette dernière est effectué afin de la perturber et la rendre illisible. Pour ce fait l'administrateur entre la clé d'embrouillage par la suite il effectue l'insertion et il a le droit de calculer le PSNR du document tatoué, de l'enregistrer et aussi d'enregistrer la clé secrète afin de l'utilisée pour l'extraction. La figure II.4 illustre le diagramme de cas d'utilisation « Insertion dans le domaine spatial ».

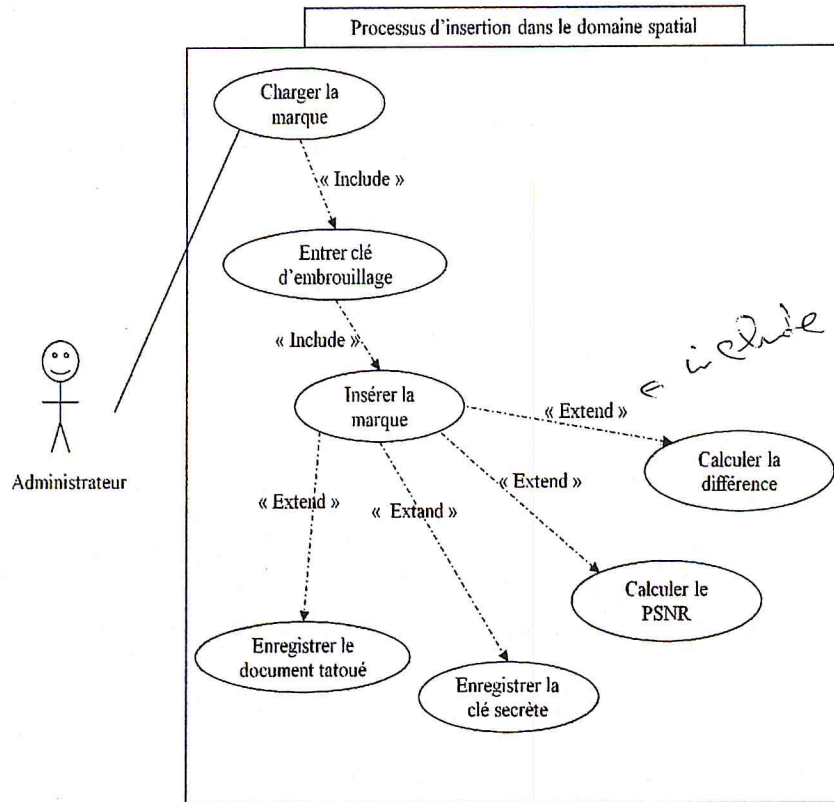


Figure II.4. Diagramme de cas d'utilisation « Insertion dans le domaine spatial »

II.2.4. Cas d'utilisation « Insertion dans le domaine fréquentiel »

Dans ce cas d'utilisation, l'administrateur charge la marque, il entre la clé d'embrouillage afin de rendre la marque illisible et l'embrouille. Ainsi il entre les facteurs d'insertion,  $k$  la position d'insertion dans un bloc de DCT et la force du marquage  $\alpha$  pour enfin effectuer l'insertion. La figure II.5 illustre le diagramme de cas d'utilisation « Insertion dans le domaine fréquentiel ».

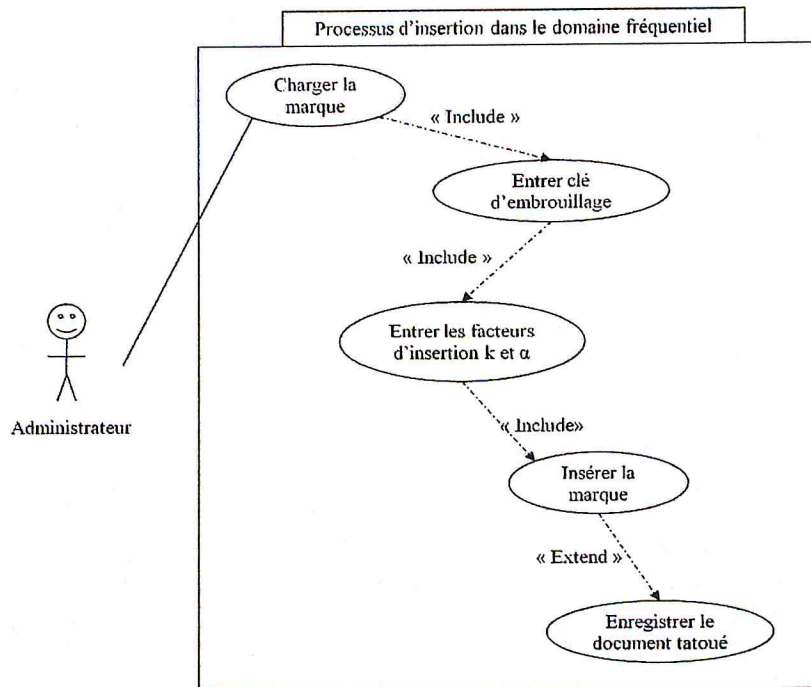


Figure II.5. Diagramme de cas d'utilisation « Insertion dans le domaine fréquentiel »

II.2.5. Cas d'utilisation « Extraction »

Dans ce cas d'utilisation, l'administrateur effectue cette opération selon le domaine d'insertion, il choisi le domaine (spatial ou fréquentiel) selon le type du document tatoué (en niveaux de gris ou en couleurs) et puis il effectue l'extraction. La figure II.6 illustre le diagramme de cas d'utilisation « Extraction ».

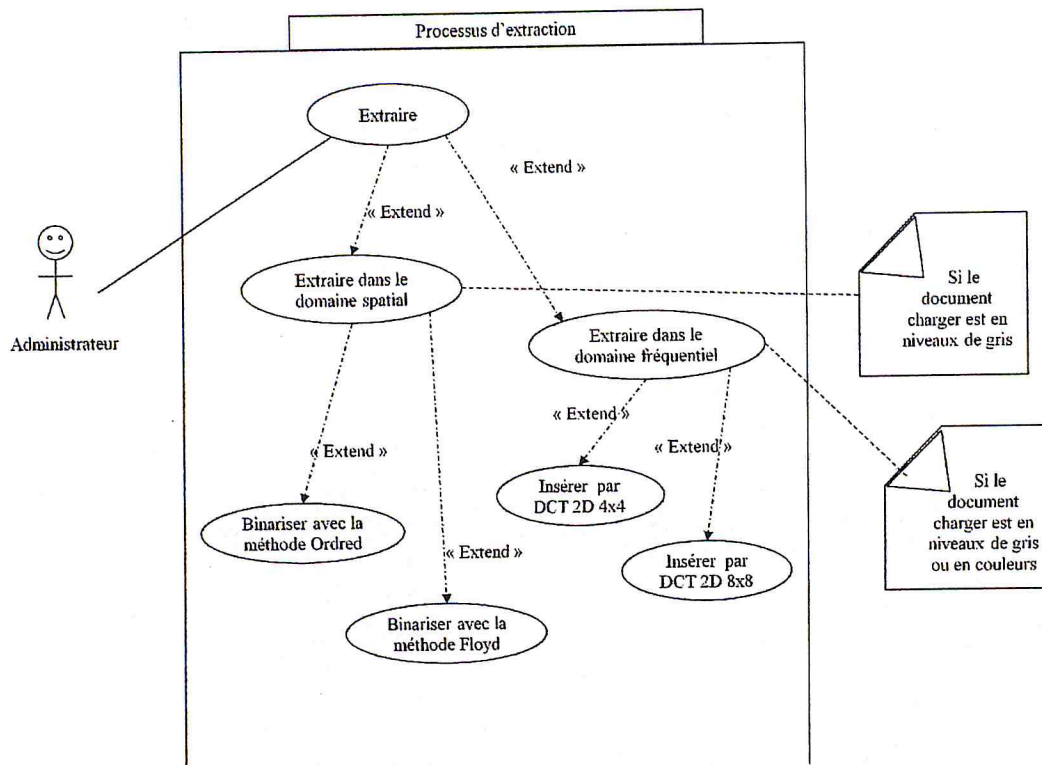


Figure II.6. Diagramme de cas d'utilisation « Extraction »

II.2.6. Cas d'utilisation « Extraction dans le domaine spatial »

Dans ce cas d'utilisation, l'administrateur commence à charger la clé secrète, il effectue l'extraction, puis il obtient la marque embrouillée et il entre la clé de désembrouillage afin de la rendre lisible. La figure II.7 illustre le diagramme de ce cas d'utilisation.



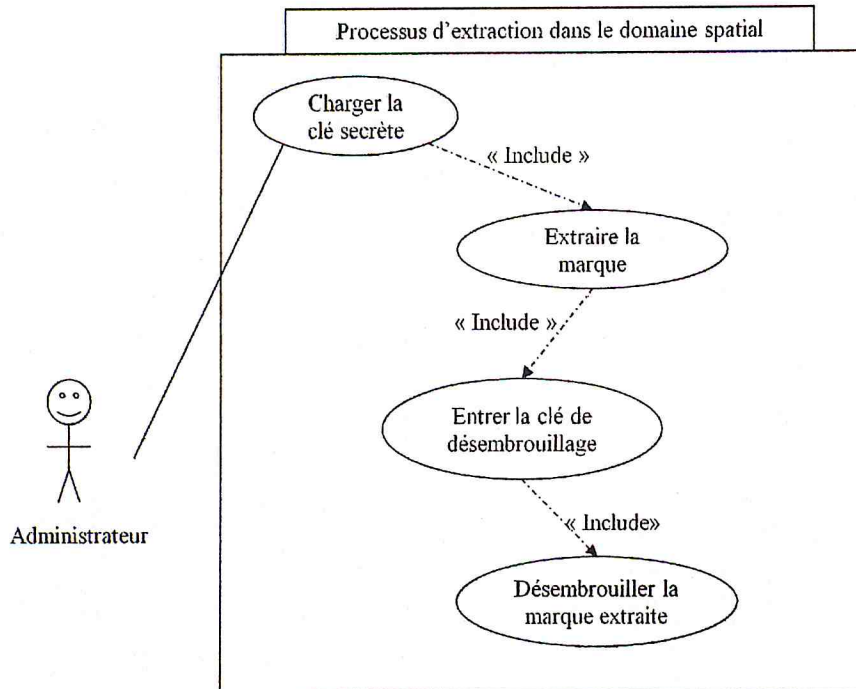


Figure II.7. Diagramme de cas d'utilisation « Extraction dans le domaine spatial »

II.2.7. Cas d'utilisation « Extraction dans le domaine fréquentiel »

Dans ce cas d'utilisation, l'administrateur entre la position K où la marque a été insérée, il extrait la marque embrouillée et à l'aide de la clé de désembrouillage il obtient la marque d'origine. La figure II.8 illustre le diagramme de cas d'utilisation « Extraction dans le domaine fréquentiel ».

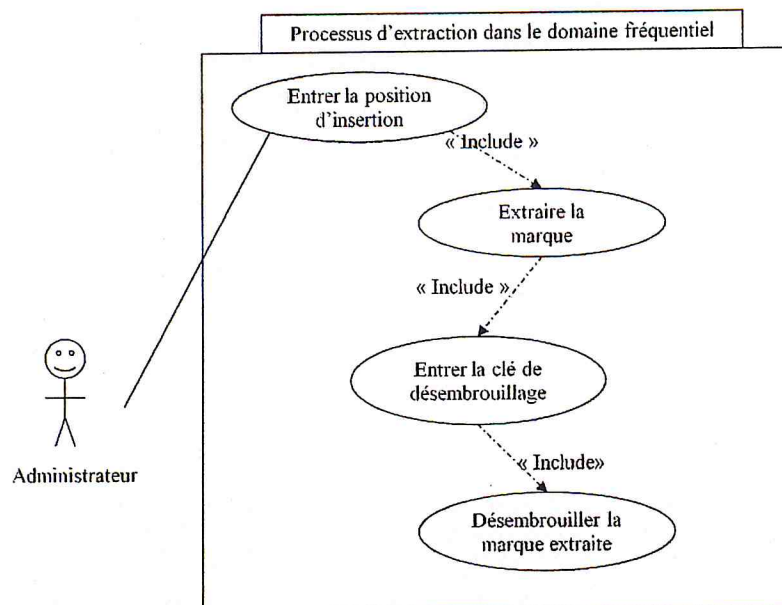


Figure II.8. Diagramme de cas d'utilisation « **Extraction dans le domaine fréquentiel** »

### II.3. Diagramme de séquence

Un diagramme de séquence est une description du diagramme de cas d'utilisation, un diagramme de séquence représente l'interaction entre les objets du système en respectant la chronologie, et la communication entre acteurs se fait par envois de messages.

#### II.3.1. Diagramme de séquence du cas d'utilisation « Insertion dans le domaine spatial »

L'opération « **Insertion dans le domaine spatial** », comporte les actions suivantes :

- L'administrateur charge le document à tatouer.
- Le système effectue la binarisation du document à tatouer.
- L'administrateur charge la marque à insérer.
- Le système effectue la binarisation de la marque à insérer.
- L'administrateur entre la clé d'embrouillage.
- Le système duplique la marque.
- Le système embrouille la marque.
- Le système génère une clé secrète.

- Le système génère une clé publique.
- Le système insère la marque.
- L'administrateur enregistre la clé secrète.
- L'administrateur enregistre le document tatoué.
- L'administrateur calcul le PSNR.
- L'administrateur calcul la différence entre le document tatoué et celui d'origine.

Cette chronologie est schématisée dans la Figure II.9 ci-dessous :

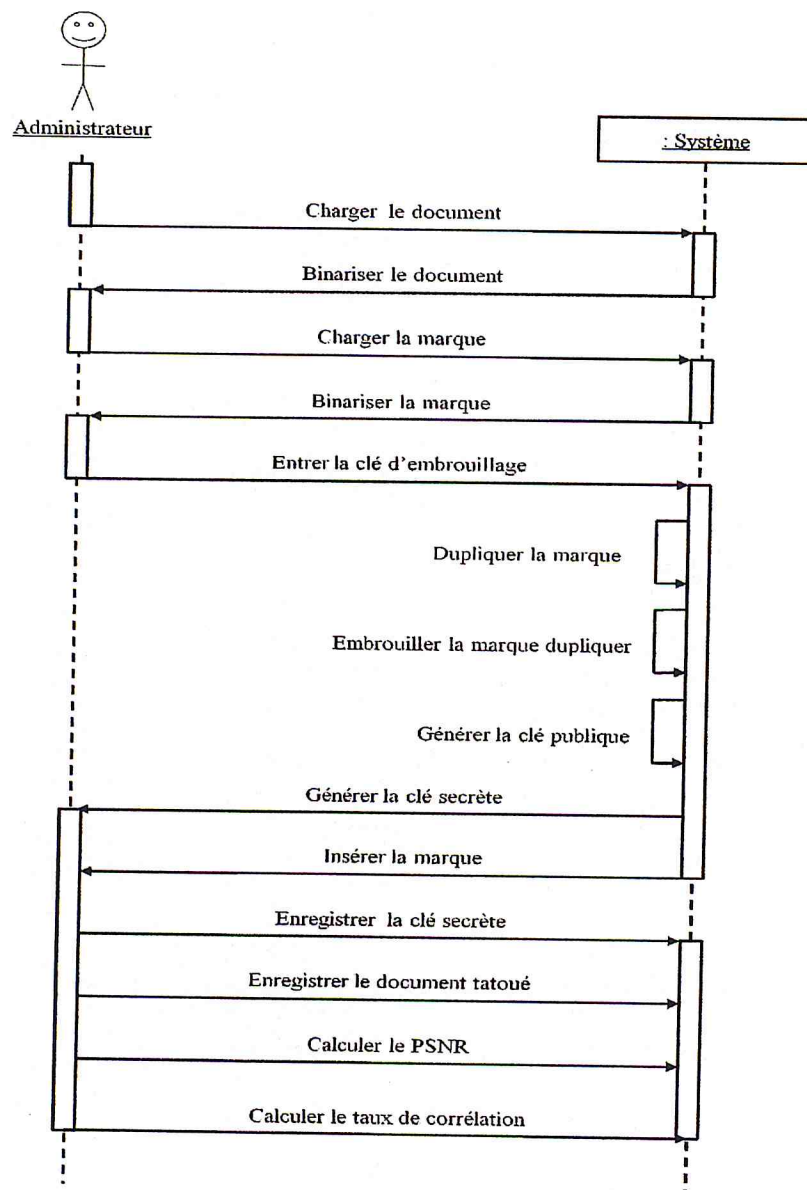


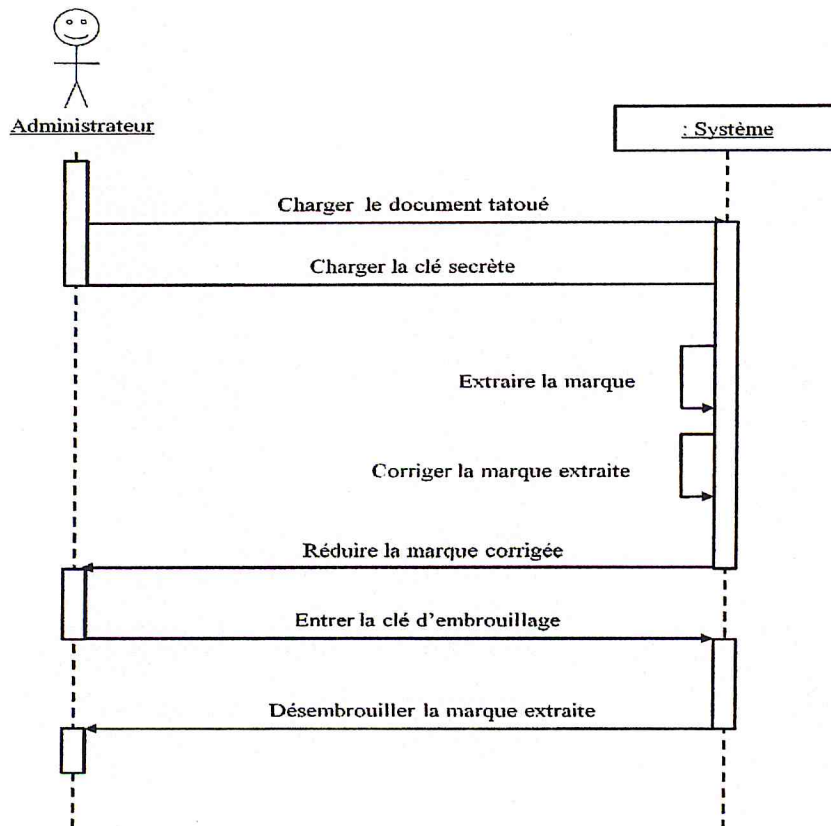
Figure II.9. Digramme de séquence du cas d'utilisation « Insertion dans le domaine spatial »

**II.3.2. Diagramme de séquence du cas d'utilisation « extraction dans le domaine spatial »**

L'opération «Extraction dans le domaine spatial», comporte les actions suivantes :

- L'administrateur charge le document tatoué.
- L'administrateur charge la clé secrète.
- Le système extrait la marque embrouillée.
- Le système corrige la marque embrouillée.
- Le système réduit la marque embrouillée.
- L'administrateur entre la clé d'embrouillage.
- Le système embrouille la marque.
- L'administrateur calcul le taux de corrélation de la marque extraite.

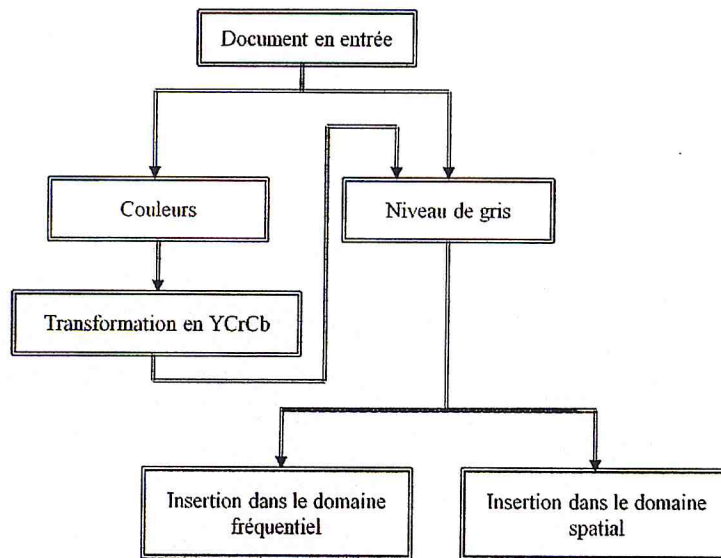
Cette chronologie est schématisée dans la figure II.10 ci-dessous :



**Figure II.10. Diagramme de séquence du cas d'utilisation « Extraction dans le domaine spatial »**

**III. Fonctionnement du système selon le type du document à tatoué**

Le système proposé prend en compte deux types de documents imprimés (figure II.11) : les documents en niveaux de gris et les documents couleurs.



**Figure II.11 :** Fonctionnement du système selon le type du document à tatoué

Dans le cas où les documents sont en couleurs généralement construits des trois plans correspondant aux composantes rouge, verte et bleue (*RVB*), une transformation vers les plans *YCrCb* est effectuée.

La composante *Y* correspond à l'information de la luminance [17]:

$$Y = 0.299 R + 0.587 G + 0.114 B \tag{II.1}$$

Et les composantes *Cr* et *Cb* aux informations de la chrominance :

$$Cr = - 0.14713 R - 0.28886 G + 0.436 B \tag{II.2}$$

Et 
$$Cb = 0.615 R - 0.5199 G - 0.10001 B \tag{II.3}$$

Seulement la composante *Y* est considérée pour l'insertion, vu que, l'œil humain est moins sensible aux modifications des informations de chrominance que la luminance. Ce phénomène est par ailleurs exploité en codage de séquence vidéo où les composantes de chrominance sont souvent sous-échantillonnées par rapport aux informations de luminance.

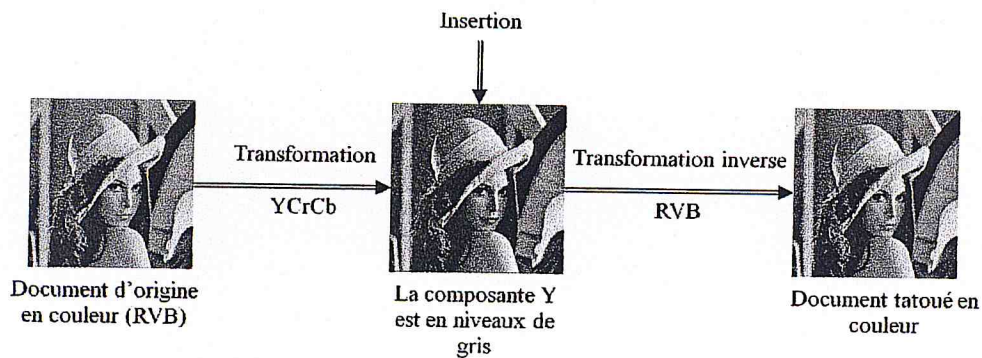
D'autre part, la détection de la signature est plus simple car l'énergie contenue dans les composantes de chrominance est bien plus faible que l'énergie contenue dans les composantes de luminance [18].

Après que le document soit tatoué, il est retransformé vers le plan RVB d'origine (figure II.12) selon les formules suivantes :

$$R = Y + 1.13983 Cb. \quad \text{II.4}$$

$$V = Y - 0.39465 Cr - 0.58060 Cb. \quad \text{II.5}$$

$$B = Y + 2.03211 Cr \quad \text{II.6}$$



**Figure II.12.** Illustration de la transformation de RVB vers le plan YCrCb et la transformation inverse.

## I. Description du processus d'insertion et d'extraction

### IV.1. Processus d'insertion dans le domaine spatial

La méthode de tatouage adoptée est basée sur l'algorithme publié dans [19]. L'insertion dans ce domaine comprend les étapes suivantes : tout d'abord, une transformation en demi-teinte du document d'origine et de la marque doit être effectuée, ensuite la marque est dupliquée selon la taille du document d'origine. La marque ainsi dupliquée est embrouillée afin d'augmenter la sécurité d'insertion. D'autre part, le document à protéger subit une subdivision en blocs de tailles 2x2 pixels pour une classification. A partir des différents types de blocs et des bits de la marque, deux clés publique et secrète sont respectivement générées.

Ces étapes sont parfaitement schématisés ci-dessous (figure II.13).

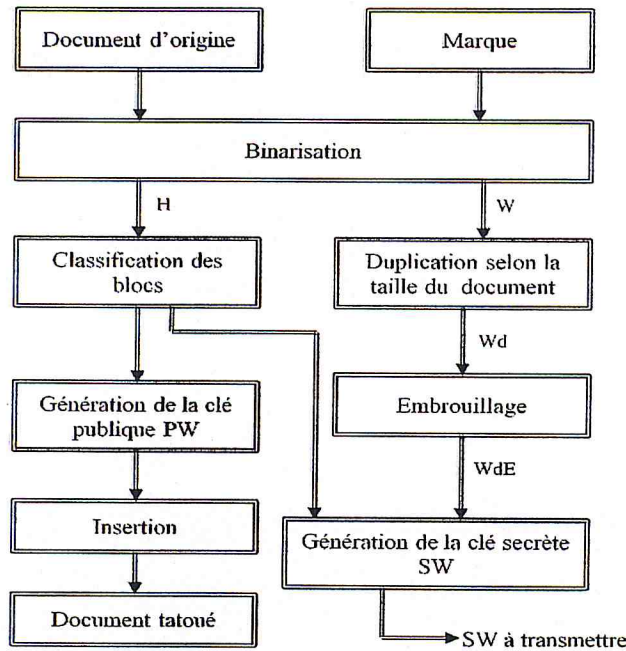


Figure II.13. Processus d'insertion dans le domaine spatial

Avec :

$H$  : document d'origine.

$W$  : la marque.

$Wd$  : la marque dupliquée.

$SW$  : la clé secrète.

$PW$  : la clé publique.

$WdE$  : la marque dupliquée embrouillée.

#### IV.1.1. La binarisation

La plupart des dispositifs de reproduction de documents, en particulier les périphériques d'impression, ont une capacité d'impression limitée à peu de couleurs, tandis que les documents imprimés sont principalement constitués de millions de couleurs [20]. Dans le cas de niveaux de gris, un document se compose de 256 niveaux de gris, tandis que les imprimantes noir et blanc utilisent un seul encres de couleur, c'est bien le noir. Ces 256 niveaux de gris devraient en quelque sorte être représentés par la couleur noir et blanc du substrat, pour cette raison le processus de binarisation est mis en œuvre.

La binarisation est un processus qui convertit un document en niveaux de gris à un document en demi-teinte tout en essayant de préserver son aspect visuel. Il existe différents algorithmes de binarisation adaptés aux différentes technologies d'impression et applications tels que *Patterning (Ordred)*, *Dithering*, *Diffusion d'erreur (Floyd-Steinberg)*, *Ordred* c'est l'une des meilleures méthodes pour transformer des documents du niveau de gris en demi-teinte, vu que les documents obtenus ont une résolution spatiale plus haute que les documents d'origine.

Nous avons opté pour les méthodes de binarisation qui sont : *Ordred* et *Floyd-Steinberg*.

#### IV.1.1.1. La méthode *Ordred*

Cette méthode consiste à générer un document en demi-teinte à partir du document d'origine qui a le même nombre de pixels que ce dernier [20]. Techniquement, chaque valeur de gris du document d'origine est représentée par une cellule en demi-teinte qui est d'origine représentée par une matrice de binarisation constituée d'un nombre variable de points noirs.

L'équation générale de la génération d'une matrice de binarisation pour une cellule demi-ton est donnée par :

$$B_3 = \begin{pmatrix} 4 \times B - \{n/2\} + 3 \times U - \{n/2\} & 4 \times B - \{n/2\} + 1 \times U - \{n/2\} \\ 4 \times B - \{n/2\} & 4 \times B - \{n/2\} + 2 \times U - \{n/2\} \end{pmatrix} \quad \text{II.7}$$

Où  $n \geq 2$ .

Différentes tailles de la matrices de binarisation génèrent différentes valeurs de niveaux de gris en fonction de la valeur de  $n$  (l'ordre de la matrice). Par exemple si  $n = k$  alors  $(k*k)+1$  différents niveaux de gris peuvent être générés.

En se basant sur les caractéristiques du système visuel humain (SVH) les matrices  $B_2$  et  $B_3$  qui représentent respectivement le second et le troisième ordre de binarisation sont suggérées comme suit :

$$B_2 = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{II.8}$$



$$B_3 = \begin{bmatrix} 7 & 2 & 6 \\ 4 & 0 & 2 \\ 3 & 8 & 5 \end{bmatrix} \quad \text{II.9}$$

En utilisant ces deux dernières matrices primitives, la matrice d'ordre supérieur est générée.

A partir de  $B_2$ , une matrice à cinq (05) niveaux de gris allant de 0 à 4 est générée. 0, 1, 2, 3 représentent quatre positions distinctes à remplir par des points noirs. Dans le cas où aucune position dans la matrice n'est remplie par un point noir, elle pourrait être utilisée pour représenter une région blanche absolue qui peut être défini comme un premier niveau de gris.

Lorsque la position 0 est remplie par un point noir alors le deuxième niveau de gris est généré. Lorsque les positions 0 et 1 sont toutes les deux remplies, alors le troisième niveau de gris est généré, etc.

Par conséquent, lorsque toutes les positions sont remplies le cinquième niveau de gris est généré. De la même façon, à partir de la matrice  $B_3$ , 10 niveaux de gris peuvent être générés. Le tableau II.1 représente la matrice utilisée pour générer 65 niveaux de gris en se basant sur les deux matrices  $B_2$  et  $B_4$ . N'importe quelle matrice de binarisation de n'importe quel ordre peut être générée à partir des matrices primitives telles que  $B_2$ ,  $B_3$ ,  $B_5$  etc.

**Tableau II.1.** Génération de la matrice  $B_8$

63	31	55	23	61	29	53	21
15	47	07	39	13	45	05	37
51	19	59	27	49	17	57	25
03	35	11	43	01	09	09	41
60	28	52	20	62	30	54	22
12	44	04	36	14	46	06	38
48	16	56	24	50	18	58	26
00	32	08	40	02	34	10	42

A partir de ce qui précède, un document généré en utilisant la méthode *Ordred* aura une taille égale à la taille du document d'origine multipliée par l'ordre de la matrice de binarisation manipulée. L'ordre de la matrice de binarisation dépend du nombre initial de niveaux de gris du document d'origine.

Pour un nombre plus grand de niveaux de gris, tel que 256 niveaux de gris, l'ordre de la matrice de binarisation nécessaire pour représenter un document en demi-teinte et égal à 16, donc la taille du document en demi-teinte est égal à 16 fois la taille du document d'origine. Cela va exiger un temps de calcul énorme, et afin de le réduire il faut diminuer la représentation de niveaux de gris du document d'origine.

Par exemple, nous pouvons considérer pour le document d'origine 64 niveaux de gris à la place de 256. Cela permet d'économiser considérablement l'espace mémoire et réduire le temps de calcul sans perte appréciable de qualité. Dans cette circonstance, la correspondance entre le niveau de gris (x) du document d'origine et sa représentation correspondante en demi-teinte (*Cellno*) est représentée par l'équation II.10 :

$$Cellno = ((256 - x) / (256 / (n \times n))) \quad \text{Tel que } 0 \leq x \leq 255 \quad \text{II.10}$$

**IV.1.1.2. La méthode de diffusion d'erreur (Floyd-Steinberg)**

Dans cette méthode, le niveau de gris de chaque pixel de l'image de départ est comparé à un seuil fixe ; s'il est inférieur à ce seuil, on met en noir le pixel qui lui correspond dans l'image résultat (noir et blanc), sinon il reste blanc. L'erreur de binarisation produite sur ce pixel est compensée par des erreurs en sens inverse sur les pixels voisins : on diffuse ainsi l'erreur sur quatre pixels voisins. La somme des quatre erreurs doit être égale exactement à l'erreur de binarisation, aucune erreur d'approximation ne doit être introduite [19][21].

Les coefficients déterminés par Floyd-Steinberg sont déterminés dans le tableau II.2 : 7 / 16 pour le pixel à droite du pixel courant, 5 / 16 pour le pixel en-dessous du pixel courant, 3 / 16 pour le pixel en-dessous et à gauche, 1 / 16 pour le pixel en-dessous et à droite.

**Tableau II.2. Filtre d'erreur de Floyd & Steinberg**

	<i>X</i>	7/16
3/16	5/16	1/16

Où *X* représente pixel central dans un voisinage d quatre pixels.

Le principe de l'approche est donné comme suit :

Pour  $i$  allant du haut vers le bas de la taille du document

**Faire**

Pour  $j$  allant de la gauche vers la droite de la taille du document

**Faire**

$$\text{Oldpix} = \text{Pix}[i,j] ;$$

$$\text{Newpix} = (\text{Oldpix} + 128) / 256 ;$$

$$Q_e = \text{Oldpix} - \text{Newpix} ;$$

$$\text{Pix}[i+1,j] = (\text{Pix}[i+1,j] + 7/16) * Q_e ;$$

$$\text{Pix}[i-1,j+1] = (\text{Pix}[i-1,j+1] + 3/16) * Q_e ;$$

$$\text{pix}[i,j+1] = \text{Pix}[i,j+1] + 5/16 * Q_e ;$$

$$\text{Pix}[i+1,j+1] = \text{Pix}[i+1,j+1] + 1/16 * Q_e ;$$

**Fait ;**

**Fait ;**

Où

*Oldpix* : ancien pixel du centre ;

*Newpix* : nouveau pixel du centre ;

*Q<sub>e</sub>* : erreur de quantification.

#### IV.1.2. Duplication de la marque

L'étape qui suit le processus de binarisation est la duplication de la marque selon la surface du document à protéger, pour cela le nombre de duplication de cette dernière ( $W$ ) est un facteur très important dans l'algorithme.

Si  $(N \times M)$  est la taille du document d'origine  $H$  et  $(n \times m)$  la taille de la marque  $W$ , alors le nombre de duplication de la marque  $k$  est donné par :

$$\text{Si } \text{mod}(N \times M, 4 * n * m) = 0$$

$$\text{alors } k = N * M / (4 * n * m) ;$$

$$\text{si non } k = \text{floor}(N * M / (4 * n * m)) + 1 ;$$

**finsi.**

Où *floor* permet d'obtenir la partie entière du résultat de la division.

Le résultat de la duplication de la marque selon la taille du document d'origine est représenté dans la figure II.14 ci-dessous :

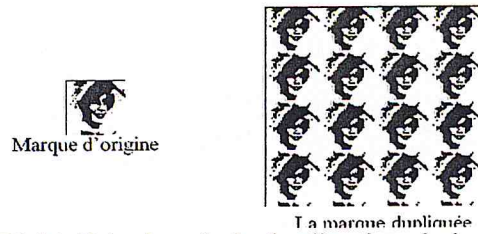


Figure II.14. Résultat de la duplication de la marque

IV.1.3. Embrouillage de la marque

Afin d'augmenter la sécurité d'insertion, la marque dupliquée est embrouillée avant d'être insérée.

Dans ce contexte, le principe utilisé est de modifier aléatoirement via un générateur congruenciel linéaire (GCL), LCG en anglais (Linear Congruential Generator), selon une clé les positions des bits de la marque, en effectuant des permutations dans les directions horizontales et verticales respectivement. Il s'agit de l'algorithme le plus utilisé pour produire des nombres aléatoires depuis qu'il a été inventé en 1948 par D. H. Lehmer. Il est déterminé par quatre valeurs entières définies dans le tableau II.3 ci-dessous [22][23].

Tableau II.3. Les quatre paramètres utilisés pour le calcul du GCL

Paramètre	signification	valeurs
$M$	<i>mod (reste entier)</i>	$m > 0$
$M$	<i>le multiplicateur</i>	$0 \leq a < m$
$Cr$	<i>l'incrément</i>	$0 \leq c < m$
$G(i-1)$	<i>la valeur initiale</i>	$0 \leq X0 < m$

La séquence est donnée par :

$$G(i) = (m * G(i - 1) + Cr) \bmod M \tag{II.11}$$

Avec :

- $m$  et  $Cr$  sont des valeurs secrètes et peuvent être changées suivants les résultats des tests.
- $mod$  : le reste de la division entière (le reste entier de la valeur entre parenthèses divisé par  $M$ ). Ces générateurs sont désignés par GCL ( $m, Cr, M, G(0)$ ).
- La formule est simple mais le choix des trois paramètres  $m, Cr$  et  $M$  ne doit pas être fait à la légère. Ils sont choisis afin de maximiser la période qui ne peut excéder  $M$  et

qui est égal à la dimension de la marque selon la taille verticale ou l'horizontale.  $G(0)$  représente la clé d'insertion et d'extraction de la marque.

Un exemple du résultat de la procédure d'embrouillage est illustré sur la figure II.15:

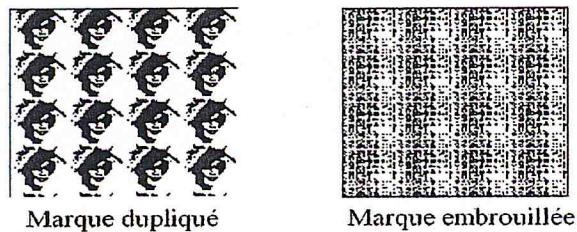


Figure II.15. Résultat de l'embrouillage de la marque dupliquée

Après l'embrouillage de la marque dupliquée selon la taille du document, nous procédons à la classification des blocs et la génération des deux clés publique et secrète.

































#### IV.1.4. Classification de blocs

Le processus de classification joue un rôle très important dans l'algorithme. Il est basé sur le codage de troncature des blocs [19]. Le but de cette approche est de classifier les blocs du document en blocs de texture et blocs uniformes, et par la suite sélectionner les blocs de textures pour l'insertion de la marque.

A partir du processus de classification des blocs, deux clés sont générées à savoir une clé publique qui est caché dans le document en demi teinte  $H$  et une clé secrète qui est transmise avec le document tatoué. La génération de la clé publique est dépendante des blocs du document d'origine et la génération de la clé secrète est dépendante des blocs du document d'origine et aussi des bits de la marque à insérer. Pour cela le document d'origine subit une subdivision en blocs de taille  $2 \times 2$ .

Les blocs sont classifiés suivant le tableau II.4 : Les blocs homogènes sont les blocs appartenant aux types de blocs  $BT(0)$  à  $BT(3)$ , alors que les blocs texturés sont les blocs appartenant aux autres blocs de  $BT(4)$  à  $BT(9)$ . Les étapes de la classification du document d'origine  $H$  et de la génération des deux clés  $PW$  et  $SW$  sont présentées dans l'organigramme de la figure II.16. Le resultat de la génération des deux clés publique et secrète est présenté dans la figure II.17 ci-dessous :

Tableau II.4. Table de classification de blocs

Type de bloc d'entrée	Caractéristiques de bloc d'entrée	Bloc de sortie de la clé publique	Bloc de sortie de la clé secrète lorsque le bit de la marque est blanc	Bloc de sortie de la clé secrète lorsque le bit de la marque est noir
BT(0)	0 Pixel blanc 4 Pixels noirs (homogène)	 CT(0)	 CT(0)	choisir au hasard à partir CT(1), CT(2), CT(3)
BT(1)	1 Pixel blanc 3 Pixels noirs (homogène)	 CT(1)	 CT(1)	choisir au hasard à partir CT(0), CT(2), CT(3)
BT(2)	3 Pixels blancs 1 Pixel noir (homogène)	 CT(2)	 CT(2)	choisir au hasard à partir CT(0), CT(1), CT(3)
BT(3)	4 Pixels blancs 0 Pixel noir (homogène)	 CT(3)	 CT(3)	choisir au hasard à partir CT(0), CT(1), CT(2)
BT(4)	 Texture	 CT(4)	 CT(4)	 CT(9)
BT(5)	 Texture	 CT(5)	 CT(5)	 CT(8)
BT(6)	 Texture	 CT(6)	 CT(6)	 CT(7)
BT(7)	 Texture	 CT(7)	 CT(7)	 CT(6)
BT(8)	 Texture	 CT(8)	 CT(8)	 CT(5)
BT(9)	 Texture	 CT(9)	 CT(9)	 CT(4)

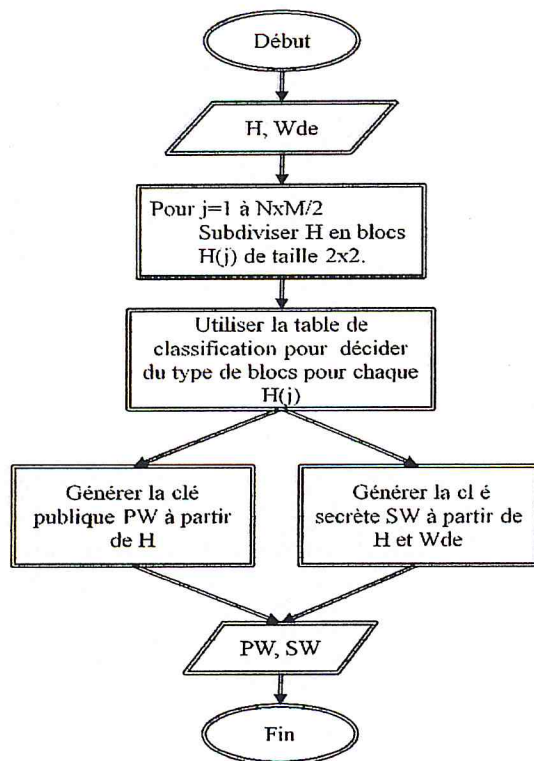


Figure II.16. Organigramme de la classification et la génération des clés

Avec :

*H* : le document d'origine.

*Wde* : la marque dupliquée et embrouillée.

*H(j)* : le bloc (j) du document d'origine.

*PW* : la clé publique.

*SW* : la clé secrète.

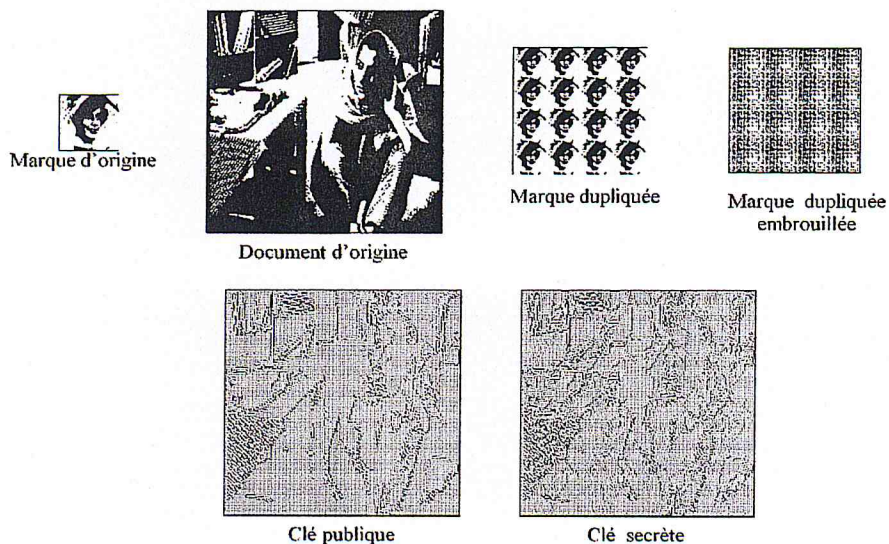


Figure II.17. Résultat de la génération des clés publique et secrète.

IV.1.5. Insertion de la marque

L'insertion est effectuée uniquement au niveau des blocs texturés appartenant au blocs BT(4) à BT(9). Les différents étapes d'insertion de la marque dans le document d'origine sont représentées dans l'organigramme de la figure II.18 ci-dessous :

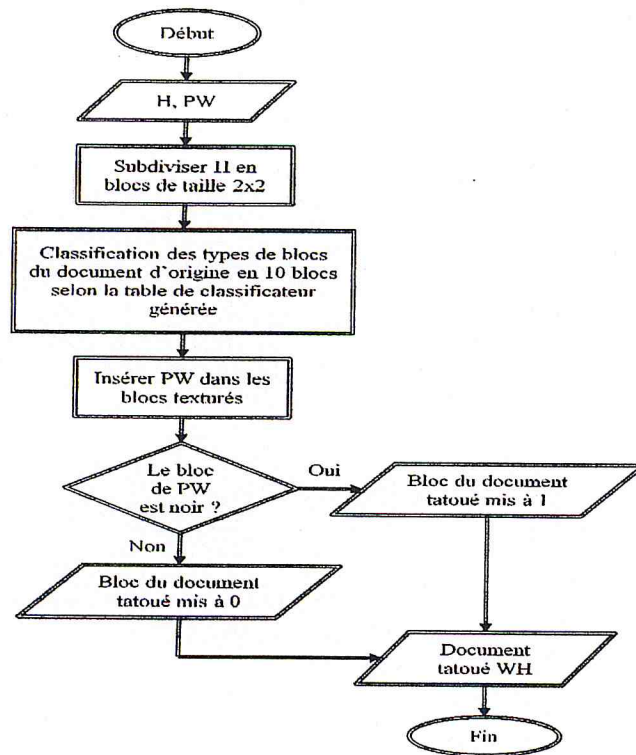


Figure II.18. Organigramme d'insertion de la clé publique

Avec :

*H* : le document d'origine.

*WP* : la clé publique.

*WH* : le document tatoué.

Le résultat d'insertion de la marque dans le document d'origine est représenté dans la figure II.19 ci-dessous :



Document d'origine  
En demi-teinte



Document tatoué

Figure II.19. Résultat du processus d'insertion



## IV.2. Processus d'extraction dans le domaine spatial

Le processus d'extraction est le processus inverse de l'insertion, il s'agit d'extraire la marque insérée dans le document tatoué afin de vérifier le droit d'auteur. Dans le système proposé l'opération d'extraction nécessite la présence du document tatoué et la clé secrète. A partir du document tatoué nous extrayons la clé publique insérée après avoir classifié les blocs de ce dernier. Puis une porte NXOR est appliquée entre la clé publique extraite et la clé secrète reçue afin d'extraire la marque brouillée.

Les différentes étapes d'extraction sont parfaitement schématisées dans la figure II.20 ci-dessous :

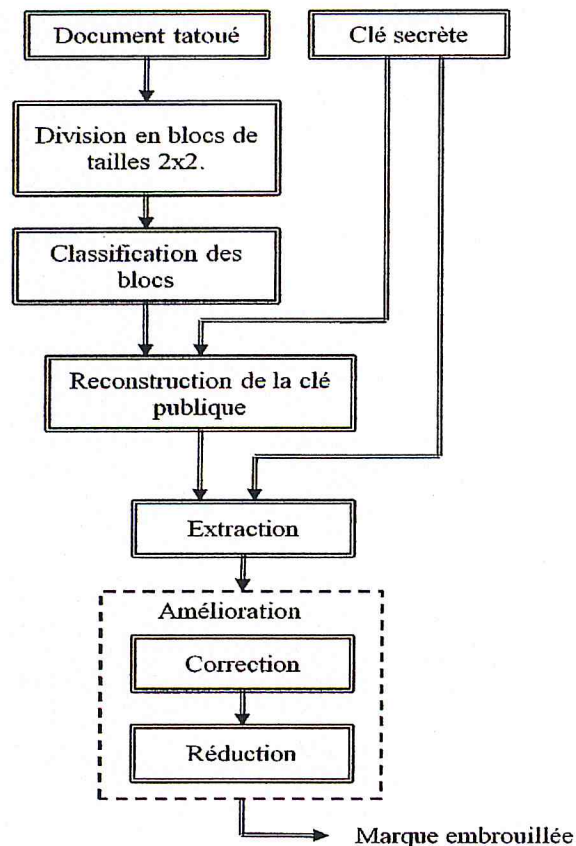
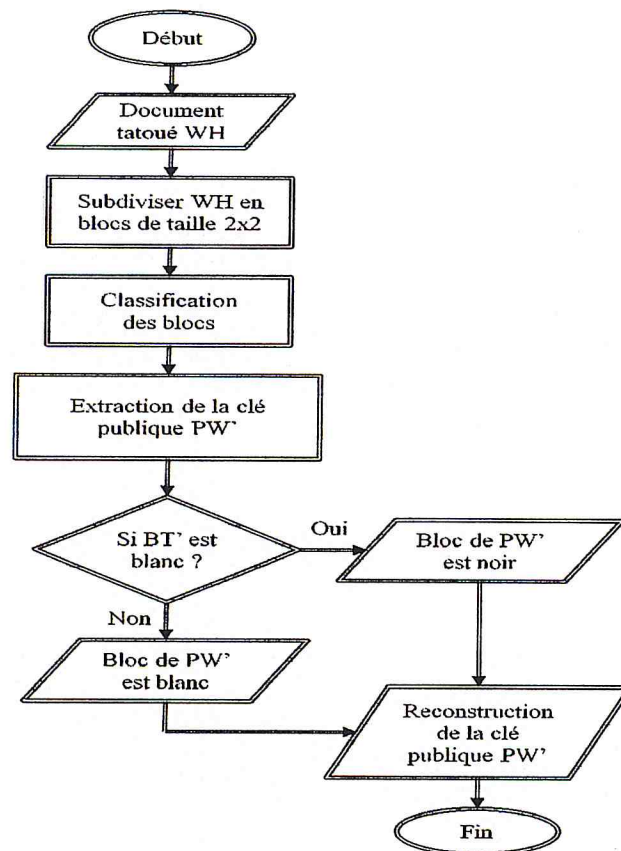


Figure II.20. Processus d'extraction dans le domaine spatial

Une fois la marque est extraite, le processus d'amélioration de la qualité de cette dernière est effectué.

### IV.2.1. Classification des blocs du document tatoué et reconstruction de la clé publique

Les différentes étapes pour reconstruire la clé publique  $PW'$  sont présentées dans l'organigramme de la figure II.21 ci-dessous :



**Figure II.21.** Organigramme des différentes étapes pour la reconstruction de la clé publique

Avec :

$WH$  : le document tatoué.

$BT'$  : le bloc du document tatoué.

$PW'$  : la clé publique reconstruite.

### IV.2.2. Extraction de la marque insérée

L'extraction de la marque insérée est faite en appliquant la porte NXOR entre la clé secrète et la clé publique qui a été reconstruite à partir du document tatoué. L'organigramme dans la figure II.22 exprime clairement cette étape :

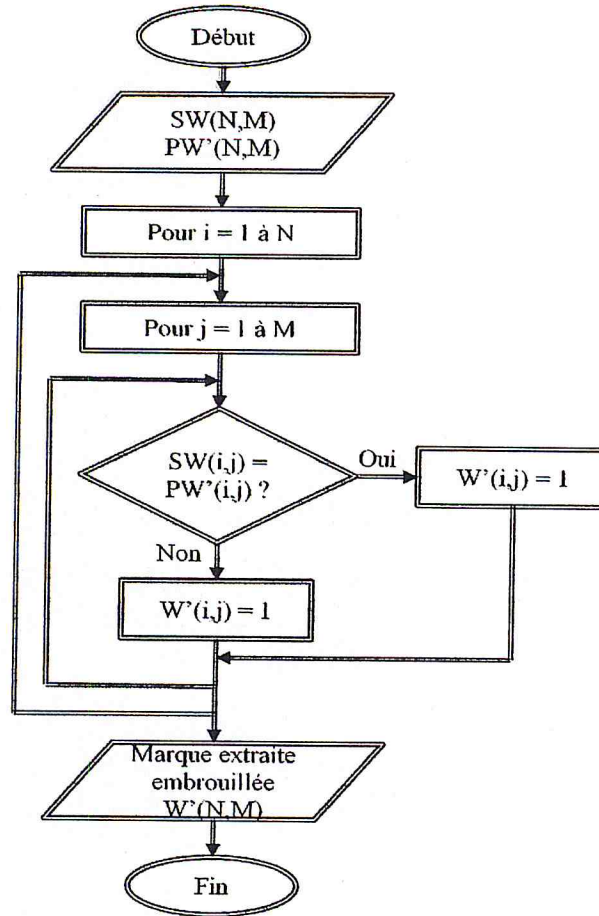


Figure II.22. Organigramme d'extraction de la marque.

Avec :

$SW(N,M)$  : la clé secrète de taille  $(N \times M)$ .

$PW'(N,M)$  : la clé publique de taille  $(N \times M)$ .

$W'(N,M)$  : la marque extraite avant l'amélioration de taille  $(N \times M)$ .

#### IV.2.3. Amélioration de la qualité de la marque

Afin d'améliorer la qualité de la marque extraite, nous utilisons un processus appelé « **correction** » afin de corriger les détériorations introduites dans le document tatoué. Pour récupérer la taille d'origine de la marque insérée, nous utilisons une opération de prétraitement appelé « **réduction** » qui permet de réduire la redondance des données introduites par l'insertion. Les opérations de « **Correction et Réduction** » sont basées sur le tableau II.5. Ce tableau montre que les 16 motifs possibles d'une image tatouée sont divisés en deux cas  $NB > 1$  et  $NB \leq 1$  (où  $NB$  représente le nombre de pixels noirs) avec quatre pixels pour chaque motif.

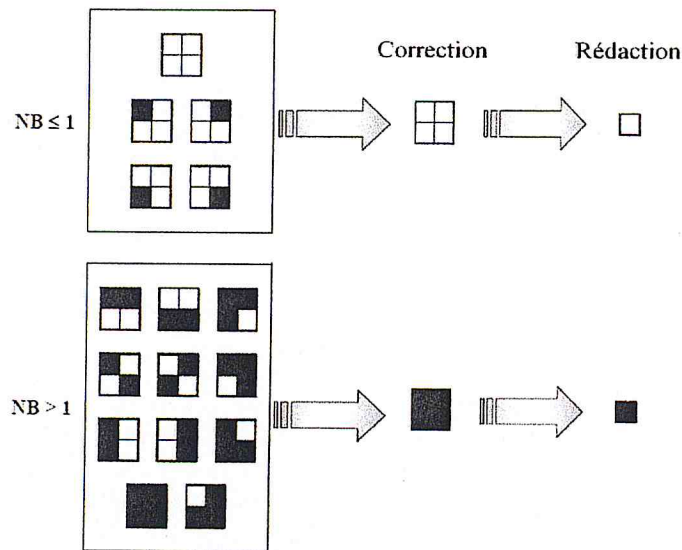


Tableau II.5. Table de référence pour les opérations de correction et réduction de la marque

Où NB représente le nombre de pixels noirs dans un bloc.

La marque extraite est embrouillée avec une clé, afin d'obtenir la marque insérée au début dans le document à protégé. Une opération inverse d'embrouillage est effectuée. Nous appliquant cette dernière (opération de désembrouillage avec 'utilisation de la même clé, celle d'embrouillage).

Les résultats d'extraction de la marque résultantes des opérations de correction et de réduction sont présentés dans la figure II.23 ci-dessous :

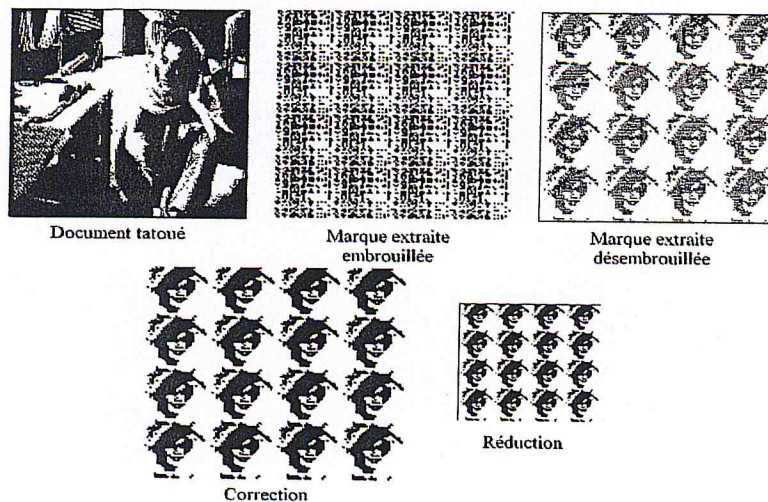


Figure II.23. Résultat des opérations de correction et réduction.

## IV.3. Analyse expérimentale des résultats

Dans cette section, nous avons effectué des tests sur notre système de protection des droits d'auteurs des documents imprimés, nous allons présenter quelques résultats obtenus lors d'une série de tests sur l'image Barbara de taille 256x256 pixels codés par 8 bits par pixel.

Les performances de la méthode développée sont évaluées en termes de qualité document en se basant sur le calcul du PSNR (*Peak Signal to Noise Ratio*) et d'intégrité de la marque (taux d'extraction).

Le tatouage numérique engendre une distorsion entre le document d'origine et le document tatoué. La mesure de cette distorsion est fondée sur la différence qu'il y a entre ces deux documents, c'est l'erreur quadratique moyenne MSE (*Mesan Squar error*) [11].

$$MSE = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [I(i,j) - I'(i,j)]^2 \quad \text{II.12}$$

$I'(m,n)$  : Pixel (m,n) du document.

$I(m,n)$  : Pixel (m,n) du document d'origine.

$M, N$  : dimensions de l'image.

Le MSE permet de calculer le PSNR du document en décibels (dB) qui est le critère le plus usuel pour l'évaluation de la qualité donné par l'équation II.13. nous obtenons un rapport signal sur bruit crête dont le maximum est  $2^R - 1$ , où R est le nombre de bit codant un pixel de l'image en demi-teinte. Lorsque les deux documents sont identiques, la MSE est nulle et le PSNR tend vers l'infini.

$$PSNR = 10 \log_{10} \left( \frac{(2^R - 1)^2}{MSE} \right) \text{ dB} \quad \text{II.13}$$

Un autre critère important qu'il faut prendre en compte consiste en la ressemblance entre la marque insérée et la marque extraite. Cette dernière est exprimée par  $\sigma$  le facteur de corrélation donné par [19] :

$$\sigma = 100 - \left( \frac{\sum_{i=1}^n \sum_{j=1}^m W(i,j) \text{ XOR } W'(i,j)}{n \times m} \right) \times 100 \quad \text{II.14}$$

Où

$W(i,j)$  : le pixel (i,j) de la marque inseree.

$W'(i,j)$  : le pixel (i,j) de la marque extraite.

$n,m$  : les dimensions de la marque.

Nous avons effectue des tests sur des documents image de reference Barbara de taille 256x256 pixels et code sur 8 bits par pixels et Femme de taille 256x256 pixels et code sur 8 bits par pixels, respectivement. La marque a inserer consiste en une petite image binaire de reference femme de taille 32x32 pixels.

Les tableaux II.6 et II.7 representent les resultats du tatouage effectue sur les images *Barabra* et *Femme* en se basant sur les deux techniques de binarisation *Ordred* et *Floyd*. Les resultats obtenus sont exprimes en fonction du PSNR, taux de correlation  $\sigma$ , et la difference entre le document d'origine et le document tatoue.

Tableau II.6. Table des resultats obtenus dans le domaine spatial (Barbara 256x256)



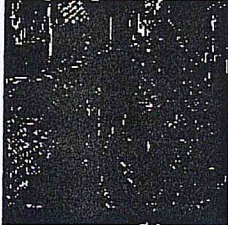


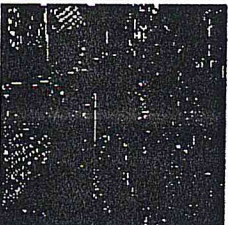


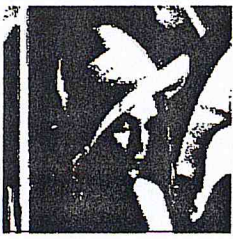
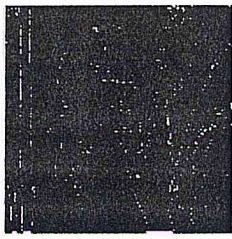


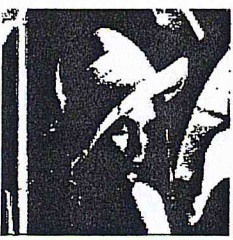
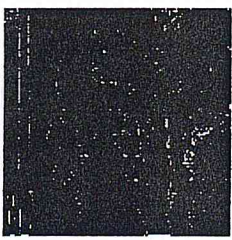

Document d'origine	Document tatoue	La difference	Marque extraite	PSNR (dB)	$\sigma$ (%)
				59,39	92,5 2
	Methode Floyd				
				59,47	92,6 5
	Methode Ordred				

Tableau II.7. Table des resultats obtenus dans le domaine spatial (Femme 256x256)

Document d'origine	Document tatoué	La difference	Marque extraite	PSNR (dB)	$\sigma$ (%)
				62,32	96,20
	Methode Floyd				
				62,13	95,85
	Methode Ordred				

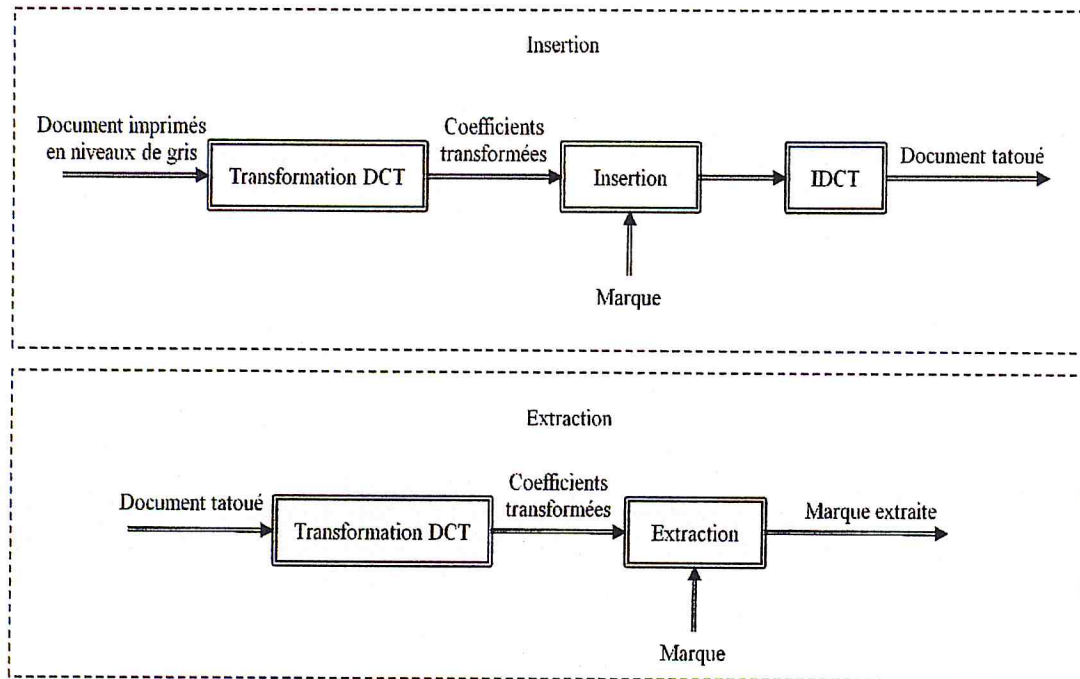
A partir des tableaux ci-dessus, nous remarquons que les résultats obtenus par les deux méthodes sont pratiquement identiques au point de vue qualité de l'image (PSNR) et aussi par taux de ressemblance de la marque extraite et la marque insérée.

La différence entre le document d'origine et le document tatoué montre que les modifications sont apportées au niveau des contours. Mais malgré la présence de la marque, aucune dégradation n'est vraiment perceptible sur le document tatoué par les deux méthodes

**II. Insertion dans le domaine fréquentiel**

La littérature sur le tatouage des documents imprimés n'est pas riche. Les quelques travaux publiés récemment utilisent la version en demi-teinte du document. Dans notre travail, la deuxième alternative de la sécurisation des documents imprimés est l'insertion de la marque dans le domaine fréquentiel. En effet, même si le tatouage dans le domaine fréquentiel est généralement destiné pour sécuriser les documents numériques, ça ne nous empêche pas de l'utiliser pour la sécurisation des documents imprimés.

La synoptique générale du tatouage numérique dans le domaine fréquentiel est illustrée sur la figure II.24 ci-dessous :



**Figure II.24.** Synoptique générale du tatouage numérique dans le domaine fréquentiel.

### V.1. Processus d'insertion

Dans cette partie du travail, l'information qui est constituée par un logo est insérée dans la version transformée (DCT) du document imprimé.

L'insertion dans le domaine fréquentiel s'effectue selon les étapes suivantes :

1. Segmentation du document d'origine en blocs de taille 8x8.
2. Application de la DCT à 2D sur chaque bloc 8x8.
3. L'insertion d'un bit de la marque dans chaque bloc de 8x8 dans les coefficients appartenant à la première diagonale du bloc.

Ces étapes sont parfaitement schématisés ci-dessous (figure II.25).



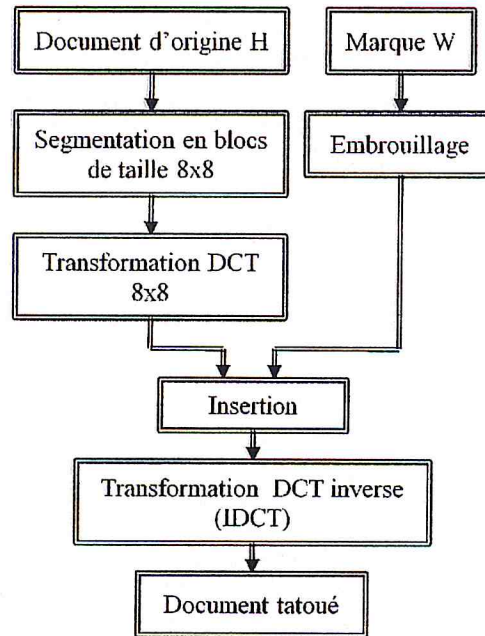


Figure II.25. Processus d'insertion dans le domaine fréquentiel.

### V.1.1. La segmentation du document en blocs de taille 8x8

Pour réaliser la DCT, il est nécessaire de découper l'image en blocs de taille  $8 \times 8$ . Pour cela, les dimensions du document d'origine et de la marque à insérer doivent être adéquates, donc chacune des dimensions verticale et horizontale du document d'origine doit être divisible sur 8, et le nombre de blocs obtenus doit être divisible sur la dimension correspondante de la marque à insérer.

### V.1.2. Application de la DCT 2D

La matrice transformée par DCT présente la propriété de regrouper les valeurs les plus élevées dans le coin supérieur gauche de la matrice (les valeurs devenant d'autant plus faibles que l'on s'approche du coin inférieur droit). Ainsi le maximum d'information sur l'image se trouve concentré sur la partie supérieure gauche de la matrice et le traitement du document sera facilité. La DCT est donc effectuée sur chaque matrice  $8 \times 8$  pixels, et elle donne une matrice  $8 \times 8$  de coefficients significatifs de fréquences spatiales : l'élément (0,0) représente la valeur moyenne du bloc (appelé coefficient DC), les autres (coefficients AC) fournissent la puissance spectrale pour chaque fréquence spatiale [SW3].

Cette transformation est généralement donnée par la formule suivante :

$$DCT(i, j) = \frac{1}{\sqrt{2}} C(i)C(j) \sum_{x=0}^7 \sum_{y=0}^7 P(x, y) \cos\left(\frac{(2x+1)i\pi}{16}\right) \cos\left(\frac{(2y+1)j\pi}{16}\right) \quad \text{II.15}$$

Avec

$$C(x) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } x = 0 \\ 1 & x > 0 \end{cases}$$

et  $P$  représente la valeur du pixel.

L'implémentation de la DCT 2D sur des blocs 8x8 est réalisée en calculant d'abord la DCT à 1D sur la direction horizontale puis une 2<sup>ème</sup> transformée à 1D sur la verticale ou l'inverse (figure II.26).

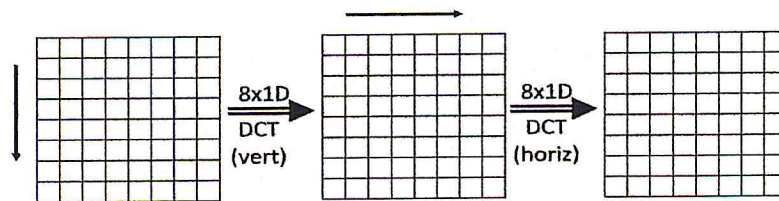


Figure II.26. Implémentation de la DCT 8x8

Pour l'implémentation de cette opération, nous avons utilisé une méthode de calcul rapide. Les équations utilisées sont :

$$\begin{pmatrix} X_0 \\ X_2 \\ X_4 \\ X_6 \end{pmatrix} = \begin{bmatrix} C_4 & C_4 & C_4 & C_4 \\ C_2 & C_6 & -C_6 & -C_2 \\ C_4 & -C_4 & -C_4 & C_4 \\ C_6 & -C_2 & C_2 & -C_6 \end{bmatrix} \begin{pmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{pmatrix}$$

$$\begin{pmatrix} X_1 \\ X_3 \\ X_5 \\ X_7 \end{pmatrix} = \begin{bmatrix} C_1 & C_3 & C_5 & C_7 \\ C_3 & -C_7 & -C_1 & -C_5 \\ C_5 & -C_1 & C_7 & C_3 \\ C_7 & -C_5 & C_3 & -C_1 \end{bmatrix} \begin{pmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{pmatrix} \quad \text{II.16}$$

Où  $X_i$  représente la valeur du pixel du document d'origine.

### V.1.3. L'insertion de la marque [20]

Comme pour la première méthode, la marque est embrouillée avant qu'elle ne soit insérée, afin d'assurer l'invisibilité de la marque, l'insertion est effectuée dans les coefficients de moyenne fréquence. Le principe général de l'algorithme consiste à insérer un bit de la marque  $W(nxm)$  ( $n=N/8$  et  $m=M/8$ ) dans chaque bloc de DCT 8x8 du document  $H(NxM)$ , dans une composante  $k$  choisit à partir des coefficients appartenant à la première diagonale du bloc. En prenant en considération la valeur  $\alpha$  qui

représente la force de marquage. On peut résumer l'opération d'insertion de la marque par l'organigramme de la figure II.27 suivant :

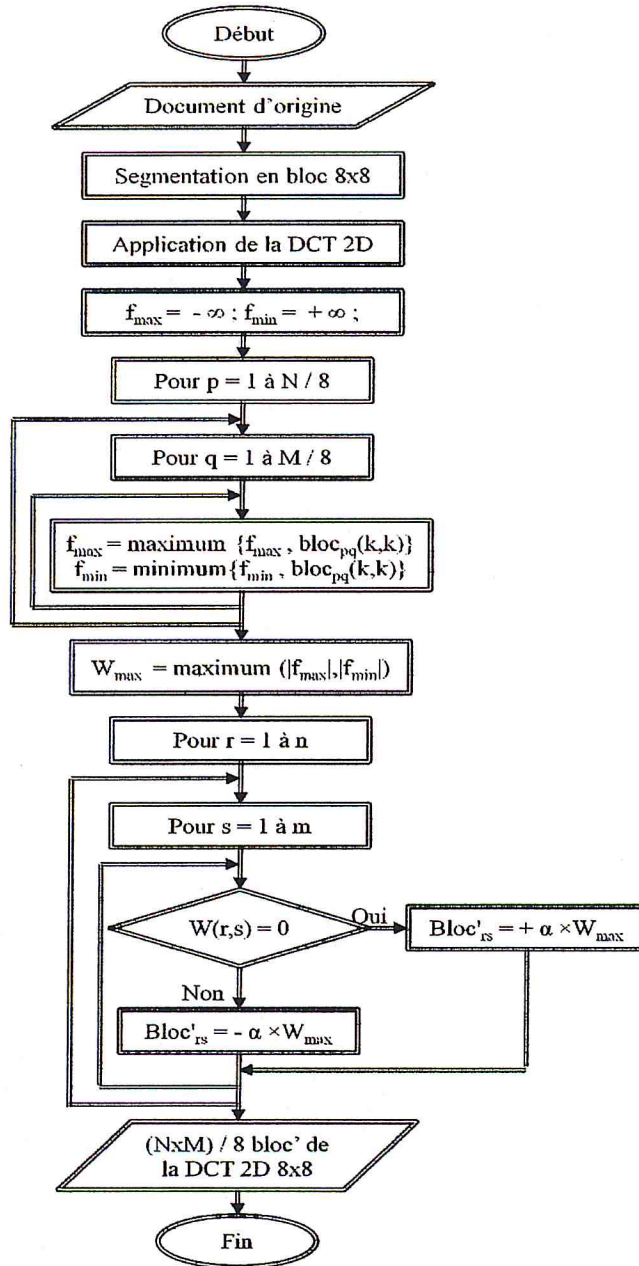


Figure II.27. Organigramme de l'opération d'insertion dans le domaine fréquentiel

Le document imprimé tatoué est reconstruit en utilisant la DCT inverse (IDCT) donnée par la formule ci-dessous :

$$P(x, y) = \frac{1}{4} \sum_{i=0}^7 \sum_{j=0}^7 C(i)C(j) DCT(i, j) \cos\left(\frac{(2x+1)i\pi}{16}\right) \cos\left(\frac{(2y+1)j\pi}{16}\right) \tag{II.17}$$

Avec

$$C(x) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } x = 0 \\ 1 & x > 0 \end{cases}$$

et  $P$  représente la valeur du pixel.

### V.2. Processus d'extraction de la marque [20]

Afin d'extraire la marque insérée, le document tatoué subit une subdivision de blocs de taille 8x8, ensuite, une transformation DCT 2D est appliquée sur chaque bloc afin d'extraire les bits de la marque insérée suivant le principe d'extraction décrit sur la figure II.28.

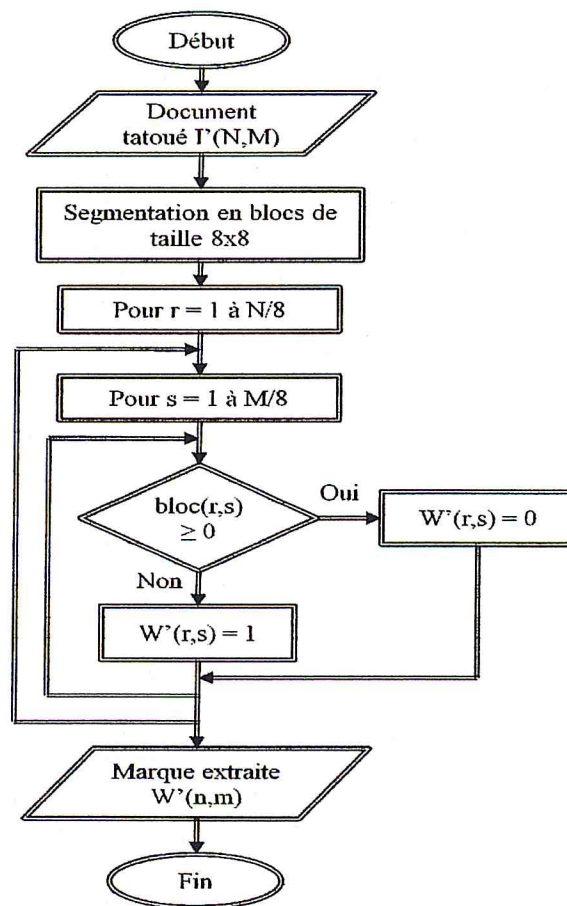


Figure II.28. Opération d'extraction dans le domaine fréquentiel

### V.3. Analyse expérimentales des résultats

En faisant référence au schéma de tatouage implémenté, le choix de la position d'insertion  $k$  est très important. Dans ce qui suit, nous allons spécifier d'une part ce paramètre et d'autre part la valeur de la force de marquage  $\alpha$ . La valeur de cette dernière

qui varie de 1 à 5 de telle façon à obtenir un bon compromis entre l'imperceptibilité de la marque et son intégrité dans le document tatoué.

Nous effectuons des tests sur un document image (Barbara), qui est de dimension 256x256 codée sur 8 bits par pixel. Nous insérons une marque de référence (Femme) de taille 32x32 pixels.

Les tableaux II.8, II.9, II.10, II.11, II.12, II.13, II.14 représentent les résultats obtenus pour les différentes positions d'insertion le long de la diagonal ( $k$  varie de 1 à 7).

**Tableau II.8.** Variation du PSNR selon  $\alpha$  pour la position  $k=1$

$\alpha$	PSNR (dB)	$\sigma$ (%)
5	19,0974	100
4	20,5424	100
3	22,9779	100
2	27,2679	100
1	37,9386	100

**Tableau II.9.** Variation du PSNR selon  $\alpha$  pour la position  $k=2$

$\alpha$	PSNR (dB)	$\sigma$ (%)
5	25,4770	100
4	27,8299	100
3	32,0218	100
2	32,0218	100
1	50,1338	100

**Tableau II.10.** Variation du PSNR selon  $\alpha$  pour la position  $k=3$

$\alpha$	PSNR (dB)	$\sigma$ (%)
5	30,1596	100
4	33,5552	100
3	38,1530	100
2	44,8447	100
1	56,5916	100

**Tableau II.11.** Variation du PSNR selon  $\alpha$  pour la position  $k=4$

$\alpha$	PSNR (dB)	$\sigma$ (%)
5	25,1202	100
4	28,5519	100
3	33,1208	100
2	39,9319	100
1	51,6525	100

**Tableau II.12.** Variation du PSNR selon  $\alpha$  pour la position  $k=5$

$\alpha$	PSNR (dB)	$\Sigma$ (%)
5	34,9898	100
4	38,6088	100
3	43,3695	100
2	50,3137	100
1	62,2	100

**Tableau II.13.** Variation du PSNR selon  $\alpha$  pour la position  $k=6$

$\alpha$	PSNR (dB)	Taux de corrélation (%)
5	43,9966	100
4	47,8	100
3	52,7797	100
2	59,7806	100
1	71,6514	100

**Tableau II.14.** Variation du PSNR selon  $\alpha$  pour la position  $k=7$

$\alpha$	PSNR (dB)	Taux de corrélation (%)
5	45,3182	100
4	49,1631	100
3	54,1351	100
2	61,1384	100
1	73,0229	100

Les résultats de variation du PSNR en fonction du facteur d'insertion  $\alpha$ , sont représentés dans la (Figure II.29) ci-dessous

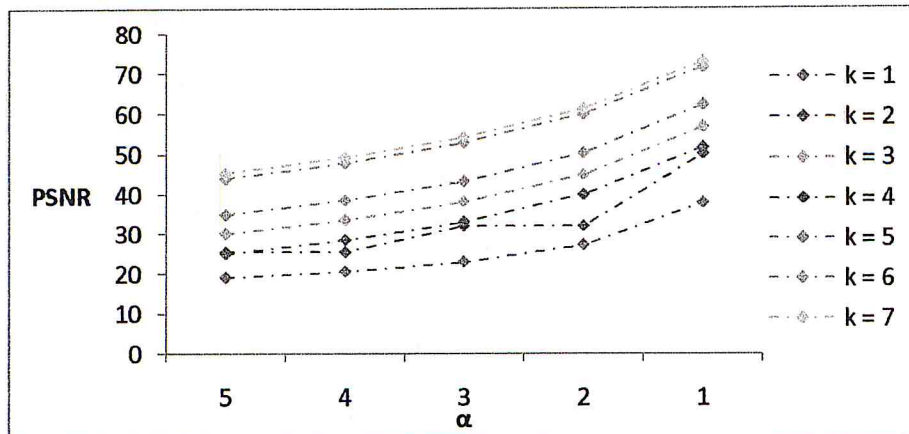







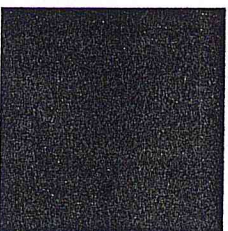
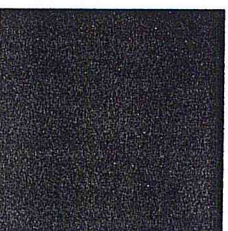


Figure II.29. Courbe de variation du PSNR en fonction de  $\alpha$  pour différentes  $k$  (DCT 2D)

**Choix de la valeur de la force d'insertion  $\alpha$**

D'après l'analyse de la courbe (figure II.29), il apparait bien à travers les valeurs du PSNR qui diminuent avec l'augmentation de la valeur de  $\alpha$  pour tous les positions de  $k$  qui varie de 1 à 7, que nous avons un meilleur résultat objectif et subjectif pour  $\alpha=1$  et les positions d'insertion  $k$  varie de 4 à 7, pour lequel nous effectuons l'extraction de la marque.

Tableau II.15. Les documents tatoués ainsi les différences pour  $k$  varie de 4 à 7

Document d'origine	k=4	k=5	k=6	k=7
	 PSNR=51,6525 (dB)	 PSNR= 62,2 (dB)	 PSNR=71,6514 (dB)	 PSNR=73,0229 (dB)
				

Afin d'augmenter la capacité d'insertion et remédier au problème d'effet de blocs. Nous avons réduit la taille des blocs de la DCT 2D à 4x4. Nous nous sommes inspirés de la dernière norme de compression vidéo H.264/AVC où DCT 4x4 est employée pour

Nous avons effectué des tests sur le même document utilisé dans l'insertion en employant la taille des blocs 8x8. Par contre la taille de la marque (femme) a augmenté jusqu'à 64 64x64 pixels. Les tableaux II.16, II.17, II.18 représentent les résultats tenus pour les différentes positions d'insertion le long de la diagonal (k varie de 1 à 3).

**Tableau II.16.** Variation du PSNR selon  $\alpha$  pour la position k=1

$\alpha$	PSNR (dB)	Taux de corrélation (%)
5	14,6826	100
4	16,3893	100
3	24,5426	100
2	19,2405	100
1	35,1294	100

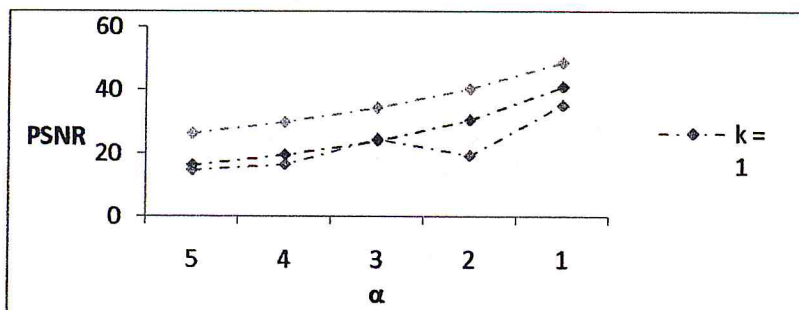
**Tableau II.17.** Variation du PSNR selon  $\alpha$  pour la position k=2

$\alpha$	PSNR (dB)	Taux de corrélation (%)
5	16,3125	100
4	19,5093	100
3	23,9549	100
2	30,3584	100
1	40,835	100

**Tableau II.18.** Variation du PSNR selon  $\alpha$  pour la position k=3

$\alpha$	PSNR (dB)	Taux de corrélation (%)
5	26,3671	100
4	29,8027	100
3	34,2427	100
2	40,2341	100
1	48,5113	100

Les résultats de variation du PSNR en fonction du facteur d'insertion  $\alpha$ , sont représentés dans la (Figure II.30) ci-dessous :





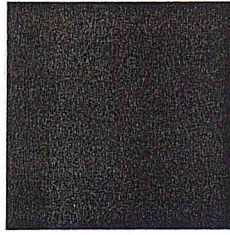
**Figure II.30.** Variation du PSNR en fonction de  $\alpha$  (DCT 2D 4x4)

**Choix de la valeur de la force d'insertion  $\alpha$**



De même façon que pour la DCT 2D 8x8, en faisant varier  $\alpha$  de 1 à 5 de telle manière à obtenir l'imperceptibilité de la marque. D'après l'analyse de la courbe (figure II.30), il apparait bien à travers les valeurs du PSNR qui diminuent avec l'augmentation de la valeur de  $\alpha$  pour tous les positions de  $k$  qui varie de 1 à 3, que nous avons un meilleur résultat objectif et subjectif pour  $\alpha=1$  et la position d'insertion  $k = 3$ .

**Tableau II.19.** Le document tatoué ainsi la différence pour  $k = 3$

Document d'origine	$k=3$	Différence
	 PSNR = 48,5113 (dB)	

**III. Conclusion**

Dans ce chapitre, les différentes techniques utilisées dans deux différents domaines de tatouage (spatial et fréquentiel) selon le type du document à protéger ont été décrites où le système du tatouage proposé est aveugle (ne nécessite pas la présence du document d'origine lors du processus d'extraction). Pour chaque domaine, les processus d'embrouillage, d'insertion et d'extraction de la marque sont bien détaillés. Les performances du système ont été évaluées en termes d'imperceptibilité, intégrité et qualité visuelle du document tatoué. L'analyse des résultats obtenus a permis d'avoir un aperçu sur la variation des performances en fonction des paramètres du système et ainsi aider au choix de ces paramètres.

Dans le prochain chapitre, nous allons appliquer quelques attaques sur les documents tatoués dans les deux domaines, afin d'analyser la robustesse de chaque technique et de sortir avec une étude comparative entre les résultats obtenus.

# CHAPITRE III

## ATAQUES

Une once de bon esprit vaut mieux  
qu'une livre de science.

## I. Introduction

L'objectif de ce projet est la sécurisation des documents imprimés par tatouage numérique, donc nous nous intéressons à la robustesse des méthodes développées face aux différentes attaques, afin de choisir les bonnes valeurs des paramètres d'insertion tel que la position d'insertion satisfaisant au compromis entre à l'imperceptibilité et la robustesse face aux différentes attaques.

Afin d'évaluer la robustesse des méthodes développées pour la protection des documents imprimés cinq différentes attaques ont été appliquées :

- L'impression-numérisation.
- Le recadrage (centré).
- Le remplacement (centré).
- Filtrage numérique (filtre gaussien et filtre médian).

### • Attaque par impression-scan

C'est l'attaque la plus essentielle pour évaluer la robustesse d'une méthode de protection des documents imprimés par tatouage numérique.

### • Attaque par recadrage

Cette attaque consiste à recadrer une surface du document. Nous avons recadré plusieurs régions dans le document tatoué selon le choix par exemple : région centrée, région haut-gauche, bas-droite, etc...

### • Attaque filtrage numérique

Les attaques bienveillants généralement utilisés sont issus généralement du filtrage numérique. Les filtres sont des produits de convolution qui mettent en jeu le voisinage de chaque pixel. L'opération de filtrage est défini pour chaque pixel de l'image par :

- 1) Une fenêtre carré centrée sur le pixel considéré de dimension impaire  $(2n+1) \times (2n+1)$ .
- 2) Une matrice  $M$  de coefficients de même dimension que la fenêtre. C'est cette matrice qui fait la différence d'un filtre à un autre.

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{III.1}$$

- 3) Une transformation combinant les valeurs recouvertes par la fenêtre et les valeurs de la matrice.

Parmi les filtres appliqués comme attaques, nous avons utilisé le filtre gaussien et le filtre médian.

#### • Filtre Gaussien

Le filtre gaussien consiste à faire une convolution de la fenêtrés de taille  $(2n+1) \times (2n+1)$  avec une matrice M dont les coefficients sont obtenus en utilisant la fonction gaussienne  $G(i,j,\sigma)$  à deux dimensions définie comme suit :

$$G(i, j, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} - \frac{i^2+j^2}{\exp(2\sigma^2)} \quad \text{III.2}$$

Le principe de ce filtre est de faire en premier lieu une convolution de l'image initial avec la gaussienne suivant la variation i puis suivant la variation j. L'intérêt du filtre gaussien est qu'il est possible de régler le degré de filtrage à travers le paramètre  $\sigma$ .

#### • Filtre Médian

Son principe est simple. Il consiste à classer, pour chaque pixel de l'image initial, les valeurs des intensités des pixels voisins de la fenêtre selon un ordre croissant ou bien décroissant, puis choisir la valeur médiane de ces pixels classés.

#### Exemple de filtrage

Les intensités d'une fenêtre de taille  $(3 \times 3)$  dans une image avant d'appliquer un filtre médian sont :

$$M = \begin{bmatrix} 40 & 11 & 20 \\ 35 & (12) & 30 \\ 10 & 20 & 20 \end{bmatrix}$$

En ordonnant les valeurs nous obtenons le vecteur suivant : (10 11 12 20 (20) 20 30 35 40) ; la médiane obtenue est 20.


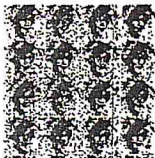

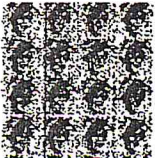


La fenêtre de l'image après filtrage est : 
$$\begin{bmatrix} 40 & 11 & 20 \\ 35 & (20) & 30 \\ 10 & 20 & 20 \end{bmatrix}$$


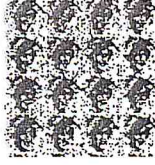






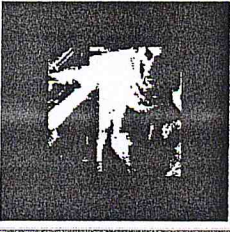



Ce filtre donne de très bons résultats vu son principe mais son inconvénient est qu'il supprime les détails fins surtout lorsque nous utilisons des fenêtres de grande taille.

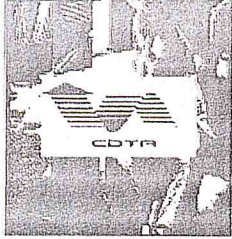

## II. Robustesse de la méthode développée dans le domaine spatial

Nous avons appliqué les cinq attaques mentionnés sur l'image de référence *Barbara* (256x256) qui est tatouée dans le domaine spatial après une binarisation par les deux méthodes *Ordred* et *Floyd*. Les résultats obtenus en terme de taux de corrélation sont illustrés sur le tableau III.1.

Tableau III.1. Résultats d'extraction après attaques dans le domaine spatial

Attaque	Document tatoué	Marque extraite	$\sigma$ (%)
Impression-scan			85,35
	Floyd		
			83,02
Ordred			
Gaussien			88,16
	Floyd		

Attaque		Document tatoué	Marque extraite	$\sigma$ (%)
Gaussien				88,23
		Ordred		
Médian				92,51
		Floyd		
				92,5
		Ordred		
Recadrage	Centré			74,32
		Floyd		
				75,81
	Ordred			
Remplacement	Centré			84,13
		Floyd		

Attaque		Document tatoué	Marque extraite	$\sigma$ (%)
Remplacement	Centré			84,08
	Ordred			

D'après le tableau ci-dessus, nous remarquons que le taux de corrélation de la marque extraite est supérieur à 75 pour toutes les différentes attaques appliquées, ce qui montre que la méthode d'insertion dans le domaine spatial suivant les deux méthodes de binarisation est robuste face aux différentes attaques mentionnées.







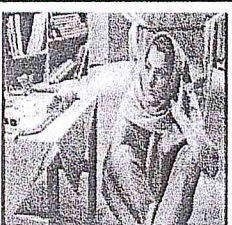

### III. Robustesse des méthodes développées dans le domaine fréquentiel

#### III.1. Méthode basée sur la DCT 8x8

Afin de définir la position adéquate d'insertion tout en satisfaisant le compromis entre l'imperceptibilité du document tatoué et le taux de corrélation de la marque extraite après l'attaque par l'opération impression-scan qui est primordiale pour les documents imprimés, les tests sont effectués sur les quatre positions  $k$  variant de 4 à 7 selon la première diagonale du bloc 8x8. Les résultats obtenus sont représentés sur le tableau III.2.

Pour mieux présenter les résultats obtenus pour cette attaque et déduire la bonne position d'insertion garantissant les critères d'imperceptibilité et robustesse, nous avons présenté les valeurs du PSNR ainsi le taux de corrélation sur l'histogramme illustré sur la figure III.1

Tableau III.2. Résultats d'extraction après l'attaque d'impression-scan dans le domaine fréquentiel pour k varie de 4 à 7

Attaque	k	Document tatoué	Marque extraite	$\sigma$ (%)
Impression-scan	4			73,43
	5			83,1
	6			39,84
	7			35,74

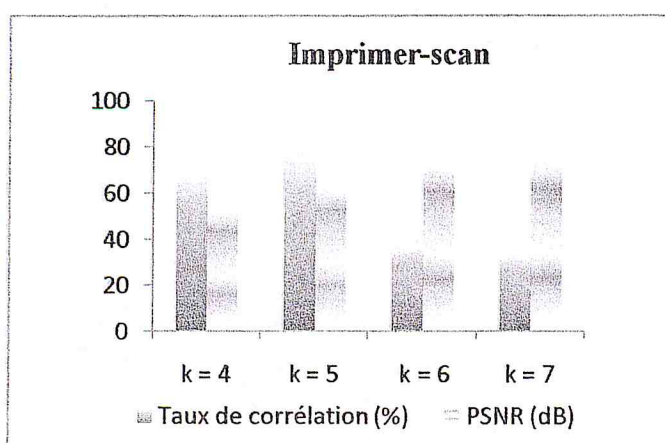


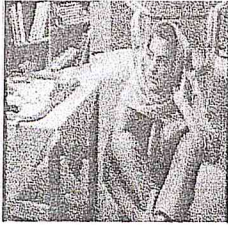







Figure III.1. Histogramme de variation du PSNR et du taux de corrélation après l'attaque d'impression pour les différentes positions k (DCT 8x8)



Après l'analyse de l'histogramme, la position d'insertion adéquate pour satisfaire le compromis robustesse/imperceptibilité est obtenu pour  $k = 5$  avec un PSNR égale à 62,2 dB et un taux de corrélation calculé  $\sigma = 83,10\%$ .

Les autres attaques citées sont effectuées pour la position d'insertion  $k=5$ . Les résultats obtenus sont représentés dans le tableau III.3 ci-dessous :

Tableau III.3. Résultats d'extraction après attaques dans le domaine fréquentiel ( $k=5$ )











Attaque		Document tatoué	Marque extraite	$\sigma$ (%)
Gaussien				98,73
Médian				100
Recadrage	Centré			65,61
Remplacement	Centré			61,80

D'après les résultats obtenus nous concluons que la méthode développée dans le domaine fréquentiel en appliquant la DCT 8x8 est robuste face aux attaques par filtrage numérique alors que pour les autres elle n'est pas robuste.

## III.1. Méthode basée sur la DCT 4x4

Les résultats d'évaluation de la robustesse de la méthode développée dans le domaine fréquentiel basée sur la DCT 4x4 pour la seule position d'insertion  $k=3$  sont présentés par le tableau III.4.

Tableau III.4. Résultats d'extraction après attaques dans le domaine fréquentiel (DCT 4x4,  $k=3$ )

Attaque		Document tatoué	Marque extraite	$\sigma$ (%)
Impression-scan				76,7
Gaussien				74,95
Médian				100
Recadrage	Centré			76,51
Remplacement	Centré			90,74

D'après les résultats obtenus nous concluons que la méthode développée dans le domaine fréquentiel en appliquant la DCT 4x4 est robuste face aux attaques par filtrage numérique, opération impression-scan, recadrage et remplacement avec une

augmentation de la taille de la marque jusqu'à 64x64 pixels et un taux de corrélation entre la marque extraite et la marque d'origine supérieur à 75%.

#### **IV. Conclusion**

Dans ce chapitre, les performances des différentes techniques d'insertion utilisées ont été évaluées en termes d'imperceptibilité et de robustesse en utilisant des documents imprimés images. L'analyse des résultats obtenus a permis d'avoir un aperçu sur la variation des performances selon le domaine d'insertion (spatial ou fréquentiel). Nous avons également eu de bons résultats dans le domaine spatial.

Dans le chapitre suivant, nous allons présenter notre logiciel avec quelque prise d'écran afin de faciliter son utilisation.

# CHAPITRE IV : IMPLEMENTATION

Celui qui a peur de demander est  
honteux d'apprendre.

**I. Introduction**

Après avoir présenté dans le chapitre précédent la modélisation UML de notre application. Nous allons dans ce chapitre mettre en œuvre cette modélisation. En présentant tout d'abord l'environnement de programmation utilisé, ensuite nous exposons l'interface de l'application ainsi que les fonctionnalités que permet d'accomplir notre application afin de faciliter son utilisation et cela par le biais des prises d'écran.

**II. Environnement de programmation**

Pour implémenter notre application, nous avons utilisé le langage de programmation Visuel C# 2008 qui est un environnement de programmation riche en outils comportant toutes les fonctionnalités nécessaires pour créer des projets C# de toute taille.

Microsoft Visual C# est un puissant langage orienté composant créé par Microsoft. C# joue un rôle essentiel dans l'architecture de Microsoft .NET Framework, et certaines personnes ont comparé son rôle à celui joué par C dans le développement d'UNIX. Son syntaxe est très proche de celles des langages comme C, C++ ou Java.

**III. Interface de l'application**

Notre application comporte une fenêtre « menu principal » (figure IV.2) et des sous fenêtres « forms ». Le menu principale de l'application nommé « Sécurisation des documents imprimés par tatouage numérique » permet le chargement du document en choisissant son type (en niveaux de gris ou en couleurs) aussi le processus qu'on va effectuer sur le document : insertion ou extraction et attaques (figure IV.4).

Pour chaque processus, les méthodes correspondantes sont associées à la barre de menu, se qui permet de les spécifier au préalable.

Ainsi, on peut accéder directement à la fenêtre « calcul du taux de corrélation » (figure IV.3) sans faire charger un document, ce qui offre la possibilité de calculer le taux sur des marques déjà extraites.

Le schéma illustré dans la figure IV.1 montre les différents processus et techniques réalisés dans notre application :

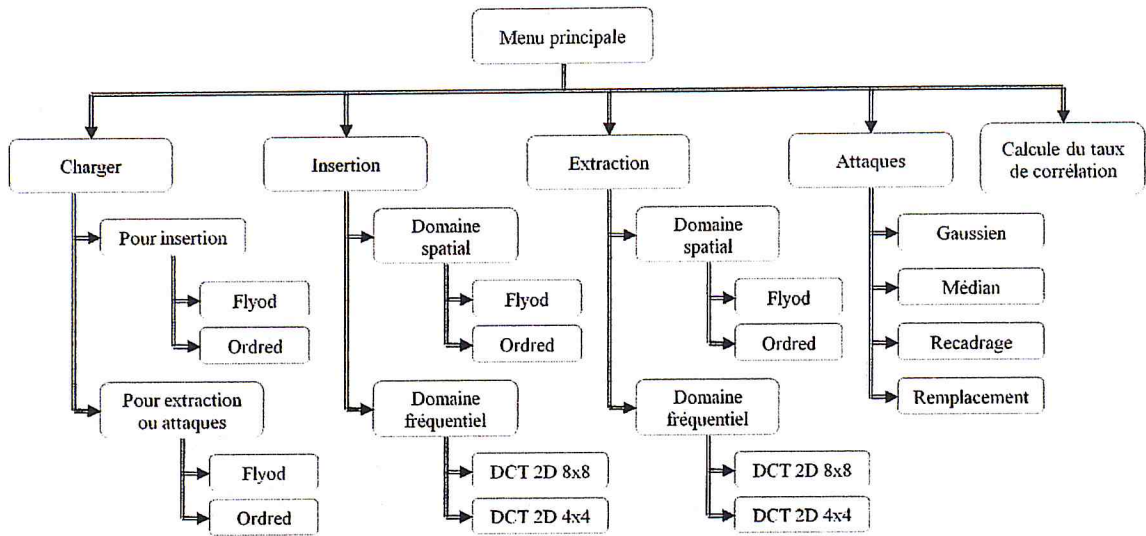


Figure IV.1. Schéma générale de l'application

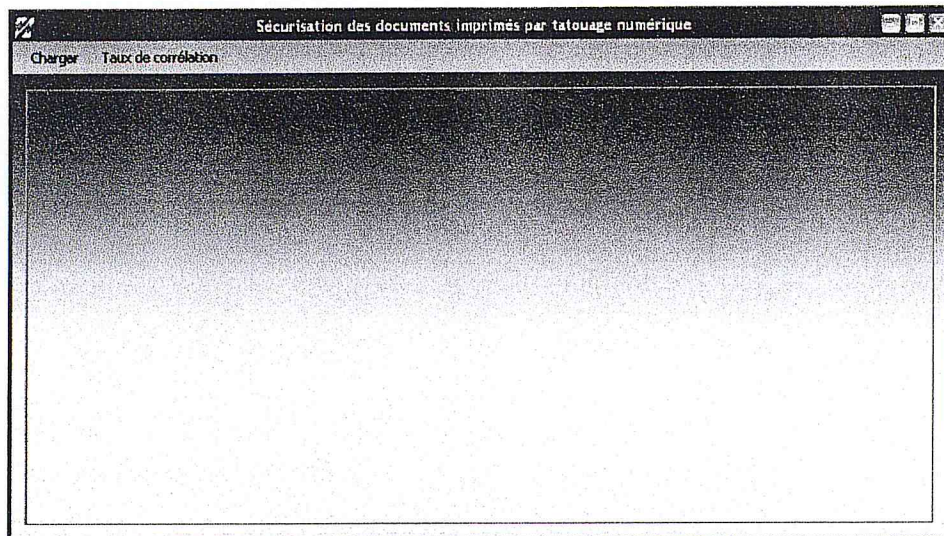


Figure IV.2. Menu principale

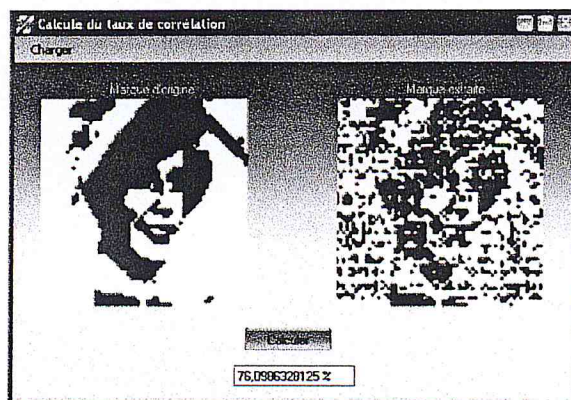
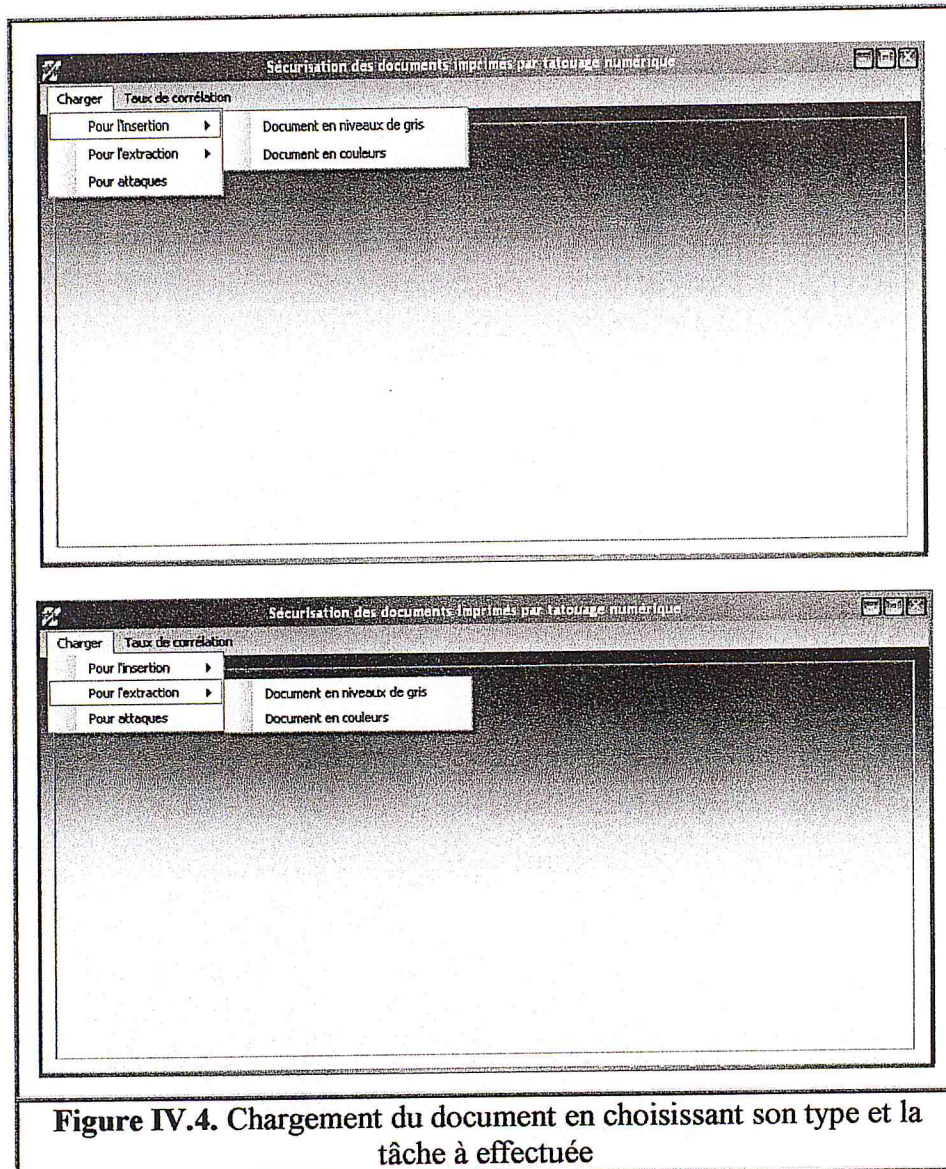
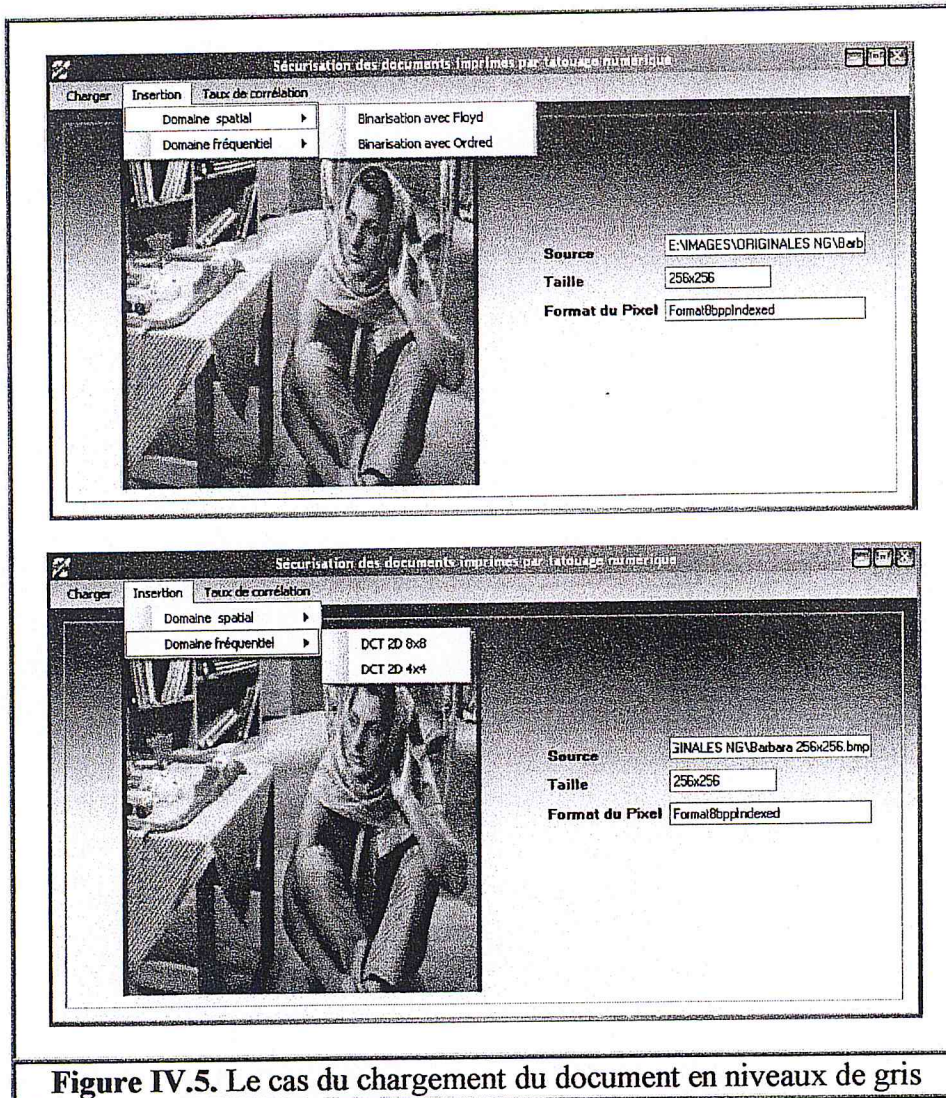


Figure IV.3. Calcul du taux de corrélation de l'extraction

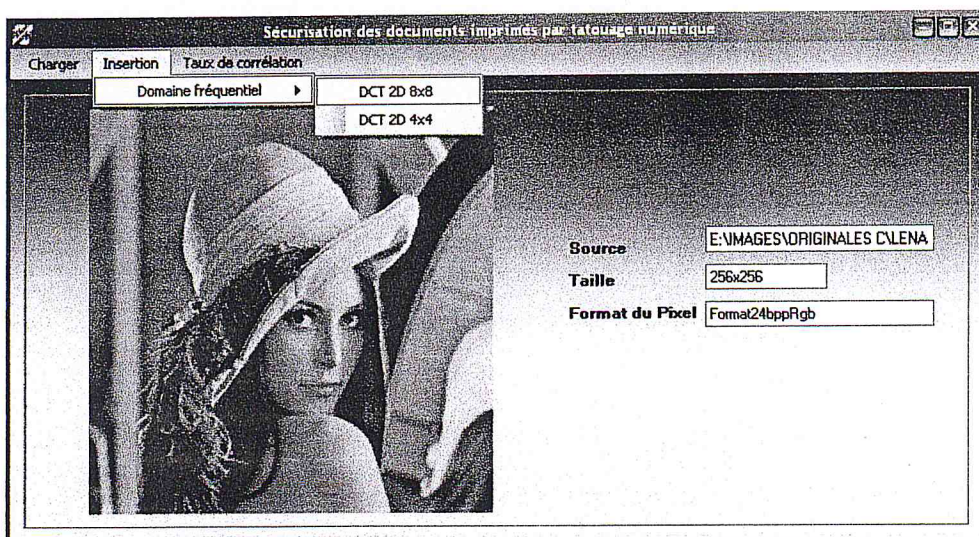


**Figure IV.4.** Chargement du document en choisissant son type et la tâche à effectuée

Pour l'insertion il existe deux choix du domaine : domaine spatial et domaine fréquentiel (figure IV.5). Dans le cas où le document chargé est en niveaux de gris, les deux domaines sont disponibles, si le domaine spatial est choisi pour effectuer l'insertion, on peut choisir entre deux méthodes de binarisation : *Floyd* ou *Ordred*. Si non, si le domaine fréquentiel est choisi, l'insertion peut être effectuée en utilisant l'une des deux transformées : DCT 8x8 ou DCT 4x4.



D'autre part, si le document chargé est en couleurs, seul le domaine fréquentiel est disponible (figure IV.6).





Quand l'utilisateur choisit d'effectuer l'insertion dans le domaine spatial (par exemple, utilisant la méthode de binarisation *Ordred*), une fenêtre correspondante est apparue (figure IV.7) contenant le document d'origine convertit en demi-teinte. Après, la marque à insérer doit être chargée, et la clé d'embrouillage est ensuite donnée. Enfin, en cliquant sur le bouton « Insérer marque », une fenêtre résultats d'insertion (figure IV.8) qui contient le document d'origine, la marque dupliquée et dupliquée avec embrouillage, les deux clés publique et secrète et le document tatoué est apparue.

Tous ces résultats peuvent être enregistrés, on peut aussi calculer le PSNR et la différence entre le document d'origine et celui tatoué, cette dernière est affichée dans une autre fenêtre (figure IV.9) et peut aussi être enregistrée.

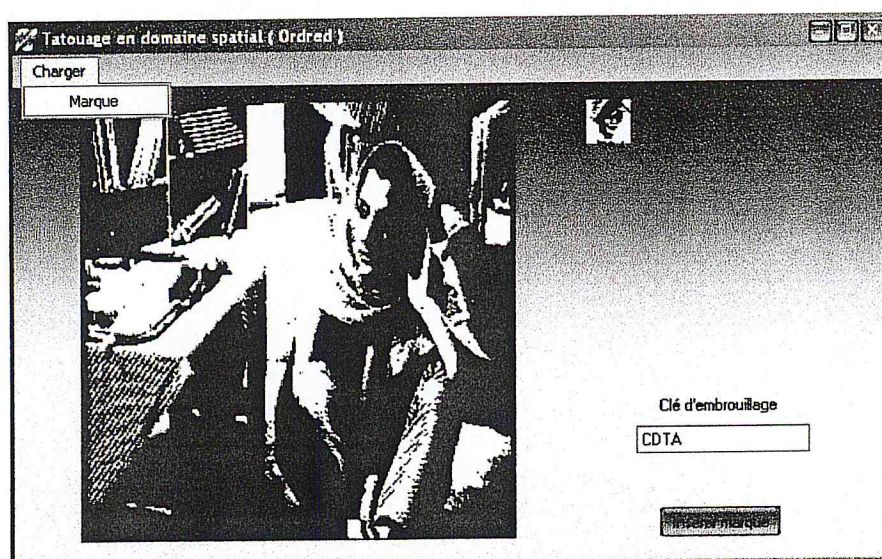


Figure IV.7. Insertion dans le domaine spatial (méthode de binarisation *Ordred*)

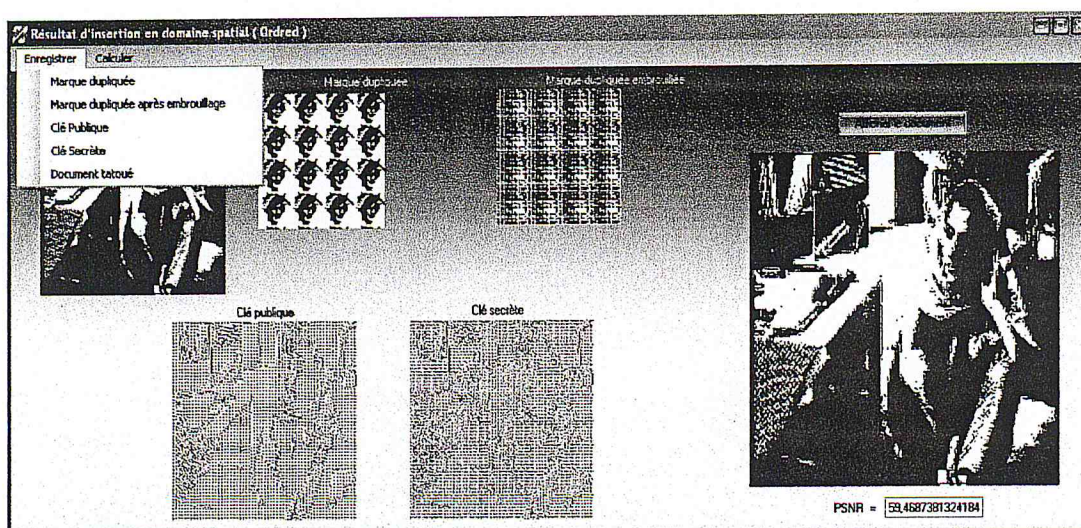


Figure IV.8. Résultats d'insertion dans le domaine spatial (méthode de binarisation *Ordred*)

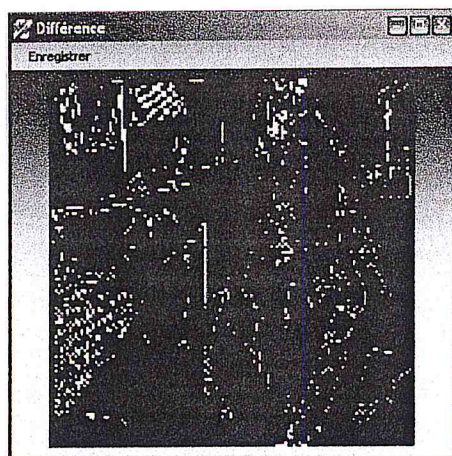


Figure IV.9. La différence entre le document d'origine et le document tatoué

D'autre part, si l'utilisateur choisit d'effectuer l'insertion dans le domaine fréquentiel (par exemple, utilisant la transformée DCT 4x4), une fenêtre correspondante est apparue (figure IV.10) contenant le document d'origine en niveaux de gris. Après, la marque à insérer doit être chargée et embrouillée par clé d'embrouillage avant l'insertion, la force d'insertion  $\alpha$  et la position d'insertion  $k$  sont ensuite données. Enfin, en cliquant sur le bouton « Insérer marque », une fenêtre résultats d'insertion (figure IV.11) qui contient le document d'origine et le document tatoué est apparue qui peut être enregistré.

La même chose que le domaine spatial, on peut calculer le PSNR et la différence entre le document d'origine et celui le tatoué à partir de cette fenêtre.

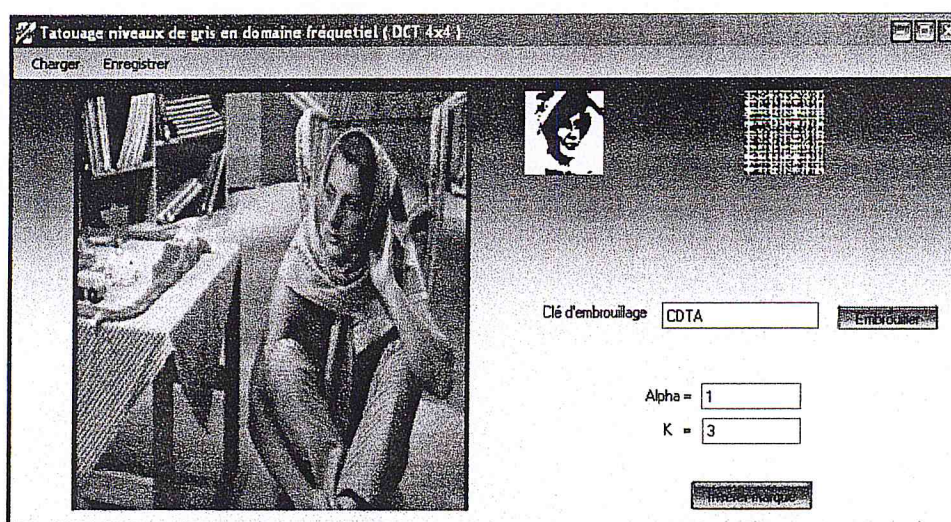


Figure IV.10. Insertion dans le domaine fréquentiel (transformée DCT 4x4)

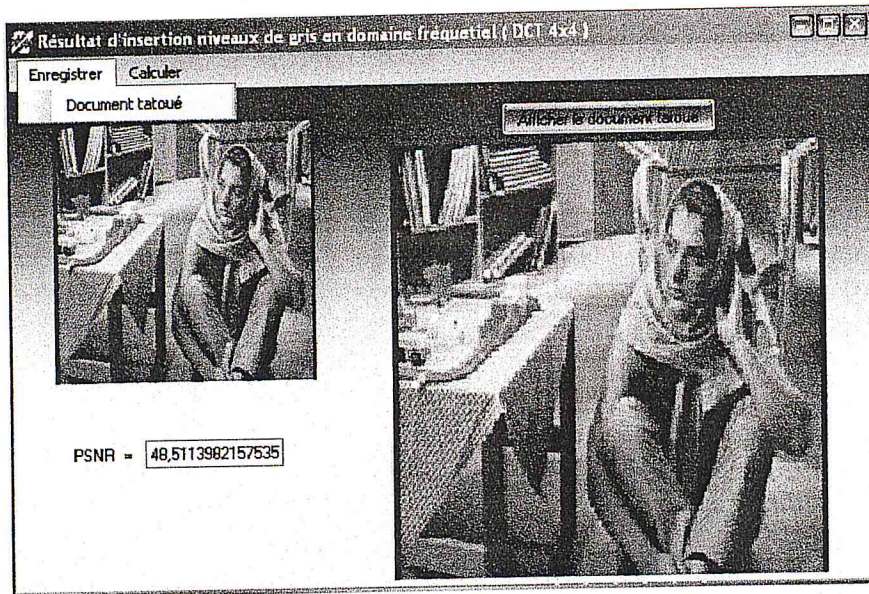


Figure IV.11. Résultats d'insertion dans le domaine fréquentiel (transformée DCT 4x4)

Pour l'extraction, la même manière de choix de domaine est suivie. Pour effectuer l'extraction dans le domaine spatial, l'utilisateur doit charger la clé secrète dans la fenêtre (figure IV.12) qui contienne le document tatoué au préalable, après l'extraction, la marque extraite embrouillée est désembrouillée en utilisant une clé de désembrouillage, les deux marques peuvent être enregistrées.

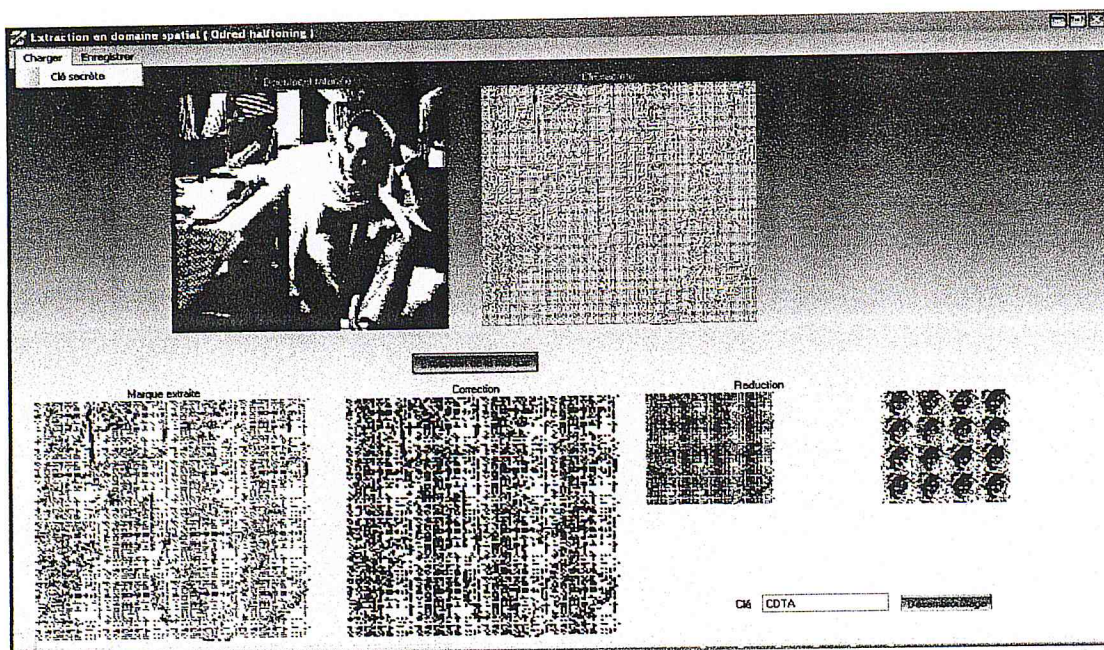


Figure IV.12. Extraction dans le domaine spatial (méthode de binarisation *Ordred*)

Si l'extraction est dans le domaine fréquentiel, une fenêtre qui contient le document tatoué est apparue (figure IV.13). Pour extraire la marque, la position d'insertion k doit

être mentionnée, après on applique le même processus de désembrouillage appliqué dans le domaine fréquentiel sur la marque extraite pour obtenir la marque désembrouillée.



Figure IV.13. Extraction dans le domaine fréquentiel (transformée DCT 4x4)

Finalement, notre application offre la possibilité dévaluer la robustesse des méthodes développées faces à quelques attaques usuelles telles que l'attaque de recadrage et l'attaque de remplacement.

Pour l'attaque de recadrage, l'utilisateur peut spécifier son format en choisissant un des choix présentés dans la fenêtre (figure IV.14).

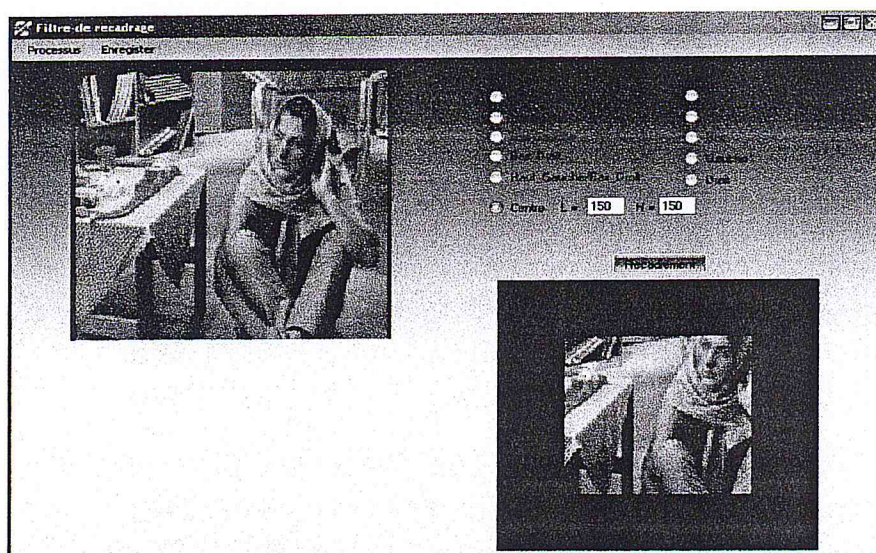


Figure IV.14. L'attaque de recadrage

**IV. Conclusion**

Dans ce chapitre nous avons donné une brève présentation de l'environnement de programmation. Aussi, nous avons décrit l'interface de l'application avec toutes les fonctionnalités qu'elle permet d'accomplir à travers des prises d'écrans et cela dans le but de faciliter son utilisation.

## Conclusion générale

Le développement des technologies et moyens informatiques et matériels, tels que les photocopieurs couleur, les scanners et les imprimantes, de même que les logiciels de retouches d'images a ouvert de grandes perspectives et possibilités de création et de manipulation des contenus des documents imprimés. Mais, ce confort matériel a également ouvert le champ à la copie et à la distribution illégale. On estime au niveau mondial à plus d'une vingtaine de milliards de dollars les pertes en matière de droits d'auteur. Face à cette situation alarmante, les chercheurs se sont intéressés aux solutions de protection numérique. Parmi les solutions les moins coûteuses et les plus efficaces on trouve le tatouage numérique, plus connu sous le nom « *watermarking* ».

Cette technique est inspirée de la méthode d'insertion d'une signature qui peut être une marque, un logo ou un identifiant du propriétaire. Il s'agit d'effectuer l'insertion dans deux différents domaines selon le type du document imprimé à protéger (en niveaux de gris ou en couleurs).

Pour ce faire, nous avons appliqué différentes techniques pour chaque domaine à savoir, la duplication de la marque à insérer selon la taille du document à tatouer, et la génération d'une clé publique qui est utilisé lors de l'insertion et une clé secrète qui est transmise pour effectuer l'extraction pour le tatouage des documents en niveaux de gris dans le domaine spatial.

D'autre part, pour le tatouage des documents en couleurs et aussi en niveaux de gris dans le domaine fréquentiel, une transformée en cosinus discrète DCT-2D ainsi que sa transformée inverse sont calculées sur des blocs de tailles 8x8. L'insertion de la marque est effectuée en ayant en entrée plusieurs paramètres tels que (la force du marquage  $\alpha$ , la position d'insertion  $k$  dans le bloc).

Pour augmenter la capacité d'insertion nous avons appliqué la transformé en cosinus discrète DCT-2D sur des blocs de taille 4x4 à la place des blocs de taille 8x8. Ainsi nous avons augmenté les performances du système face aux attaques appliquées.

Ainsi pour augmenter la sécurisation du système du tatouage proposé nous avons effectué un processus d'embrouillage de la marque avant qu'elle ne soit insérée avec l'utilisation d'une clé d'embrouillage.

Quelques tests ont été réalisés afin d'évaluer les performances des techniques implémentés en terme de capacité d'insertion, d'imperceptibilité et de robustesse. Le premier mesure la qualité d'information requise par la marque numérique, le deuxième est directement lié à la qualité du document tatoué et enfin, le troisième mesure la résistance aux différentes attaques.

Toutefois, ce système peut être amélioré en envisageant les perspectives suivantes :

- Appliquer une méthode de binarisation inverse (du binaire vers le niveau de gris) sur les documents tatoués dans le domaine spatial afin de reconstruire le document d'origine.
- Appliquer la méthode binarisation sur les trois composantes de couleurs dans le cas des documents en couleurs
- Utiliser d'autres méthode de transformation vers le fréquentiel comme DWT (*Discret Wavelet Transform* ou *TOD Transformée en ondelettes*).

Le but attendu de ce travail est de développer un système de tatouage numérique visant à protéger le droit de propriété des documents imprimés. En fait les résultats obtenus seront appliqués aux vignettes d'affranchissement d'Algérie Post.

## Bibliographie

- [1] F.A.P Peticolas, R.J. Anderson et M.G. Kuhn, " Information Hiding- A Survey, " Proceeding of the IEEE, special issue on Protection of Multimedia Content, vol. 87, no. 7, pp. 1062- 1078, July 1999.
- [2] J.Cox et Morgan Kaufmann, " Digital watermarking and stéganography," deuxieme edition, 2007.
- [3] Vincent Martin, " Contribution des filtres LPTV et des techniques d'interpolation au tatouage numérique," thèse de doctorat, École doctorale : Informatique et Télécommunications, 28 novembre 2006. ← Université
- [4] Yann bodo, " Elaboration d'une technique d'accès conditionnel par tatouage et embrouillage vidéo basée sur la perturbation des vecteurs de mouvement," thèse de doctorat, Ecole National Supérieur des Télécommunications, 09 septembre 2004.
- [5] A. Parisi, P. Carré and A. Trémeau " Introduction au tatouage d'images couleur," Laboratoire SIC - FRE-CNRS 2731, Université de Poitiers, 2004.
- [6] E.Khelifi, " Image Compression and Watermarking in the Walevet Transform Domain," Phd, Queen's University of Belfast, UK, Mai 2007.
- [7] N. Ratha, J. H. Conell et R.M. Bolle, " Enhancing Security and Privacy in Biometrics-Based Authentication Systems," IBM Systems Journal, vol. 40, no. 3, pp. 614-634, 2001.
- [8] F. Raynal, " Etudes d'outils pour la dissimulation d'information: approches fractales, protocoles d'évaluation et protocoles cryptographiques," thèse de doctorat, Université Paris XI, mars 2002.
- [9] P. Bas, J. M. Chassery, " Tatouage d'images résistant aux transformations géométriques," dix-septième colloque GRETSI, Vannes, 13-17 septembre 1999.
- [10] Patrick Bas, Docteur de l'INPG, " Méthode de tatouage d'images fondées sur le contenu," spécialité : signal, image, parole, telecom. Laboratoire LIS de Grenoble. Directeur de thèse : jean-Marc Chassery, le 05 octobre 2000.
- [11] Belmouloud Hichem et Benkaci Sofiane " Tatouage d'images médicales en vue d'une transmission de données," Mémoire d'ingénieur d'état en Electronique, Faculté d'Electronique et d'Informatique, Département Télécommunications, USTHB, Algérie, 2005.

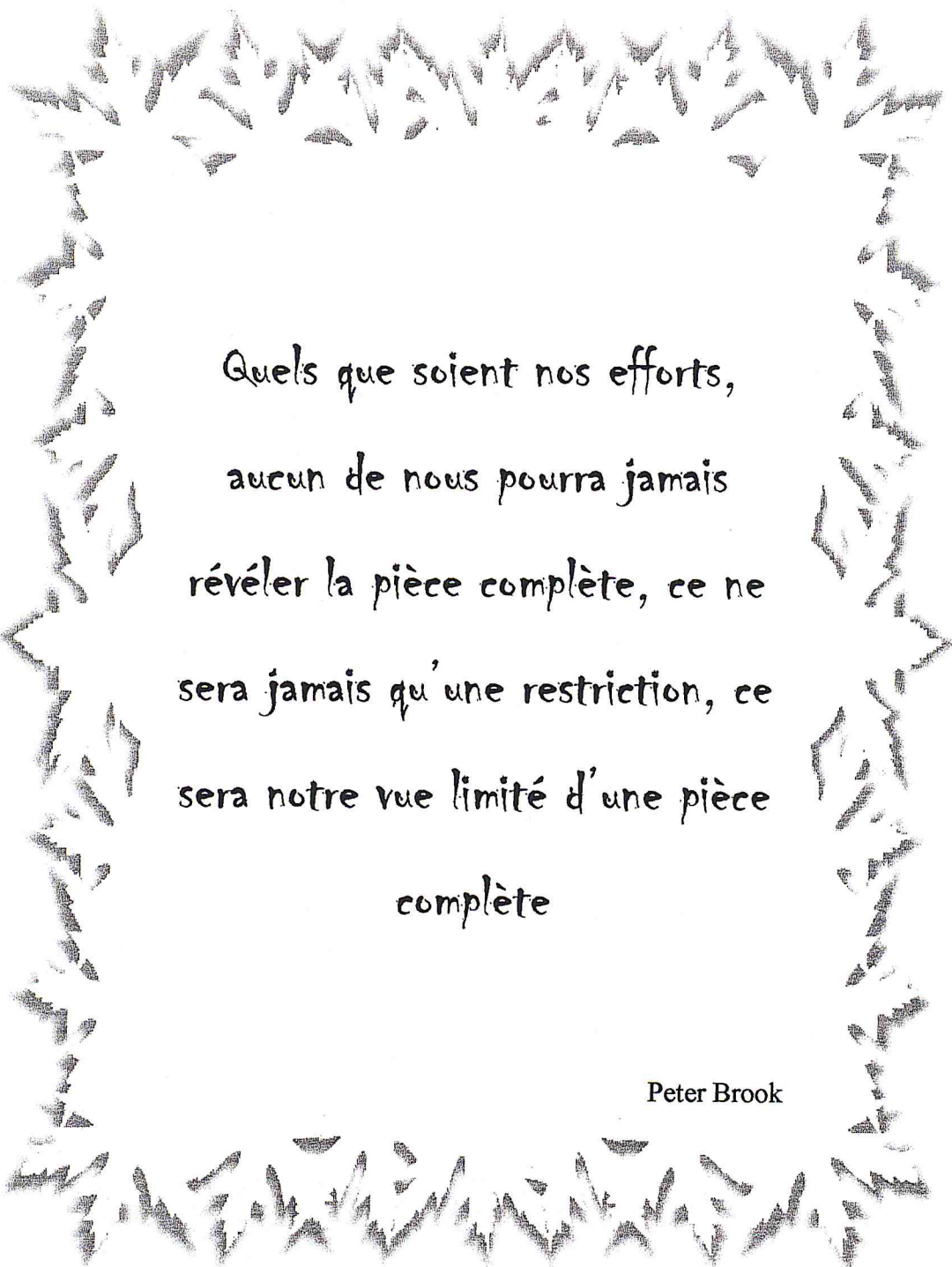


- [12] Naornita Corina " Filigranage dans le domaine des ondelettes," Mémoire de diplôme pour obtenir le degré de M.Sc, L'université "Politehnica" Timisoara, Faculté d'Electronique et Télécommunications, 2004.
- [13] Khalil Zebbiche, " Data hiding for securing fingerprint data access," thèse Phd, Queen's University of Belfast, juillet 2008.
- [14] Guekhakhma Said, " Authentification du contenu H.264/AVC le tatouage numérique et la signature numérique," Mémoire pour l'obtention D'un diplôme d'ingénieur d'état en informatique, Option : Système d'informatique (SI), 2008-2009.
- [15] VU Duc Minh et NGUYEN Thi Hoang Lan Maiti " Tatouage des images dans un domaine fréquentiel," pp 6-14, 15 janvier 2006.
- [16] Minya Chen, Edward K. Wong, Nasir Memon et Scott Adams, " Recent Developments in Document Image Watermarking and Data Hiding," Department of Computer and Information Science, Brooklyn.
- [17] Bourouba Rafik et Igroupfa Hiba " Mise sous réseau internet d'un système de vidéosurveillance suivant la norme MPEG-4/AVC basée objet," Mémoire de fin d'études pour l'obtention du diplôme d'ingénieur d'état en informatique, Option : Système informatique, 2008/2009.
- [18] Patrick Bas et Jean-Marc Chassery, " Tatouage couleur adaptatif fondé sur l'utilisation d'espaces perceptifs uniformes ", Laboratoire des Images et des Signaux, France, 2004.
- [19] Zohra Tifedjadjine," Halftone Image Watermarking based on Visual Cryptography," Dissertation for master in Electronics, Option Micro-waves for Telecommunication, pp 1, 5, Novembre 2005.
- [20] Malay K. Kundu and Arpan K. Maiti "An Inexpensive Digital Watermarking Scheme for Printed Document," Machine Intelligence Unit, Indian Statistical Institute, pp 2.
- [21] R.W.Floyd et L.Steinberg," An Adaptative Algorithm for Spatial Gray Scale," SID International Symposium Digest of Technical Papers, 1975.
- [22] Lee Raymond Lam et Henry," A Chaotic Real-time Cryptosystem using a Switching Algorithmic-based Linear Congruential Generator (SLCG) ", IJCSNS International Journal of Computer Science and Network Security, vol.6, No.8, pp.116-123 August 2006.

- [23] S. Wegenkittl, " On empirical testing of pseudorandom number generators". In G. De Pietro, A. Giordano, M. Vajtersic, and P. Zinterhof, editors, Proceedings of the international workshop Parallel Numerics '95. CEI-PACT Project, pp.59-71, 1995.
- [24] T. Wiegand, G.J. Sullivan, G. Bjontegaard, et A. Luthra, Overview of the H.264/AVC Video Coding Standard, IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560-576, 2003.
- [25] Muhammed Yusuf Khan, Ekram Khan et M. Salim Beg "Performance Evaluation of 4x4 DCT Algorithms for Low Power Wireless Applications ," Electrical Engineering Section, University Polytechnic , India

### **Sites Web**

- SW1.** <http://www-ljk.imag.fr/membres/Valerie.Perrier/SiteWeb/node9.html>.
- SW2.** <http://www.univ-brest.fr/lest/tst/projets/atb/atb.htm>.
- SW3.** [http://home.nordnet.fr/~jpbaey/tipe/compression\\_jpeg/transformation\\_dct.htm](http://home.nordnet.fr/~jpbaey/tipe/compression_jpeg/transformation_dct.htm).
- SW4.** <http://www-ljk.imag.fr/membres/Valerie.Perrier/SiteWeb/node9.html>.



Quels que soient nos efforts,  
aucun de nous pourra jamais  
révéler la pièce complète, ce ne  
sera jamais qu'une restriction, ce  
sera notre vue limitée d'une pièce  
complète

Peter Brook