Travail d'étude – Licence informatique

Le tatouage d'images ou "Watermarking"

(Cédric Piovano & Julien Pugliesi)

JUIN 2004

Encadrement: Pierre Crescenzo

Université Nice SOPHIA ANTIPOLIS

1. Introduction

- Le watermarking, qu'est-ce donc ?
- En quoi cela diffère de la cryptographie ?
- Mais, à quoi cela peut bien servir ?

….. de gros problèmes de droits d'auteurs !

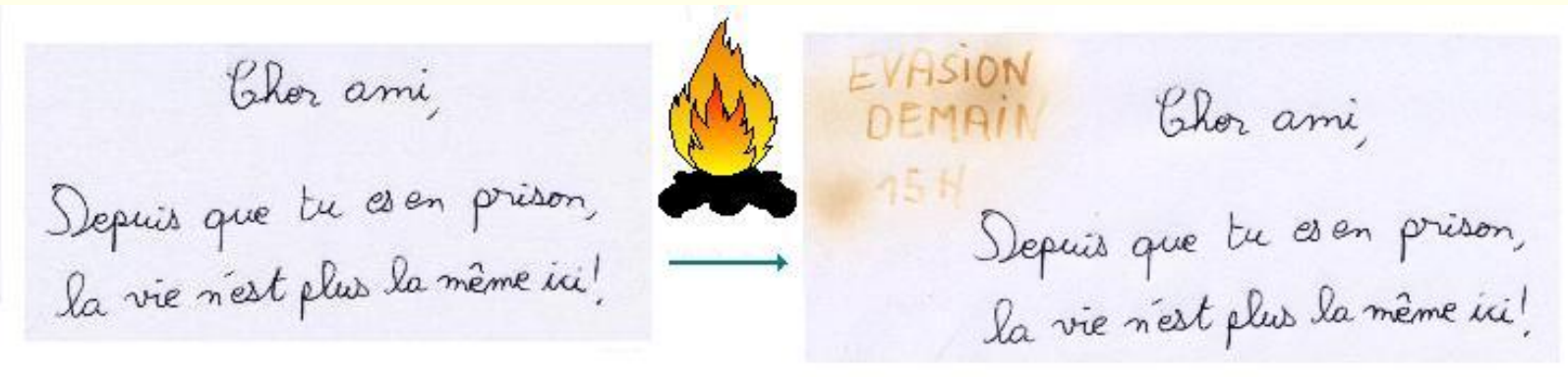
Le watermarking peut y répondre : copyright, numéro de licence

1. Histoire de l'art …

Stéganographie ……. Watermarking .

-Premières utilisations de la stéganographie
- chez les grecs
- encres sympathiques (voir schéma)

-De la Stéganographie au watermarking
- Madame Tatcher
- Le DVD, les firmes Américaines (JASRAC, RIAJ)



1. Quelques notions de base

- **Notion de pixels, et de valeurs**
- **Notion de domaine Spatial**
- **Notion de domaine Fréquentiel, DCT**
- **Un peu de terminologie**

Un nom plus approprié… **…le « data-embding »**

1. Degrés et formes de tatouages

- **Notion de visibilité**
- **Notion de robustesse et de fragilité**
- **Un cas intermédiaire le « semi-fragile »**
- **Notion de ratio**

Exemple Visible ci-dessous

Exemple Invisible ci-dessous

2. Les algorithmes

1. **Domaine spatial:**

   - **Bit de poids faible**

   - **PatchWork**

2. **Domaine fréquentiel:**

   - **Algorithme de Koch et Zao**

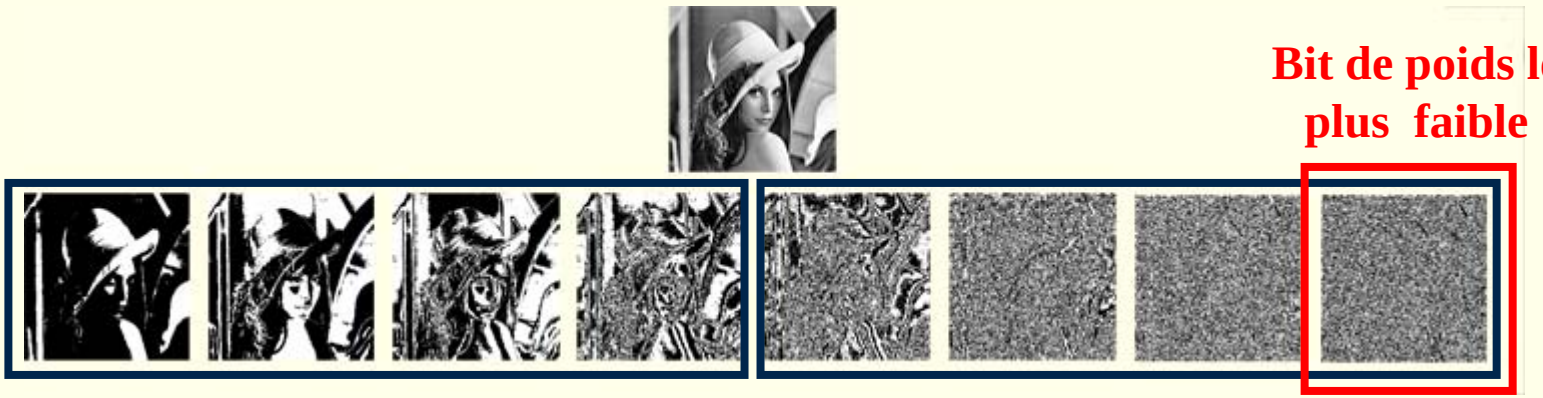
   - **Etalement de spectre**

Domaine Spatial

2.1 Bit de poids faible (1)

- L'information est caché dans le **Bit De Poids Faible.**

⇒ **Pas de changement perceptible** (Watermark invisible)



Exemple simplifié :

- **Inserer un 'A' (en Binaire 01000001, en Decimal 65)**

Avant : 10000000 , 00100100 , 10110101 , 00110101 , 11110011 , 10110111 , 11100111 , 10110011

Aprés : 1000000**0** , 0010010**1** , 1011010**0** , 0011010**0** , 1111001**0** , 1011011**0** , 1110011**0** , 10110011

Bit de poids Le plus faible

2.1 Bit de poids faible (2)

Extrement ***sensible aux modifications et peu sûr !!!***

- Mise en page, rotation, compression, ….

- Le fait d'enlever tout les derniers bits (**zero)** éfface le marquage.

- **Très bon ratio :**

Ex) une image 8-bits de 300x300

300x300 = 90000 bits

2.1 L'algorithme du Patchwork

- **Clé** ⇒ deux parties aléatoires de l'image (**patches**): A & B
- n pixels dans chaque parties (ai & bi sont pair dans A & B)
- **Supposons** (pour *n* suffisament grand): $S = \sum_{i=1}^{n} a_i - b_i = 0$
- A ⇒ ai' = ai + 1
- B ⇒ bi' = bi + 1

⇕ ***Le contraste code l'information***

Decodons avec la même **Clé**

$$S = \sum_{i=1}^{n} a_i' - b_i' = 2n \quad \Rightarrow \textit{Le marquage est present}$$

- Bonne **Invisibilité**
- **Très mauvais ratio** (seulement quelques parties de l'image)
- Robuste aux **Changements d'intensité** (contraste, luminance, Gamma, …)
- Vulnerable aux **Transformations géometriques** (rotation, decoupage, …)

2.2 Domaine Frequentiel

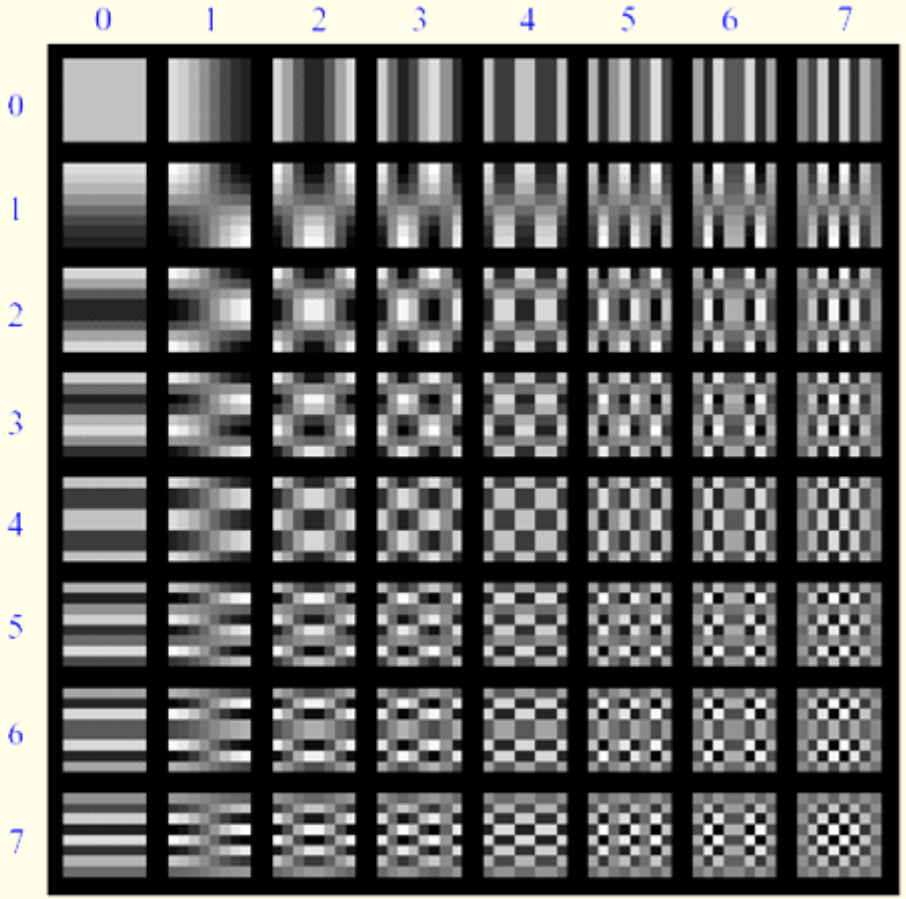
Domaine Fréquentiel

2.2 Domaine Frequentiel

- Plus robuste contre les **COMPRESSION A PERTE (JPG)**
- Robuste contre **LES TRANSFORMATION GEOMETRIQUE** (redimensionement, translation…)

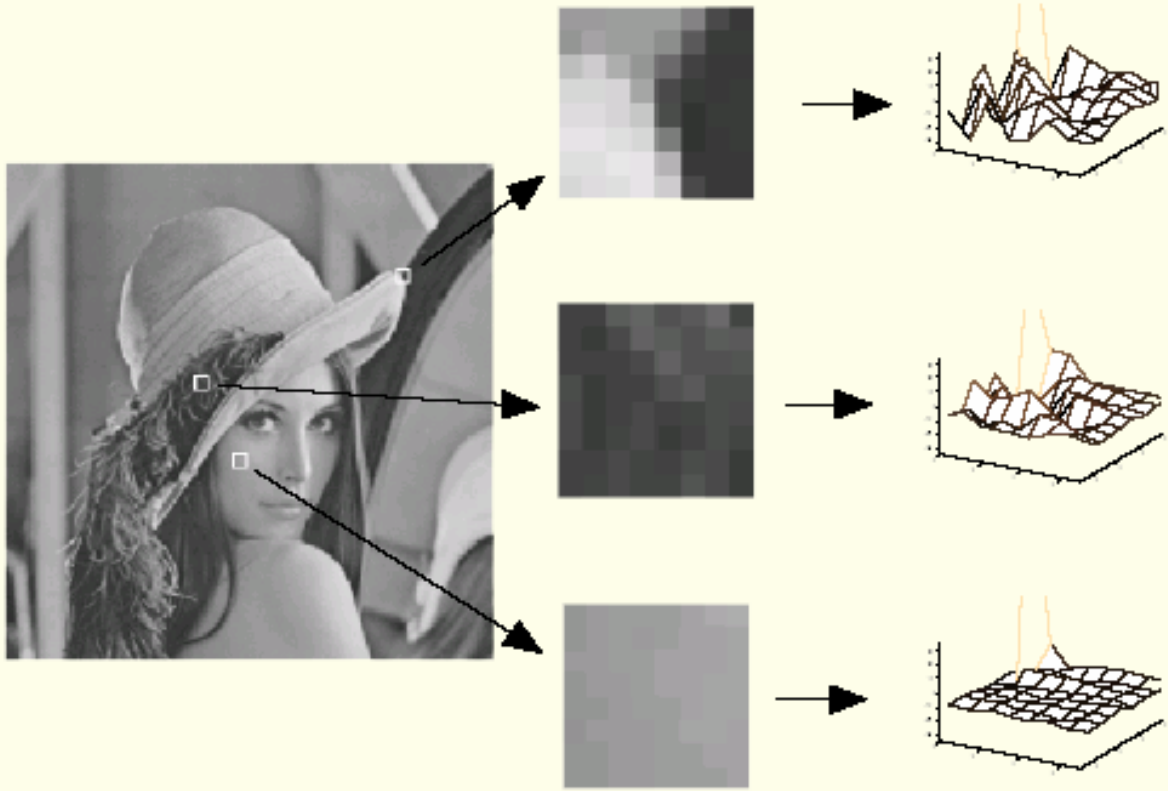
Discrete Cosine Transform

2.2 Domaine DCT (1)



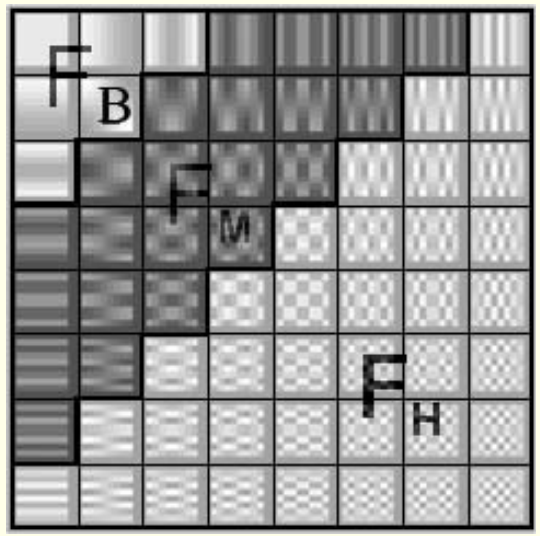
Quelle combinaison linéaire des 8x8 fonctions basiques produisent des blocs de 8x8 pixels dans l'image ?

2.2 Domaine DCT (2)



Exemple de representation fréquentielle de blocs 8x8 d'une image

2.2 Algorithme de Koch & Zhao (1)



• S'utilise sur des **blocks 8x8 DCT** d'une image.

• Marquage sur les bits de frequences **Moyennes**

Pourquoi?

• Les fréquences basses correspondent aux grandes zones homogène

- Robuste mais visible

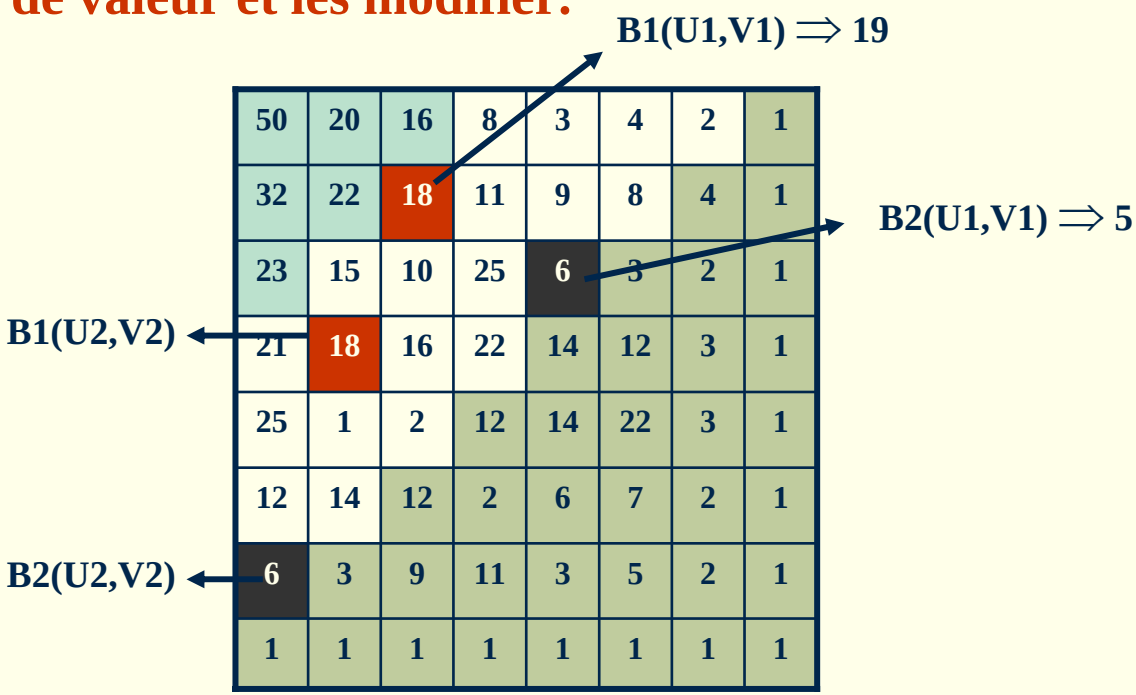
• Les frequences les plus hautes correspondent aux pixels.

- Invisible mais fragile

2.2 Algorithme de Koch & Zhao (2)

Idée basique

• Choisir des zones des blocs frequentiels avec la **même amplitude de valeur et les modifier.**



• **Bi(U1,V1) > Bi(U2,V2) ⇒ "1" sinon"0"**

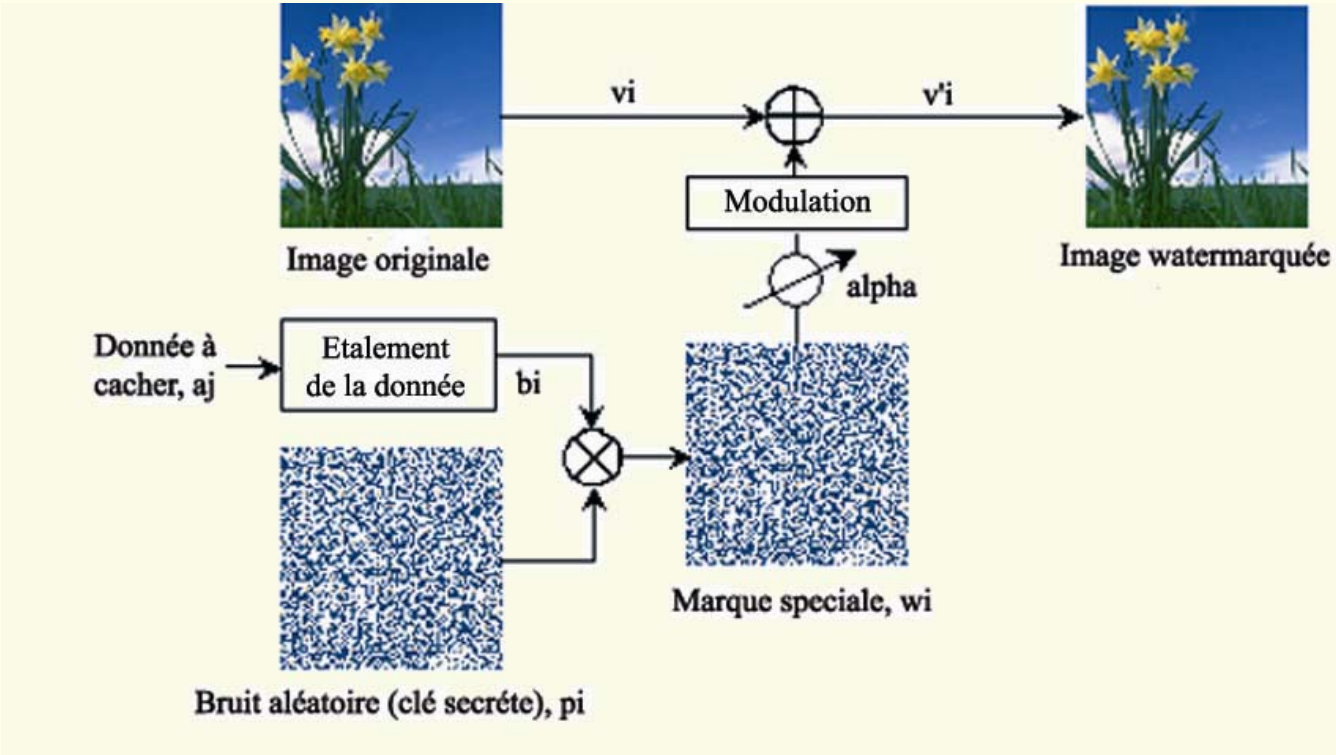
2.2 Algorithme de Koch & Zhao (3)

- Conflit **ROBUSTESSE / VISIBILITE**
- Utilisation de blocks => Ratio **Faible** :

Ex) une image 8-bits de 300x300 (blocks 8*8 = 64)
Bits = 90000/64 = 1400 bits

2.2 L'Etalement de Spectre (1)

- **Technique utilisée dans les telecomunications radio.**
- La donnée a cacher (signal) est **étalée sur une bande de fréquences plus large.**
- La clé secrete : **très peu de chance** d' avoir la même valeur, **unique** pour chaque utilisateur.

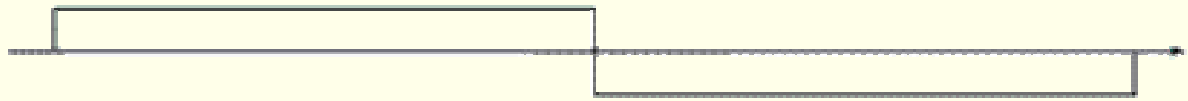


2.2 L'Etalement de Spectre (2)

Insertion

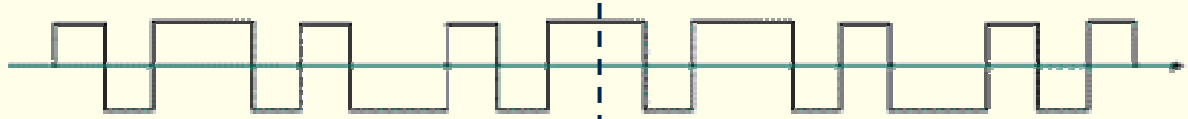
Donnée à cacher originale

Donnée étalée



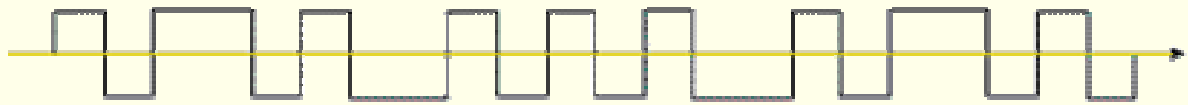
11 fois le signal de base => chiprate de 11

Clé aléatoire (Bruit)



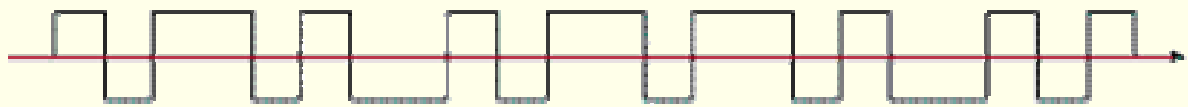
XOR

Marque spéciale

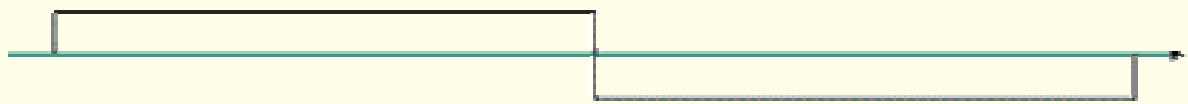


Extraction

Clé aléatoire (Bruit)



Donnée retrouvée (sous forme étalée)



2.2 L'Etalement de Spectre (3)

- Amelioration de la **ROBUSTESSE**

  et de la **SECURITE**

- Conflit **SECURITE / RATIO**

- Ratio **Moyen** :

Ex) une image 8-bits de 300x300 (chip rate = 50)

Bits = 90000/50 = 1800 bits

3. Les attaques

- Qu’est-ce qu’une attaque ?
- Comment cela fonctionne ?
- Les différents types rencontrés :
  - Attaques fréquentielles
  - Attaques géométriques
  - Attaques volontaires

3. Attaques fréquentielles

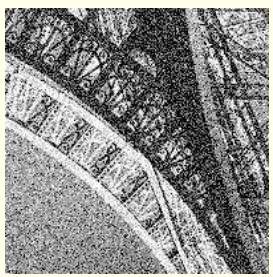
Conversion de format

- **Changement du format des fichiers : Image - JPG, TIFF, GIF, BMP…**
  Video – NTSC/PAL, Frame manipulation…

Compression à perte

Bruit & Filtrage

Bruit Gaussien



Sel&Poivre 30%

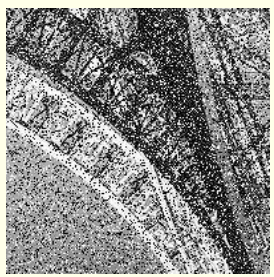
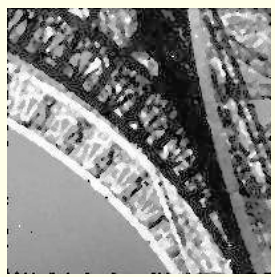


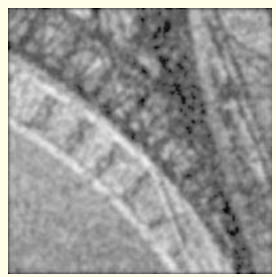
Image originale

⇒

Filtrage Non-Lineaire



Filtrage Lineaire

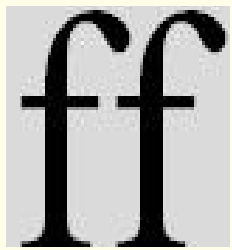


3. Attaques Géométriques

Original

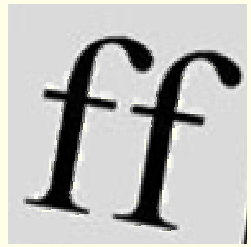
Decoupage

Zoom



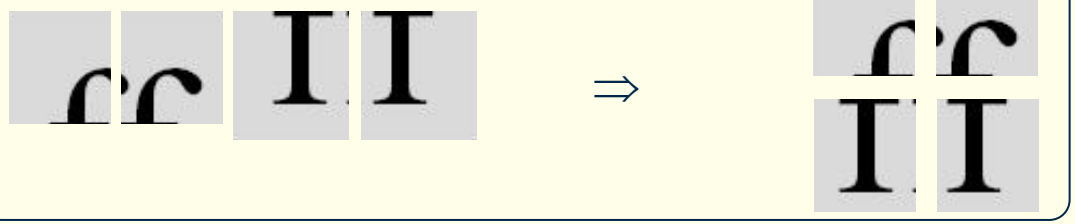
Translation

Rotation



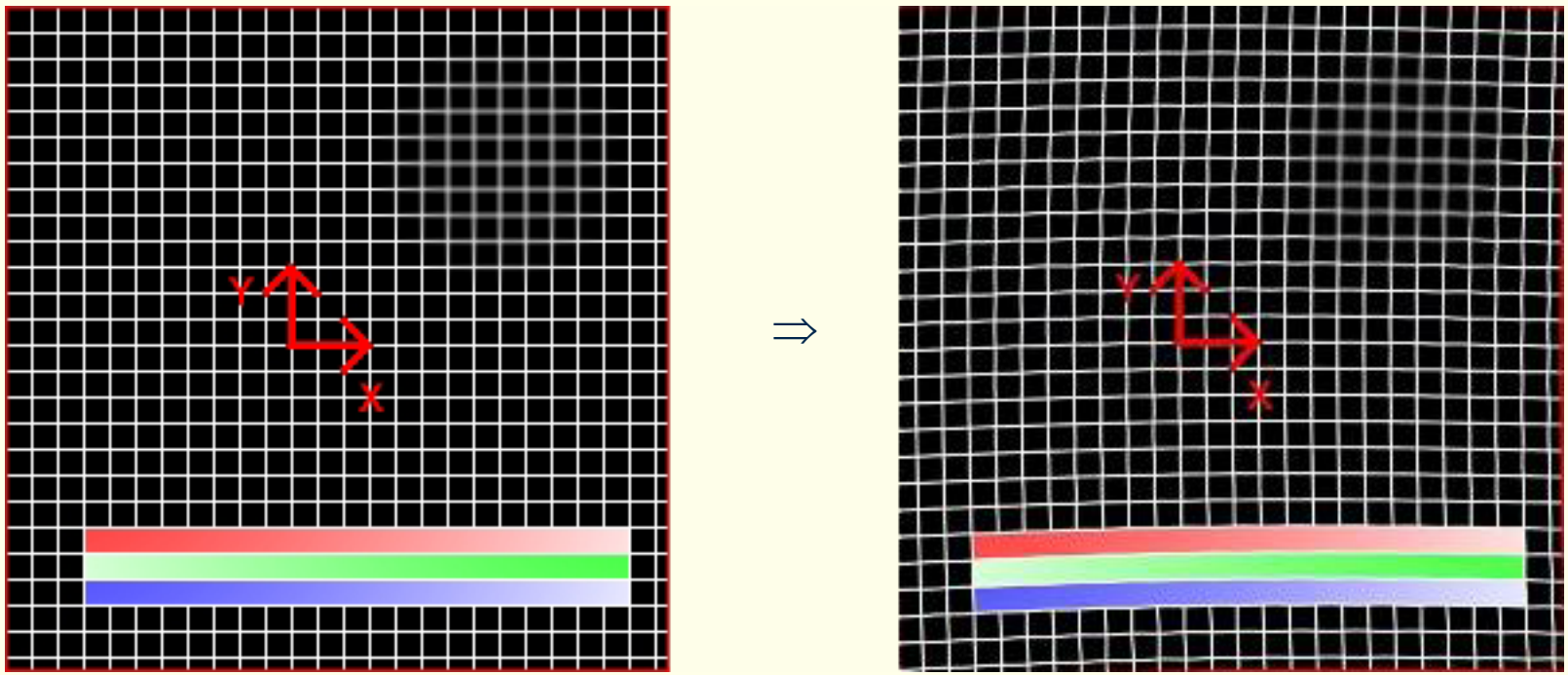
Mosaique d'image

1) Diviser l'image marquée en plusieurs petites images
2) Les rassembler pour recuperer l'information



3. Attaques Volontaires

- **STIRMARK :** Transformation géometrique aléatoire.



4. Conclusion

- **Explosion du domaine**
  - Amélioration des algorithmes
- **Les pirates ont toujours une longueurs d'avance**
- **Difficultés d'appliquer la tatouage d'images au grand public**

Fin