Stéganographie
Tatouage

Lionel Fillatre
Polytech Nice Sophia
2013-2014

Sommaire

- Introduction
- Stéganographie
  - Historique
  - Technique et exemples
- Tatouage
  - Applications
  - Techniques
- L’aspect sécurité
- Conclusion

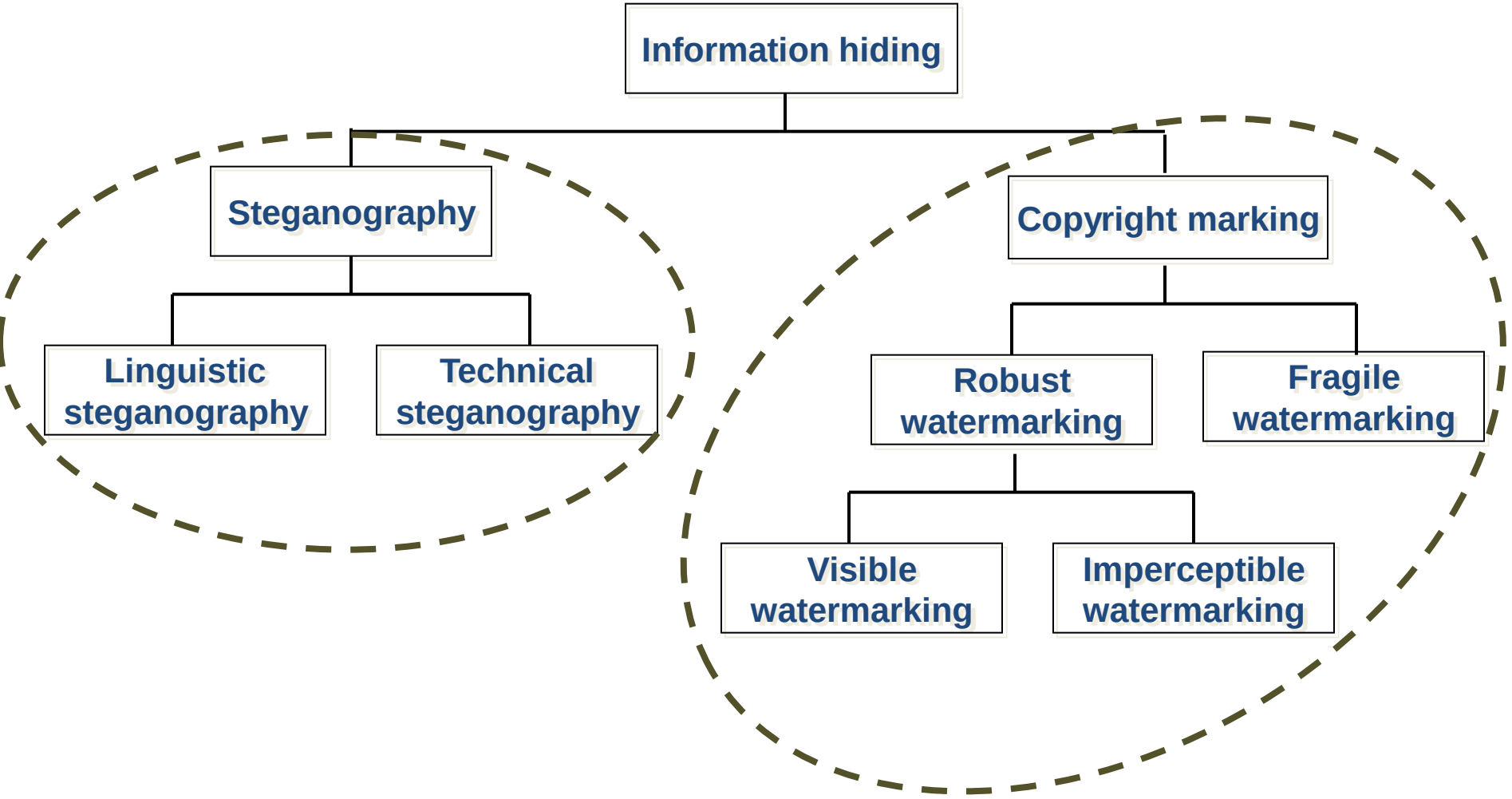
Introduction : motivations

- Transmission d'informations secrètes
  - Information hiding : procédé qui permet de dissimuler des informations à l'intérieur d'une autre source de données
- Avancée rapide des technologies multimédias
- Besoin d'établir des canaux de communications secrets
- Besoins en copyright, authentification

Un peu de vocabulaire

- Trois éléments pour la dissimulation d'informations :
  - Médium de couverture
  - Données
  - Stego-object ou stego-medium

Information Hiding



Crypto, Stégano et Tatouage

- Alice communique avec Bob
- Charlie espionne…
  - Cryptographie : communication sécurisée entre Alice et Bob en chiffrant le message.
    - ✓ *message indéchiffrable*
  - Stéganographie : On dissimule le message dans un autre document. Charlie ne se doute pas qu'ils discutent.
    - ✓ *message imperceptible*
  - Tatouage (Watermarking) : On fait la même chose mais en plus le message est indélébile
    - ✓ *message imperceptible et indélébile*

Des objectifs différents…

- Stéganographie
  - Dissimuler un message secret
  - Médium de couverture n'ayant aucun rapport avec le message
- Watermarking
  - Dissimuler une faible quantité d'information
  - Informations en rapport (direct ou indirect) avec le médium de couverture

Des objectifs différents… (2)

- Stéganographie
  - Clé utilisée pour insérer et **décrypter** un message
  - Opération de sortie : extraire les données du stego-medium
- Watermarking
  - Clé utilisée pour insérer et **détecter** un message
  - Opération de sortie : aucune, ou alors détecter sa présence…

… mais des critères similaires

- Imperceptibilité
- Capacité
- Robustesse
  - Variables suivant l'application visée

Stéganographie

steganós (« étanche »)
et *graphế* (« écriture »).

Des méthodes ancestrales…

- Hérodote :
  - -5000 : Histiée et son esclave
  - Démarate et ses tablettes de cire
- Enée le Tacticien
  - « piquer » les lettres dans un texte
- XVI° siècle : G. Porta
  - encre sympathique (vinaigre sur œuf dur)

… Encore utilisées !

- Épistoliers anglais
  - Faire des économies !
- 1ère guerre mondiale
  - Même procédé pour communiquer…
- 2nde guerre mondiale
  - Encre sympathique dans les journaux
- Billets de banque…

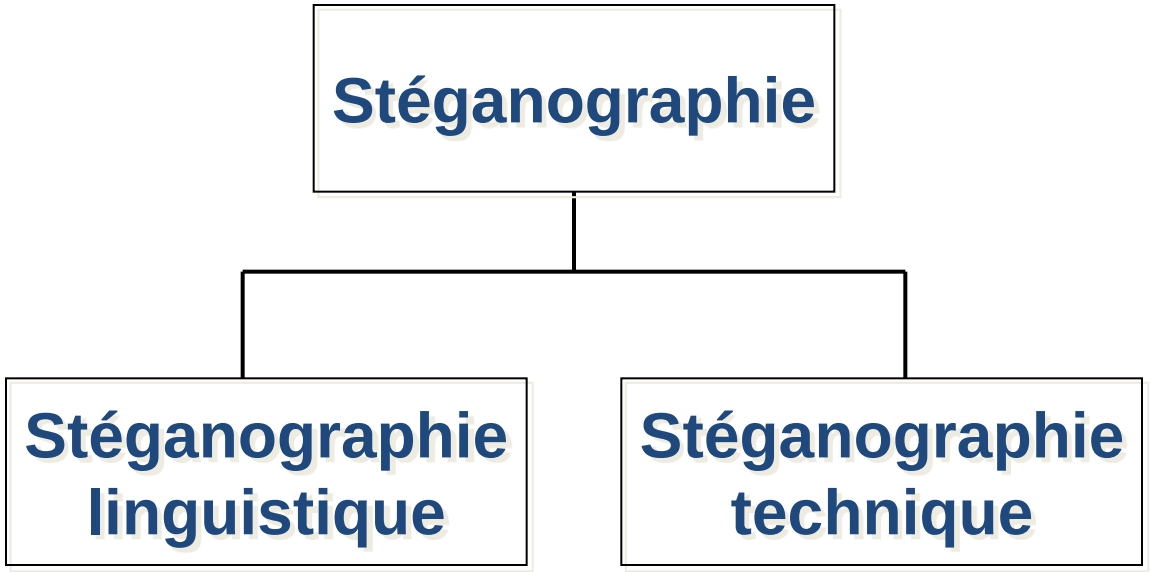
Les systèmes de stéganographie

- Stéganographie pure :
  - même algorithme pour Alice et Bob, pas d'entente préalable.
- Stéganographie à clef secrète :
  - Alice et Bob conviennent d'une clef secrète.
- Stéganographie à clef publique :
  - Utilisation de la clef publique de Bob par Alice pour cacher son message. Bob utilise sa clef privée pour l'extraire.

Les critères de la stéganographie

- La clef : elle détermine la position du message caché.
- Possibilité de cacher un message préalablement crypté.
- Robustesse : peu importante
  - médium non modifié
  - imperceptibilité et capacité importantes

Les types de stéganographie



Stéganographie linguistique

- 1313 : les 3 sonnets de Boccaccio
  - Les lettres des 3 sonnets : premières lettres d'autres poèmes.
- Georges Sand et Alfred de Musset
- Tintin

Stéganographie technique

- Toutes les techniques qui ne jouent pas sur les mots
- Les exemples de l'historique en font partie
- Permet la dissimulation de données dans différents types de média

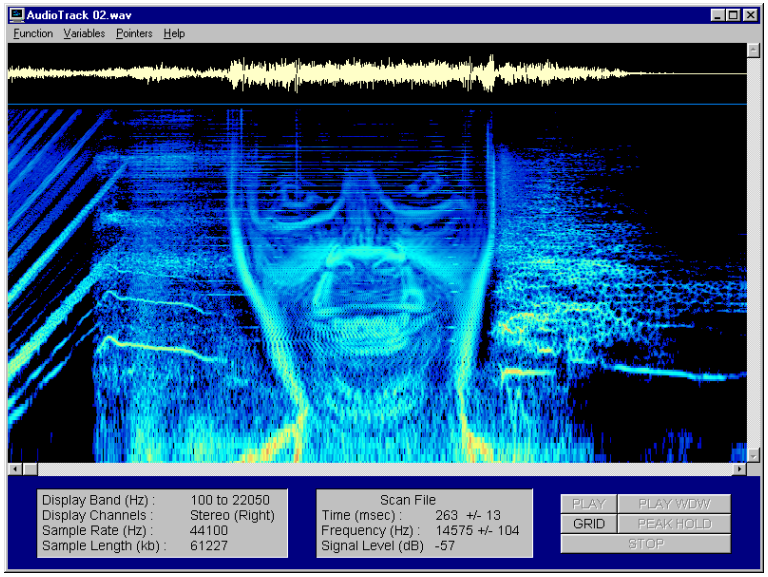
Dans le texte…

- Trous sous les lettres important*es*
- Changement du type d'écriture : Franc*is* Bacon
- Utilisation des synonymes
- Changement des règles de grammaire
- Jeu sur les espaces
- Inconvénients :
  - Faibles capacités
  - Fastidieuses…
- V*a*riante : formatage du texte… peu sû*r*

Dans le son

- Problème : oreille humaine plus sensible que les yeux…
- Médium beaucoup moins employé que la vidéo ou l'image.
- Ex : Groupe techno

  Aphex Twin



Dans les images

- 6 grandes approches :
  - Substitution
  - Transformations du support (JPEG par exemple)
  - Étalement de spectre
  - Méthodes statistiques pour modifier les caractéristiques du support
  - Distorsion du support de dissimulation
  - Génération du support de dissimulation en fonction du secret

Modification des bits de poids faible

image originale

R=254 = 11111110_2

G=183 = 10110111_2

B=156 = 10011100_2

modification des LSB

R=255 = $1111111\ \underline{\mathbf{1}}_2$

G=182 = $1011011\ \underline{\mathbf{0}}_2$

B=157 = $1001110\ \underline{\mathbf{1}}_2$

modification du MSB vert

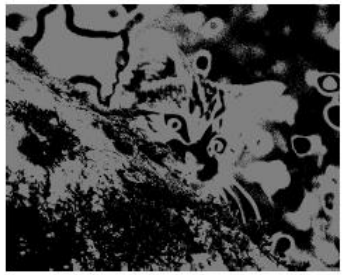
R=254 = 11111110_2

G=55 = $\underline{\mathbf{0}}0110111_2$

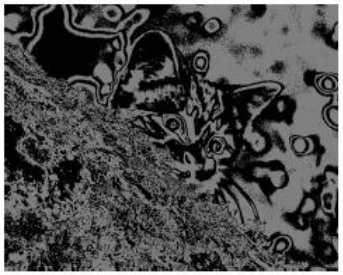
B=156 = 10011100_2

Plan de bits

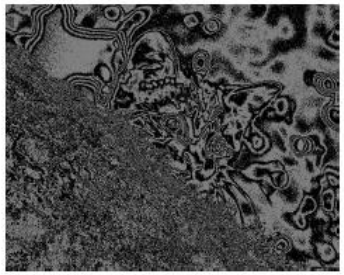
Plan de bit $i = 7$



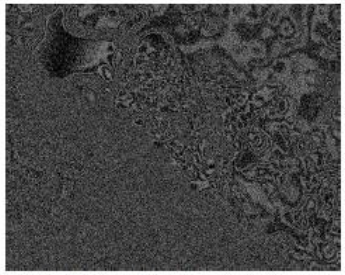
Plan de bit $i = 6$



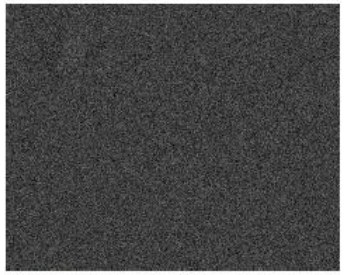
Plan de bit $i = 5$



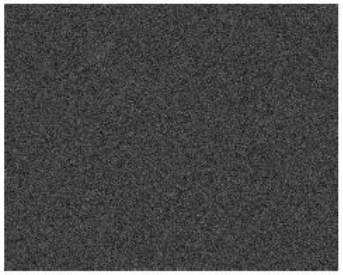
Plan de bit $i = 4$



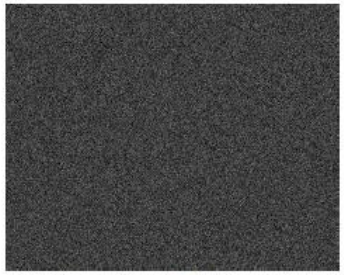
Plan de bit $i = 3$



Plan de bit $i = 2$



Plan de bit $i = 1$



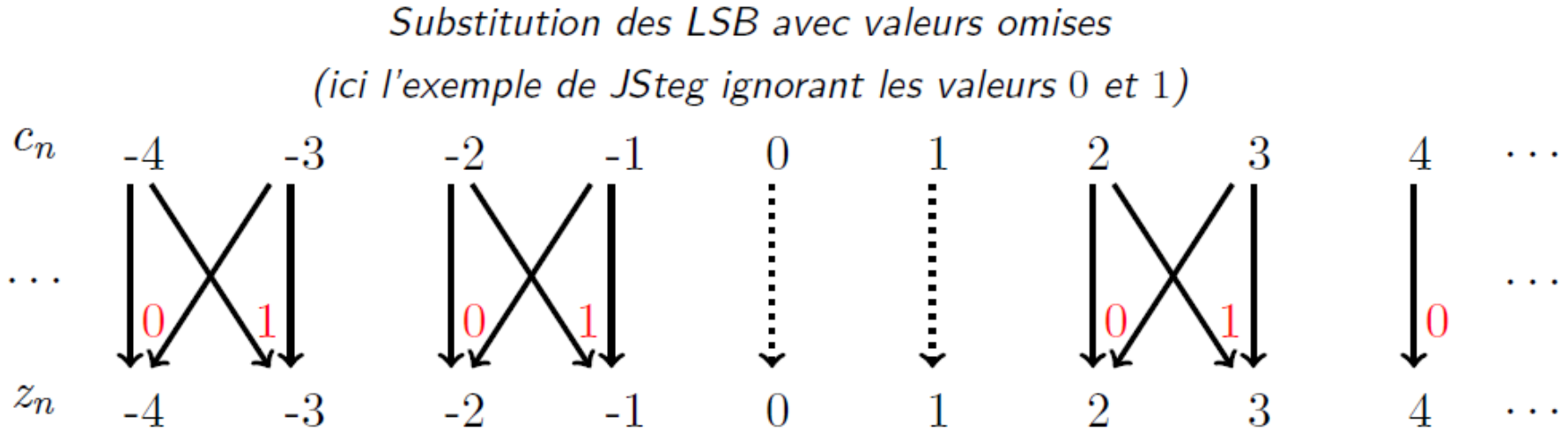
Plan de bit $i = 0$

Deuxième exemple

- Par DCT :
  - La clef détermine 2 coefficients d'un tableau 8x8 obtenu par DCT, modification de ces valeurs, DCT inverse

70	-34	20	-8	-11	7	-3	3	186	-18	15	-9	23	-9	-14	-19	70	-34	20	-6	-11	9	-3	3
68	-25	22	-11	1	0	-7	-6	21	-34	26	-9	-11	11	14	7	68	-25	22	-13	1	1	-7	-6
83	-15	-1	4	17	-8	-5	-15	-10	-24	-2	6	-18	8	-20	-1	83	-15	-1	2	17	-9	-5	-15
74	9	2	-1	17	1	16	-5	-8	-5	14	-15	-8	-3	-3	8	74	9	2	1	17	-1	16	-5
62	7	17	-12	9	5	5	4	-3	10	8	6	-11	18	18	15	62	7	17	-10	9	3	5	4
63	6	4	-9	13	-19	-11	-10	4	-2	-18	8	8	-4	1	-7	63	6	4	-11	13	-20	-11	-10
53	-5	15	3	17	-11	-22	-17	9	1	-3	4	-1	-7	-1	-2	53	-5	-15	1	17	-10	-22	-17
54	6	-7	9	2	0	-13	-7	0	-8	-2	2	1	4	-6	0	54	6	-7	11	2	2	-13	-7

Exemple de JSTEG



Troisième exemple : stégocode

Utilisation du code de Hamming [3,1,3] aussi appelé code à triple répétition.

Ce code contient deux mots 000 et 111.

Transmission de 01110001

C	5	2	3	6	4	1	2	1	3	5	3	1
LSB	1	0	1	0	0	1	0	1	1	1	1	1

$$\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}\times\begin{pmatrix}1\\0\\1\end{pmatrix}\oplus\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \qquad \begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}\times\begin{pmatrix}0\\0\\1\end{pmatrix}\oplus\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

$$\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}\times\begin{pmatrix}0\\1\\1\end{pmatrix}\oplus\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix} \qquad \begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}\times\begin{pmatrix}1\\1\\1\end{pmatrix}\oplus\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$$

Décodage

C	5	2	3	6	4	1	2	1	3	5	3	1
LSB av. modif	1	0	1	0	0	1	0	1	1	1	1	1
LSB ap. modif	1	0	1	0	0	1		1	1	1		1
S	5	2	3	6	4	1		1	3	5		1

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Dans la vidéo

- Même techniques que pour les images

- Avantages :
  - Durée…
  - Souvent plus bruitées
    - facilite l'imperceptibilité

Tatouage

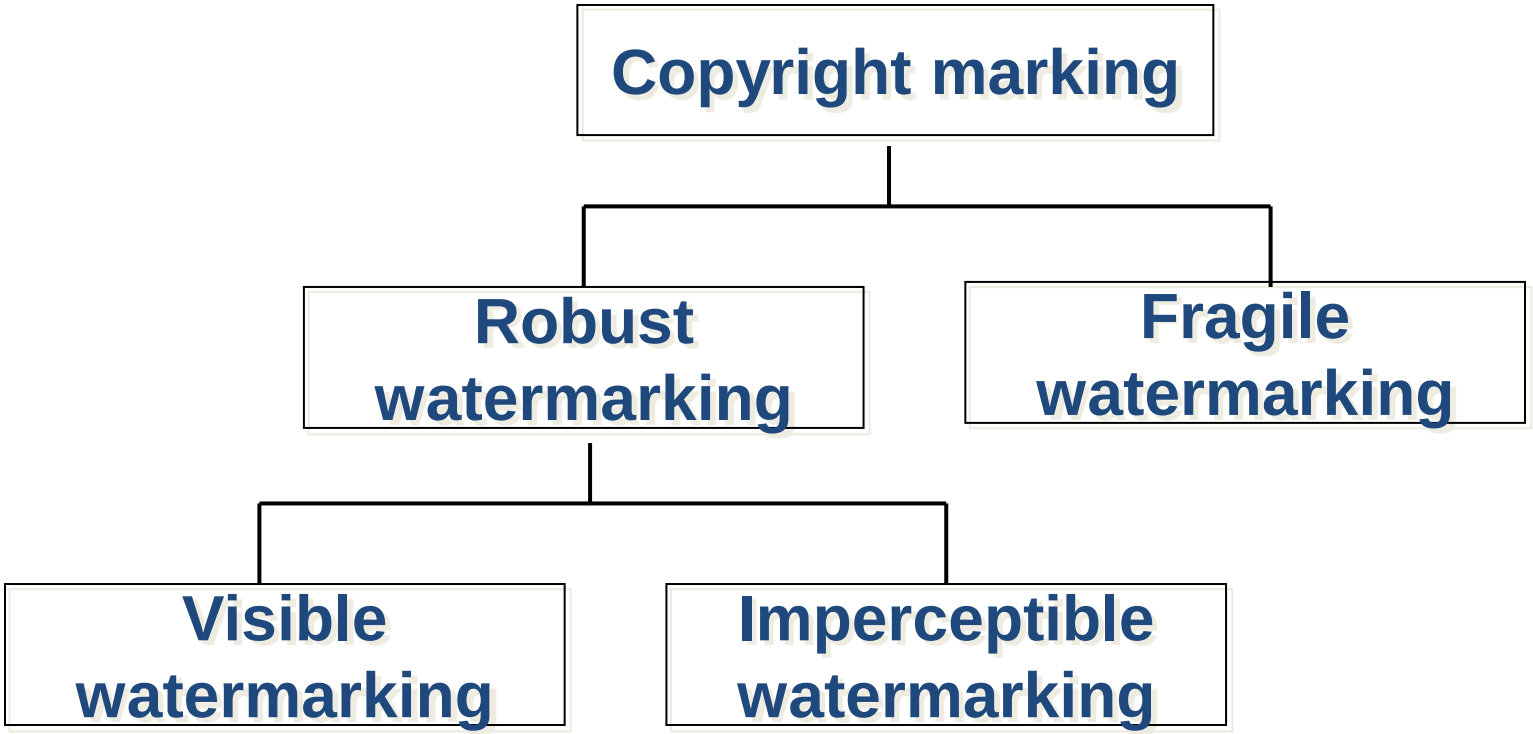
ou Watermarking

Tatouage

- Premiers exemples au XVII° siècle
  - Fabricants de papier

- Aujourd'hui :
  protection du copyright

Les types de watermarking



La robustesse

- Caractérisée par la résistance aux attaques
- Bon tatouage = on ne peut pas le supprimer sans dégrader le support
- Attention à l'algorithme utilisé

Applications du tatouage

- Indexation, extension d'un média
  - Ex : sous-titrage
- Intégrité
  - Tatouage fragile pour détecter les modifications du médium
- Protection des droits d'auteurs : « fingerprinting »
  - Prouver l'appartenance d'une œuvre
  - La plus utilisée, mais aussi la plus attaquée

Les schémas de tatouage

- Privé
  - Médium initial, marque, clef fournie
  - Comparaison original / stego objet
- Semi aveugle
  - Fonction de détection : marque et clef
- Aveugle
  - Connaît uniquement la clef secrète
- Asymétrique
  - Ne nécessite aucune connaissance particulière

Détection de tatouage

- **faux positif :** Survient lorsque le détecteur indique qu'une marque est présente alors qu'il y en a pas.
- **faux négatif :** Survient lorsque le détecteur indique qu'une marque est abscente alors qu'il y en a une.

Dans une application de **surveillance de diffusion télévisuelle** où les publicités sont tatouées, un **faux négatif** mène à la conclusion que la publicité n'a pas été diffusée...

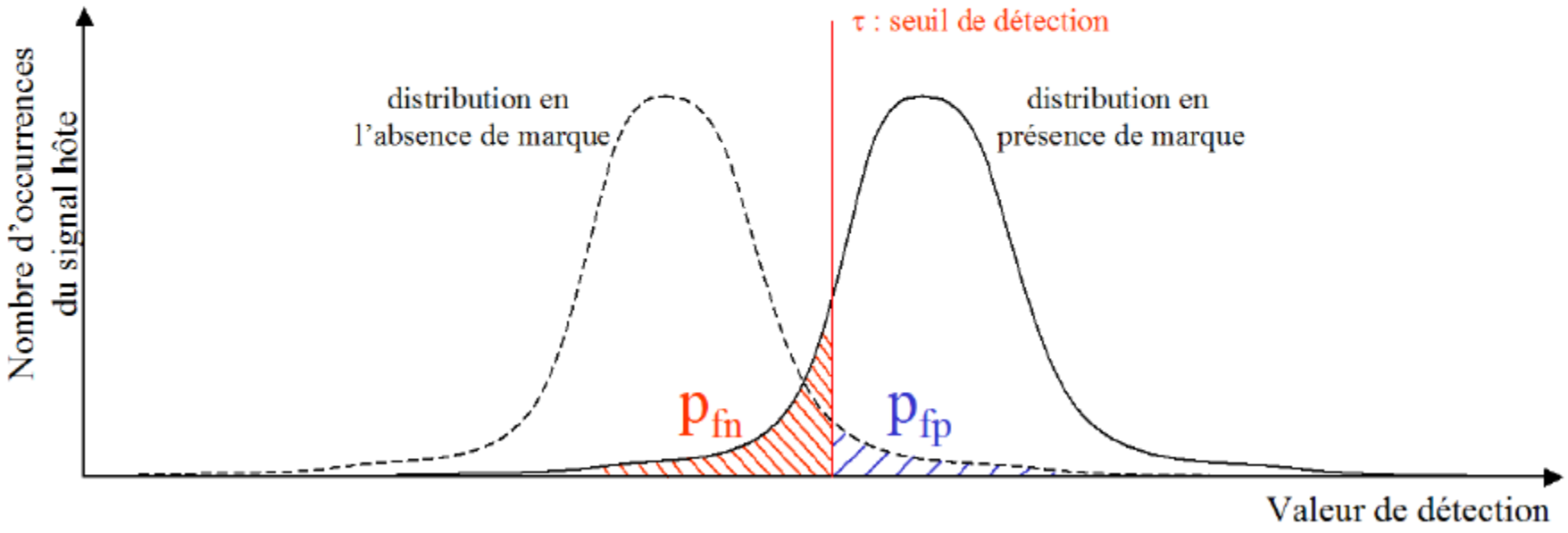
Dans une application où la présence de la marque autorise la **lecture d'un DVD**, un **faux négatif** empêche la lecture d'un DVD censé être autorisé.

Dans une application de preuve d'appartenance, la détection d'un faux positif mène à une accusation de vol de l'œuvre d'un ayant-droit (alors que la marque n'a pas été insérée).

Valeurs typiques de faux-positif

- Dans le cas d'une application de preuvre d'appartenance, le détecteur étant rarement utilisé, une probabilité de **faux-positif** de 10^{-6} doit suffire.
- Dans le cas d'une application de contrôle de copies, des millions de détecteurs tournent constament sur des millions de documents. La probabilité de **faux-positif** doit être de 10^{-12} (1 erreur pour 1000 années de calcul continu).

Principe de détection

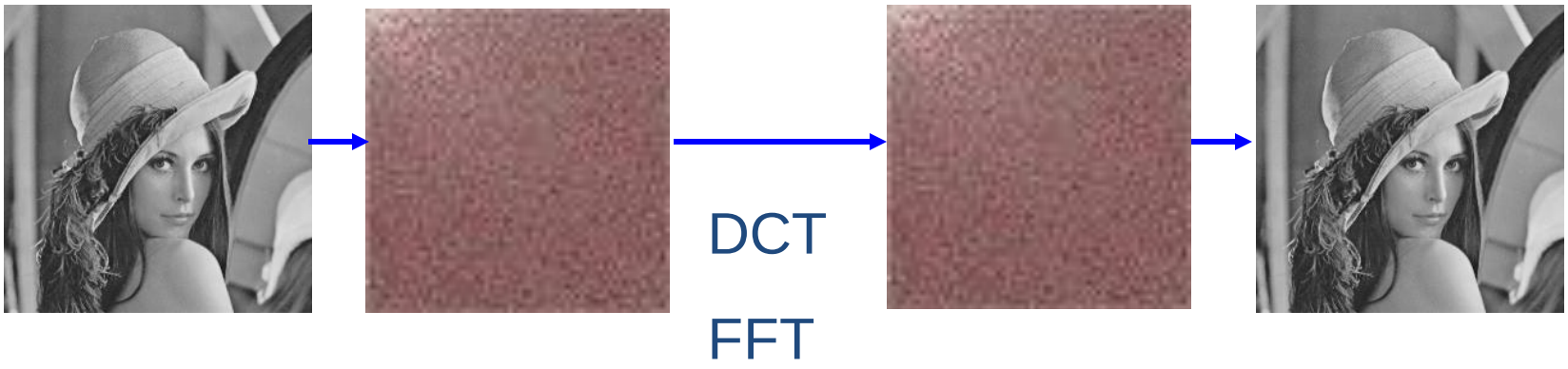


Deux domaines de tatouage

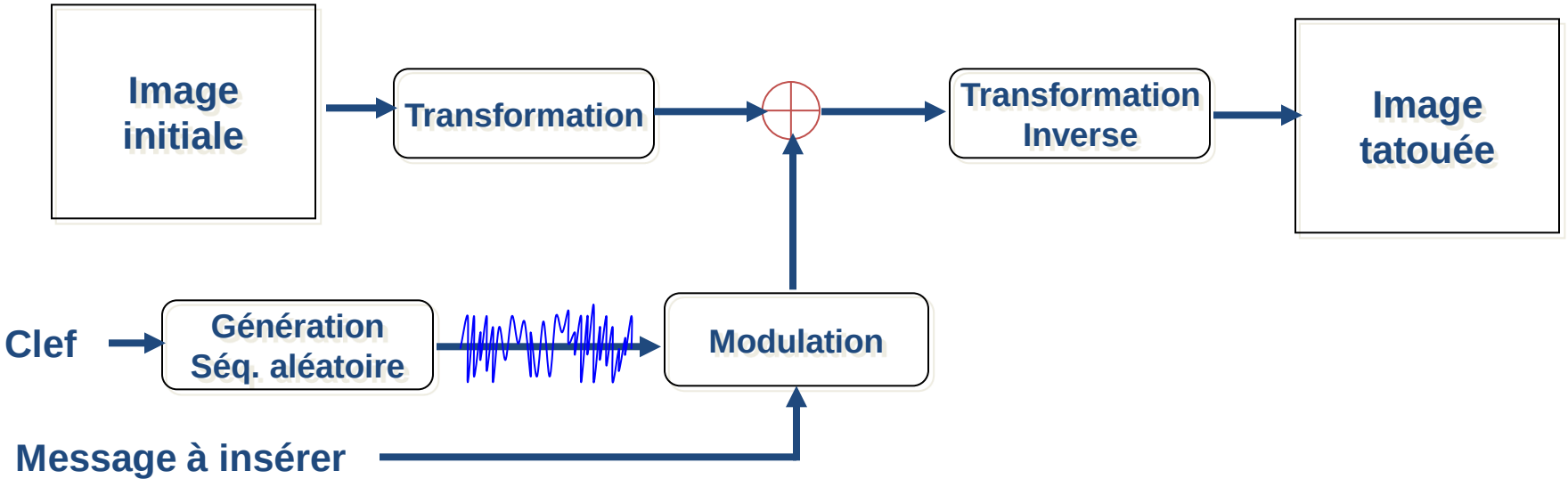
- Domaine spatial :
  - Agit directement sur les pixels
  - Méthode rapide

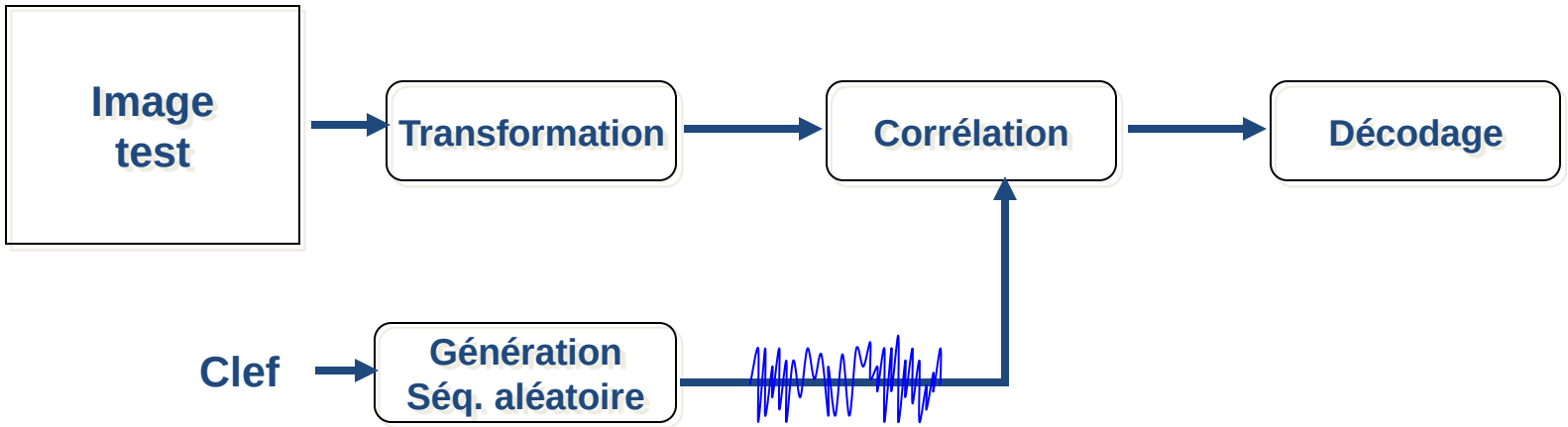
Deux domaines de tatouage

- Domaine fréquentiel
  - Utilisation de DCT ou FFT
  - Méthode plus lourde

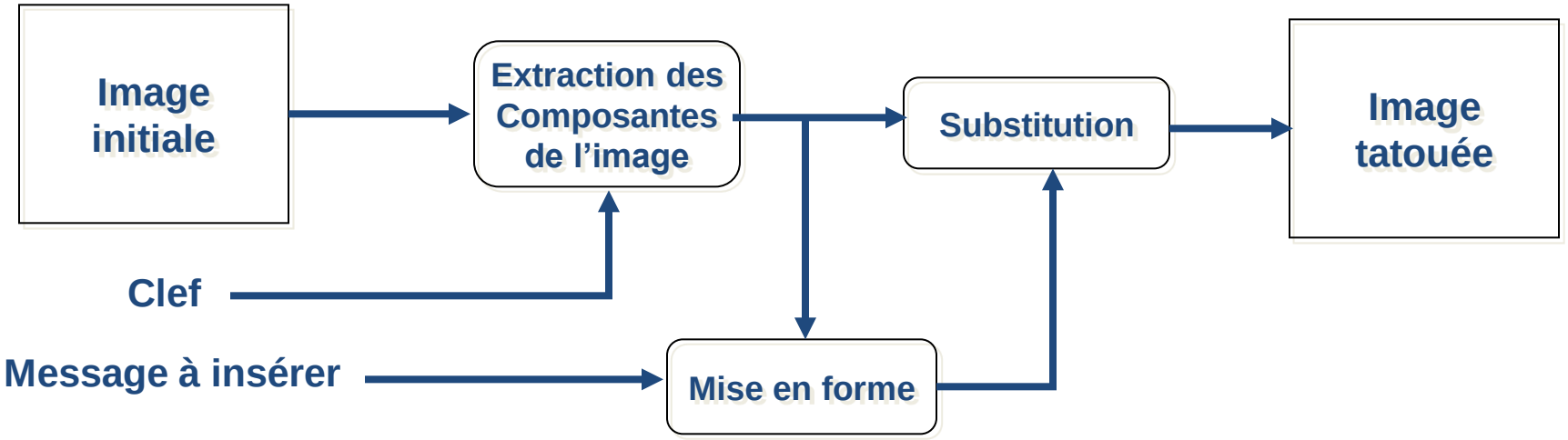


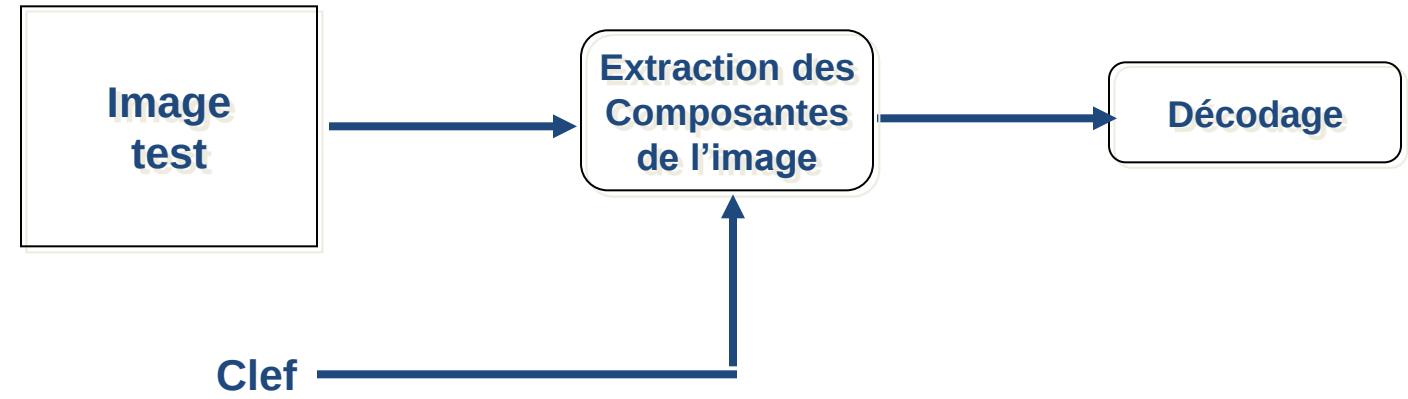
Les tatouages additifs



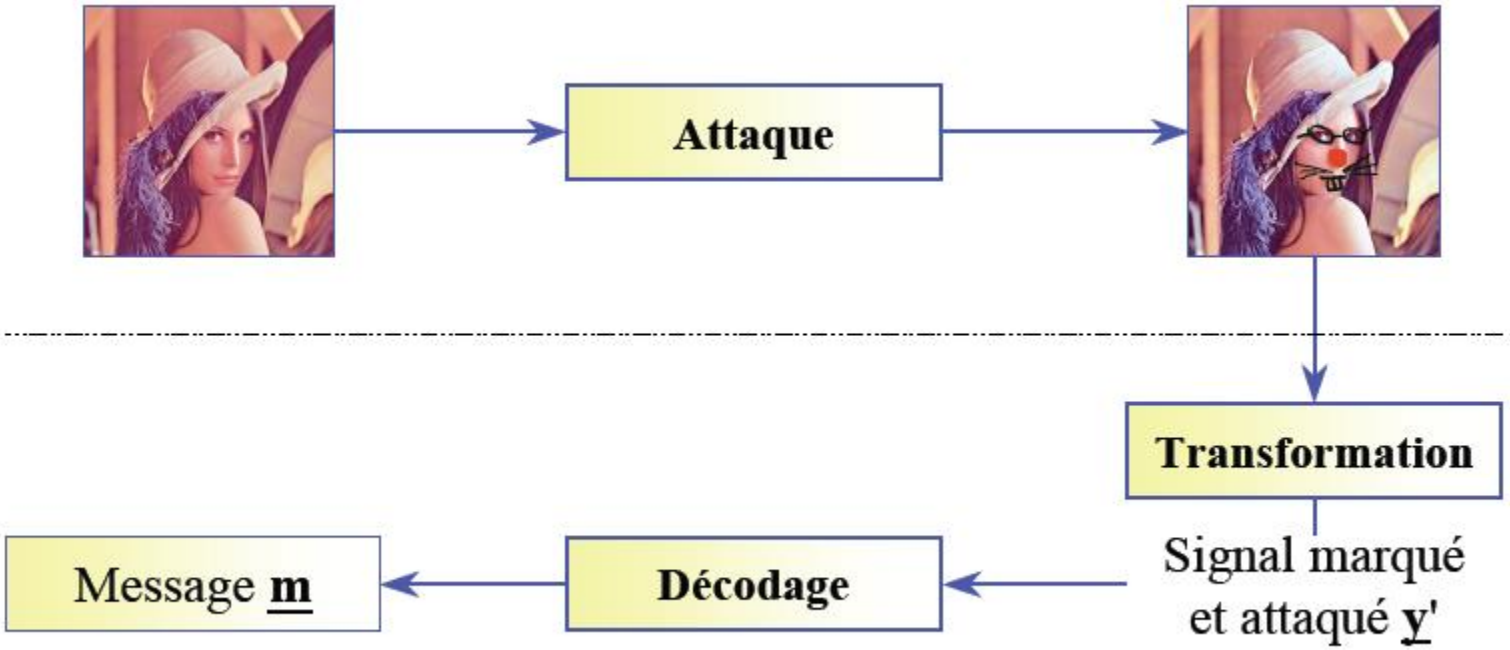


Les tatouages substitutifs



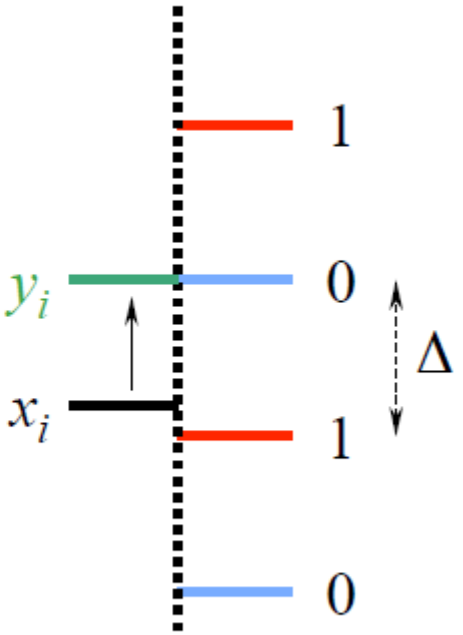


Attaque et extraction



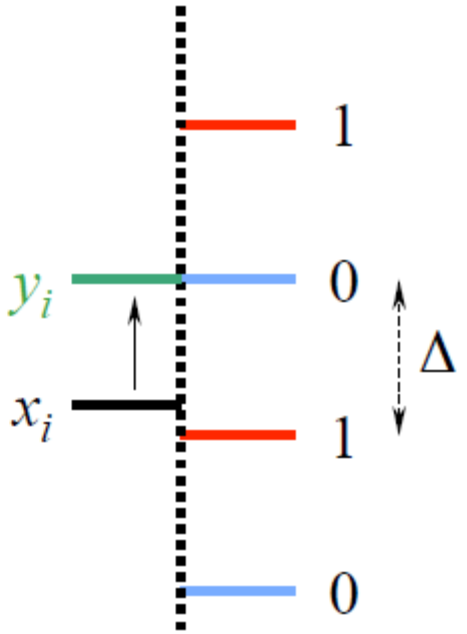
Technique « basique » : quantification

- Les données x_i sont quantifiées et à chaque étape est associé un symbole
- La donnée est modifiée pour correspondre à l'étape la plus proche
- Extraction = recherche de l'étape la plus proche



Technique « basique » : quantification

- Le pas Δ doit être connu à l'extraction
- Pas de résistance au changement d'échelle
  $\mathbf{y}' = \alpha\,\mathbf{y}$
  - Filtrages
  - Contraste, volume…
  - Compression



Insertion aveugle de 1 bit

- Le message m est un unique bit (0 ou 1).
- Soit w_m générée à partir d'un unique pattern (porteuse) w_r de la même taille que l'image c_o. Ce pattern w_r est généré pseudo-aléatoirement via une clef secrête. On a :

$$w_m = \begin{cases} w_r \ si \ m = 1 \\ -w_r \ si \ m = 0 \end{cases}$$

- La marque est alors définie par $w_a = \alpha w_m$. Le scalaire α permet de contrôler la **force d'insertion** de la marque.
- Finalement, le tatouage est réalisé comme ceci : $c_w = c_o + w_a$.

Détection aveugle

Détection aveugle :

Pour détecter la marque, il faut détecter $\pm w_r$ en présence du bruit causé par le signal hôte c_o et le bruit n. La manière optimale pour détecter ce signal en présence de bruit additif Gaussien est de calculer la corrélation linéaire entre l'image reçue c_{wn} et le pattern w_r :

$$z_{lc}(c_{wn}, w_r) = \frac{1}{N} c_{wn}.w_r = \frac{1}{N} \sum_{i=1}^{N} c_{wn}[i].w_r[i]$$

Principe de la détection

Une justification ("intuitive") de l'utilisation de la corrélation linéaire :
Si $c_{wn} = c_o + w_a + n$, alors
$z_{lc}(c_{wn}, w_r) = \frac{1}{N}(c_o.w_r + w_a.w_r + n.w_r)$
Avec l'hypothèse que c_o et n suivent une distribution Gaussienne, $c_o.w_r$ et $n.w_r$ ont de fortes chances d'être de faible amplitude. Au contraire, $w_a.w_r \pm \alpha w_r.w_r$ doit être de forte amplitude. D'où :

$$z_{lc}(c_{wn}, w_r) \approx \alpha w_r.w_r/N \ si \ m = 1$$
$$z_{lc}(c_{wn}, w_r) \approx -\alpha w_r.w_r/N \ si \ m = 0$$

Sortie du détecteur

Sortie du détecteur :

$$m_n = \begin{cases} 1 & si\ z_{lc}(c_{wn}, w_r) > \tau_{lc} \\ pas\ de\ marque & si\ -\tau_{lc} \leq z_{lc}(c_{wn}, w_r) \leq \tau_{lc} \\ 0 & si\ z_{lc}(c_{wn}, w_r) < -\tau_{lc} \end{cases}$$

Exemple de résultats (4000 images)

le pattern (porteuse) est obtenu avec un distribution uniforme normalisé et α est mis à 1.

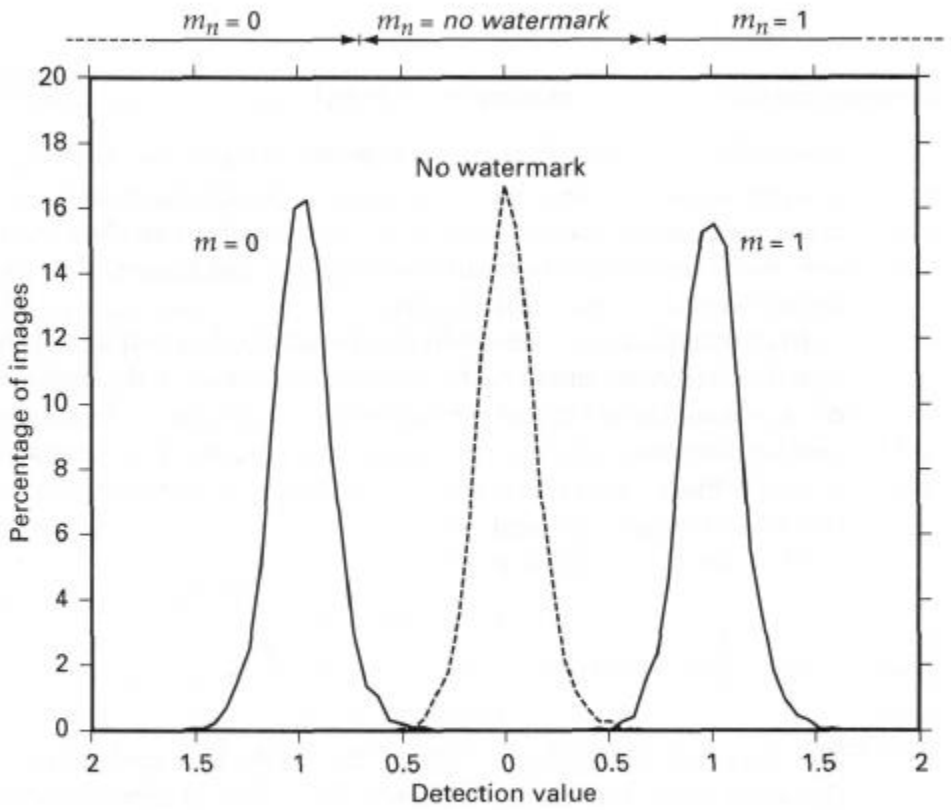


Fig. 3.6 Distributions of linear correlation values resulting from `System 1` (`E_BLIND/D_LC`) with a white noise reference pattern. The left-hand curve is the distribution when the embedded message was 0. The right-hand curve is the distribution when it was 1. The dashed curve is the distribution when no watermark was embedded. The legend at the top of the graph shows how the linear correlation decoder (`D_LC`) maps correlation values into messages.

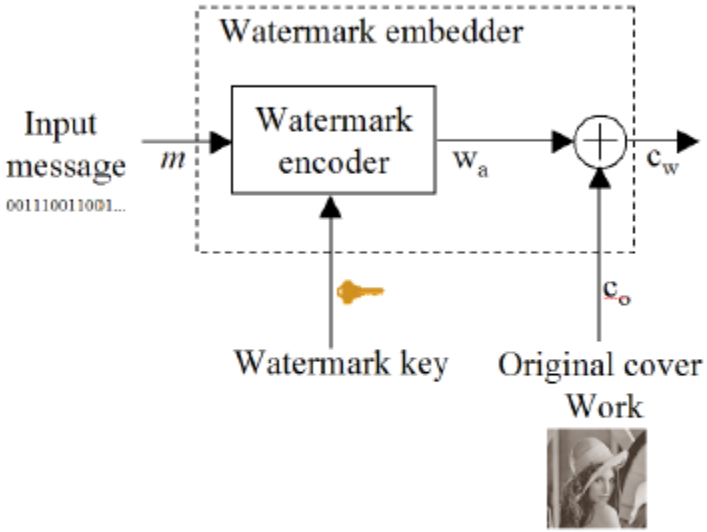
Etalement de spectre

La technique provient du monde des télécommunications. Un message m est composé d'un ensemble de symboles $m[i]$ ($m[i]$ vaut bien souvent 0 ou 1). Chaque symbole $m[i]$ est transmis à travers un signal appelé **porteuse** et noté u_i. Une porteuse est un signal pseudo-aléatoire (obtenu par un GNPA) pouvant être composé de 0 et de 1 ou bien distribué suivant une loi Gaussienne normale $\mathcal{N}(0, 1)$. On peut également contraindre les porteuses à être orthogonales ($\forall i$, $\forall j$, $u_i.u_j = 0$).

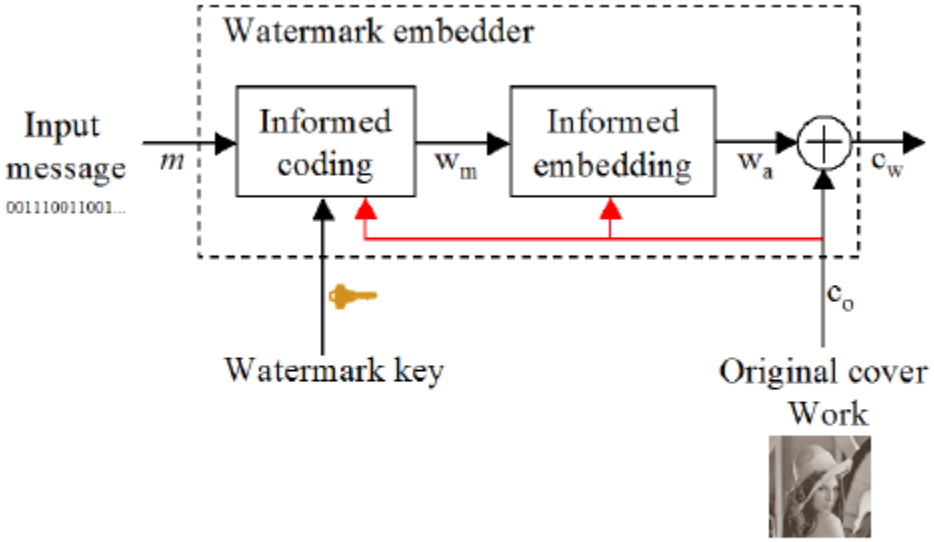
Formalisation

- Insertion
  - u_i : un vecteur (porteuse) de la taille du signal hôte N,
  - m : un message composé de N_c bits,
  - s : une fonction (appelée modulation) $\{0,1\} \rightarrow \mathbb{R}$.
    Par exemple, $s(m[i]) = \gamma(-1)^{m[i]}$ avec γ un facteur réglant l'ampleur de la distorsion,
  - La marque est alors $w_r = \sum_{i=1}^{N_c} u_i.s(m[i])$,
  - L'insertion est $c_w = c_0 + w_r$.
- Détection
  - Soit c_{wn} un signal tatoué attaqué : $c_{wn} = c_w + n$,
  - Le message extrait est $\hat{m}[i] = \text{sign}(c_w.u_i)$
    où $\text{sign}(x) = 0$ si $x > 0$ et $\text{sign}(x) = 1$ si $x \leq 0$.

Tatouage aveugle contre tatouage informé



Tatouage aveugle



Tatouage informé

- la solution par étalement de spectre a été vue dans le premier exemple du cours (insertion d'un seul bit 0 ou 1).
- la solution par étalement de spectre n'est pas une solution informée. Il existe des solutions informées (Informed Spread Spectrum, ...)
- Dans le domaine fréquentiel, les porteuses vont modifier un grand nombre d'échantillons fréquentiels du signal hôte. Si le signal (tatoué) subit des dommages sur une fraction des fréquences (filtre passe-bande, ...), les porteuses restent identifiable et le message embarqué peut-être extrait correctement.
- caractérisation du spread spectrum : l'energie insérée dans 1 fréquence est très faible. Il y a donc peu de dégradations perceptuelles et il y a une bonne robustesse aux distortions grâce à la dispersion.

Mesure de distorsion (Watermak to Content Ratio)

- On peut quantifier la distorsion d'insertion du tatouage par le rapport signal-à-bruit entre la marque et le vecteur hôte

$$\mathrm{WCR[dB]} = 10 \log_{10} \left(\frac{||\mathbf{w}||^2}{||\mathbf{x}||^2} \right)$$

- On utilise aussi l'erreur quadratique moyenne...

$$\mathrm{MSE}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m} (\mathbf{x}[i] - \mathbf{y}[i])^2$$

- ... À partir de laquelle on calcul le PSNR

$$\mathrm{PSNR[dB]} = 10 \log_{10} \left[\frac{d^2}{\mathrm{MSE}} \right]$$

Distorsion perceptuelle

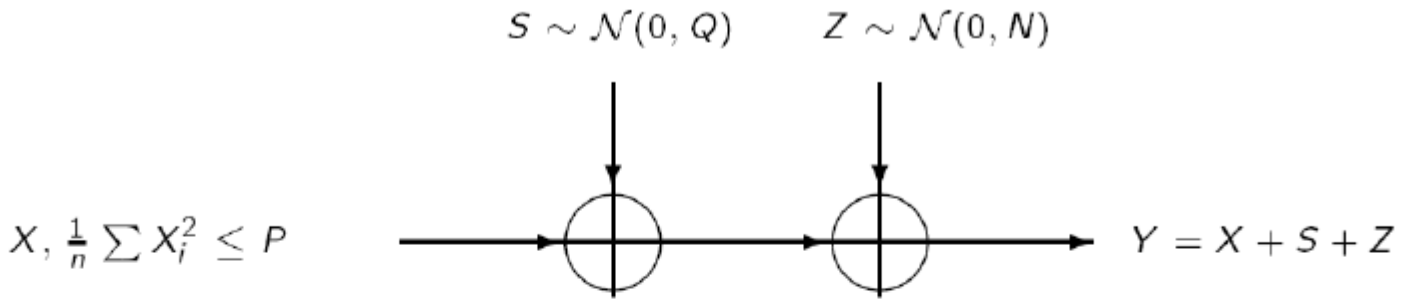
- Mesure classique de l'erreur quadratique moyenne entre vecteur hôte et vecteur marqué :

$$\mathrm{EQM} = \frac{1}{m} \sum_{i=1}^{m} (\mathbf{x}[i] - \mathbf{y}[i])^2$$

- On y ajoute une pondération perceptuelle sous la forme d'un vecteur $\mathbf{p}$, ce qui donne la distorsion d'insertion :

$$\mathrm{D}_{\mathrm{xy}} = \frac{1}{m} \mathbb{E} \left[\sum_{i=1}^{m} \mathbf{p}[i]^2 (\mathbf{x}[i] - \mathbf{y}[i])^2 \right] = \frac{n}{m} \sum_{i=1}^{m} \mathbf{p}[i]^2 \mathbf{a}[i]^2$$

Tatouage informé (Costa)



Si S est inconnu de l'émetteur et du récepteur, alors

$$C = \frac{1}{2}\log(1 + \frac{P}{N+Q})$$

Si S est connu de l'émetteur, alors

$$C = \frac{1}{2}\log(1 + \frac{P}{N})$$

Analyse et sécurité

Stéganographie

- Application des principes de Kerschoff :
  - La sécurité doit **reposer sur la clef et non sur l'algorithme**
- Objectifs :
  - Effacer les données cachées
  - Lire les données
  - Changer les données

Stéganographie

- Les types d'attaques :
  - Stego-only attack
  - Known cover attack
  - Known message attack
  - Chosen steganography attack
  - Chosen message attack
  - Known steganography attack

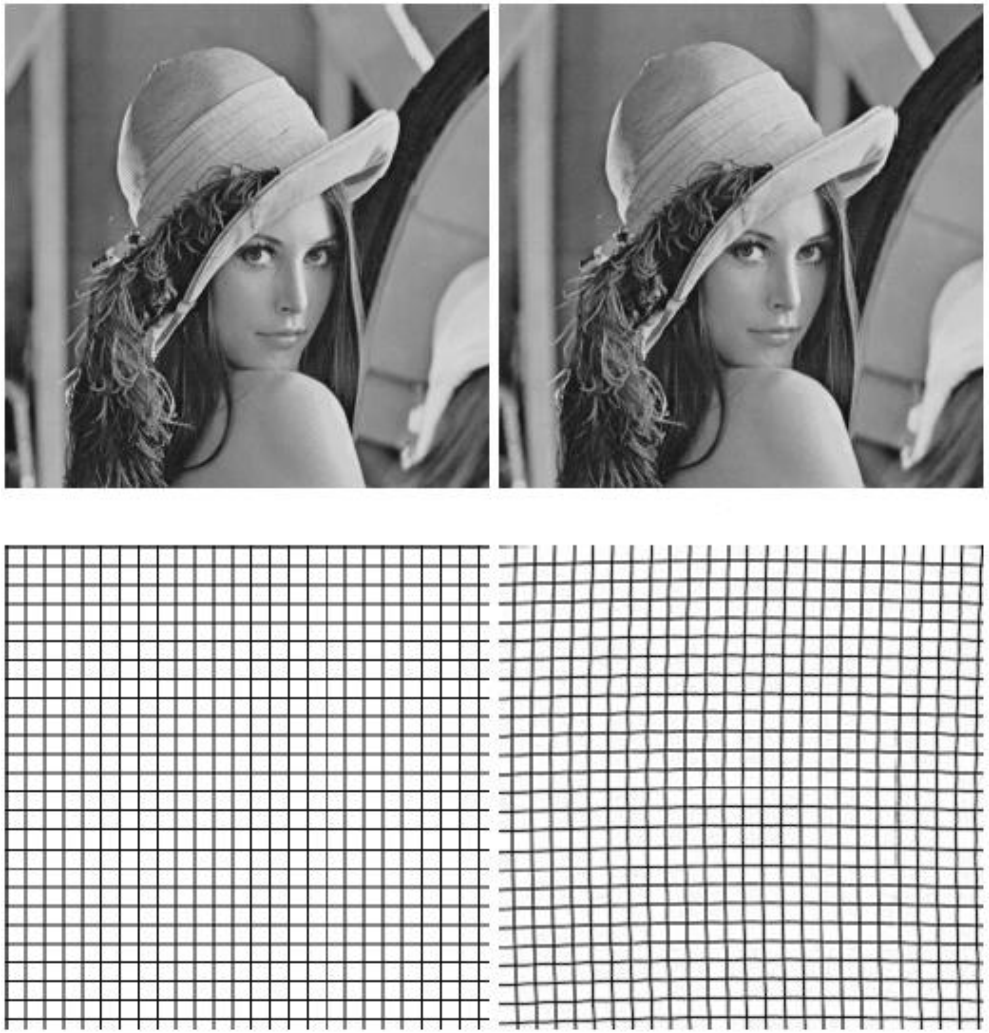
Stéganographie

- Méthode d'attaque
  - Savoir si le document contient des données cachées
  - Détection des différences par rapport à l'original
  - Analyse des discontinuités
  - Recherche de schémas répétitifs (déformation)
    - Repérage des données
    - Reconstitutions des données de départ

Watermarking sur image

- Essentiellement tatouages de copyright
  - Suppression, dégradation
- Types d'attaque
  - Compression JPEG (domaine spatial)
  - Scan/Numérisation (domaine spatial)
  - Transformations géométriques (DCT)
  - Découpage (tous)
  - Moyennage, déformation, filtrage blur, ajout de bruit, ajout de watermark
  - Distorsion

Attaque par distorsion



Conclusion

- Watermarking *vs* Stéganographie
- À la mode
  - Réseau échelon
  - Piratage
  - Terrorisme
  - P2P
- Évolutions techniques avec les tatouages 2nde génération (adaptation au contenu)