Le tatouage de documents numériques

Marc Chaumont

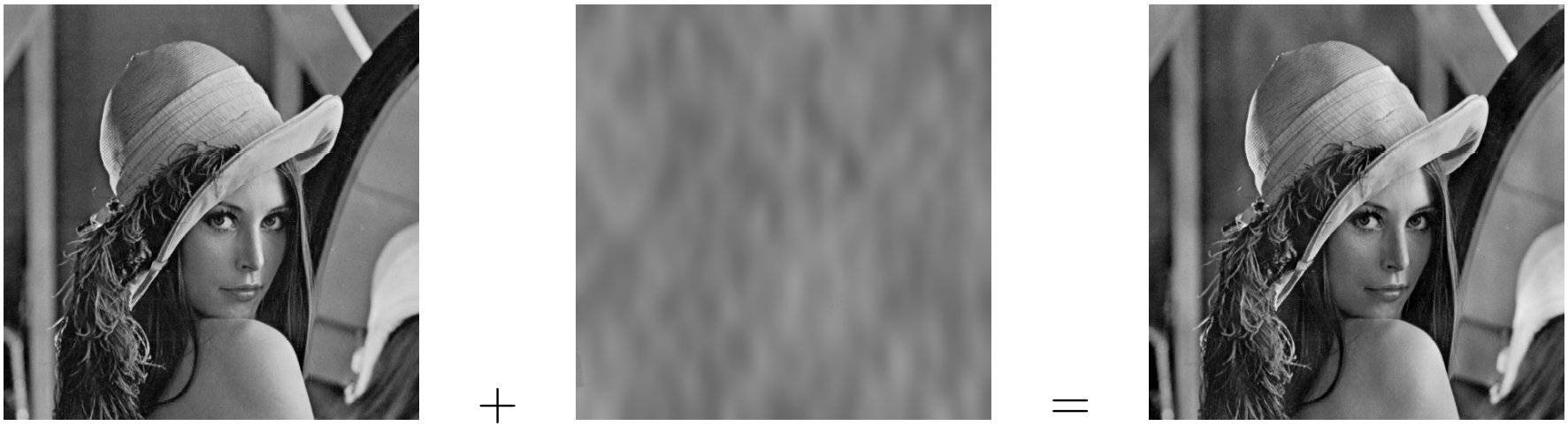
6 mars 2009

Plan

Tatouage visible et tatouage invisible pour une image

Fig.: Le tatouage visible et invisible sur une image

Exemple de système de tatouage : Cox et al. 97



I. J. Cox, J. Kilian, T. Leighton, et T. G. Shamoon. "Secure spread spectrum watermarking for multimedia". ICIP '97.

Définition du tatouage (watermarking)

Le tatouage - Insertion de données cachées

Le tatouage est l'art d'altérer un média (une image, un son, une vidéo ...) de sorte qu'il contienne un message le plus souvent en rapport avec le média, le plus souvent de manière imperceptible et le plus souvent de manière robuste et sûre.

La stéganographie

La stéganographie est l'art de dissimuler au sein d'un support anodin une information qui bien souvent est sans rapport avec le support. Cette dissimulation se fait de sorte qu'il soit difficile pour un observateur extérieur de se rendre compte qu'il y a eu dissimulation.

La cryptographie

La cryptographie est l'art de rendre indéchiffrable un message et ceci au sus de tout le monde.

Le tatouage, une science jeune

Year	1992	1993	1994	1995	1996	1997	1998
Publications	2	2	4	13	29	64	103

Table: Number of publications on digital watermarking during the years 1992-1998 according to [PETITCOLAS1999IEEE]

F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn. "Information Hiding - A Survey". 1999.

Tatouage : La motivation première

- facilité de stockage, de copie et de redistribution (disque dur, CD)
- 1993 : Navigateur Web Mosaic et début de l'ère Internet,

Réticence des grands et petits possesseurs et diffuseurs de données numériques envers internet, le DVD , ...

Il faut des solutions pour protéger les ayant droits de ces documents.
Remarque : le cryptage protège tant que le support est crypté mais plus une fois qu'il est en clair.

Les propriétés qu'apporte le tatouage

1. Le tatouage est invisible ; l'esthétique est conservée,
2. Le tatouage est inséparable de son support quand le tatouage est robuste ; Un changement de format ne fait pas disparaître le message caché ;
3. Le tatouage subit les mêmes transformations que le support.

Les médias numériques

- texte,
- programme,
- image,
- son,
- vidéo,
- modèle 3D,
- ...

Les applications possibles

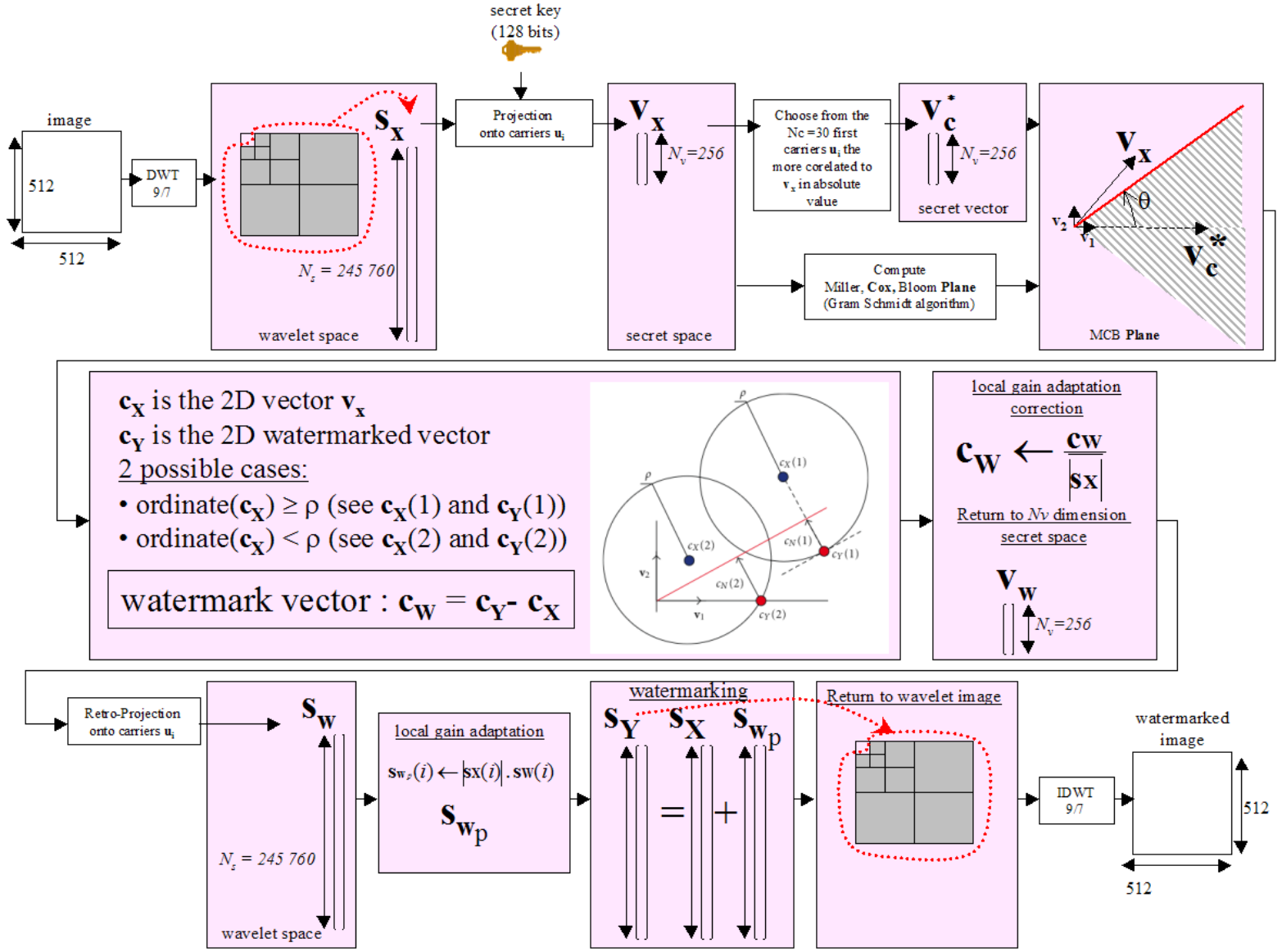
- **contrôle de diffusion** - "broadcast monitoring",
- **identification du propiétaire** - "copyright identification",
- **preuve de propiété** - "copyright proof",
- **suivi de transaction** - "fingerprinting",
- **authentification du support** - "authentication",
- **contrôle de copie** - "copy control",
- **contrôle de périphérique** - "device control",
- **enrichissement** - "enhancement"

Plan

Le tatouage 0-bit

Illustration avec l'algorithme de "Broken Arrows" utilisé lors de la compétition BOWS-2 (Break Our Watermarking System - 2)

"Broken Arrows", Teddy Furon and Patrick Bas, EURASIP Journal on Information Security, 2008.


secret key
(128 bits)
image
512
512
DWT
9/7
S_X
N_s = 245 760
wavelet space
Projection
onto carriers u_i
V_X
N_v=256
secret space
Choose from the
Nc =30 first
carriers u_i the
more corelated to
v_x in absolute
value
V_c^*
N_v=256
secret vector
Compute
Miller, Cox, Bloom Plane
(Gram Schmidt algorithm)
V_X
θ
v_2
v_1
V_c^*
MCB Plane
c_X is the 2D vector v_x
c_Y is the 2D watermarked vector
2 possible cases:
• ordinate(c_X) ≥ ρ (see $c_X(1)$ and $c_Y(1)$)
• ordinate(c_X) < ρ (see $c_X(2)$ and $c_Y(2)$)
watermark vector : $c_W = c_Y - c_X$
ρ
$c_X(1)$
$c_N(1)$
$c_Y(1)$
$c_X(2)$
$c_N(2)$
$c_Y(2)$
v_2
v_1
local gain adaptation
correction
$c_W \leftarrow \frac{c_W}{|s_X|}$
Return to Nv dimension
secret space
V_W
N_v=256
Retro-Projection
onto carriers u_i
S_W
N_s = 245 760
wavelet space
local gain adaptation
$s_{W_p}(i) \leftarrow |s_X(i)| . s_W(i)$
S_{W_p}
watermarking
S_Y S_X S_{W_p}
=
+
Return to wavelet image
IDWT
9/7
watermarked
image
512
512

Boken Arrows : illustration de la robustesse

original

tatoué

Boken Arrows : marque présente (1)

tatoué

bruit

Boken Arrows : marque présente (2)

tatoué

flou

Boken Arrows : marque présente (3)

tatoué

luminosité

Boken Arrows : marque présente (4)

tatoué

luminosité

Boken Arrows : marque présente (5)

tatoué

contour

Plan

Le tatouage non informé par étalement de spectre

- Des porteuses $\mathbf{u}_i \in \mathbb{R}^N$, avec $i \in [1, N_{sec}]$;
- Un message $\mathbf{m}$ composé de N_{sec} bits,
- Une fonction de modulation $s : \{0, 1\} \longrightarrow \mathbb{R}$; On peut prendre par exemple $s(b) = (-1)^b$.

Le signal de tatouage est :

$$\mathbf{w} = \sum_{i=1}^{N_{sec}} \mathbf{u}_i . s(\mathbf{m}[i])$$

Le signal original est noté $\mathbf{x}$ et le signal tatoué $\mathbf{y}$ est tel que :

$$\mathbf{y} = \mathbf{x} + \mathbf{w}$$

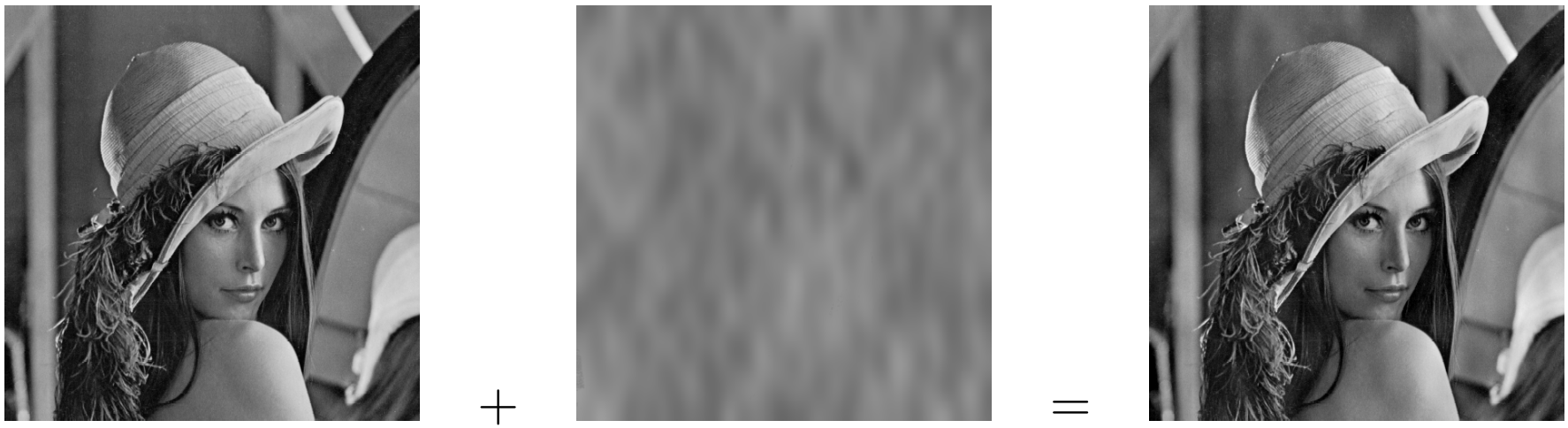
Le tatouage non informé par étalement de spectre

- Des porteuses $\mathbf{u}_i \in \mathbb{R}^N$, avec $i \in [1, N_{sec}]$;
- Un message $\mathbf{m}$ composé de N_{sec} bits,
- Une fonction de modulation $s : \{0, 1\} \rightarrow \mathbb{R}$; On peut prendre par exemple $s(b) = (-1)^b$.

À la détection on recoit un signal $\mathbf{z}$ ayant possiblement subi des dégradations (attaques) et on calcule la corrélation avec chaque porteuse pour retrouver les bits du message m :

$$\mathbf{m}[i] = \begin{cases} 0 \ si \ (< \mathbf{z} | \mathbf{u_i} >= \sum_{j=1}^{N} z[j].u_i[j]) > 0, \\ 1 \ si \ (< \mathbf{z} | \mathbf{u_i} >= \sum_{j=1}^{N} z[j].u_i[j]) < 0. \end{cases}$$

Exemple de tatouage par étalement de spectre : Cox et al. 97



I. J. Cox, J. Kilian, T. Leighton, et T. G. Shamoon. "Secure spread spectrum watermarking for multimedia". ICIP '97.

Plan

Le tatouage informé basé quantification

Illustration avec l'algorithme QIM de B. Chen and G. Wornell, "Quantization index modulation : A class of provably good methods for digital watermarking and infromation embedding", 2001.

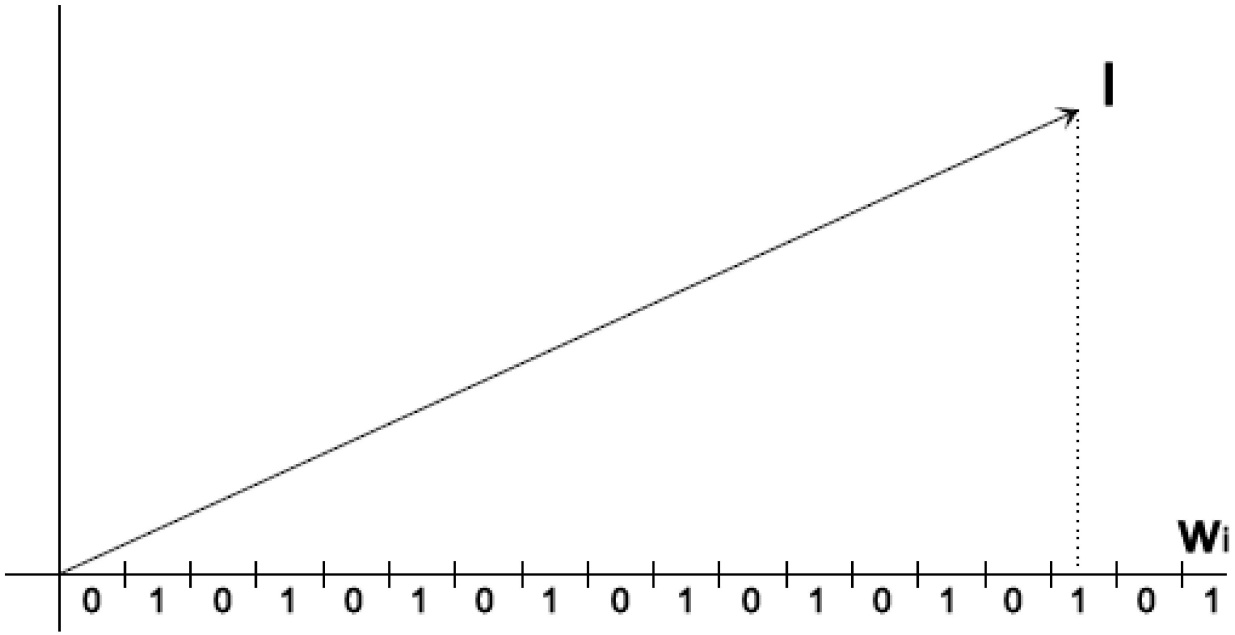


Fig.: Représentation dans le plan du vecteur image I et d'une porteuse w_i

Le tatouage informé basé quantification

Illustration avec l'algorithme QIM de B. Chen and G. Wornell, "Quantization index modulation : A class of provably good methods for digital watermarking and infromation embedding", 2001.

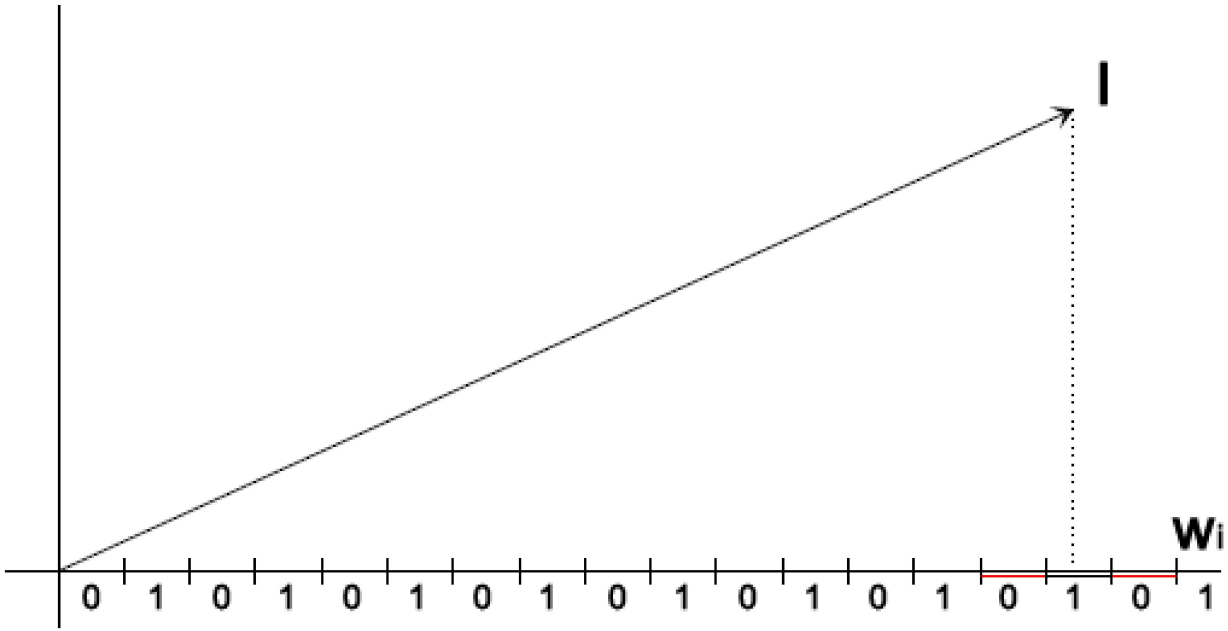


Fig.: Représentation dans le plan du vecteur image I et d'une porteuse w_i

Le tatouage informé basé quantification

Illustration avec l'algorithme QIM de B. Chen and G. Wornell, "Quantization index modulation : A class of provably good methods for digital watermarking and infromation embedding", 2001.

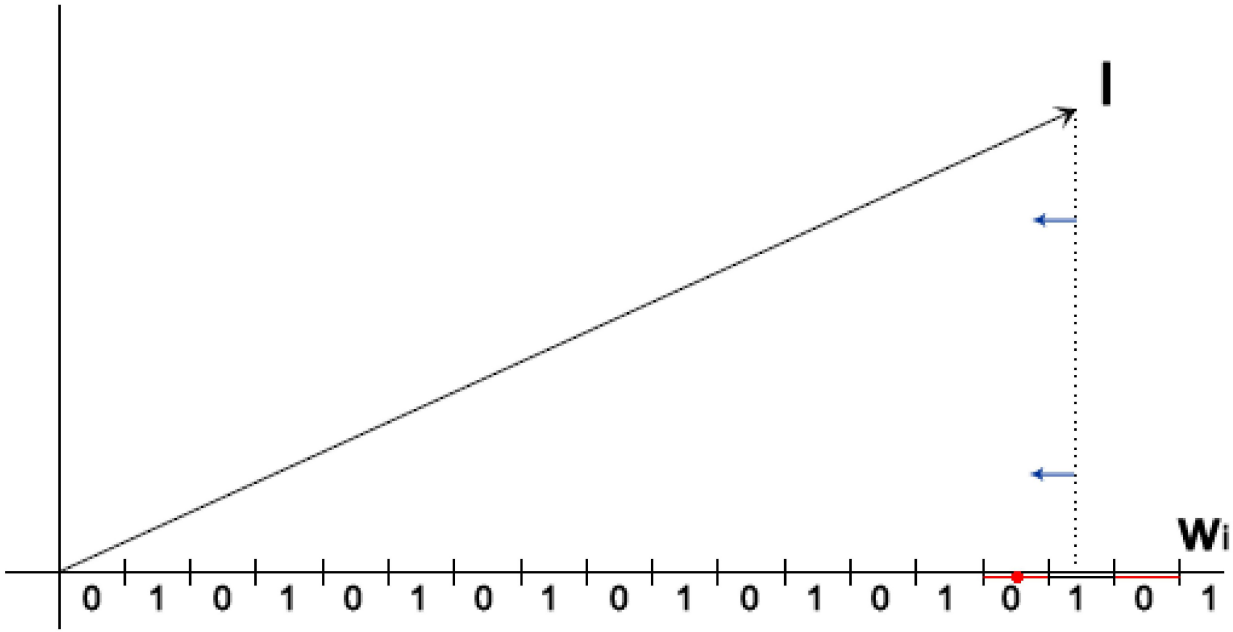


Fig.: Représentation dans le plan du vecteur image I et d'une porteuse w_i

Le tatouage informé basé quantification

Illustration avec l'algorithme QIM de B. Chen and G. Wornell, "Quantization index modulation : A class of provably good methods for digital watermarking and infromation embedding", 2001.

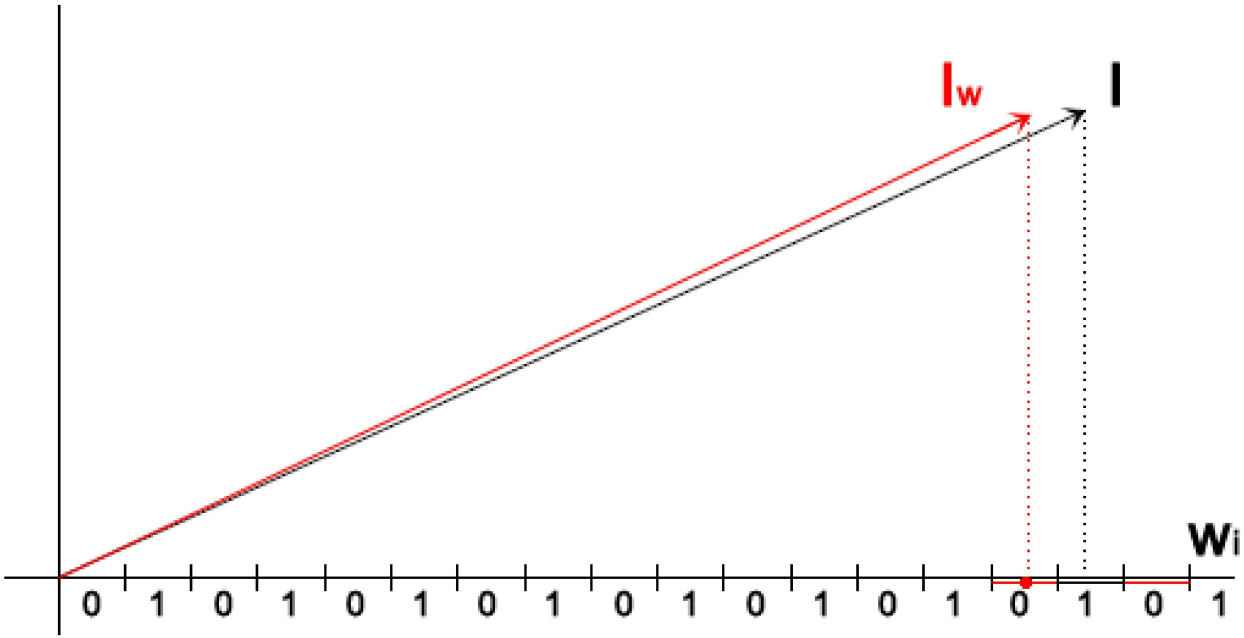


Fig.: Représentation dans le plan du vecteur image I et d'une porteuse w_i

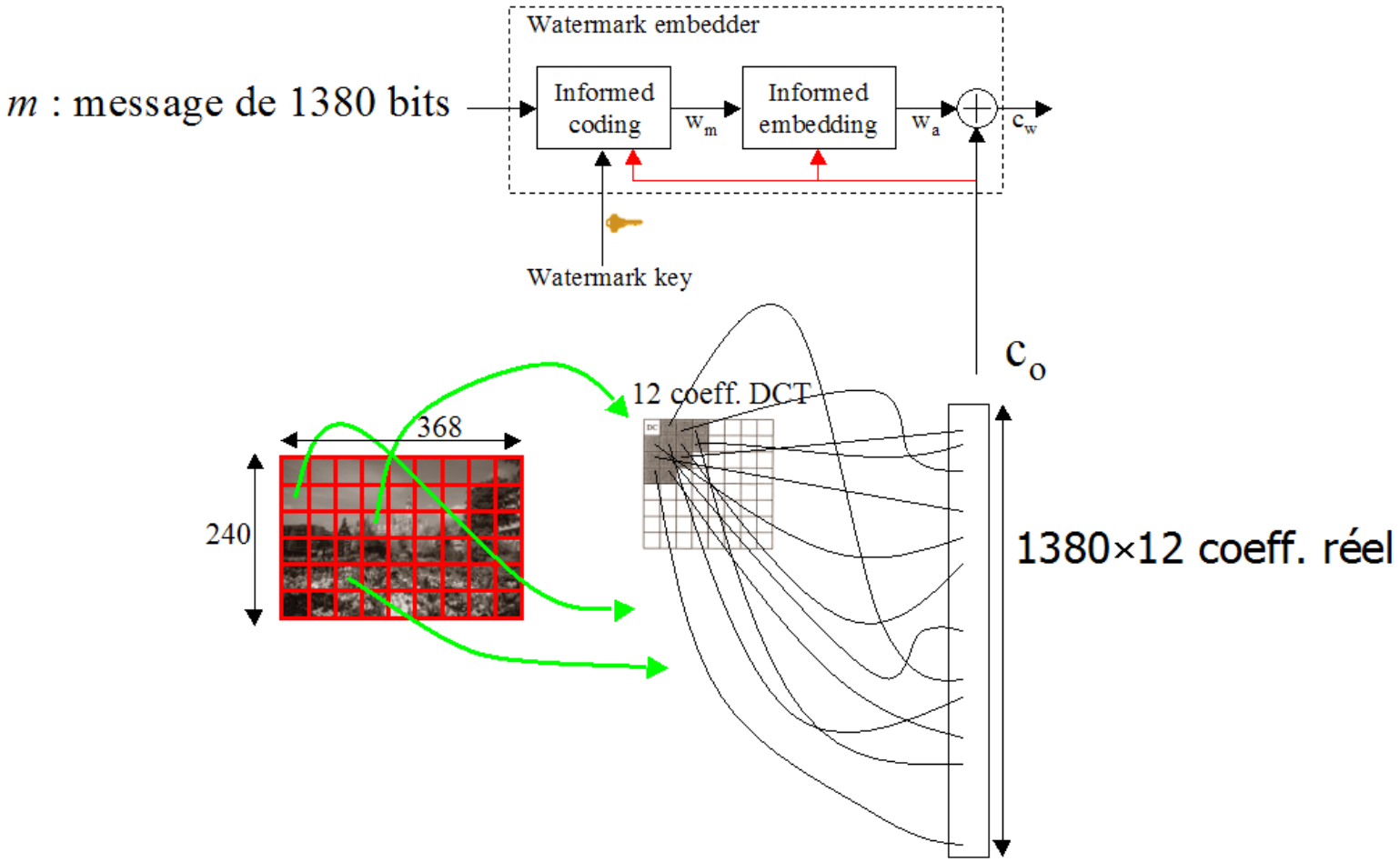
Plan

Le tatouage informé basé treillis

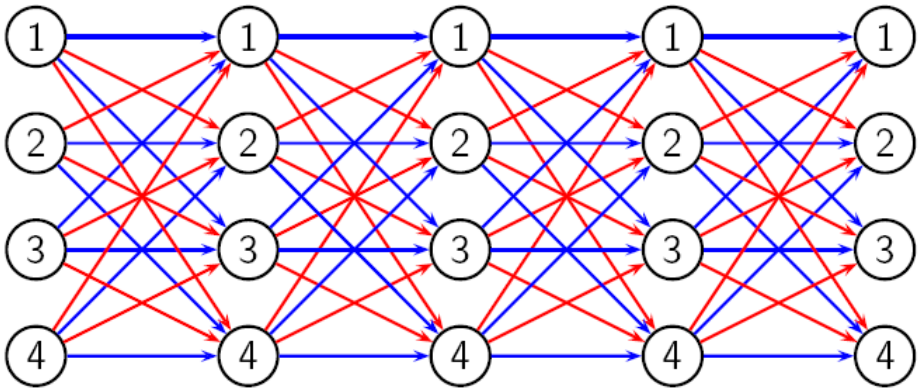
"**Applying Informed Coding and Informed Embedding to Design a Robust, High Capacity Watermark**". M. L. Miller, G. J. Doërr and I. J. Cox, In IEEE Transactions on Image Processing, 2004.

On souhaite embarquer un message m de taille L (ex : $L = 1380$ bits) dans une image de taille N (ex : $N = 240 \times 368$).

Le tatouage informé basé treillis



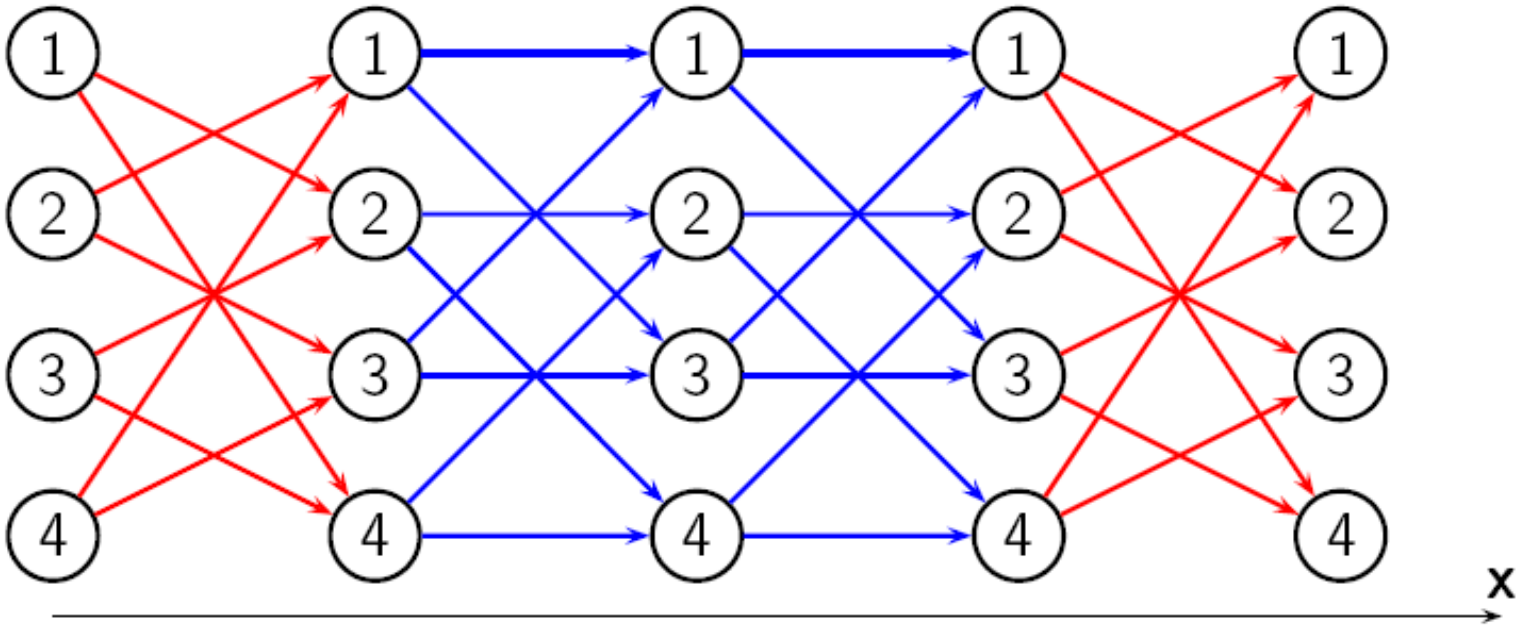
Le tatouage informé basé treillis : le codage informé



Définition du treillis :

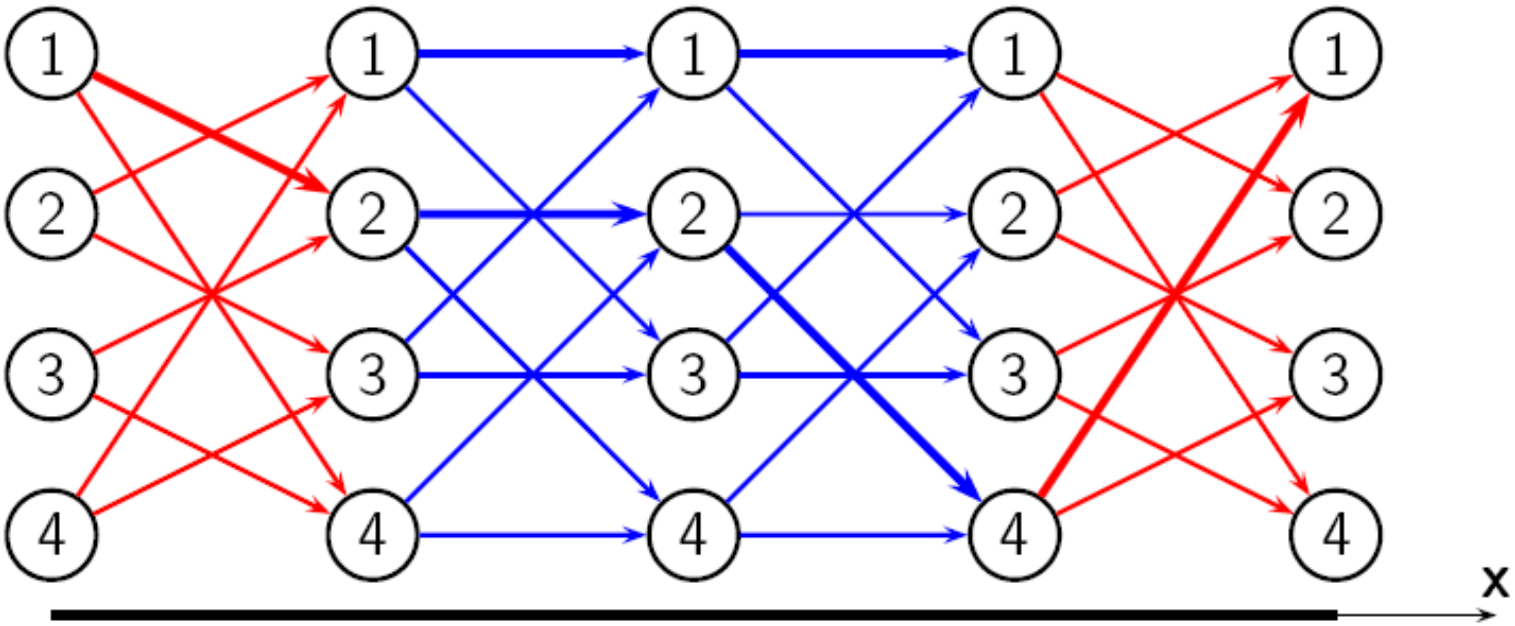
- 64 états,
- 64 arcs sortant d'un état,
- L étapes (exemple $L = 1380$),
- arc bleu = entrée 0, arc rouge = entrée 1,
- chaque arc est valué (sortie du codage) par une séquence pseudo-aléatoire de 12 coefficients réels.

Le tatouage informé basé treillis : le codage informé



Encoding of $\mathbf{m} = (1001)$

Le tatouage informé basé treillis : le codage informé



Encoding of $\mathbf{m} = (1001)$

On détermine le chemin le plus corrélé (produit scalaire) entre les arcs de sortie et le vecteur c_o (noté x ici) : Algorithme de Viterbi modifié.

Le tatouage informé basé treillis : rappel des 2 étapes

1. Affectuer le **codage informé** en traversant un treillis élagué.
2. Effectuer **l'insertion informée** par une approche Monte Carlo par exemple (alternance d'attaque et de contre attaque).