

# Sécuriser la diffusion des documents multimédia grâce à la dissimulation d'information ? Securing multimedia content distribution with information hiding?

Caroline.Fontaine@irisa.fr

January, 29th 2008

Information hiding : generalities and focus on watermarking of still image

Digital watermarking : which security level ?

Conclusion and future work

# Outline

## Information hiding : generalities and focus on watermarking of still image

Introduction

Simple embedding strategies for still images

Informed embedding

How to embed more than one bit ?

How to ensure robustness ?

# Outline

Information hiding : generalities and focus on watermarking of still image

## Introduction

Simple embedding strategies for still images

Informed embedding

How to embed more than one bit ?

How to ensure robustness ?

# At the (very) beginning

First use mentioned :

Antique Greece → **analogical steganography**

Bank notes → **analogical watermarking**

More recent needs for the protection of digital contents ( $\simeq$  1993) :

- ▶ steganography ?
- ▶ encryption, digital signature, authentication ?
- ▶ access control (smart card, biometry) ?

→ **digital watermarking, fingerprinting**

Also a parallel development of **digital steganography**

# Definitions

**steganography** : hiding the transmission itself, using a common/regular cover document (the cover document is not important)

**watermarking** : hiding a watermark in a document, once for all, e.g. to identify its copyright owner (anybody gets the same copy of the content, the document is important, the transmission is not secret)

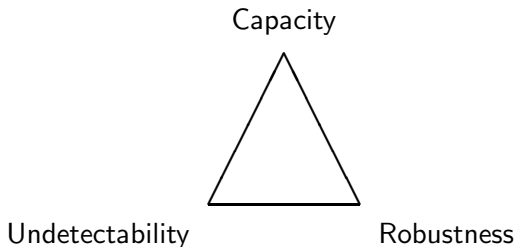
**fingerprinting** : hiding in a document a fingerprint that identifies its user (each distributed copy of the same content is different, the document is important, the transmission is not secret)

# Example



**Documents** : texts, still images, video, audio, (2D,3D) virtual objects, comic strips, maps, data bases, programs ...

# Properties, tradeoffs



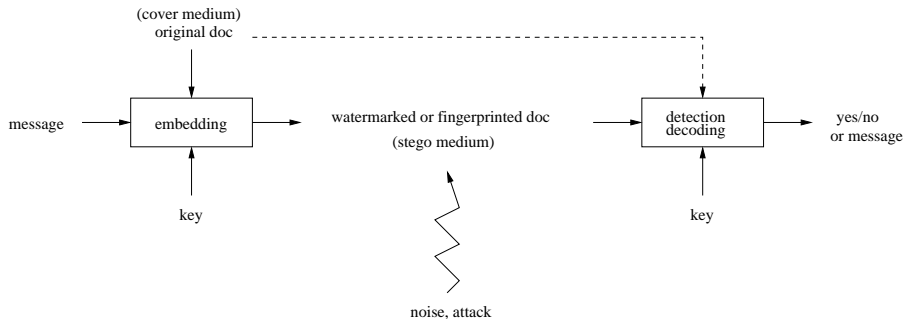


# Applications and added values

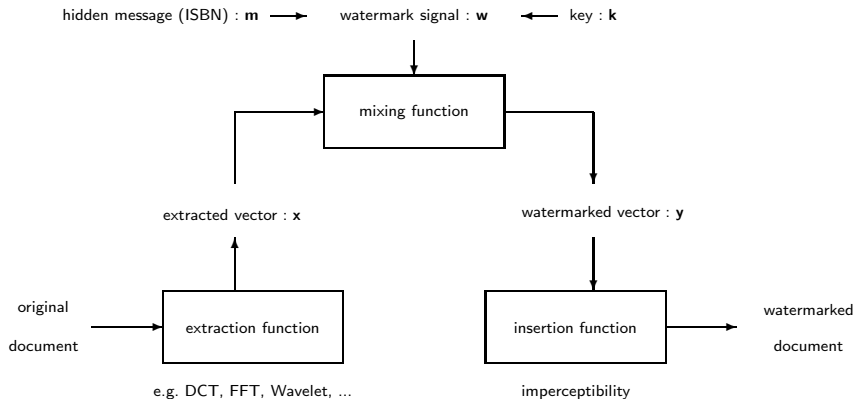
- ▶ transmission : hidden, synchro, auto-correction
- ▶ copyright protection
- ▶ integrity
- ▶ traceability
- ▶ copy protection
- ▶ enriched content (meta-data, indexing, etc)
- ▶ ...

Some applications deal with security, not all.

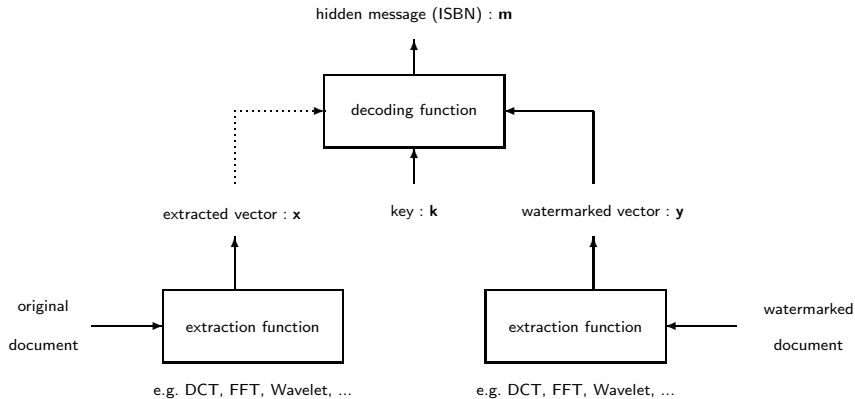
# General viewpoint



# Embedding



# Detection/decoding



# Outline

## Information hiding : generalities and focus on watermarking of still image

Introduction

Simple embedding strategies for still images

Informed embedding

How to embed more than one bit ?

How to ensure robustness ?

# What is an image ?

It is an array of pixels, each of them being encoded on, say, 3 bytes :

- ▶  $(R, G, B)$  : Red, Green, Blue ;
- ▶  $(Y, U, V)$  :  $Y$  represents the luminance  
 $U$  and  $V$  chrominance.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,144 \\ 0,596 & -0,274 & -0,322 \\ 0,211 & -0,523 & 0,312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

# Substituting LSBs

- ▶ select some pixels (key) and replace them by the message to hide
- ▶ select a geometric pattern (key) and embed it (similar to the one-time-pad)

REMARK : really easy to perform, but those bits will be destroyed by any transformation (compression, filtering, noise, etc)

## Patchwork Bender [95], Pitas [96]

Let  $A, B$  be two sets of  $n$  pixels (key = choice), with luminances  $\{a_1, \dots, a_n\}, \{b_1, \dots, b_n\}$ .

Statement :

$$S = \frac{1}{n} \sum_{i=1}^n (a_i - b_i) \simeq 0$$



## Patchwork Bender [95], Pitas [96]

Let  $A, B$  be two sets of  $n$  pixels (key = choice), with luminances  $\{a_1, \dots, a_n\}, \{b_1, \dots, b_n\}$ .

Statement :

$$S = \frac{1}{n} \sum_{i=1}^n (a_i - b_i) \simeq 0$$

Idea : we modify the luminances :  $a'_i := a_i + C, b'_i := b_i - C$ .

$$S' = \frac{1}{n} \sum_{i=1}^n (a'_i - b'_i) = S + 2C \simeq 2C$$

## Patchwork **Bender [95], Pitas [96]**

Let  $A, B$  be two sets of  $n$  pixels (key = choice), with luminances  $\{a_1, \dots, a_n\}, \{b_1, \dots, b_n\}$ .

Statement :

$$S = \frac{1}{n} \sum_{i=1}^n (a_i - b_i) \simeq 0$$

Idea : we modify the luminances :  $a'_i := a_i + C, b'_i := b_i - C$ .

$$S' = \frac{1}{n} \sum_{i=1}^n (a'_i - b'_i) = S + 2C \simeq 2C$$

**Introducing a bias into statistics.** To measure it, you need to use the right key.

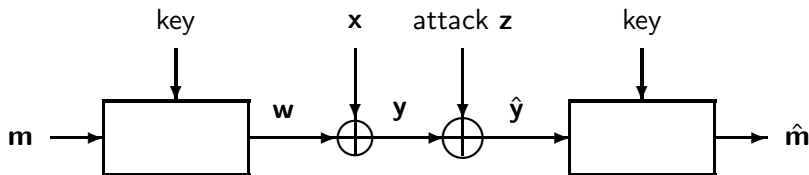
# JPEG related Koch [95]

Explanations given on blackboard.

# Telecommunications : transmitting through a channel

1998 : statement that we are in fact transmitting a message through a channel ...

Example of an additive watermarking scheme :



# Modulation of a random sequence/carrier by a message : a toy example

Example 1 given on blackboard.

It is the opportunity to speak about detection strategies and tradeoffs.

# Outline

## Information hiding : generalities and focus on watermarking of still image

Introduction

Simple embedding strategies for still images

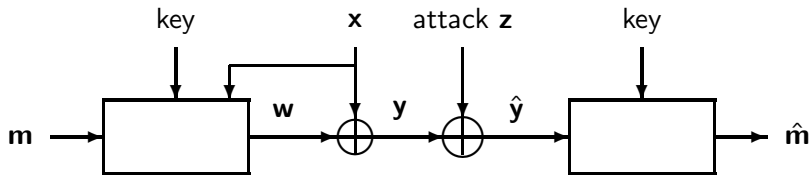
**Informed embedding**

How to embed more than one bit ?

How to ensure robustness ?

# Informed embedding : introduction

Transmission with side information at the transmitter has been initially proposed by **Shannon [58]** .



Example 2 given on blackboard.

# Informed embedding : an optimization problem

**Imperceptibility** : optimization to stay in the acceptable fidelity region (MSE, SNR).

**Robustness** : optimization with respect to a detection statistic to ensure a given robustness level.



# Informed embedding : a revolution in watermarking

Writing on dirty papers Costa [83] : considering the previous scheme (with side information on  $S$  available to the transmitter), Costa showed that when  $\mathbf{x}$  and  $\mathbf{z}$  are Gaussian and the watermark power is upper-bounded the channel capacity is the same as if there were no noise  $S$ ! Instead of fighting the original document (noise), it is better to exploit it. This will be a good point for invisibility and increase the capacity to the maximum !

Costa's proof gives a constructive strategy.

This resulted in numerous schemes Eggers [00 :SCS,QIM], Chen [01 :QIM], Miller [02], Furon [02], Le Guelvouit [03]<sup>1</sup>, Eggers [03], Miller [04], ... and to a study of the more realistic non-Gaussian case (information theory).

---

<sup>1</sup>also with game theory

# Outline

## Information hiding : generalities and focus on watermarking of still image

Introduction

Simple embedding strategies for still images

Informed embedding

**How to embed more than one bit ?**

How to ensure robustness ?

# How to embed more than one bit ?

Mapping messages into vectors in the marking space.

- ▶ direct message coding
- ▶ multi-symbol message coding (see Example 3 on blackboard)

Error Correcting Codes (more efficient)

## Spread spectrum, a famous approach

**Spread Spectrum** at a glance : telecommunications (WWII), modulation of the message by a **secret carrier**. It enables the transmission of several signals at the same time, hides the signal and disables the access without the knowledge of the secret carrier.

**Secret carrier** : pseudo-random temporal sequence (Direct Sequence Spread Spectrum); sinusoid whose frequency is moving quickly (Frequency Hopping Spread Spectrum).

**90% of watermarking techniques are DSSS** (in spatial, Fourier, wavelet domains).

# Outline

## Information hiding : generalities and focus on watermarking of still image

Introduction

Simple embedding strategies for still images

Informed embedding

How to embed more than one bit ?

How to ensure robustness ?

# How to ensure robustness ?

Force a robustness level (see above)

Add redundancy with Error Correcting Codes

Fight specific synchronous modifications (compression, etc)

Fight specific asynchronous modifications (jitter, Stirmark, etc)  
with synchronization patterns, invariant data or embedding  
domain, etc

# Outline

## Digital watermarking : which security level ?

Introduction

Example of the first "cryptanalysis" published in watermarking

Context, statements

Methodology

Theoretical results

What about practical tools ?

Conclusion

Security : watermarking vs. cryptography

How to trace users ?

References on watermarking security

# Outline

## Digital watermarking : which security level ?

### Introduction

Example of the first "cryptanalysis" published in watermarking

Context, statements

Methodology

Theoretical results

What about practical tools ?

Conclusion

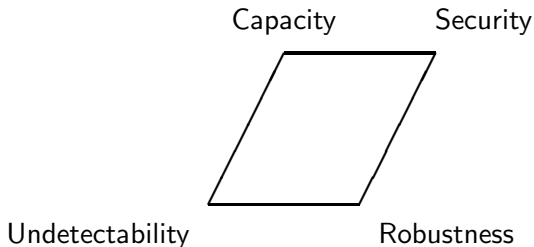
Security : watermarking vs. cryptography

How to trace users ?

References on watermarking security



# Properties, tradeoffs



# What is precisely security ?

**Definition ?** not easy, as robustness and security are sometimes really close.

Robustness	Security
non-malicious attacks common processing one step	intentional attacks dedicated hacks two steps (learn and hack)

# Information hiding and security

Kerckhoffs principle.

**Steganography** : undetectability estimation, via theoretical and heuristic tools; studied since the very beginning [Cachin \[98\]](#), [Zollner \[98\]](#), [Mittelholzer \[99\]](#), ... .

**Watermarking** : watermark estimation (e.g. oracle sensitivity attacks [Cox \[97\]](#), [Linnartz \[98\]](#), [Kalker \[98\]](#), ..., [Comesana \[06\]](#) , "collusion" attacks [Doërr \[05\]](#) ), search for key recovering [Cayre \[04,05\]](#), [Pérez-Freire \[06,07\]](#) .

**Fingerprinting** : collusion attacks (more or less realistic models and solutions), under study.

# Outline

## Digital watermarking : which security level ?

Introduction

Example of the first "cryptanalysis" published in watermarking

Context, statements

Methodology

Theoretical results

What about practical tools ?

Conclusion

Security : watermarking vs. cryptography

How to trace users ?

References on watermarking security

# Context

Let us consider a **robust watermarking technique**.

**Is this robustness sufficient ?**

Since 2002, this point has been addressed through a few papers.

Cayre [04,05] provides the **first concrete study**.

# Security vs. robustness

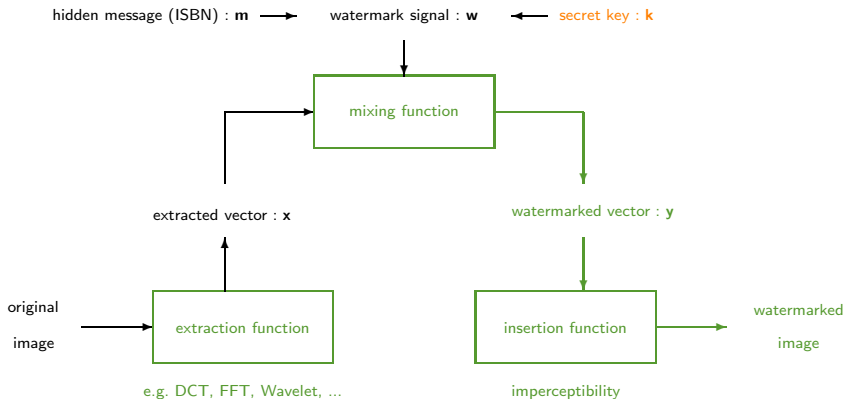
**Robustness** is the ability to resist alteration and/or invalidation of the watermark.

⇒ blur, compression, geometrical transformations, ...

**Security** is the ability to resist key/message recovery.

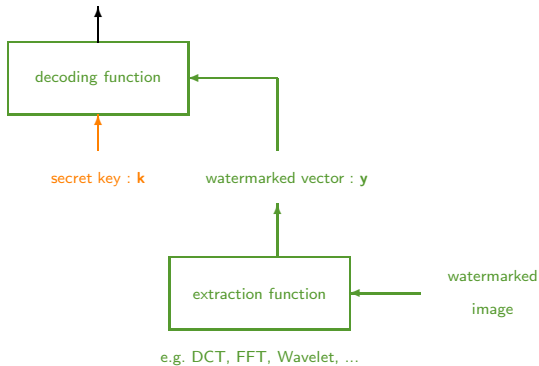
⇒ cryptanalysis

# Embedding



# Detection/decoding

hidden message (ISBN) :  $m$





# Methodology

$N_o$  images : different messages, but the same secret key.

We follow [Shannon's methodology](#) [Sha49] :

What is **theoretically possible** ?

Which **practical tools to succeed** ?

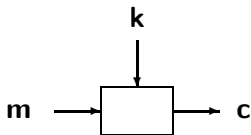
and [Diffie-Hellman's classification](#) [DH76] :

Known Original Attack – **KOA**

Known Messages Attack – **KMA**

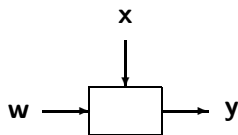
Watermark Only Attack – **WOA**

# "perfect covering"



*encryption*

**m** plaintext, **k** key,  
**c** ciphertext



*watermark embedding*

**w** watermark, **x** original content,  
**y** watermarked content

A watermark embedding makes a **perfect covering** if

$$p_{\mathcal{W}}(\mathbf{w}) = p_{\mathcal{W}}(\mathbf{w}|\mathbf{y}) \quad \forall (\mathbf{y}, \mathbf{w}) \in \mathcal{Y} \times \mathcal{W}$$

## Information leakage Shannon [49]

Entropy :  $H(\mathbf{K}) = - \sum_{\mathbf{k}} p(\mathbf{k}) \log p(\mathbf{k})$

Equivocation :  $H(\mathbf{K} | \mathbf{Y}_1, \dots, \mathbf{Y}_{N_o}) = H(\mathbf{K}) - I(\mathbf{K}; \mathbf{Y}_1, \dots, \mathbf{Y}_{N_o})$

Both measure the ignorance on the secret key.

The equivocation is a non-increasing function, going from  $H(\mathbf{K})$  "down to 0", as  $N_o$  increases.

The smallest  $N_o$  giving a null equivocation is denoted  $N_o^*$ , called the **unicity distance**, and corresponds to a **security level**. Perfect covering is denoted by  $N_o^* = +\infty$ .

# Substitutive scheme e.g. Koch [95]

$\mathbf{x}, \mathbf{y}$  : binary vectors of length  $N_v$

$\mathbf{k}$  : list of  $N_c$  integers,  $1 \leq k(j) \leq N_v$

$\mathbf{m}$  : binary vector of length  $N_c$

$$\begin{aligned} y(j) &\leftarrow x(j) & \forall 1 \leq j \leq N_v \\ y(k(j)) &\leftarrow m(j) & \forall 1 \leq j \leq N_c \end{aligned}$$

**Example :**

$$\begin{array}{rcl} \mathbf{m} & = & (1101) \\ \mathbf{x} & = & (01001011) \end{array} \quad \begin{array}{rcl} \mathbf{k} & = & [2, 8, 5, 3] \\ \mathbf{y} & = & (01100011) \end{array}$$

## Substitutive scheme e.g. Koch [95]

$\mathbf{x}, \mathbf{y}$  : binary vectors of length  $N_v$

$\mathbf{k}$  : list of  $N_c$  integers,  $1 \leq k(j) \leq N_v$

$\mathbf{m}$  : binary vector of length  $N_c$

$$\begin{aligned} y(j) &\leftarrow x(j) & \forall 1 \leq j \leq N_v \\ y(k(j)) &\leftarrow m(j) & \forall 1 \leq j \leq N_c \end{aligned}$$

**Example :**

$$\begin{array}{lcl} \mathbf{m} & = & (1101) \\ \mathbf{x} & = & (01001011) \end{array} \quad \begin{array}{lcl} \mathbf{k} & = & [2, 8, 5, 3] \\ \mathbf{y} & = & (0\mathbf{1}100011) \end{array}$$

# Substitutive scheme e.g. Koch [95]

$\mathbf{x}, \mathbf{y}$  : binary vectors of length  $N_v$

$\mathbf{k}$  : list of  $N_c$  integers,  $1 \leq k(j) \leq N_v$

$\mathbf{m}$  : binary vector of length  $N_c$

$$\begin{aligned} y(j) &\leftarrow x(j) & \forall 1 \leq j \leq N_v \\ y(k(j)) &\leftarrow m(j) & \forall 1 \leq j \leq N_c \end{aligned}$$

**Example :**

$$\begin{array}{lcl} \mathbf{m} & = & (1101) \\ \mathbf{x} & = & (01001011) \end{array} \quad \begin{array}{lcl} \mathbf{k} & = & [2, 8, 5, 3] \\ \mathbf{y} & = & (0\mathbf{1}1000\mathbf{1}1) \end{array}$$

# Substitutive scheme e.g. Koch [95]

$\mathbf{x}, \mathbf{y}$  : binary vectors of length  $N_v$

$\mathbf{k}$  : list of  $N_c$  integers,  $1 \leq k(j) \leq N_v$

$\mathbf{m}$  : binary vector of length  $N_c$

$$\begin{aligned} y(j) &\leftarrow x(j) & \forall 1 \leq j \leq N_v \\ y(k(j)) &\leftarrow m(j) & \forall 1 \leq j \leq N_c \end{aligned}$$

**Example :**

$$\begin{array}{lcl} \mathbf{m} & = & (1101) \\ \mathbf{x} & = & (01001011) \end{array} \quad \begin{array}{lcl} \mathbf{k} & = & [2, 8, 5, 3] \\ \mathbf{y} & = & (0\mathbf{1}100\mathbf{0}1\mathbf{1}) \end{array}$$

# Substitutive scheme e.g. Koch [95]

$\mathbf{x}, \mathbf{y}$  : binary vectors of length  $N_v$

$\mathbf{k}$  : list of  $N_c$  integers,  $1 \leq k(j) \leq N_v$

$\mathbf{m}$  : binary vector of length  $N_c$

$$\begin{aligned} y(j) &\leftarrow x(j) & \forall 1 \leq j \leq N_v \\ y(k(j)) &\leftarrow m(j) & \forall 1 \leq j \leq N_c \end{aligned}$$

**Example :**

$$\begin{array}{lcl} \mathbf{m} & = & (1101) \\ \mathbf{x} & = & (01001011) \end{array} \quad \begin{array}{lcl} \mathbf{k} & = & [2, 8, 5, 3] \\ \mathbf{y} & = & (01100011) \end{array}$$



WOA :  $N_o^* = +\infty$ , perfect covering

KMA :  $N_o^* = \log_2 N_v$

$m_1 = (1101)$

$y_1 = (01100011)$

$m_2 = (1011)$

$y_2 = (11101000)$

$m_3 = (1000)$

$y_3 = (11010010)$

KOA :  $N_o^* = \log_2 N_c$  (up to permutation of the indices)

$x_1 = (01001011)$

$y_1 = (01100011)$

$x_2 = (10001001)$

$y_2 = (11101000)$

WOA :  $N_o^* = +\infty$ , perfect covering

KMA :  $N_o^* = \log_2 N_v$

$m_1 = (1101)$

$y_1 = (0\mathbf{1}100011)$

$m_2 = (1011)$

$y_2 = (\mathbf{1}1101000)$

$m_3 = (1000)$

$y_3 = (\mathbf{1}1010010)$

KOA :  $N_o^* = \log_2 N_c$  (up to permutation of the indices)

$x_1 = (01001011)$

$y_1 = (01100011)$

$x_2 = (10001001)$

$y_2 = (11101000)$

WOA :  $N_o^* = +\infty$ , perfect covering

KMA :  $N_o^* = \log_2 N_v$

$m_1 = (1101)$

$y_1 = (01100011)$

$m_2 = (1011)$

$y_2 = (11101000)$

$m_3 = (1000)$

$y_3 = (11010010)$

KOA :  $N_o^* = \log_2 N_c$  (up to permutation of the indices)

$x_1 = (01001011)$

$y_1 = (01100011)$

$x_2 = (10001001)$

$y_2 = (11101000)$

WOA :  $N_o^* = +\infty$ , perfect covering

KMA :  $N_o^* = \log_2 N_v$

$m1 = (1101)$

$y1 = (01100011)$

$m2 = (1011)$

$y2 = (11101000)$

$m3 = (1000)$

$y3 = (11010010)$

KOA :  $N_o^* = \log_2 N_c$  (up to permutation of the indices)

$x1 = (01001011)$

$y1 = (01100011)$

$x2 = (10001001)$

$y2 = (11101000)$

WOA :  $N_o^* = +\infty$ , perfect covering

KMA :  $N_o^* = \log_2 N_v$

$m_1 = (1101)$

$y_1 = (01100011)$

$m_2 = (1011)$

$y_2 = (11101000)$

$m_3 = (1000)$

$y_3 = (11010010)$

KOA :  $N_o^* = \log_2 N_c$  (up to permutation of the indices)

$x_1 = (01001011)$

$y_1 = (01100011)$

$x_2 = (10001001)$

$y_2 = (11101000)$

WOA :  $N_o^* = +\infty$ , perfect covering

KMA :  $N_o^* = \log_2 N_v$

$m_1 = (1101)$

$y_1 = (01100011)$

$m_2 = (1011)$

$y_2 = (11101000)$

$m_3 = (1000)$

$y_3 = (11010010)$

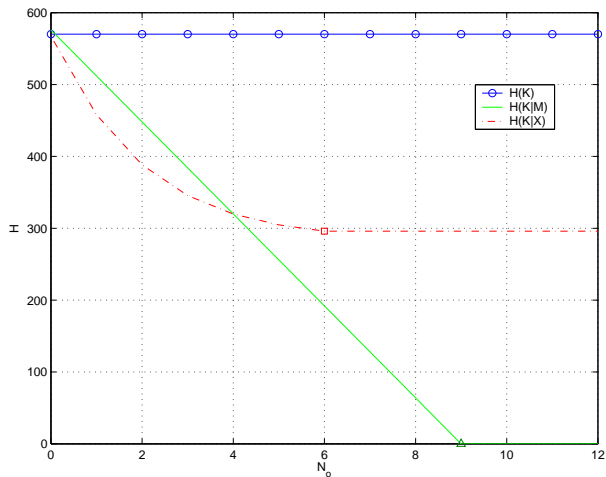
KOA :  $N_o^* = \log_2 N_c$  (up to permutation of the indices)

$x_1 = (01001011)$

$y_1 = (01100011)$

$x_2 = (10001001)$

$y_2 = (11101000)$



## Spread spectrum based schemes

We have

$$\mathbf{y} = \mathbf{x} + \mathbf{w}, \quad \text{with } \mathbf{w} = \frac{\gamma}{\sqrt{N_c}} \sum_{\ell=1}^{N_c} m(\ell) \mathbf{u}_\ell$$

modulation of  $N_c$  private carriers  $\mathbf{u}_\ell$ ,  $\|\mathbf{u}_\ell\| = 1$ ;

$\gamma > 0$  being a small gain fixing the embedding strength.

The carriers are two-by-two orthogonal vectors.

The message symbols usually belong to  $\{-1, +1\}$  Pateux [03] .



First result : no perfect covering !

How much information is leaking ?

# How to measure information leakage ?

**Problem** : Shannon's tools are not well suited for continuous real valued data ... Mutual information still makes sense, but Entropy does no more have a physical interpretation.

Fisher's tools suit better.

# How to measure information leakage with Fisher's tools

**Principle** : estimation of an unknown parameter (here, the secret key).

**Fisher Information Matrix** :

$$\text{FIM}(\theta) = E\psi\psi^T \quad \text{with} \quad \psi = \nabla_{\theta} \log p_{\mathbf{X}}(\mathbf{y} - \mathbf{w}_{\theta}).$$

**Cramér-Rao's theorem** gives a lower bound on the covariance matrix of an unbiased estimator, whenever FIM is invertible :

$$\mathcal{R}_{\hat{\theta}} \geq \text{FIM}(\theta)^{-1},$$

## Remark

In WOA and KOA, the estimation of the carriers can only be up to sign and permutation !

We have

$$\mathbf{w}_j = \frac{\gamma}{\sqrt{N_c}} \mathcal{U} \mathbf{m}_j,$$

where  $\mathcal{U}^T \mathcal{U} = \mathcal{I}_{N_c}$ , and  $\mathbf{m}_j \in \{-1, +1\}^{N_c}$ .

For whatever  $N_c \times N_c$  unitary matrix  $\mathcal{P}$ , we have

$$\mathbf{w}_j = \frac{\gamma}{\sqrt{N_c}} \tilde{\mathcal{U}} \tilde{\mathbf{m}}_j, \text{ with } \tilde{\mathcal{U}} = \mathcal{U}\mathcal{P} \text{ and } \tilde{\mathbf{m}}_j = \mathcal{P}^{-1} \mathbf{m}_j$$

## Security levels definition

$$\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \sigma_x^2 \mathcal{I}_{N_v})$$

KOA :  $N_o^* = O(N_c)$ , up to sign and permutation

$$\text{KMA} : N_o^* = O(N_c \sigma_x^2 / \gamma^2)$$

WOA :  $N_o^* = O(N_c \sigma_x^2 / \gamma^2)$ , up to sign and permutation

## What about practical tools ?

We are dealing with a **Blind Source Separation (BSS)** problem :

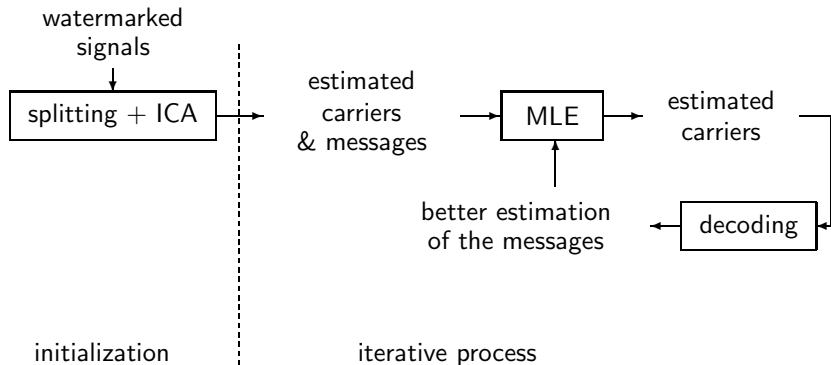
$$\mathbf{d}_j = \frac{\gamma}{\sqrt{N_c}} \mathcal{U} \mathbf{m}_j \text{ (+noise)}$$

**KOA** = without noise ; quite easy : ICA gives a correct basis for  $\text{Span}(\mathcal{U})$ , up to sign and permutation.  $N_o > N_c$  obs.

**WOA** = with noise ; more difficult (see next slide).

**KMA** = the easiest : the MLE converges to the Cramér-Rao bound. Complexity =  $O(N_v N_o^2 N_c) + O(N_c^3)$ .

# An hybrid strategy for the WOA case



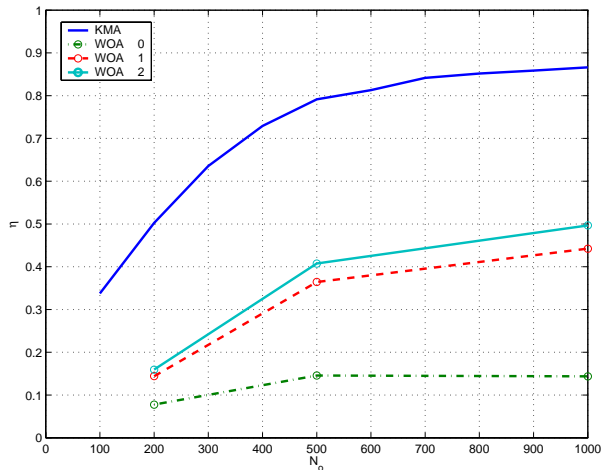
# Application to still images

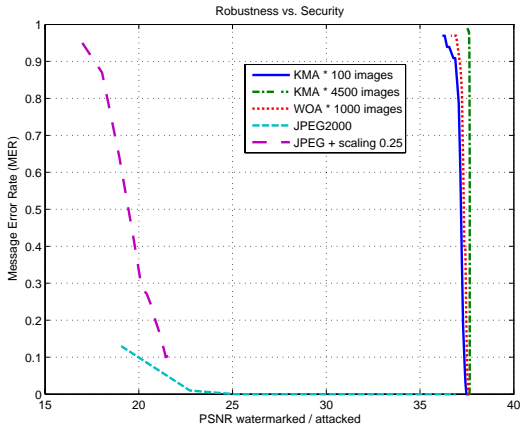
**Embedding** : proportional embedding on some wavelet coefficients **Pateux [03]** .

**Quality** : 38 dB.

**Data** :  $N_v = 258058$ ,  $N_c = 8$









(a) Best quality obtained for a successful blind attack : PSNR=21.8 dB.



(b) Best quality obtained for a successful cryptanalysis : PSNR=35.8 dB.

# Conclusion

This was the first cryptanalytic study in watermarking **Cayre [04,05]**

This led to new recommendations/requirements :

- ▶ practically speaking, it is easy to recover enough information on the insertion keys to perform an efficient attack, if we can observe a sufficiently large number of watermarked documents ;
- ▶ **Keys must be changed often !**

This kind of attack is real and powerful (perhaps even more in video). It has to be taken into account in the design of practical systems.

# Outline

## Digital watermarking : which security level ?

Introduction

Example of the first "cryptanalysis" published in watermarking

Context, statements

Methodology

Theoretical results

What about practical tools ?

Conclusion

Security : watermarking vs. cryptography

How to trace users ?

References on watermarking security

# Links between watermarking/signal processing and cryptography

- ▶ key management
- ▶ symmetric/asymmetric : e.g. Van Schyndel [99], Eggers [99,00], Smith [99], Sylvestre [01], Stern [01], Furon [99,01,03], ...
- ▶ authentication (ZK : e.g. Craver [99], Lévy-dit-Véhel[04] )
- ▶ integrity ((semi-)fragile wm ; robust hash : e.g. Lefèbvre [04] )
- ▶ cryptanalysis : Cayre [04,05](SS), Pérez-Freire [06,07](QIM,lattices)
- ▶ joint watermarking and encryption : e.g. Puech [04]
- ▶ processing encrypted documents : e.g. Kalker [06]

# Links between watermarking/signal processing and cryptography

- ▶ key management
- ▶ symmetric/asymmetric : e.g. Van Schyndel [99], Eggers [99,00], Smith [99], Sylvestre [01], Stern [01], Furon [99,01,03], ...
- ▶ authentication (ZK : e.g. Craver [99], Lévy-dit-Véhel[04] )
- ▶ integrity ((semi-)fragile wm ; robust hash : e.g. Lefèbvre [04] )
- ▶ cryptanalysis : Cayre [04,05](SS), Pérez-Freire [06,07](QIM,lattices)
- ▶ joint watermarking and encryption : e.g. Puech [04]
- ▶ processing encrypted documents : e.g. Kalker [06]

But watermarking/signal processing is not cryptography ! Cox [06]

# The keyspace analogy, the spread spectrum example

1. Let us consider a symmetric-key encryption scheme that uses a secret key  $k$  of length  $n$ .



# The keyspace analogy, the spread spectrum example

1. Let us consider a symmetric-key encryption scheme that uses a secret key  $k$  of length  $n$ .

$$\#\mathcal{K} = 2^n$$

$$P(\text{pick at random a key leading to a successful attack}) = 2^{-n}$$

# The keyspace analogy, the spread spectrum example

1. Let us consider a symmetric-key encryption scheme that uses a secret key  $k$  of length  $n$ .

$$\#\mathcal{K} = 2^n$$

$$P(\text{pick at random a key leading to a successful attack}) = 2^{-n}$$

2. Let us now consider a symmetric-key SS-like watermarking scheme that uses a secret key  $k$  of length  $n$ .

# The keyspace analogy, the spread spectrum example

1. Let us consider a **symmetric-key encryption scheme** that uses a **secret key  $k$**  of length  $n$ .

$$\#\mathcal{K} = 2^n$$

$$P(\text{pick at random a key leading to a successful attack}) = 2^{-n}$$

2. Let us now consider a **symmetric-key SS-like watermarking scheme** that uses a **secret key  $k$**  of length  $n$ .

$$\#\mathcal{K} \neq 2^n$$

$$P(\text{pick at random a key leading to a successful attack}) \neq 2^{-n}$$

# What is going on in the watermarking case ? (1/2)

a **key** generates a set of carriers, i.e. a set of binary sequences.  
Often the **key is identified with these carriers**. Let  $n$  denote the length of the whole **secret sequence**.

# What is going on in the watermarking case? (1/2)

a **key** generates a set of carriers, i.e. a set of binary sequences.  
Often the **key is identified with these carriers**. Let  $n$  denote the length of the whole **secret sequence**.

Not all such sequences are of interest to perform a good embedding!

$$\#\{\text{eligible sequences of length } n\} = 2^{n - \frac{1}{2} \log_2 n}$$

## What is going on in the watermarking case ? (2/2)

An attack may be successful even if the estimated/tested sequence is not exactly the same as the secret one !

A correlation of  $\rho_{\min} = 0.4$  is sufficient.

If at least  $k_{\min} = \lceil \frac{n(\rho_{\min}+1)}{2} \rceil$  samples of the estimated carrier match the secret one, the attack will be successful.

$P(\text{pick at random an eligible carrier leading to a successful attack})$

$$= \sum_{k_{\min} \leq k \text{ even} \leq n} \frac{\binom{n/2}{k/2}^2}{\binom{n}{n/2}} \simeq 2^{-0.12n}$$

# The public key analogy

**public key in cryptography** : to initiate a secure communication without having to share a secret / to enable anybody to check a signature while not being able to forge any fraudulent one.

**public key in watermarking** : to enable anybody to check a watermark while not being able to remove it.

# The public key analogy

**public key in cryptography** : to initiate a secure communication without having to share a secret / to enable anybody to check a signature while not being able to forge any fraudulent one.

**public key in watermarking** : to enable anybody to check a watermark while not being able to remove it.

**But proposals are not convincing** : Hartung [97], Furon [99], Eggers [00] provide a better robustness against average attack, PCA, oracle attacks, but the disclosure of the detection key permits specialized closest-point attacks that prevent detection while maintaining a good perceptual quality Furon [01,03] .



# The public key analogy

**public key in cryptography** : to initiate a secure communication without having to share a secret / to enable anybody to check a signature while not being able to forge any fraudulent one.

**public key in watermarking** : to enable anybody to check a watermark while not being able to remove it.

**But proposals are not convincing** : Hartung [97], Furon [99], Eggers [00] provide a better robustness against average attack, PCA, oracle attacks, but the disclosure of the detection key permits specialized closest-point attacks that prevent detection while maintaining a good perceptual quality Furon [01,03] .

**Asymmetry is not sufficient and perhaps not necessary** Miller [02] .

## A more reasonable approach : layers


Both fields are clearly separated, leading to a cleaner security analysis, and reducing the risk of applying an inappropriate technique to solve a specific security issue.

## A more reasonable approach : layers

Both fields are clearly separated, leading to a cleaner security analysis, and reducing the risk of applying an inappropriate technique to solve a specific security issue.

**Content authentication** : we embed a signature/MAC of the content into itself, but without modifying the data used to compute this signature/MAC. <sup>2</sup>

---

<sup>2</sup>We can locate the tampering, and be robust to some manipulations: 


## A more reasonable approach : layers

Both fields are clearly separated, leading to a cleaner security analysis, and reducing the risk of applying an inappropriate technique to solve a specific security issue.

**Content authentication** : we embed a signature/MAC of the content into itself, but without modifying the data used to compute this signature/MAC. <sup>2</sup>

**Traitor tracing/fingerprinting** : we introduce some user-dependent modifications in the content : one crypto/codes layer to design the user-dependent information to resist collusion attacks on the messages, one embedding layer to resist signal processing attacks.

---

<sup>2</sup>We can locate the tampering, and be robust to some manipulations: 

# Outline

## Digital watermarking : which security level ?

Introduction

Example of the first "cryptanalysis" published in watermarking

Context, statements

Methodology

Theoretical results

What about practical tools ?

Conclusion

Security : watermarking vs. cryptography

**How to trace users ?**

References on watermarking security

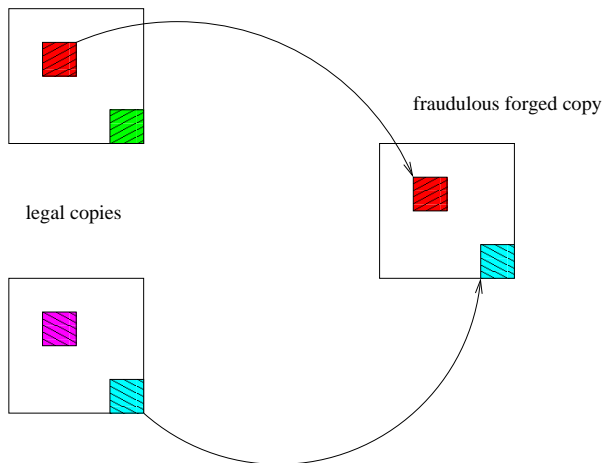
# Generalities

Each version of the document contains a different identifying message (fingerprint).

Same kind of embedding strategies than for robust watermarking.  
Same kind of attacks or security requirements.

But we also have to face **collusion attack** and **specific security considerations**.

# Example of collusion attack



## Solutions ?

Using appropriated combinatorial structures, we try to trace the members of a collusion.

The "marking assumption" model Boneh [98] : IPP (Identity Parent Property) error correcting codes) Van Lint [98], Cohen [00], Barg [01], ...

→ strong vs. weak traceability.

A more realistic model : Somekh-Baruch [05], Galand [06]

Use of "random" codes : Tardos [03], ...



# Outline

## Digital watermarking : which security level ?

Introduction

Example of the first "cryptanalysis" published in watermarking

Context, statements

Methodology

Theoretical results

What about practical tools ?

Conclusion

Security : watermarking vs. cryptography

How to trace users ?

References on watermarking security

## Some references on watermarking security

*Watermarking Security : Theory and Practice*, F. Cayre, C. Fontaine, T. Furon. IEEE Transactions on Signal Processing, vol. 53, num. 10, pp. 3976-3987, 2005.

*Watermarking security : a survey*, L. Pérez-Freire, P. Comesana, J.R. Troncoso-Pastoriza, F. Pérez-González. Transactions on DHMS I, LNCS 4300, pp. 41-72, 2006.

*Watermarking is not cryptography*, I. Cox, G. Doërr, T. Furon. International Workshop on Digital Watermarking – IWDW 2005, LNCS 4283, pp. 1-15, 2006.

*Exploiting security holes in lattice data hiding*, L. Pérez-Freire, F. Pérez-González. Information Hiding – IH'07, LNCS, 2007.

# Outline

## Conclusion and future work

# Conclusion

## Is watermarking a secure primitive ?

- ▶ It depends on the applications ...
- ▶ For the schemes analysed so far, security levels are not high enough (personal).
- ▶ However, there is some hope ...

## Security is a hot issue

- ▶ Now that watermarking is pretty robust, security is worth studying (see above and e.g. Bas [07] )
- ▶ Security is generally underestimated
- ▶ Security is fun (BOWS contest <http://bows2.gipsa-lab.inpg.fr/>)
- ▶ New domain : security with signal processing tools

## New application domains

- ▶ Content based description
- ▶ Tracing the use of contents (traitor tracing/fingerprinting)
- ▶ Does this picture belong to this data basis ?
- ▶ Image Forensics
- ▶ Is this image authentic ?
- ▶ What is the model of the camera which took these pictures ?
- ▶ Did this camera take this picture ?

## More references

### Dissemination articles :

- ▶ *Le tatouage des images numériques*, Pour la Science, dossier "L'art du secret", été 2002
- ▶ *La stéganographie moderne*, M.I.S.C. numéro 18, 2005

### Books :

- ▶ *Tatouage de documents audiovisuels numériques*, Hermès-Lavoisier, 2004
- ▶ *Digital watermarking*, Academic Press, 2002

### Web sites :

- ▶ <https://www.picsi.org/accueil.html>
- ▶ <http://www.petitcolas.net/fabien/>
- ▶ <http://gleguelv.free.fr/>
- ▶ <http://www.openwatermark.org/>
- ▶ <http://www.jjtc.com/Security/>
- ▶ <http://www.watermarkingworld.org/>
- ▶ <http://jcberniere.free.fr/watermarking/WATERMK.HTM>
- ▶ <http://www.cosy.sbg.ac.at/~pmeerw/Watermarking/>