

Watermarks

Fernando Martínez
fernando.martinez@upc.edu

Departament de Matemàtica • Universitat Politècnica de Catalunya

14 de mayo de 2020

Data Hiding

- Esteganografía: El mensaje no está relacionado con la cobertura.
- Watermarking: El mensaje está relacionado con la cobertura.

Aplicaciones Watermarking

- Seguimiento de emisiones de TV y radio.
- Identificación del autor (propietario).
- Prueba de autoría (propiedad).
- Seguimiento de transacciones.
- Autentificación del contenido.
- Control de copia.
- Control de software y/o hardware.
- Identificación del todo a partir de una parte.

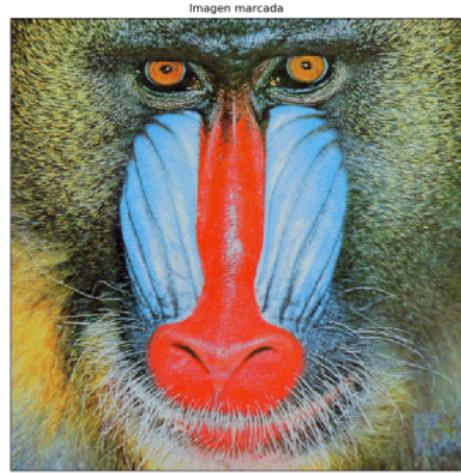
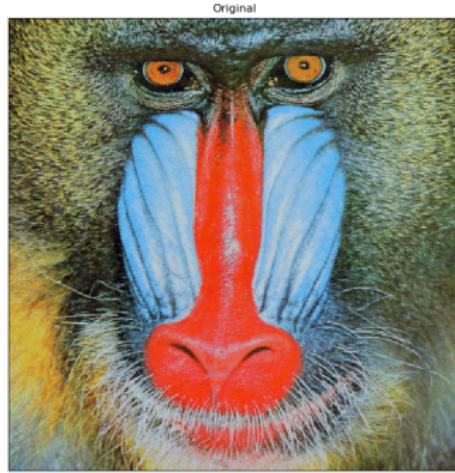
Watermarking: Propiedades

- Efectividad.
- Imperceptibilidad.
- Robustez.
- Seguridad.
 - Eliminación no autorizada.
 - Inserción no autorizada.

Watermarking: Tipos de ataques

- Robustez: eliminación de la marca.
- Presentación: no detección de la marca.
- Interpretación: dudas sobre la marca.
- Legales.

Watermarking: marca visible



Watermarking: un ejemplo sencillo (I)

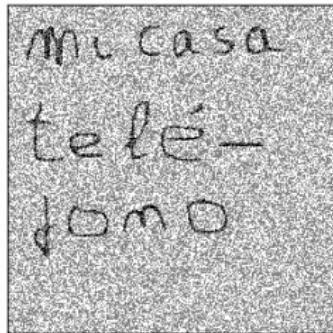
Se inserta la marca en el bit menos significativo



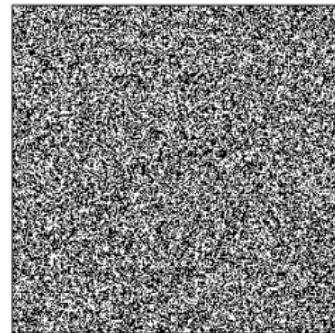
Watermarking: un ejemplo sencillo (II)

Diferencia entre imagen original e imagen:

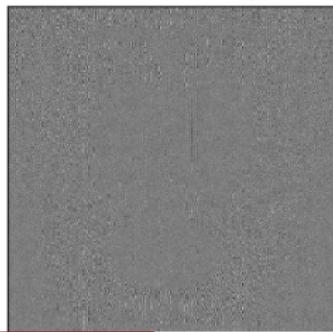
marcada



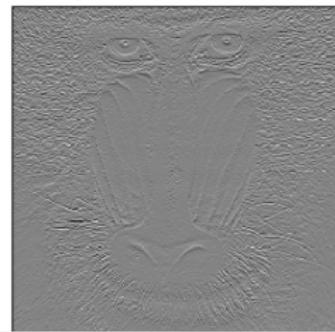
marcada con bit menos significativo a 0



marcada tras pasar JPEG



marcada tras eliminar la primera fila
y añadir otra al final



Watermarking: *Patchwork*

(a_i, b_i) N pares de píxeles elegidos a partir de una clave secreta K_s . Se modifican de acuerdo a:

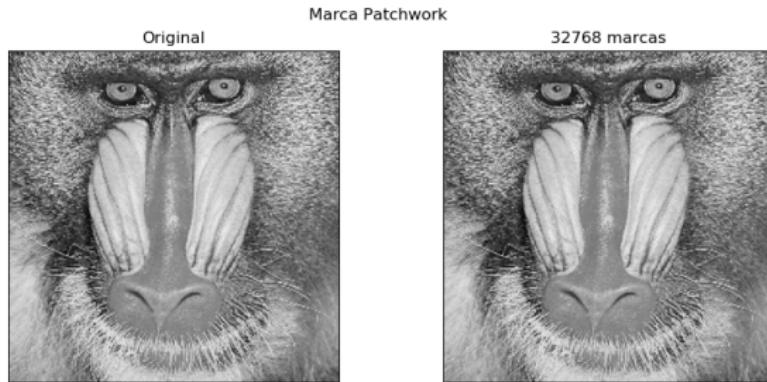
$$\tilde{a}_i = a_i + 1 \quad \tilde{b}_i = b_i - 1$$

Para extraer la marca, a partir de K_s se buscan los pares $(\tilde{a}_i, \tilde{b}_i)$ y se calcula

$$S = \sum_{i=1}^N (\tilde{a}_i - \tilde{b}_i).$$

Si $\begin{cases} S \approx 2N & \text{marca detectada,} \\ S \approx 0 & \text{no hay marca.} \end{cases}$

Watermarking: *Patchwork*



Tipo de imagen	S	ratio $\frac{S}{N}$
Marcada	58465	1.784
Marcada borrando bit menos significativo	58618	1.789
Marcada manipulada jpeg	11702	0.3579
Marcada manipulada fila	-5038	-0.154
Original	-6955	-0.212

Inserción de la marca en el dominio de frecuencias

A partir de una clave K_s se genera un vector \vec{m} con M componentes cuyos valores son números enteros.

Sean ν_i , $i = 1, \dots, M$, las M frecuencias más altas y $\alpha \in [0, 1]$.

Se modifican las frecuencias de acuerdo a:

$$\tilde{\nu}_k = \nu_k (1 + \alpha m_i)$$

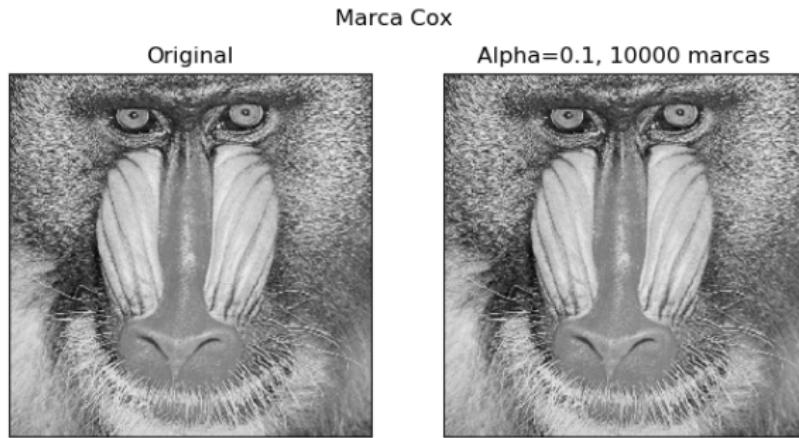
Alg. de Cox et al.: Recuperación

De la imagen original y de la imagen presuntamente marcada se extraen las M frecuencia mayores, ν_i y $\tilde{\nu}_i$ respectivamente.

Para cada una de las frecuencias se calcula:

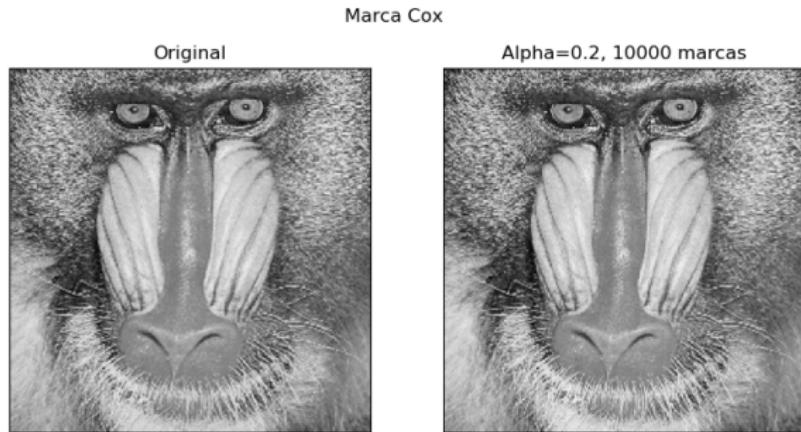
$$\tilde{m}_i = \left[\frac{1}{\alpha} \left(\frac{\tilde{\nu}_i}{\nu_i} - 1 \right) \right]$$

Imagen marcada $\alpha = 0,1$



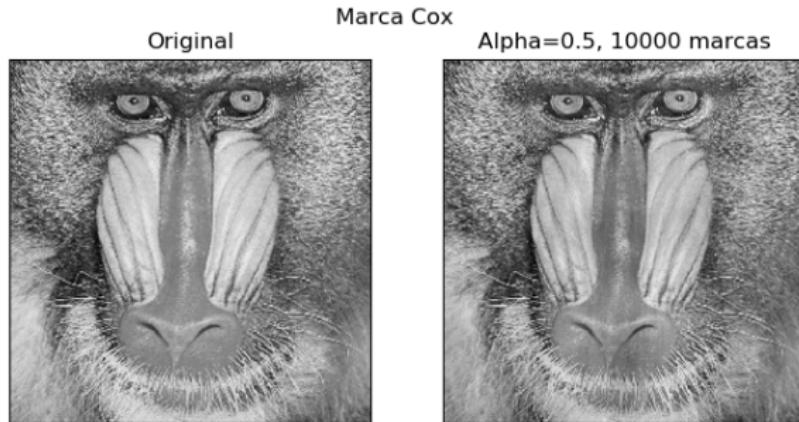
Tipo de imagen	marca recuperada
Marcada	1.000
Marcada borrando bit menos significativo	0.994
Marcada manipulada jpeg	0.606
Marcada manipulada fila	0.753
Marcada manipulada Haar	0.997

Imagen marcada $\alpha = 0,2$



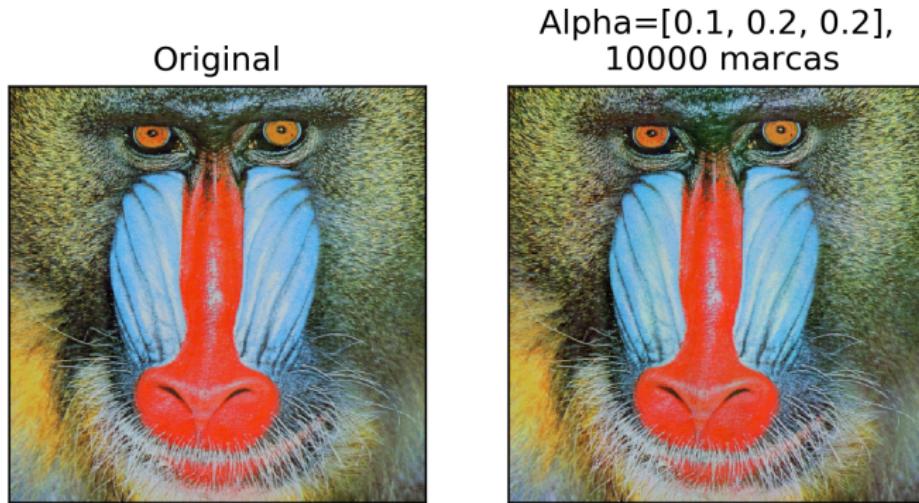
Tipo de imagen	marca recuperada
Marcada	0.999999
Marcada borrando bit menos significativo	0.998
Marcada manipulada jpeg	0.803
Marcada manipulada fila	0.869
Marcada manipulada Haar	0.998

Imagen marcada $\alpha = 0,5$



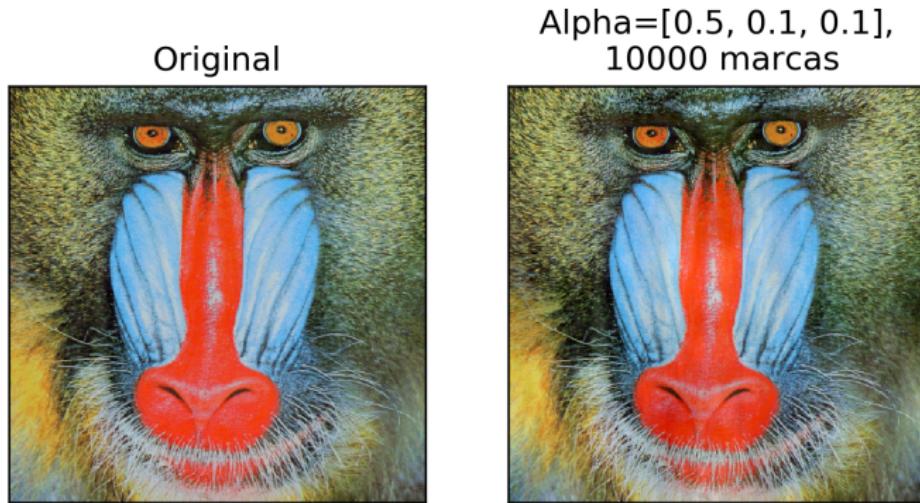
Tipo de imagen	marca recuperada
Marcada	1.0
Marcadaborrando bit menos significativo	0.9999
Marcada manipulada jpeg	0.977
Marcada manipulada fila	0.987
Marcada manipulada Haar	0.999

Imagen color marcada $\alpha = [0,1,0,2,0,2]$



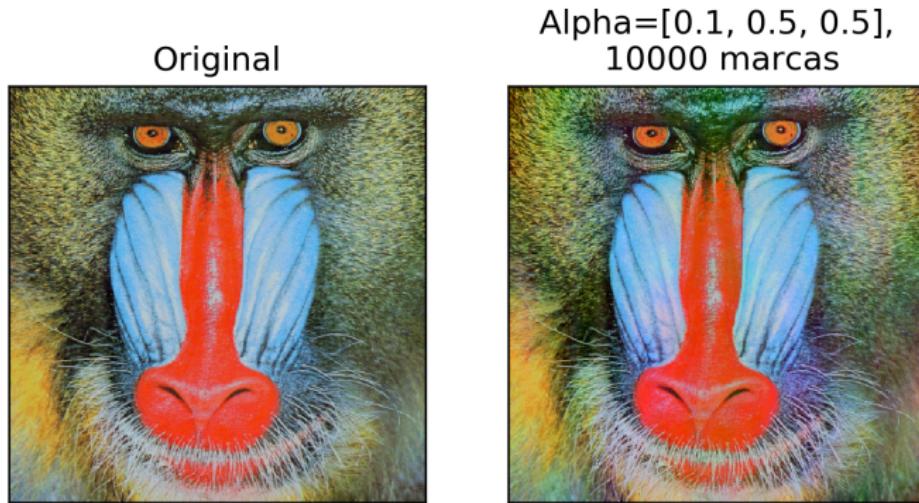
Tipo de imagen	marca recuperada
Marcada	(0.998, 0.998, 0.9995)
Marcada borrando bit menos significativo	(0.996, 0.997, 0.999)
Marcada manipulada jpeg	(0.6505, 0.690, 0.861)
Marcada manipulada fila	(0.724, 0.941, 0.961)

Imagen color marcada $\alpha = [0,5, 0,1, 0,1]$



Tipo de imagen	marca recuperada
Marcada	(0.9999, 0.979, 0.994)
Marcada borrando bit menos significativo	(0.999, 0.975, 0.992)
Marcada manipulada jpeg	(0.974, 0.367, 0.523)
Marcada manipulada fila	(0.980, 0.790, 0.817)

Imagen color marcada $\alpha = [0,1,0,5,0,5]$



Tipo de imagen	marca recuperada
Marcada	(0.9979, 0.999, 0.999)
Marcada borrando bit menos significativo	(0.996, 0.999, 0.999)
Marcada manipulada jpeg	(0.671, 0.948, 0.957)
Marcada manipulada fila	(0.733, 0.992, 0.989)