Capítulo 6. Algoritmos implementados

6.1 Introducción

En este capítulo, se detallan las características de los algoritmos de sellado digital de imágenes implementados, en una primera parte. Más tarde, se explica el protocolo de pruebas realizado, con el objeto de obtener resultados fiables y ciertos, así como los diferentes ataques implementados y banco de imágenes utilizado. Los resultados más significativos obtenidos tras la detección de la marca en función de diferentes ataques y técnicas de recuperación de la información escondida, se muestran después, estableciéndose comparaciones entre los algoritmos, (para un estudio más detallado de resultados véase anexo V). Finaliza el capítulo con las conclusiones más sobresalientes.

6.2 Algoritmos en dominios transformados de la imagen. Descripción

Se especifican las características de los algoritmos implementados, que insertan la marca de agua en diferentes dominios transformados (nos centramos en los dominios DCT y DWT) de la imagen, para conseguir una mayor robustez frente a la variedad de ataques llevados a cabo y/o una mejor característica de transparencia o invisibilidad de la marca de agua. Considerando las diferentes clasificaciones referidas en el capítulo 2 de técnicas de sellado invisible con marcas de agua, los algoritmos aquí realizados son privados, al igual que la mayoría de los mecanismos de watermarking basados en los dominios transformados de la imagen, en el sentido de que requieren del documento original para la recuperación de la información escondida.

6.2.1 Algoritmos en el dominio DCT de la imagen

Ilustramos las características de los algoritmos que insertan la marca de agua en el dominio DCT de la imagen según diferentes estrategias.

6.2.1.1 Algoritmo 1. Dominio DCT no basado en bloques

6.2.1.1.1 Descripción

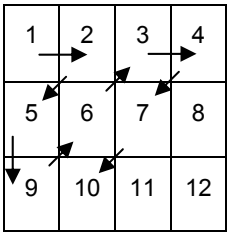
En este algoritmo se realiza la inserción de la marca de agua en los coeficientes obtenidos de aplicar la Transformada Discreta del Coseno (DCT) a

la imagen completa. Se procede, luego, a un escaneado en zig-zag de la imagen transformada de forma que resulte una secuencia unidimensional, para, finalmente, insertar los elementos de un vector de marca de agua de longitud ***n*** en los ***n*** mayores coeficientes escaneados (obviando el primer término, habitualmente referido como término de continua o DC).

El proceso de recuperación de la marca de agua es el inverso al de inserción, como se detallará más tarde.

6.2.1.1.2 Proceso de inserción de la marca de agua en la imagen

Como ya se ha indicado el proceso de inserción se realiza en los coeficientes DCT de la imagen, obtenidos de aplicar dicha transformada a la imagen completa, no por bloques. Estos coeficientes son escaneados en zig-zag de un modo similar al realizado en la compresión JPEG. Para una mejor comprensión del proceso la figura 6.1 trata de ilustrarlo. El mecanismo completo de inserción de la marca de agua en la imagen se muestra en el diagrama de bloques de la figura 6.2.



Vector resultado del escaneado

(1, 2, 5, 9,6 3, 4, 7,10, 13, 14, 11, 8, 12)

FIGURA 6.1. Escaneado realizado a los coeficientes DCT de la imagen

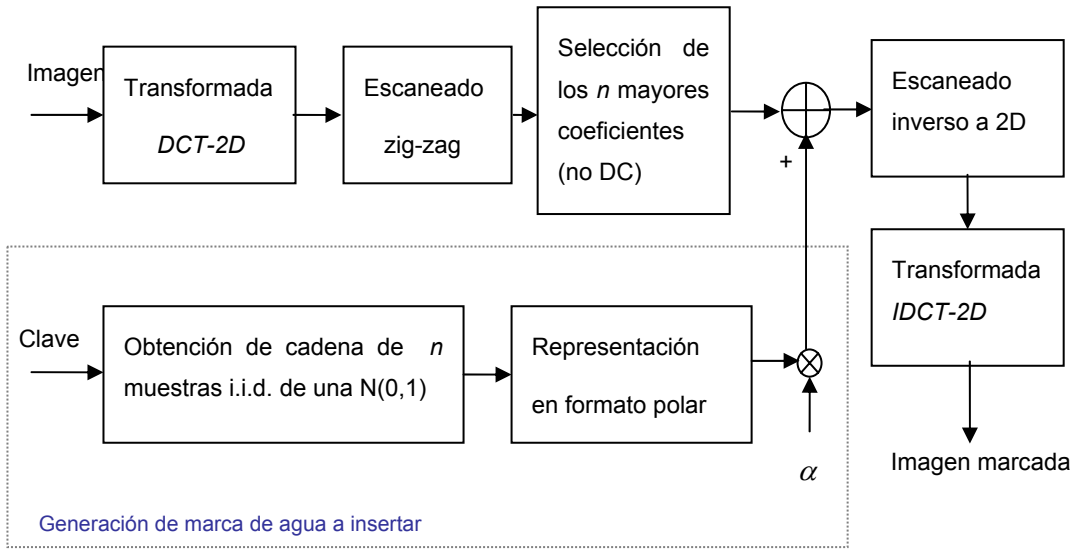


FIGURA 6.2. Esquema general del proceso de inserción de marca de agua del algoritmo 1.

De la observación de la figura se aprecia que la marca de agua generada se corresponde con una secuencia de muestras independientes e idénticamente distribuidas obtenidas a partir de una distribución normal con media cero y varianza unidad, que, luego, es mapeada a formato polar para poder insertarla en las coeficientes DCT seleccionados de la imagen.

La regla de selección de los bloque candidatos puede ser cualquiera, siempre que se consideren los criterios básicos de invisibilidad de la marca de agua ya enunciados a lo largo de los capítulos incluidos en la primera parte del presente documento. En nuestro caso, por su sencillez y excelentes características, los bloques candidatos serán los n (donde n vendrá determinado por la cantidad de bits de la marca de agua) mayores coeficientes, exceptuando, claro está, el término de continua, pues su modificación provocaría que la información escondida en la imagen marcada fuera visible.

La inserción de la marca de agua en el dominio DCT constituye una práctica habitual en sellado invisible de imágenes, pues, esta transformada elimina la correlación entre filas y columnas existente, además de concentrar la energía en unos cuantos coeficientes lo que permite distribuir la información de la marca de agua por toda la imagen con sólo marcar algunos de los coeficientes más significativos, posibilitando así su invisibilidad.

En la figura de la página siguiente se ilustra el proceso realizado en el algoritmo bajo estudio (versión similar a la de [Cox97]).

ALGORITMO DE INSERCIÓN

(1) Lectura de la imagen en color RGB y obtención de su versión en escala de grises para marcado simultáneo en color y en escala de grises. Ajuste de formato y clase.

(2) Transformada DCT en 2 dimensiones a las imágenes en color y en escala de grises.

(3) Selección de los n mayores coeficientes para las transformadas obtenidas en (2).

(4) Generación de la marca de agua (w) a partir de la clave como n muestras i.i.d de una N(0,1) y representación en formato polar de la misma.

(5) Adición de la marca de agua generada en (4), y ponderada por un factor de escala α, a los coeficientes seleccionados en (3).

(6) Transformada inversa IDCT en 2 dimensiones y ajustes de presentación de las imágenes marcadas en color y en escala de intensidades.

TABLA 6.1. Las etapas fundamentales del proceso de inserción de la marca de agua.

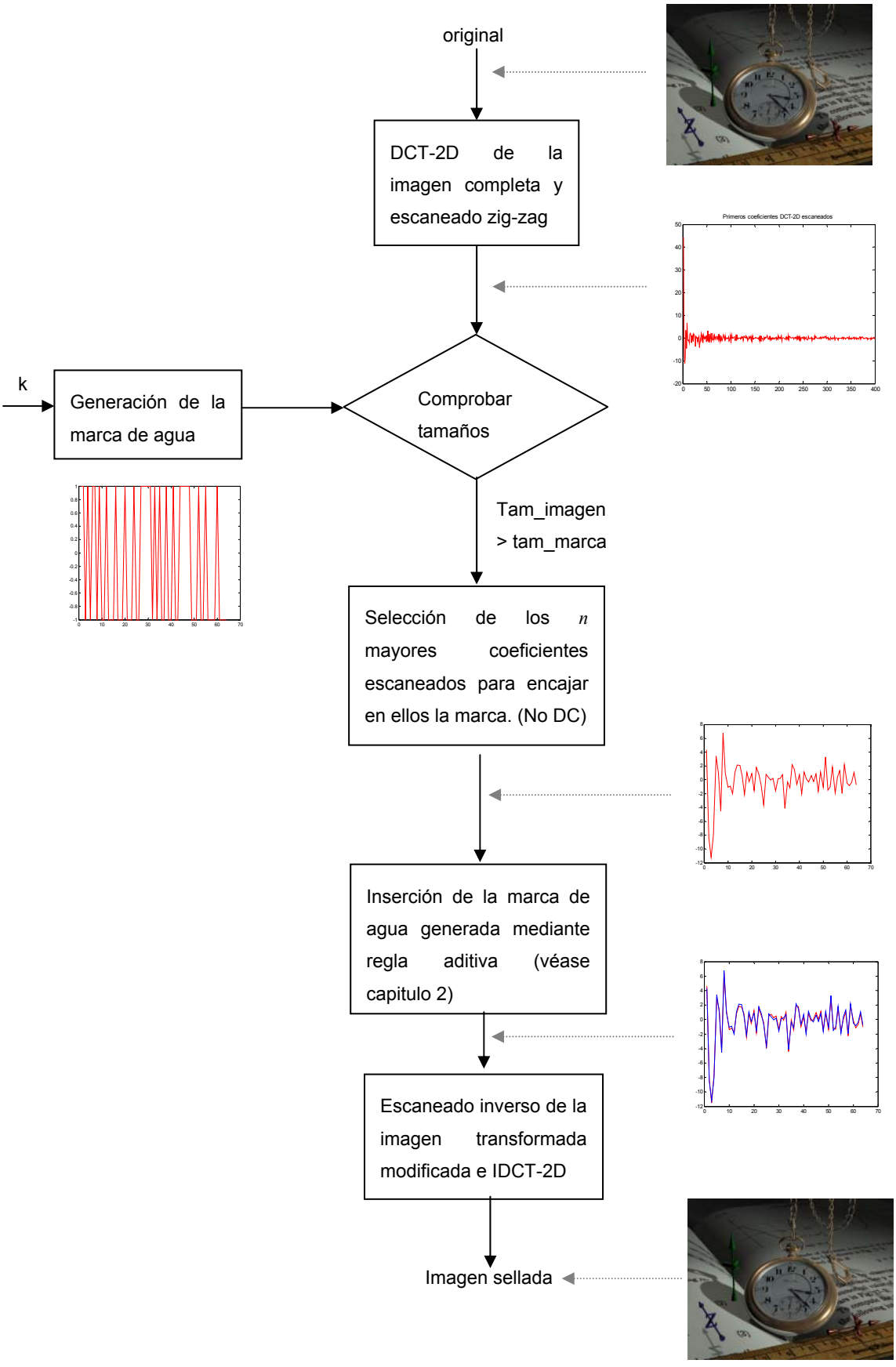


FIGURA 6.3. Esquema de marcado de la imagen en el algoritmo 1.

6.2.1.1.3 Proceso de detección/recuperación de la marca de agua

Como en todo sistema de comunicaciones, el proceso de recepción o recuperación de la información enviada es el inverso al llevado a cabo en la transmisión, que, en nuestro caso, viene representada por la inserción de la marca de agua en la imagen.

Notar que en el proceso de recuperación de la información escondida hacemos uso de dos métodos, como ya se ha señalado con anterioridad. Por un lado, empleamos el mecanismo tradicional de sustracción directa de la imagen original a la imagen marcada, lo que de aquí en adelante será referido como *estimación por sustracción o resta,* por otro lado se ha empleado un método de Análisis de Componentes Independientes (en concreto SICA) para la estimación de las señales originales, es decir, tanto de la imagen original como de la marca de agua, a partir de dos observaciones o mezclas consideradas, que para nosotros, en todos los algoritmos implementados, serán las imagen atacada y original o determinadas características de las mismas.

El problema del análisis de componentes independientes (véase segunda parte del proyecto para una lectura más detallada de las características fundamentales del mismo), aplicado a las técnicas de watermarking, en concreto al proceso de detección, encuentra una aplicación interesante puesto que es capaz de determinar la presencia de una marca de agua sin disponer de la imagen original, por ejemplo los vectores observados podrían ser la marca de agua usada en la inserción así como la imagen marcada, no disponiéndose de la versión original en la detección (no obstante según ciertas clasificaciones, las técnicas que disponen, en detección, de información acerca de la marca de agua insertada también se encuadran dentro de los métodos privados o semi privados, por lo que estrictamente hablado no se trataría de un algoritmo público). Sin embargo, no es el caso de presente documento, en el que pretendemos comparar esta técnica de detección basada en ICA con la tradicional, partiendo de un mismo conjunto de datos disponibles.

Realizar un Análisis de Componentes Independientes, consiste en encontrar una transformación lineal que pueda descomponer los vectores observados y hacer sus componentes lo más independientes posible, siempre que se cumplan ciertas condiciones (remitimos al capítulo 5 del presente documento para mayor detalle). Para la realización de ICA necesitamos disponer, como mínimo, de tantas observaciones como fuentes independientes

tengamos, es decir, si tenemos m sensores y n fuentes de señal, debe cumplirse siempre que $m \geq n$. En la siguiente tabla se ilustra la formulación general del problema ICA.

FORMULACIÓN ICA
Relación entra las fuentes y las observaciones es: $\mathbf{X}(t) = \mathbf{AS}(t)$ Donde $\mathbf{X}(t) = [x_1(t), x_2(t), ..., x_m(t)]^T$ es el vector de observaciones o mezclas simultáneas de n fuentes desconocidas $\mathbf{S}(t) = [s_1(t), s_2(t), ..., s_n(t)]^T$ donde $s_i(t),\ i = 1...n$ son n señales aleatorias mutuamente independientes. La matriz de mezcla es $\mathbf{A} = (a_{ij}),\ i = 1,2,...,m,\ j = 1,2,...,n$. La reconstrucción de las fuentes se conseguía mediante la estimación de la matriz **B**: $$\mathbf{Y}(t) = B\mathbf{X}(t) = \hat{\mathbf{S}}(t)$$

TABLA 6.2. Esquema genérico del problema de ICA

Si trasladamos la formulación general al caso concreto de watermarking, tenemos que los vectores observados, así como las componentes independientes que queremos encontrar son, en nuestra aplicación:

$$\begin{pmatrix} I \\ w \end{pmatrix} = \mathbf{B} \cdot \begin{pmatrix} (I_w)_{atacada} \\ I \end{pmatrix} \qquad \textbf{(6.1)}$$

Donde:

I: imagen original

w: marca de agua

$(I_w)_{atacada}$: imagen marcada atacada En cualquiera dominio de representación .

B: Matriz de separación.

Teniendo en cuenta todo lo visto podemos indicar las principales etapas del algoritmo de detección/recuperación de la marca de agua implementado. En el que se observa que se realiza la estimación del sello por los dos métodos propuestos de cara a poder realizar comparaciones entre ambos. Véase figura 6.4. En la tabla 6.3 se enumeran las etapas fundamentales realizadas en el proceso de detección.

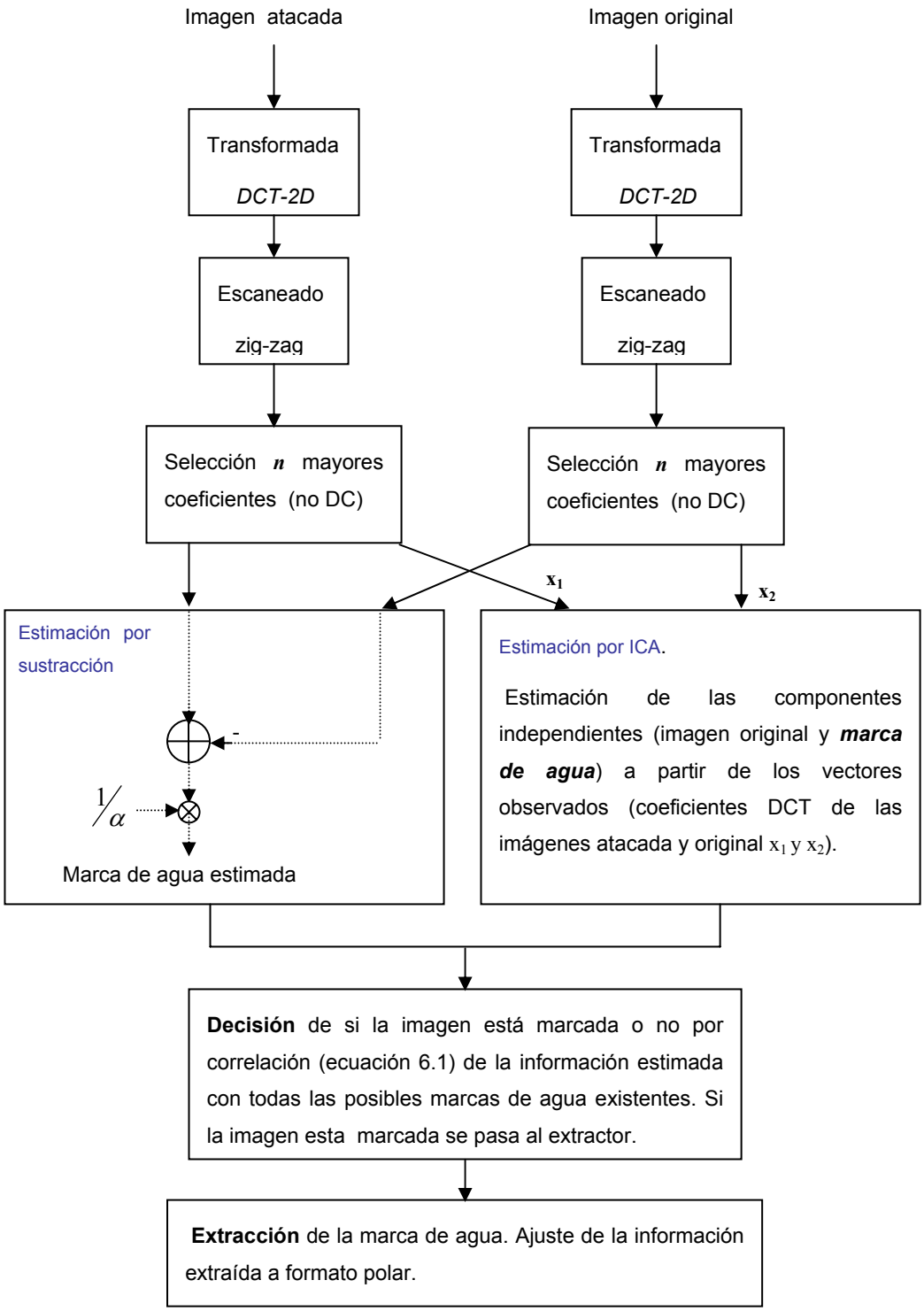


FIGURA 6.4. Diagrama de bloques del detector/extractor de marca de agua del algoritmo1.

ALGORITMO DE DETECCIÓN

(1) Lectura de las imágenes original y atacada. Ajustes.

(2) Preprocesado de la imagen atacada en cuanto a dimensión de la misma. Si el ataque sufrido por la imagen es un cortado de parte de la misma le añadimos el trozo eliminado de la imagen original sin marcar, si fue escalada, se le aplica un factor de escalado inverso.

(3) Transformada DCT de las imágenes original y atacada ajustada.

(4) Selección de los n coeficientes mayores de ambas transformadas.

(5) Sustracción de coeficientes DCT de la imagen original a los coeficientes DCT de la atacada. Ponderación inversa a la realizada en la inserción por el factor $1/\alpha$. El resultado es la estima por sustracción. Notar aquí que esta ponderación inversa no es imprescindible para la obtención de la marca de agua.

(6) Obtención de la matriz de separación aplicando algoritmo SICA basado (resultado de (4)).

(7) Selección de fuente independiente correspondiente a la marca de agua por correlación de las dos componentes obtenidas con la imagen. El resultado es la estima por ICA.

(8) Estadístico que mide la correlación entre las estimas obtenidas en (5) y (7) con todas las posibles marcas de agua. Si el valor del estadístico supera valor umbral y diferencia entre máximos global y siguiente local supera umbral relativo pasar a (9), sino detección fallida.

(9) Extracción de la marca de agua mediante presentación polar de las estimas obtenidas en (5) y (7). Obtención de probabilidades de error, de detección y de falsa alarma a partir del conocimiento de la clave.

TABLA 6.3. Algoritmo de detección/recuperación de la marca de agua realizado en el dominio DCT no basado en bloques.

6.2.1.2 Algoritmo 2. Dominio DCT completo y marca de agua en 2D

6.2.1.2.1 Descripción

En el algoritmo que detallamos en este apartado se implementa la transformada discreta del coseno (DCT en 2 dimensiones) a la imagen completa, dominio en el que se inserta una marca de agua ensanchada mediante una secuencia de máxima longitud, cuya generación depende del valor de una clave. De manera que cada bit de la marca de agua se convierte en un bloque de

sesenta y cuatro chips (trabajamos con bloques de dimensiones 8×8). La selección de los coeficientes DCT de la imagen candidatos a albergar la marca de agua está basada en las características de la DCT cuyos coeficientes de mayor energía se concentran en las zonas de baja frecuencia, luego serán estos coeficientes los que contengan a la marca de agua ensanchada, según se ilustra en la figura 6.6.

El esquema de detección y recuperación de la marca de agua consiste, en esencia, en el proceso inverso al de sellado. Notar que se trata de un algoritmo privado.

6.2.1.2.2 Proceso de inserción de la marca de agua

En el siguiente esquema se ilustra el proceso de inserción realizado en este algoritmo, en color añil se resaltan la diferencias clave entre el algoritmo aquí descrito y el de la sección anterior (apartado 6.2.1.1).

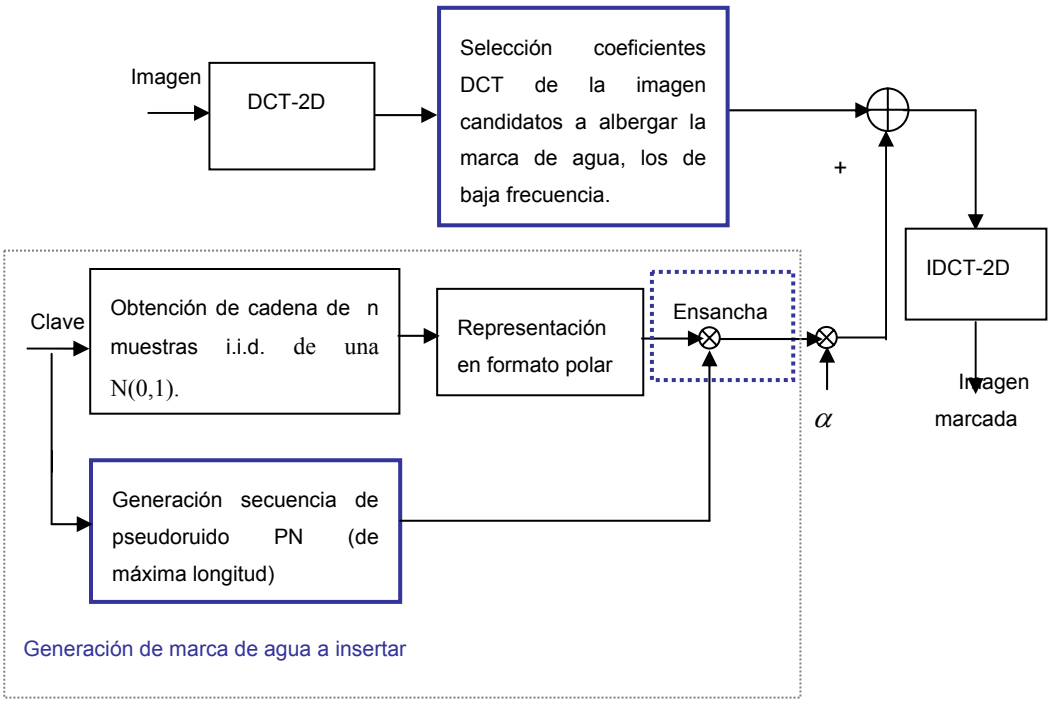


FIGURA 6.5. Proceso de inserción de marca de agua bidimensional en el dominio DCT de la imagen completa.

Las características, tanto de la estrategia de inserción utilizada como de la marca de agua ensanchada mediante el empleo de técnicas de espectro expandido de secuencia directa, donde la secuencia de ensanchado usada se obtiene como una secuencia de pseudo ruido de máxima longitud, (remitimos a

las funciones implementadas para mayor claridad, indicamos aquí únicamente que estas secuencias son obtenidas a partir de conceptos de codificación de bloques con códigos de máxima longitud, usados comúnmente para espectro expandido [Haykin97]), se muestran en la figura 6.6.

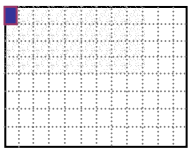
ESTRATEGIA DE SELECCIÓN DE BLOQUES EN LOS QUE INSERTAR LA MARCA DE AGUA.

(1) Descomposición en bloques de la imagen transformada. El tamaño de los bloques será el mismo que el de la secuencia de ensanchado en 2D usada.

(2) Selección de los primeros n bloques de la imagen transformada más cercanos a la esquina superior izquierda. Donde n es la longitud de la marca de agua antes de ser ensanchada (puesto que cada bit de la marca se convierte en un bloque de dimensiones determinadas por el factor de ensanchado empleado).

Ejemplo ilustrativo.

Coeficientes DCT de la imagen, divididos en bloques.



La inserción de la marca de agua se realiza de la forma indicada, el número de bloques fila y columna marcados depende del número de bits de la secuencia de la marca de agua antes de ser ensanchada. Para el ejemplo de la figura la secuencia sería de 36bits, (modelada en dos dimensiones como un bloque [4×9]), tras el ensanchado de cada bit a un bloque [8×8], las dimensiones finales de la marca de agua insertada serían de [32×72] bits.

Notar que para no marcar el término DC la secuencia de ensanchado usada (de dimensiones [8×8], presenta un valor nulo en la esquena superior izquierda).

FIGURA 6.6. Ilustración de la estrategia de inserción utilizada en el algoritmo 2. Además se aprecia que cada bit de una secuencia aletoria intermedia, generada de la misma forma que las marcas de agua del algoritmo 1, es expandido según un cierto factor de ensanchado.

La tabla 6.4 ilustra las etapas fundamentales involucradas en el proceso de generación de la marca de agua e inserción de la misma en la imagen realizado en este algoritmo.

ALGORITMO DE INSERCIÓN

(1) Lectura de la imagen en color RGB y obtención de su versión en escala de grises para marcado simultáneo en color y en escala de grises. Ajuste de formato y clase.

(2) Transformada DCT en 2 dimensiones a las imágenes en color y en escala de grises.

(3) División de la imagen transformada en bloques de las mismas dimensiones que la secuencia de ensanchado 2D utilizada.

(4) Selección de los bloques de la imagen a marcar según se ilustra en la figura 6.6. De modo que aseguremos que la marca de agua se localiza entre las componentes más significativas de la imagen.

(5) Generación de la marca de agua (w) a partir de la clave como n muestras i.i.d de una N(0,1) y representación en formato polar de la misma.

(6) Generación de una secuencia de pseudoruido para ensanchar la marca de agua obtenida en (5), a partir del valor de clave.

(7) Ensanchado de (5) con secuencia obtenida en (6) y presentación en 2 dimensiones.

(8) Ponderación o escalado de (7) con factor de energía α .

(9) Inserción de la marca de agua ensanchada en la zona de la imagen determinada en (4) mediante regla aditiva.

(10) Transformada inversa del coseno (IDCT-2D) de los coeficientes resultado de (9) junto con los que no fueron modificados.

(11) Ajustes de presentación.

TABLA 6.4. Algoritmo de sellado de imágenes en dominio DCT con marcas de agua ensanchadas.

Se marcan en color añil las diferencias esenciales entre el algoritmo descrito en esta sección y el algoritmo 1 (apartado 6.2.1.1).

6.2.1.2.3 Proceso de detección/recuperación de la marca de agua

El proceso de detección, realiza la recuperación de la marca de agua insertada mediante los dos métodos de detección propuestos, ya mencionados. A su vez, extrae la marca de agua de imágenes en color RGB y en escala de grises (al igual que el algoritmo anterior y siguientes). Un esquema ilustrativo del método se muestra en la figura 6.7, en la tabla 6.5 se enumeran las etapas principales del mecanismo de detección/recuperación de la marca de agua.

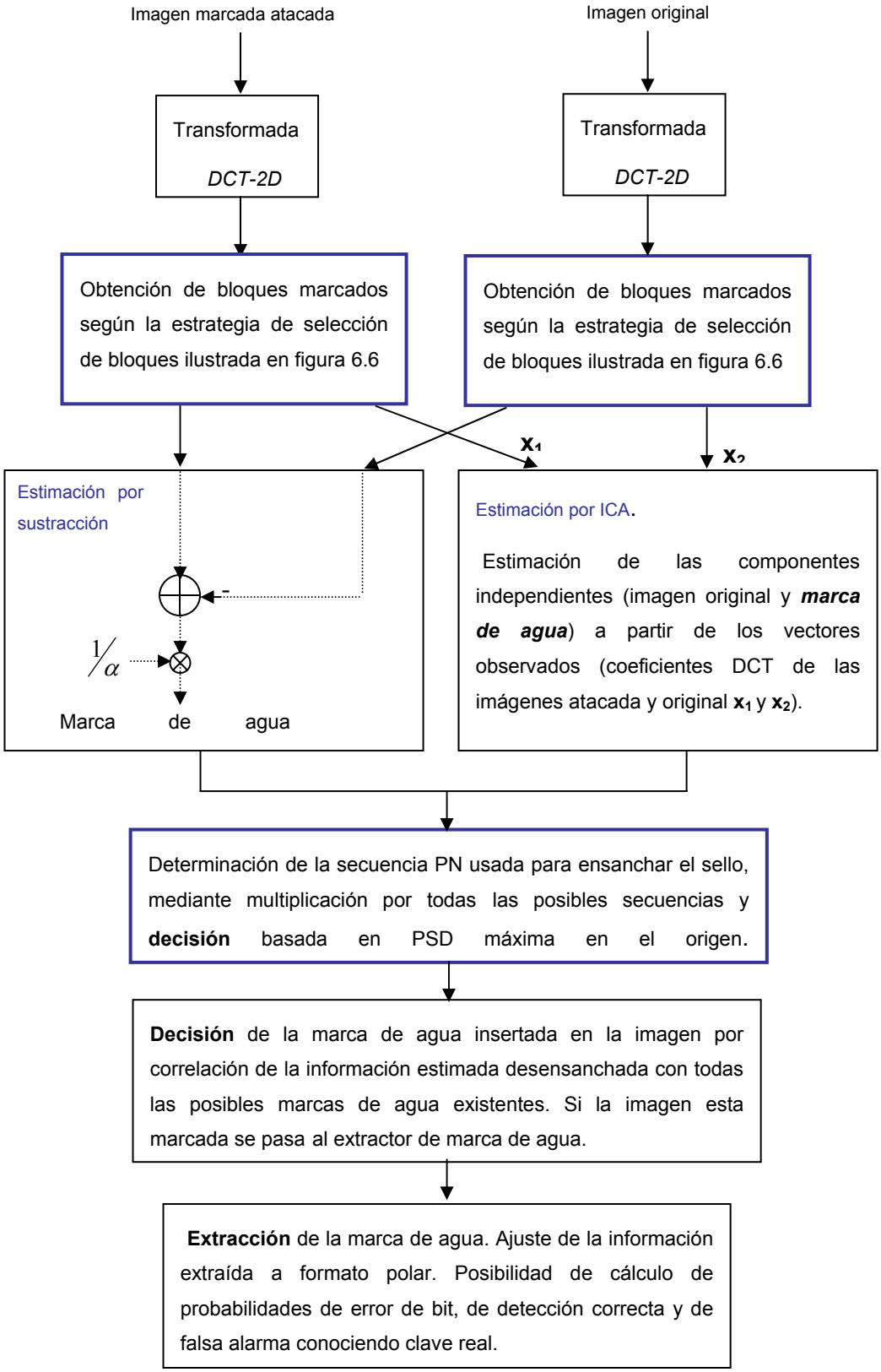


FIGURA 6.7. Diagrama de bloques del proceso de recuperación/extracción de la marca de agua.

Los cuadros en color añil de la figura 6.7 remarcan las diferencias existentes con el esquema de detección del algoritmo descrito en 6.2.1.1.

ALGORITMO DE DETECCIÓN

(1) Lectura de las imágenes original y atacada. Ajustes.

(2) Preprocesado de la imagen atacada con objeto de ajustar su dimensión. Si el ataque sufrido por la imagen es un cortado de parte de la misma le añadimos el trozo eliminado de la imagen original sin marcar, si fue escalada, se le aplica un factor inverso al usado en el ataque.

(3) Transformada DCT en 2 dimensiones de las imágenes original y atacada ajustada.

(4) Selección de los bloques que contienen la marca de agua según la regla de selección de bloques ilustrada en la figura 6.6.

(5) Sustracción de coeficientes DCT (en zona de interés) de la imagen original a los coeficientes DCT de la atacada. Ponderación inversa a la realizada en la inserción por el factor $1/\alpha$.

(6) Obtención de la matriz de separación aplicando algoritmo SICA a los vectores observados (resultado de (4)).

(7) Selección de fuente independiente correspondiente a la marca de agua por correlación de las dos componentes obtenidas con la imagen. El resultado es la estima por ICA.

(8) Determinación de la secuencia de pseudoruido usada para ensanchar la marca de agua. Decisión basada en PSD máxima en el origen.

(9) Multiplicación de cada bloque de la marca de agua ensanchada estimada (resultado de (5) y (7), según mecanismo de detección) por la secuencia PN que se estima fue usada y combinación de resultados para determinar la marca de agua original (no ensanchada).

(10) Estadístico que mide la correlación entre las estimas obtenidas en (9) con todas las posibles marcas de agua sin ensanchar. Si el valor del estadístico supera valor umbral y diferencia entre máximos global y siguiente local supera umbral relativo pasar a (11), sino detección fallida.

(11) Extracción de la marca de agua mediante presentación polar de las estimas obtenidas al realizar (9). Obtención de probabilidades de error, de detección y de falsa alarma a partir del conocimiento de la clave realmente usada.

TABLA 6.5. Etapas fundamentales del algoritmo de detección/recuperación de marca de agua en dominio DCT completo de la imagen con marca de agua en 2D. Para mayor claridad se remarcan en color añil las diferencias con el esquema de detección del algoritmo descrito en 6.2.1.1.

6.2.2 Algoritmos basados en dominios DWT y DCT de la imagen

En este apartado se van a estudiar algoritmos realizados mediante la utilización conjunta de los dominios Wavelet y DCT. En primer lugar se calcula la transformada Wavelet bidimensional de la imagen de nivel dos, con el objeto de separar del resto aquellas componentes de más baja frecuencia, que son las de mayor energía y, por tanto, en las que debemos insertar la marca de agua si deseamos disponer de un algoritmo que resulte robusto a técnicas de procesado que eliminan las componentes menos significativas de la imagen, característica general de los ataques realizados.

La transformada Wavelet divide la imagen inicial en diferentes subimágenes, como resultado de aplicar bancos distintos de filtros a la misma. Estas subbandas guardan información acerca de diferentes características de la escena procesada, así, obtenemos la denominada subbanda aproximada (conocida como LL_n, donde el subíndice n marca el nivel de descomposición de la Wavelet), que es la que almacena información referente a las componentes más significativas de la imagen, la de mayor energía, junto con diferentes subbandas de detalle (LH_n ,HL_n, y HH_n) que aportan información acerca de características de bordes, texturas, etcétera…(componentes de alta frecuencia, en general), véase apartado de anexos dedicado a esta transformada (anexo I).

Por todo lo dicho deducimos que la subbanda aproximada sería la candidata adecuada para albergar la marca de agua, pero la modificación directa de sus componentes presenta el inconveniente de que la marca no resulta invisible sino todo lo contrario, puesto que se modifican directamente las componentes de menor frecuencia de la imagen. En la figura 6.8 se muestran los resultados obtenidos de insertar la marca de agua directamente en la subimagen aproximada seleccionando para ello aquellos bloques de la misma que presentan mayor varianza (es decir tratando de optimizar la característica de invisibilidad). Se observa que la marca de agua es claramente visible, por lo que tratar de insertar la misma directamente en el dominio Wavelet de la imagen no resulta óptimo para el requisito de transparencia, que todo algoritmo de watermarking robusto debe cumplir.

Para suplir este inconveniente se puede optar por varias soluciones, en este caso se considera la posibilidad de añadir la marca de agua en los coeficientes obtenidos como resultado de aplicar la transformada discreta del coseno (*DCT-2D*) a la subbanda aproximada LL_2, (idea extraída de [Liu03]).

FIGURA 6.8. Imagen autumn.tif sellada empleando únicamente el dominio DWT, mediante marcado de los bloques de más varianza de la subbanda aproximada. Se observa claramente un efecto de bloque en las zonas marcadas.

6.2.2.1 Algoritmo 3. Dominio Wavelet y DCT de la imagen aproximada

6.2.2.1.1 Descripción

Como se ha indicado ya, en este algoritmo se realiza, en primer lugar, la transformada Wavelet bidimensional de nivel dos de la imagen, se escoge, luego, la subbanda aproximada y se le aplica a ésta la transformada discreta del coseno. A partir de aquí la estrategia de inserción, así como las características de la marca de agua insertada en los coeficientes DCT de la subimagen, son similares a las del algoritmo 1 descrito en la sección 6.2.1.1, con el objeto de poder comparar y ver si introduce alguna mejora el hecho de usar la transformada Wavelet combinada con la del coseno. Por lo que directamente mostramos los esquemas de sellado y de detección/recuperación de la marca de agua remarcando las diferencias con el algoritmo 1.

6.2.2.1.2 Descripción proceso de inserción de la marca de agua

El diagrama de bloques del algoritmo se ilustra en la figura 6.9. En la tabla 6.6 se enumeran la etapas básicas del algoritmo de sellado de la imagen, nótese que las diferencias con el algoritmo 1 se remarcan en color añil para mayor claridad.

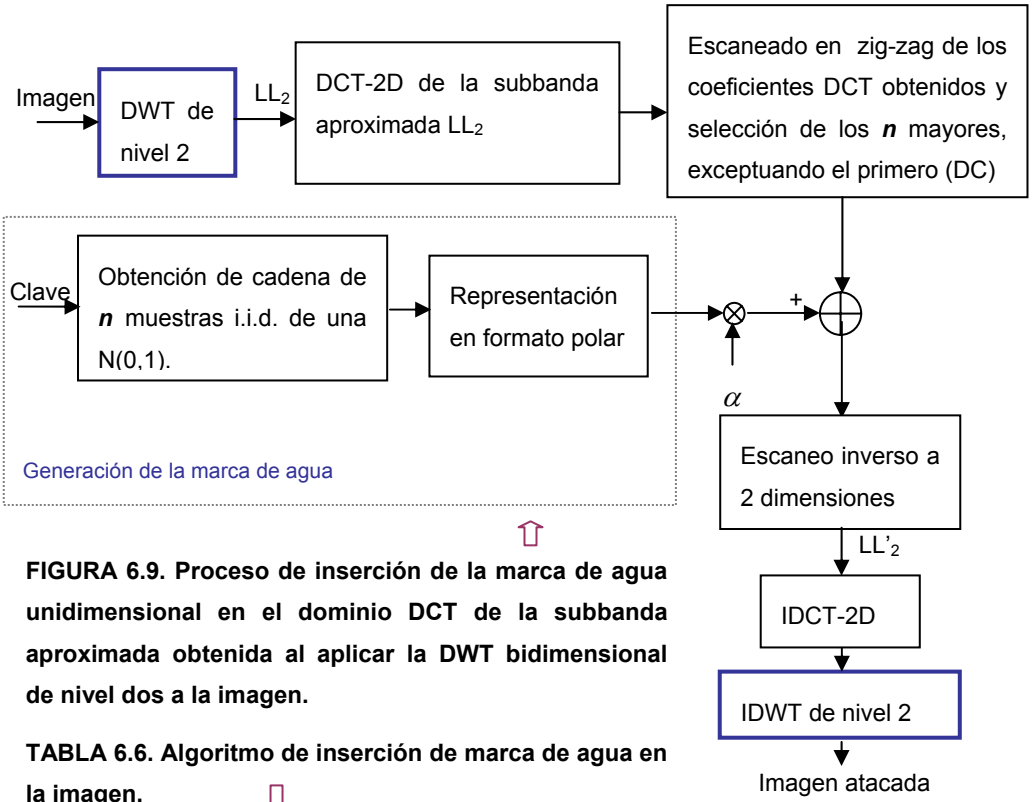


FIGURA 6.9. Proceso de inserción de la marca de agua unidimensional en el dominio DCT de la subbanda aproximada obtenida al aplicar la DWT bidimensional de nivel dos a la imagen.

TABLA 6.6. Algoritmo de inserción de marca de agua en la imagen.

ALGORITMO DE INSERCIÓN

(1) Lectura de la imagen en color RGB y obtención de su versión en escala de grises para marcado simultáneo en color y en escala de grises. Ajuste de formato y clase.

(2) Transformada *DWT* de nivel 2 a las imágenes en color y en escala de grises. Obtención de las subbandas aproximadas LL_{2BN} y LL_{2RGB}.

(3) Transformada *DCT* en 2 dimensiones a las subimágenes aproximadas determinadas en (2).

(4) Escaneado en zig-zag similar al empleado en compresión JPEG de los coeficientes obtenidos en (3).

(5) Selección de los *n* mayores coeficientes de las secuencias escaneadas para ambos tipos de imágenes.

(6) Generación de la marca de agua (*w*) a partir de la clave como secuencia de *n* muestras i.i.d de una *N(0,1)* y representación en formato polar de la misma.

(7) Adición de la marca de agua generada en (6), y ponderada por un factor de escala α, a los coeficientes seleccionados en (5).

(8) Escaneado inverso de las subimágenes modificadas.

(9) Transformada inversa *IDCT* en 2 dimensiones de subbandas obtenidas en (8).

(10) Transformada inversa *IDWT* de nivel 2 de las subimágenes marcadas junto con las subbandas de detalle no modificadas. Ajustes de presentación.

Se observa que la única diferencia con el algoritmo 1 es la inclusión de un nuevo dominio transformado de la imagen, el dominio Wavelet.

6.2.2.1.3 Descripción del proceso de detección/recuperación de la marca de agua

De la misma manera que en el proceso de inserción, un esquema genérico del proceso sería el mismo que el ilustrado para el algoritmo 1, al que habría que incluirle el uso de la transformada *DWT*, resultando en un diagrama de bloques tal como el mostrado en la figura 6.10. Las etapas fundamentales de proceso se ilustran en la tabla 6.7 (obsérvense diferencias con el algoritmo anterior, marcadas en color añil).

ALGORITMO DE DETECCIÓN

(1) Lectura de las imágenes original y atacada. Ajustes. Preprocesado de la imagen atacada en cuanto a dimensión de la misma.

(2) Transformada *DWT* de nivel 2 de las imágenes original y atacada ajustada. Selección de la subbanda aproximada.

(3) Transformada *DCT* en 2 dimensiones de las subbandas aproximadas obtenidas en (2) y escaneado en zig-zag de los coeficientes obtenidos.

(4) Selección de los *n* coeficientes mayores de ambas secuencias escaneadas transformadas.

(5) Sustracción de coeficientes *DCT* de la imagen original a los coeficientes *DCT* de la atacada. Ponderación inversa a la realizada en la inserción por el factor $1/\alpha$. El resultado es la estima por sustracción.

(6) Obtención de la matriz de separación aplicando algoritmo SICA a los vectores observados (resultado de (4)).

(7) Selección de fuente independiente correspondiente a la marca de agua por correlación de las dos componentes obtenidas con la imagen. El resultado es la estima por *ICA.*

(8) Estadístico que mide la correlación entre las estimas obtenidas en (5) y (7) con todas las posibles marcas de agua. Si el valor del estadístico supera valor umbral y diferencia entre máximos global y siguiente local supera umbral relativo pasar a (9).

(9) Extracción de la marca de agua mediante presentación polar de las estimas obtenidas en (7) y (9). Obtención de probabilidades de error, de detección y de falsa alarma a partir del conocimiento de la clave.

TABLA 6.7. Algoritmo de detección/recuperación de la marca de agua en dominios DWT y DCT.

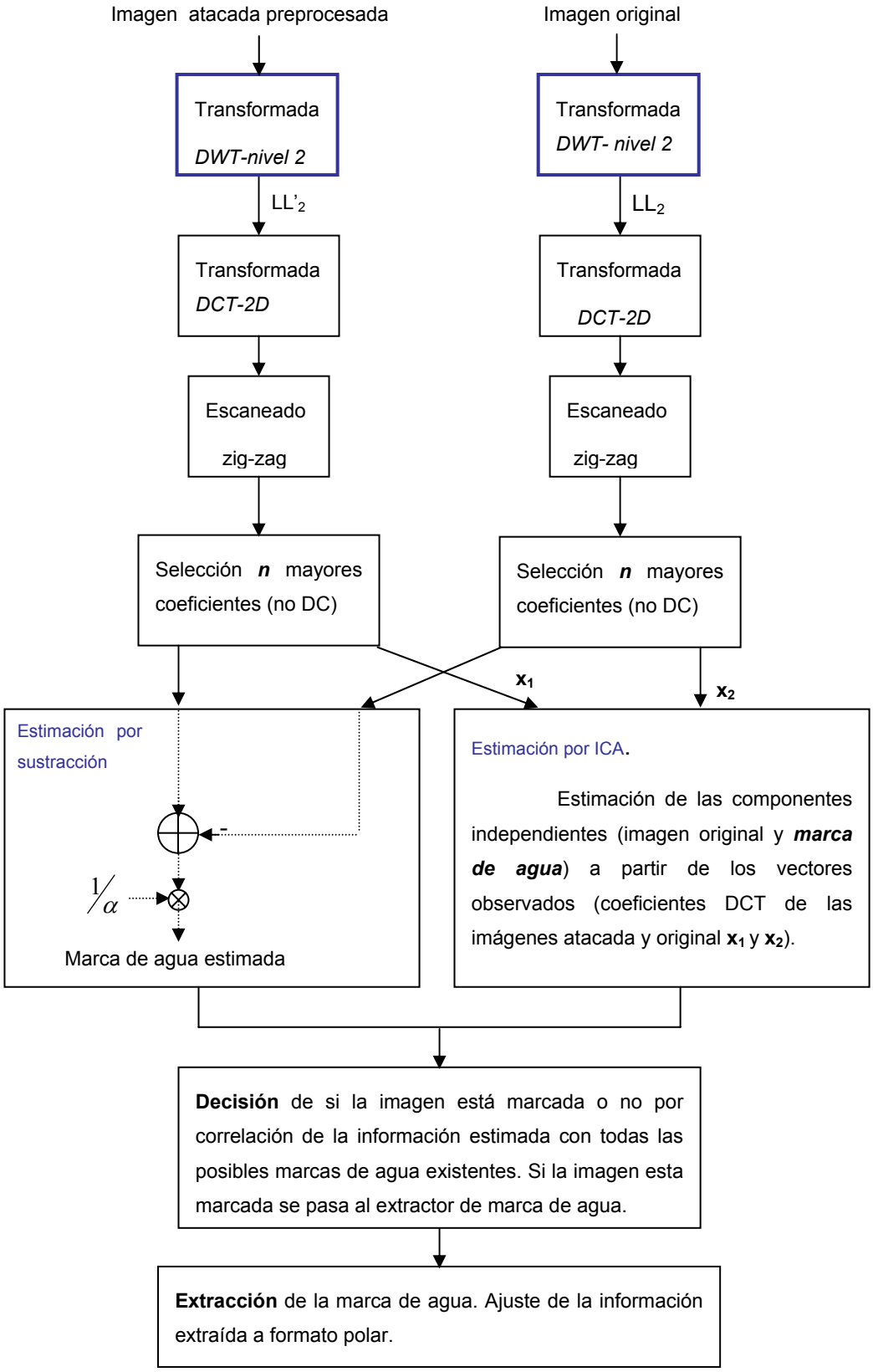


FIGURA 6.10. Esquema general de detección del algoritmo de sellado en los dominios DCT y DWT.

6.2.2.2 Algoritmo 4. Dominio Wavelet y DCT con marca de agua ensanchada en 2D

6.2.2.2.1 Descripción

El algoritmo aquí referido presenta cierto parecido con el descrito en la sección 6.2.1.2. La diferencia estriba en el empleo de la transformada Wavelet (que no se usó en el algoritmo 2), la selección de los bloques candidatos a contener la marca de agua consiste en el marcado de los primeros bloques de coeficientes DCT de la subimagen (resultado de aplicar la DWT a la imagen completa) tal y como se ilustra en la figura 6.11.

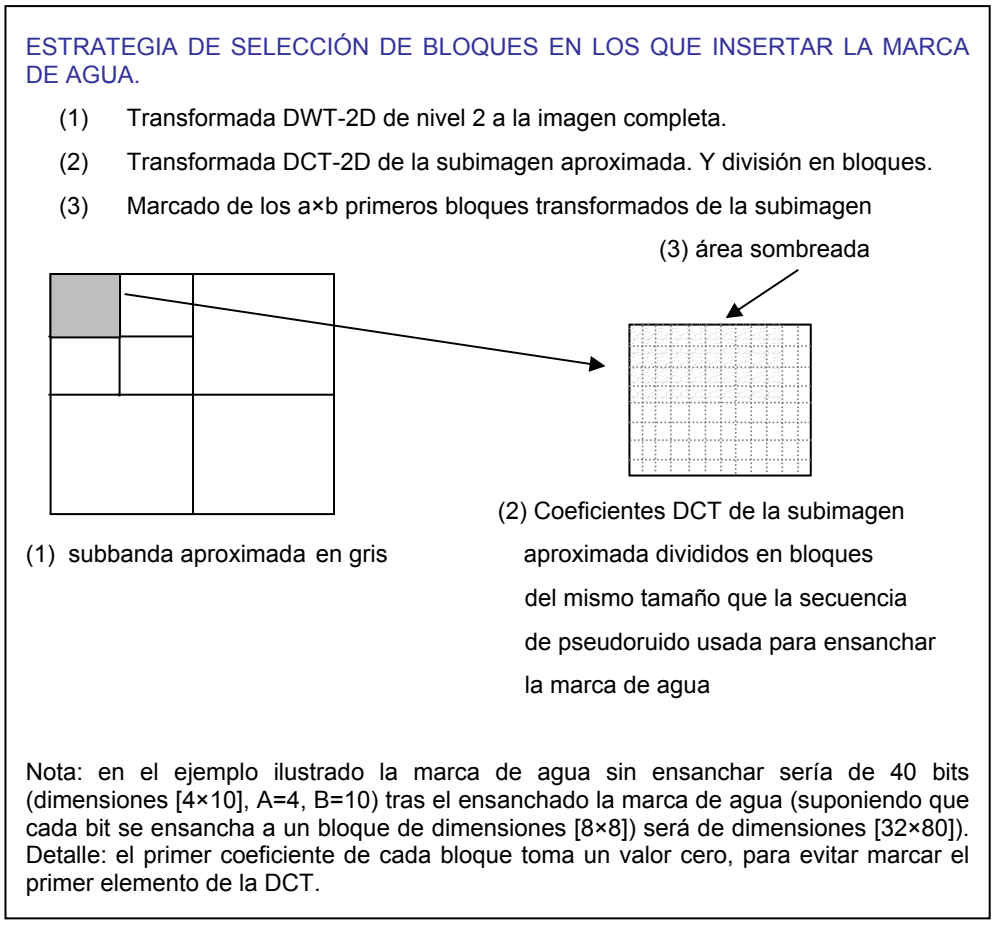


FIGURA 6.11. Ilustración de la estrategia de inserción llevada a cabo en el algoritmo 4.

En la figura anterior se ilustran las diferencias esenciales con el algoritmo 2 descrito en la sección 6.2.1.2 de modo que remitimos a mencionado apartado para conocer en detalle las características de los procesos de inserción y recuperación de la marca de agua, teniendo en cuenta las modificaciones

ilustradas en la figura 6.11, que radican en la incorporación de un nuevo dominio transformado.

6.2.3 Resumen

Con el objeto de aclarar ideas y de mantener en mente las características fundamentales de los algoritmos descritos en las secciones 6.2.1 y 6.2.2, se ilustran, en la siguiente tabla, los aspectos clave relacionados con el formato de las marcas de agua utilizadas y con la estrategia de inserción, según algoritmos.

MARCAS DE AGUA	**ALGORITMOS 1 y 3**	La marca de agua es una secuencia unidimensional pseudoaleatoria con distribución N(0,1), que mediante el uso del operador signo es convertida en una secuencia polar $\{\pm 1\}$
	ALGORITMOS 2 y 4	El proceso de generación de marca de agua parte de una secuencia intermedia similar a la de los algoritmos 1 y 3, pero realiza el ensanchado de cada bit de esta cadena mediante el uso de una secuencia de pseudoruido de máxima longitud (generada según fundamentos de códigos de bloque de máxima longitud (véase función *generar_secuencia_ensanchado* para mayor detalle)). Cada bit original se convierte en un bloque cuadrado de dos dimensiones (trabajamos con bloques de [8×8])
ESTRATEGIA DE INSERCIÓN	**ALGORITMO 1**	Se realiza el escaneado en zig-zag de los coeficientes DCT-2D de la imagen completa y se seleccionan los n mayores coeficientes para insertar en ellos la marca de agua
	ALGORITMO 2	Se realiza la DCT-2D a la imagen completa, pero no se escanean los coeficientes, sino que se divide la imagen transformada en bloques, seleccionándose los bloque de mayor energía para albergar la marca de agua.
	ALGORITMO 3	Idem algoritmo 1 pero antes de la DCT se introduce un nuevo domino transformado la DWT. La DCT se realiza a la subbanda aproximada fruto de esta transformación.
	ALGORITMO 4	Se realiza la DWT de nivel 2 a la imagen, la subbanda aproximada que se obtiene como resultado de esa transformación es de nuevo modificada al dominio de la transformada del coseno (DCT-2D). Los coeficientes transformados se dividen en bloques y se encaja la marca de agua bidimensional ensanchada en los primeros bloques de coeficientes transformados.

TABLA 6.8. Resumen de las características fundamentales de los algoritmos implementados.

6.3 Protocolo de pruebas

Ya se ha indicado que, independientemente del algoritmo considerado, la puesta en marcha para la obtención de resultados fiables, requiere de la

experimentación con imágenes de diferentes características (de texturas, bordes, y gamas reducidas o amplias de color) y del empleo de diversas marcas de agua para sellar dichas imágenes, pues, podría ocurrir que se insertara una determinada marca de agua cuyas características, en relación con la imagen en la que se inserta, la hicieran o bien muy robusta o extremadamente débil. Es por ello que deben seleccionarse varias marcas de agua para sellar cada una de las imágenes y luego promediar los resultados. Mostramos, a continuación el método de prueba común, realizado a los diferentes algoritmos.

- **Proceso de sellado.** En primer lugar se seleccionan de manera pseudoaleatoria diez claves para la generación de diez marcas de agua, que luego se insertan en cada una de las cinco imágenes usadas para probar los algoritmos (véanse en figura 6.12 sus dimensiones se indican en la tabla 6.9), los algoritmos están preparados para trabajar con imágenes en escala de grises y en color RGB (en cuyo caso se selecciona el plano de color azul de la imagen, pues nuestro sistema de visión es menos sensible a los cambios que se producen en dicho plano, lo que ya se apuntó en el capítulo 2, sección 2.2.3). De manera que tras realizar el proceso de sellado en cada algoritmo implementado se dispondrán de cien imágenes marcadas preparadas para ser atacadas (cincuenta en escala de grises y cincuenta en color verdadero RGB).

- **Proceso de ataque.** En esta etapa se realiza el ataque de las diferentes imágenes ya marcadas. Es el mismo proceso para todos los algoritmos expuestos, con el fin de poder comparar resultados. Los diferentes ataques implementados, junto con las características y grado de severidad de cada uno de ellos se muestran en la tabla 6.10. Por cada imagen marcada resultan un total de ochenta y tres imágenes atacadas que luego tendrán que ser procesadas para la extracción de la marca de agua encajada. Como se procede al ataque de cien imágenes marcadas, al final de esta etapa se dispone de un total de ocho mil trescientas imágenes atacadas por cada algoritmo probado.

- **Proceso de detección/extracción de la marca de agua.** En este apartado se realiza la detección y estimación de la marca de agua insertada en las diferentes imágenes atacadas mediante dos métodos:
  - Sustracción directa de la imagen original a la atacada.
  - Uso de herramientas de Análisis de Componentes Independientes.

De manera que en esta etapa serán procesadas ***ocho mil trescientas*** imágenes por cada uno de los dos mecanismos indicados.

El banco de imágenes utilizado se muestra en la figura siguiente, en la que puede observarse la presencia de imágenes con diferentes características (de bordes, texturas, creadas por ordenador…etcétera) y dimensiones.

(a)

(b)

(c)

(d)

(e)

FIGURA 6.12. Imágenes de intensidad empleadas para probar las características de los diferentes algoritmos implementados. (a)brandyrose.jpeg, imagen con un reducido conjunto de color, (b) imagen con gran cantidad de líneas y bordes, skyline_arch.jpeg, (c) watch.jpeg, creada por ordenador, (d) detalles finos en bear.jpeg y (e) autumn.tif.

Autumn.tif	Bear.jpg	Skyline_arch.jpg	Brandyrose.jpg	Watch.jpg
206×345	305×200	297×200	287×200	150×200

TABLA 6.9. Dimensiones de las imágenes mostradas en la figura 6.20. En formato M×N, donde M es el número de píxeles por filas y N el de columnas.

Como subapartado de esta introducción indicaremos las características del test de detección utilizado, por ser también común, en esencia, a todos los algoritmos de detección usados.

TABLA 6.10 (PÁGINA SIGUIENTE). Tabla que muestra las características de los diferentes ataques realizados a las imágenes.

ATAQUES REALIZADOS A LAS IMÁGENES SELLADAS	
FILTRADO DE MEDIANA	(2×2, 3×3, 4×4)
ATAQUE FMLR	ÚNICO
COMPRESIÓN JPEG	La mayoría de los algoritmos existentes dejan de funcionar a partir de valores para el factor de calidad inferiores a 50. (Factores de calidad de 90, 80, 70, 60, 50, 40, 35, 30, 25, 20, 15 Y 10)
TRANSFORMACIÓN GEOMÉTRICA LINEAL GENERAL $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$	Con los parámetros siguientes: (a, b, c, d) = (1.010, 0.013, 0.009, 1.011), (1.007, 0.010, 0.010, 1.012), (1.013, 0.008, 0.011, 1.008)
CAMBIO DE LA RELACIÓN DE ASPECTO	Se modifica la relación de aspecto de la imagen por una pequeña cantidad, los parámetros del ataque son (x,y)=(0.8,1.0), (0.9,1.0), (1.0,0.8), (1.0,0.9), (1.0,1.1), (1.0,1.2), (1.1,1.0), (1.2,1.0).
BORRADO SIMÉTRICO Y ASIMÉTRICO DE FILAS Y COLUMNAS ELEGIDAS ALEATORIAMENTE	Posibilidades: (1, 1), (1, 5), (5, 1), (5, 17), (17, 5). La primera componente es el número de columnas borradas y la segunda el de filas
ESCALADO DE LA IMAGEN POR DIFERENTES FACTORES.	Factores de escalado usados 0.5, 0.75, 0.9, 1.1, 1.5 y 2.
ROTACION POR UN PEQUEÑO ÁNGULO Y CORTADO	Esto simula en parte lo que se obtiene cuando se escanea una imagen, es imposible la alineación perfecta. Los ángulos de rotación aplicados a las imágenes son -2º, -1º, -0.75º, -0.5º, -0.25º, 0.25º, 0.5º, 0.75º, 1º y 2º.
ROTACIÓN POR UN PEQUEÑO ÁNGULO (los mismos casos que en el anterior) **SEGUIDO DE ESCALADO Y CORTADO**	En este caso se realiza un escalado de la imagen rotada y cortada para mantener el tamaño original
ROTACIÓN POR UN ÁNGULO GRANDE	5, 10, 15, 30, 45 Y 90 grados
SHEARING SIMÉTRICO Y ASIMÉTRICO	En dirección X y/o Y: (0, 1), (0, 5) (1, 0), (5, 0), (1,1) y (5, 5). La primera componente marca el desplazamiento en la dirección X (porcentaje del ancho) y la segunda el desplazamiento en la dirección Y (porcentaje de la altura).
CORTADO CENTRADO	1%, 2%, 5%, 10%,, 15%, 20%, 25%, 50% Y 75%
FILTRADO GAUSSIANO	$H=\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$
FILTRADO SHARPENIG	$H=\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$
DISTORSIÓN GEOMÉTRICA ALEATORIA IMPLEMENTADO "STIRMARK"	Se trata de una distorsión aleatoria que realiza un desplazamiento de los píxeles no uniforme, siendo prácticamente nulo en los bordes de la imagen y máximo en el centro

6.3.1 Test de detección de la marca de agua. Justificación de su elección

Para justificar la presencia de una determinada marca de agua en la imagen recibida en el detector se obtiene una medida del parecido entre la marca extraída y las incluidas en un banco de sellos disponible en el receptor que contiene todas las posibles marcas que se pueden insertar en la imagen en relación con la clave pública utilizada. Esta medida se calcula de la siguiente manera:

$$sim(w,\hat{w}) = \frac{\sum_i w(i)\cdot \hat{w}(i)}{\sqrt{\sum_i \hat{w}(i)^2}} > T_0 \qquad \textbf{(6.2)}$$

Donde el estadístico utilizado representa, en esencia, una medida de correlación entre secuencias y T_0 se corresponde con el valor del umbral que determina si la marca de agua $w(i)$ está presente en la imagen bajo prueba, asegurando un valor máximo para la probabilidad de falsa alarma, lo denominaremos como umbral de detección por correlación. Aunque no de manera explícita el valor de este estadístico en presencia de marca de agua, presenta cierta relación directa con el tamaño de la marca, de manera que el valor del umbral debe decidirse de un modo prueba y error hasta encontrar el más adecuado.

La elección de este criterio está basada en la consideración de las aplicaciones más habituales de las técnicas de sellado con marcas de agua, es decir, autentificación, protección de los derechos de copia y de autor, etcétera…en estos casos minimizar la probabilidad de pérdida de detección no es una situación tan restrictiva como pueda ser reducir al máximo las falsas alarmas, no cabe duda de que decidir que un impostor es dueño de una obra presenta mayores riesgos que el hecho de que en un momento dado no sepamos determinar quién es el autor o dueño de un cierto documento.

6.4 Resultados y comparación entre algoritmos

En este apartado se muestran los resultados más significativos obtenidos para los diferentes algoritmos implementados. En primer lugar, se indican los parámetros de experimentación o valores asignados a los parámetros de inserción y/o detección en los diferentes algoritmos, después se ilustra la característica de transparencia o invisibilidad de la marca de agua en la imagen

sellada, según algunas medidas objetivas de distorsión habitualmente usadas en tratamiento de imágenes, como pueden ser la relación pico de señal a ruido (PSNR) o el error cuadrático medio (MSE), cuya formulación puede verse en el anexo III, relacionado con medidas de distorsión usadas en imágenes. También se muestran las características de robustez de los diferentes sistemas de watermarking probados, o lo que es lo mismo, la efectividad de los ataques realizados a las imágenes marcadas. Finalmente se indicará el tiempo promedio de computación de los procesos de sellado y detección/recuperación de la marca de agua, obtenidos para todos los algoritmos estudiados.

6.4.1 Parámetros de experimentación

Los valores asignados a los diferentes parámetros en cada uno de los algoritmos, tanto de inserción como de detección, para la realización de la simulación y la obtención de resultados se indican en las tablas 6.11 y 6.12.

PARÁMETROS DE INSERCIÓN	**ALGORITMO 1**	Factor de escalado o energía: $\alpha = 0.2$ Marca de agua: cadena de 64bits representada en formato polar $\{\pm 1\}$
	ALGORITMO 2	Factor de escalado o energía: $\alpha = 0.1$ Marca de agua de cadena de 16bits representada en formato polar $\{\pm 1\}$. Secuencia PN de máxima longitud en 2D: bloque de dimensiones 8×8 (chips), que ensancha cada uno de los bits de la marca de agua.
	ALGORITMO 3	Factor de escalado o energía: $\alpha = 0.2$ Marca de agua: cadena de 64bits representada en formato polar $\{\pm 1\}$
	ALGORITMO 4	Factor de escalado o energía: $\alpha = 0.1$ Marca de agua de 16bits representada en formato polar $\{\pm 1\}$ obtenida a partir de muestras de una N(0,1) Secuencia de pseudoruido PN de máxima longitud, adaptada a dimensiones [8 8], que ensancha cada uno de los 16 bits de la secuencia aleatoria

TABLA 6.11. Valores asignados a los parámetros de inserción en los diferentes algoritmos implementados.

En la tabla 6.11 se muestran los valores asignados a los diferentes parámetros de inserción, donde α es el factor que limita la energía con que la

marca de agua se inserta en la imagen, como se ha indicado anteriormente, y el factor de ensanchado representa la cantidad de chips que contiene la secuencia de pseudoruido empleada para expandir la marca de agua.

PARÁMETROS DE DETECCIÓN	**ALGORITMO 1**	Umbral de detección por correlación: *umbral_sim =4.50* Umbral de diferencia entre máximos: *umbral_relativo = 1*
	ALGORITMO 2	Umbral de detección por correlación: *umbral_sim =4* Umbral de diferencia entre máximos: *umbral_relativo = 1*
	ALGORITMO 3	Umbral de detección por correlación. *umbral_sim=4.50* Umbral de diferencia entre máximos *umbral_relativo=1*
	ALGORITMO 4	Umbral de detección por correlación. umbral_sim=4 Umbral de diferencia entre máximos umbral_relativo=1 Medida adicional redundante por comparación de las claves estimadas para la secuencia PN y para la secuencia desensanchada de 16 bits, que deben coincidir. Idem en algoritmo 2

TABLA 6.12. Valores asignados a los parámetros de detección/recuperación de la marca de agua en los diferentes algoritmos implementados.

Donde umbral de detección por correlación representa al parámetro To de la expresión (6.2) y umbral de diferencia entre máximos es una medida adicional de redundancia que asegura que la detección se produzca cuando el parecido entre las secuencias comparadas presente clara ventaja con respecto a la comparación de la estima con otras secuencias.

6.4.2 Transparencia

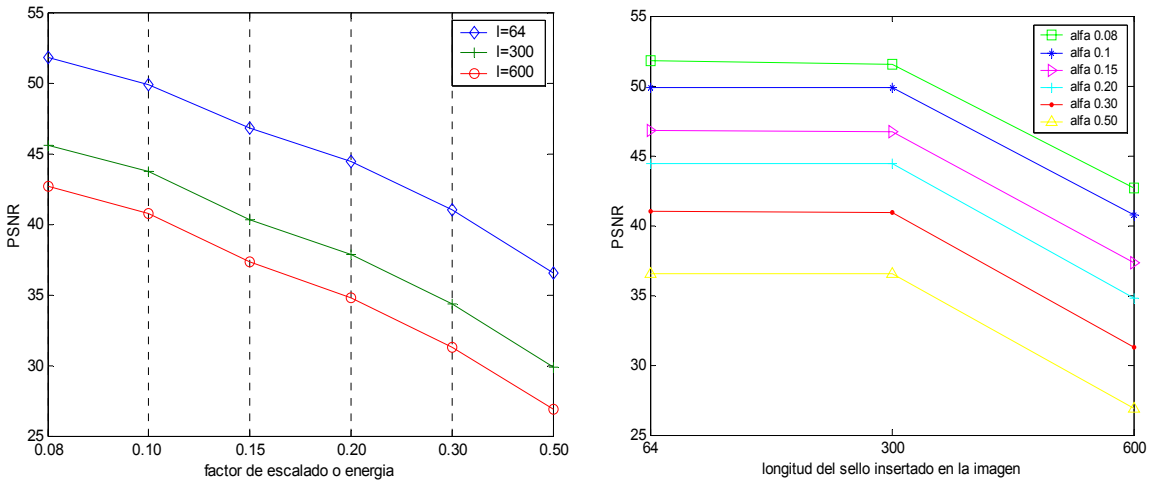
En primer lugar, tendríamos que comprobar si los valores asignados a los parámetros de inserción mostrados en la tabla 6.11, son adecuados, en el sentido de que la marca de agua insertada, por cualquiera de los métodos probados, no degrade de forma considerable la calidad de la imagen desde el punto de vista perceptual. El proceso para determinar el valor adecuado de estos parámetros, consistiría en probar con marcas de agua de diferente tamaño así como con diferentes valores del factor de escalado α. Ilustramos el proceso llevado a cabo en el algoritmo 3 (sección 6.2.2.1), similar al realizado en todos los demás.

6.4.2.1 Proceso de selección de los parámetros de inserción

El proceso para determinar los valores adecuados de los parámetros de inserción de la marca de agua en la imagen consistió en realizar el sellado de las imágenes de prueba tanto en su versión en blanco y negro como en color, para los siguientes valores de los parámetros de longitud de la marca de agua y factor de escalado o energía, midiéndose luego la PSNR de cada imagen marcada con respecto a la original.

- La longitud de la marca da agua toma los siguientes valores: $n = 64$, $n = 300$ y $n = 600$.
- El factor de escalado también se hace variar tomando los valores: $\alpha = 0.08, \alpha = 0.15, \alpha = 0.2, \alpha = 0.3$ y $\alpha = 0.5$.

Los resultados obtenidos para la PSNR muestran que dicha magnitud reduce su valor a medida que la longitud del sello insertado aumenta, fijado un determinado valor para el factor de energía, si lo que se fija es la longitud de la marca y modificamos el factor de escalado, se observa nuevamente cómo decrece el valor de la PSNR. Las siguientes gráficas muestran lo indicado.



(a) (b)

FIGURA 6.13. Variación de la PSNR en función (a) del factor de escalado para una longitud fija de la marca de agua y (b) de la longitud de la marca de agua, para factor de energía fijo.

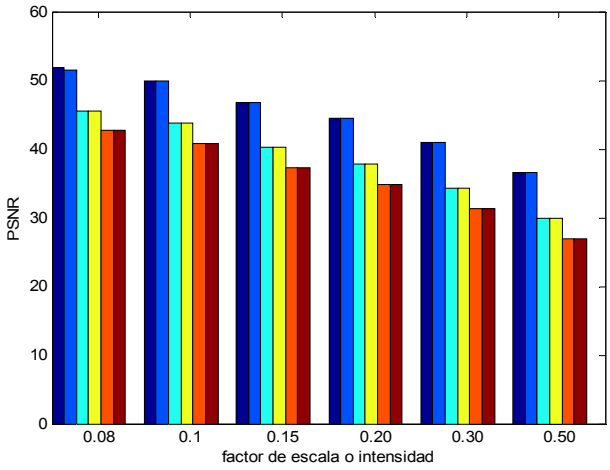


FIGURA 6.14. Valor de PSNR en función de la longitud de la secuencia y del factor de escalado. Cada barra representa el promediado realizado a las imágenes en color (barras impares) y en blanco y negro (barras pares), es decir, para cada valor de α, n = 64 para el primer par de barras, n = 300 para el segundo par de barras y n = 600 para el tercer par de barras.

Podemos observar que la elección de $\alpha = 0.2\, y\, n = 64$, parámetros seleccionados para la prueba del algoritmo (véase tabla 6.11) presenta unos buenos resultados en el valor de la PSNR.

6.4.2.2 Valores de la PSNR según parámetros de inserción seleccionados

Según el criterio de transparencia se observa que, en general, los algoritmos que añaden a la imagen marcas de agua unidimensionales, es decir los algoritmos 1 y 3, presentan mejores características en cuanto a menor degradación de la calidad visual de las imágenes marcadas. Para observar esto se emplearon dos medidas objetivas, la PSNR (relación pico de señal a ruido) y el MSE (error cuadrático medio) (véase anexo III).

Además, el gran parecido entre el nivel de degradación que produce el marcado de la imagen mediante los algoritmos 1 y 3 por un lado y los algoritmos 2 y 4 por el otro, nos permite establecer comparaciones dos a dos entre los resultados obtenidos en cuanto a robustez o resistencia a los diferentes ataques implementados.

En las gráficas de la figura 6.15 se muestran los valores obtenidos para ambas métricas desglosados según algoritmo de sellado empleado e imagen

original (promediando entre los valores obtenidos para cada una de las imágenes marcadas en cada algoritmo, es decir cien imágenes).

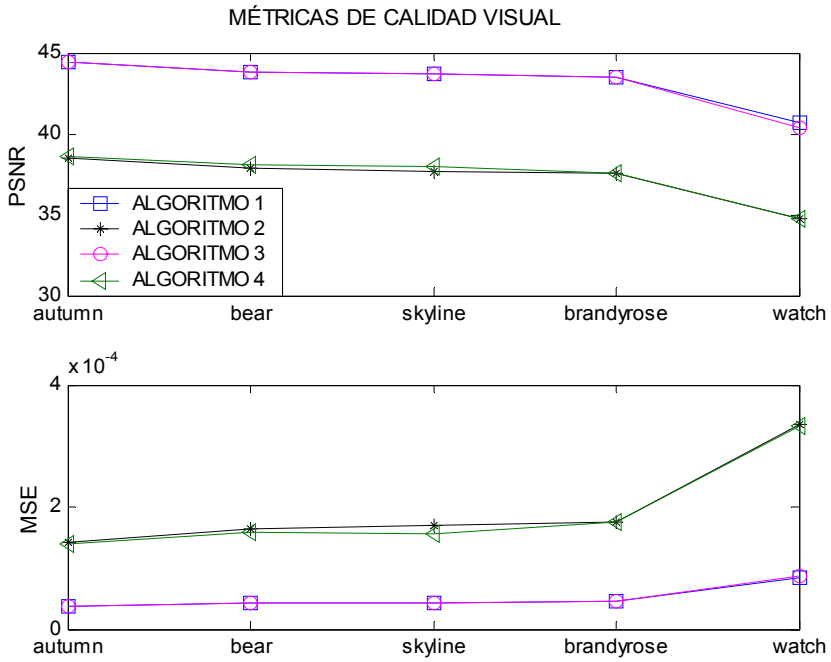


FIGURA 6.15. Ilustración de las métricas de distorsión de las imágenes marcadas con respecto a su versión original sin marcar.

Se observa en la figura que los niveles de degradación provocados por la presencia de la marca de agua, con los valores de los parámetros de inserción mostrados en la tabla 6.11 son aceptables en todos los algoritmos. Medidas subjetivas por observación directa de las imágenes corroboran lo indicado.

6.4.3 Robustez

La respuesta de los algoritmos implementados frente a los diferentes ataques que pueda sufrir la imagen marcada, constituye una medida importante de la robustez y/o debilidad de las técnicas de watermarking estudiadas. Es por ello que mostramos la característica de resistencia de la marca de agua (frente a los ataques mostrados en la tabla 6.10), es decir, la probabilidad de detección correcta de la marca insertada en la imagen atacada, según la categoría de ataque. Además se muestra el valor promedio de la probabilidad de error de bit para las estimaciones obtenidas en caso de detección correcta. Dadas las características de los métodos de sellado invisible implementados, y el parecido entre los algoritmos 1 y 3 por un lado y 2 y 4 por otro se irán presentando los resultados en ese orden con el objeto de que se puedan comparar más fácilmente. Además y puesto que los algoritmos que emplean técnicas de

espectro expandido (algoritmos 2 y 4), en caso de detección correcta, presentan probabilidades de error de bit nulas en todas las estimaciones, obtenidas tanto por ICA como por sustracción directa de la imagen original a la atacada, presentamos para ellos, únicamente las probabilidades de detección correcta en función del ataque.

6.4.3.1 Resultados generales. Promediado

Dada la cantidad de ataques implementados, que podrían hacer confundir al lector, mostramos aquí los valores promedio de los resultados obtenidos (en cada uno de los algoritmos) frente a los diversos ataques, tanto para probabilidades de detección como para probabilidades de error. Un análisis más preciso puede encontrarse en el Anexo V, en el que se incluyen los resultados obtenidos frente a los diferentes ataques para todos los mecanismos de watermarking implementados. Así mismo indicaremos aquí los resultados obtenidos para imágenes en escala de grises (puesto que los de imágenes en color resultan redundantes por ser semejantes frente a la mayoría de ataques implementados).

Los resultados se mostraran como valores promedio obtenidos frente a los diferentes ataques. Con el fin de sistematizar el método de presentación de resultados hemos procedido a la asignación de numeración a los ataques implementados. Las correspondencias entre ataques y números presentados en el eje de abscisas de las diferentes gráficas se muestran en la tabla 6.13.

Filtrado de mediana	1	**Rotación negativa con escalado**	9
Ataque FMLR	2	**Escalado**	10
Borrado filas/columnas	3	**Ataque Shearing**	11
Desplazamiento lineal general	4	**Filtrado Gaussiano**	12
Cambio de relación de aspecto	5	**Filtrado Sharp**	13
Rotación positiva	6	**Ataque STIRMARK**	14
Rotación positiva con escalado	7	**Compresión JPEG**	15
Rotación negativa	8	**Cortado**	16

TABLA 6.13. Numeración asignada a los ataques implementados sobre las imágenes marcadas. Véase tabla 6.10 para una definición más precisa de los ataques realizados.

6.4.3.1.1 Algoritmo 1. Dominio DCT de la imagen y marca de agua unidimensional

En las siguientes figuras se muestran los resultados obtenidos para los valores de probabilidad de error de bit y de probabilidad de detección promedio frente a los diferentes ataques mostrados en la tabla 6.13.

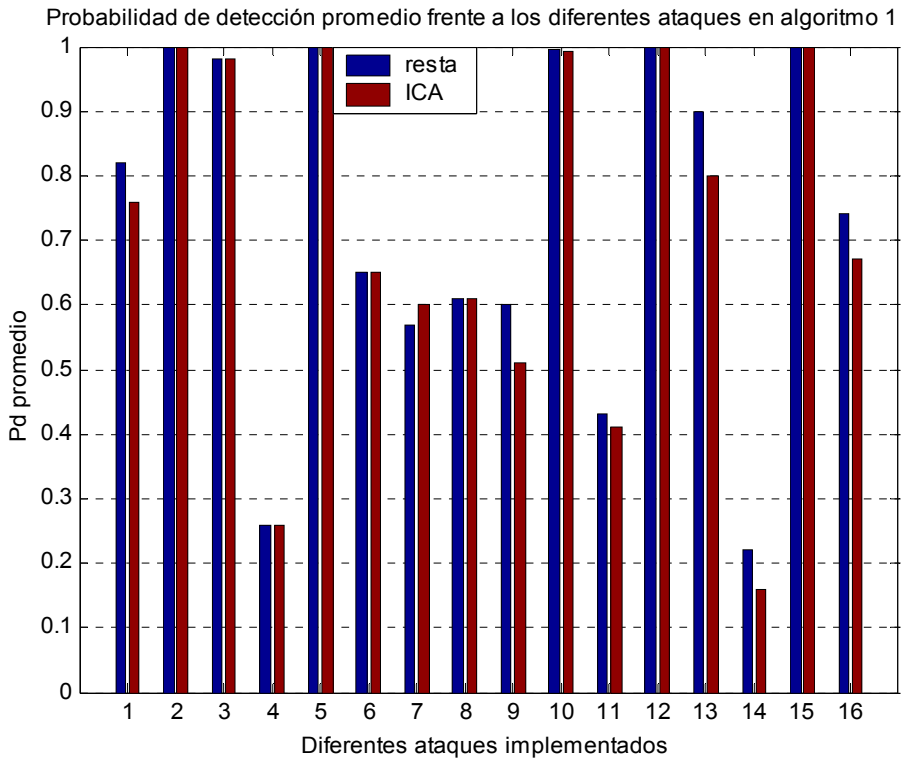


FIGURA 6.16. Probabilidad de detección en función de la modalidad de ataque implementado, resultados obtenidos de promediar los diferentes grados de severidad de cada categoría de ataque realizado en el algoritmo 1, para los dos métodos de detección propuestos (resta e ICA). Véase tabla 6.13 para correspondencia de numeración asignada en el eje de abscisas con el ataque representado.

En la figura puede observarse que los ataques frente a los que el método de watermarking implementado presenta mayores deficiencias son los correspondientes a rotación de la imagen (ya sea por un ángulo positivo o negativo), así como ataques aleatorios del tipo Stirmark y shearing. Presentando, sin embargo una clara eficiencia frente a ataques que eliminan las componentes menos significativas de la imagen, como son los filtrados de promediado y paso de baja así como los mecanismos de compresión tipo JPEG. La respuesta es también buena frente a ataques que tratan de deteriorar la sincronización del sistema como el borrado aleatorio de un reducido número de filas y/o columnas, modificación de la relación de aspecto o escalado de la imagen.

Los valores de probabilidad de error de bit promedio obtenidos en las marcas de agua recuperadas se muestran en la figura 6.17 en función de las diferentes modalidades de ataques.

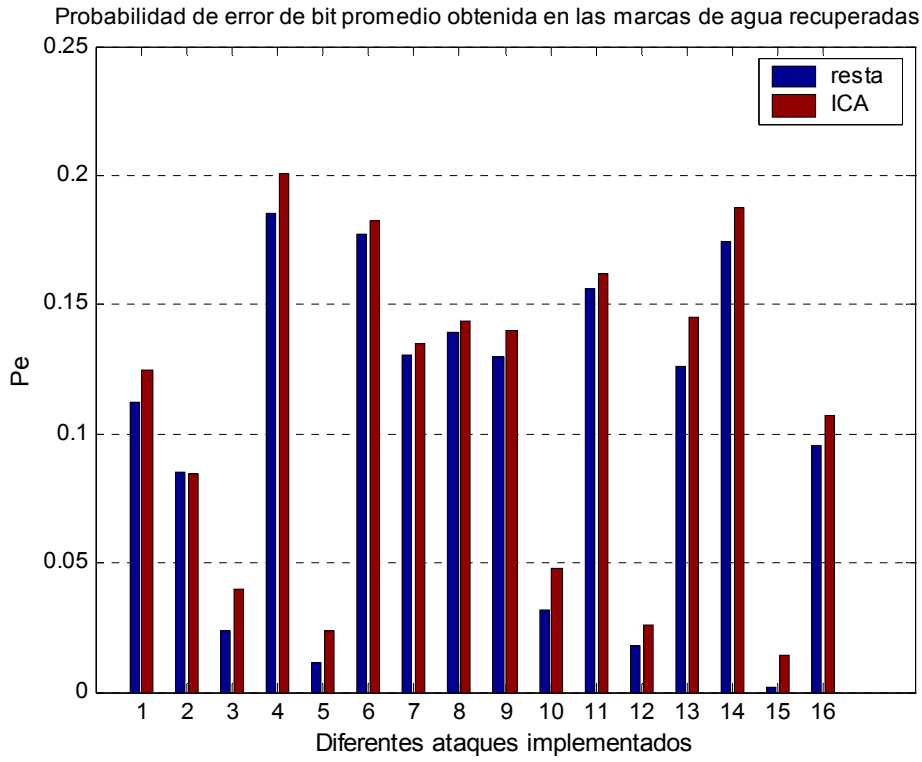


FIGURA 6.17. Valores de probabilidad de error de bit promedio obtenidos en las marcas de agua recuperadas en cada categoría de ataque realizado. Resultados obtenidos según el mecanismo de detección del algoritmo 1, tanto por sustracción directa de la imagen original a la atacada como mediante ICA.

6.4.3.1.2 Algoritmo 2. Dominio DCT de la imagen y marca de agua ensanchada mediante técnicas de SSM

Las características de robustez del algoritmo 2 se ilustran en función de la probabilidad de detección correcta de la marca de agua en función del tipo de ataque implementado. No se muestran los valores de probabilidades de error de bits obtenidas puesto que en caso de detección correcta, la marca de agua estimada coincide exactamente con la marca de agua insertada, sin ningún bit erróneo, en todas las pruebas realizadas sin excepción. De manera que la probabilidad de error de bit de la marca de agua estimada es nula cuando la estimación es correcta. En la siguiente figura se muestran las características de detección del método de watermarking estudiado en este apartado.

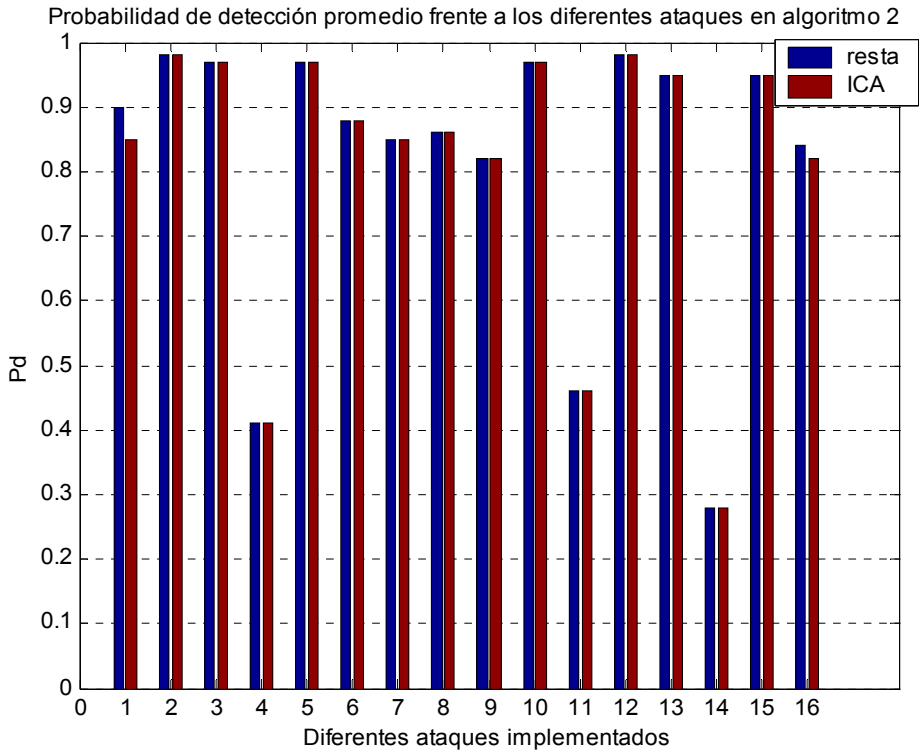


FIGURA 6.18. Probabilidad de detección correcta de la marca de agua en función de la modalidad de ataque implementado, resultados obtenidos de promediar los diferentes grados de severidad de cada categoría de ataque realizado en el algoritmo 2, para los dos métodos de detección propuestos (resta e ICA). Véase tabla 6.13 para correspondencia de numeración asignada en el eje de abscisas con el ataque representado.

En la figura de arriba se observa una notoria mejoría de las características del algoritmo con respecto al presentado en el apartado 6.4.3.1.1, que usaba una marca de agua unidimensional para sellar la imagen.

6.4.3.1.3 Algoritmo 3. Dominio DWT y DCT de la imagen y marca de agua unidimensional

En este algoritmo se combinan dos dominios transformados de la imagen para la obtención de las componentes de la misma adecuadas para ser marcadas.

Los resultados obtenidos muestran la desventaja del método con respecto a aquel que únicamente emplea la transformada DCT de la imagen (sección 6.4.3.1.1). No obstante, se observa que los resultados obtenidos por los métodos de detección empleados (sustracción directa de la imagen original a la atacada e ICA para separación de la marca de agua de la imagen marcada) se aproximan hasta hacerse prácticamente idénticos (obsérvese que en los métodos anteriores la detección por resta era levemente mejor que la que

emplea técnicas de análisis de componentes independientes). Observándose incluso situaciones en las que ICA presenta ventajas de detección. Remitimos al apartado anexos (Anexo V) para un análisis más detallado.

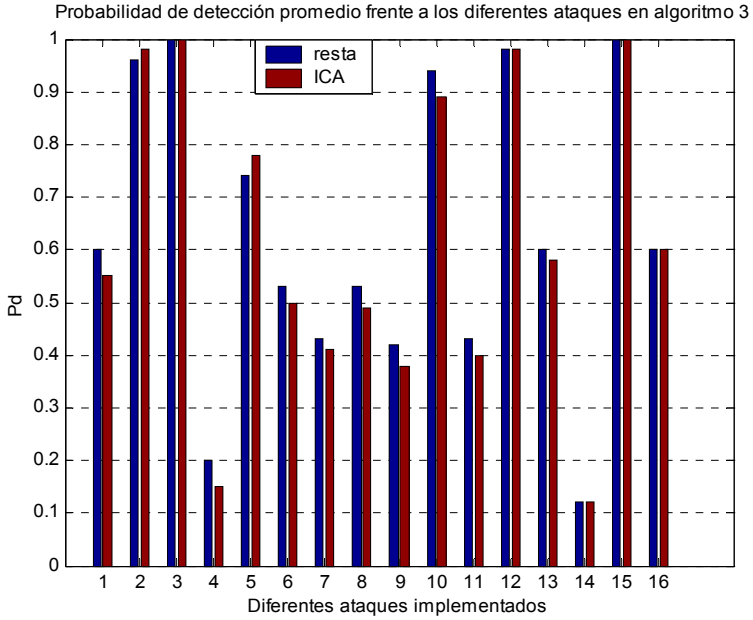


FIGURA 6.19. Características de detección correcta del algoritmo 3 frente a los diferentes ataques implementados, según los dos métodos de detección usados (resta e ICA).

Valores de probabilidad de error obtenidos se ilustran en la gráfica de la figura 6.20.

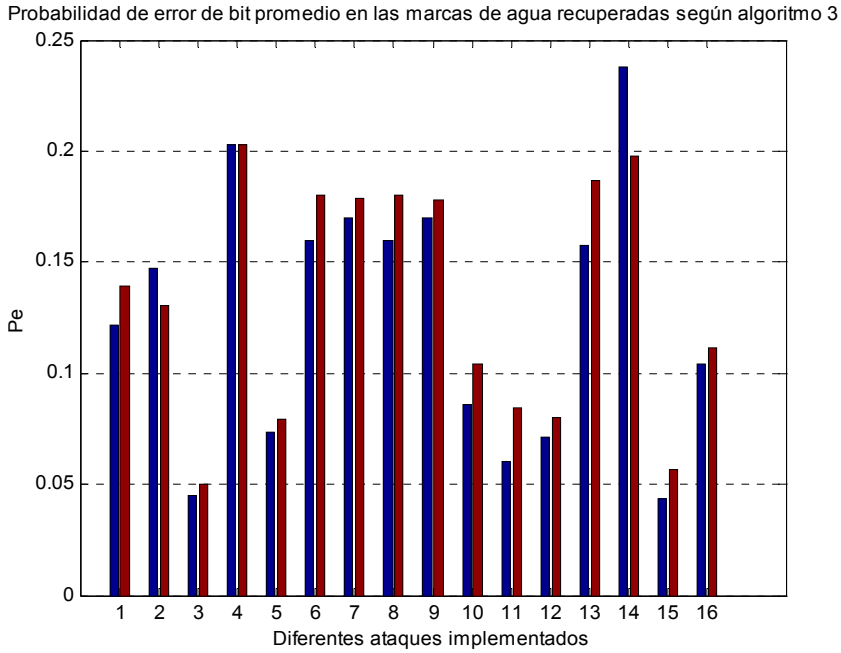


FIGURA 6.20. Probabilidad de error de bit promedio para las marcas de agua recuperadas según el algoritmo 3 (para ambos métodos de detección resta e ICA, véase leyenda fig.6.19).

6.4.3.1.4 Algoritmo 4. Dominio DWT y DCT de la imagen y marca de agua ensanchada según técnicas SSM

La mejora que introduce el usar técnicas de espectro expandido en la generación/detección de la marca de agua se puede observar en la figura 6.21, en la que se muestran las características de detección del algoritmo estudiado. Al igual que en el caso del algoritmo descrito en 6.4.3.1.2 se observa que la marca de agua estimada cuando la detección es correcta es exactamente la marca de agua insertada, de manera que en caso de funcionamiento correcto del método la probabilidad de error de bit promedio es nula en todas las pruebas realizadas, sin excepción.

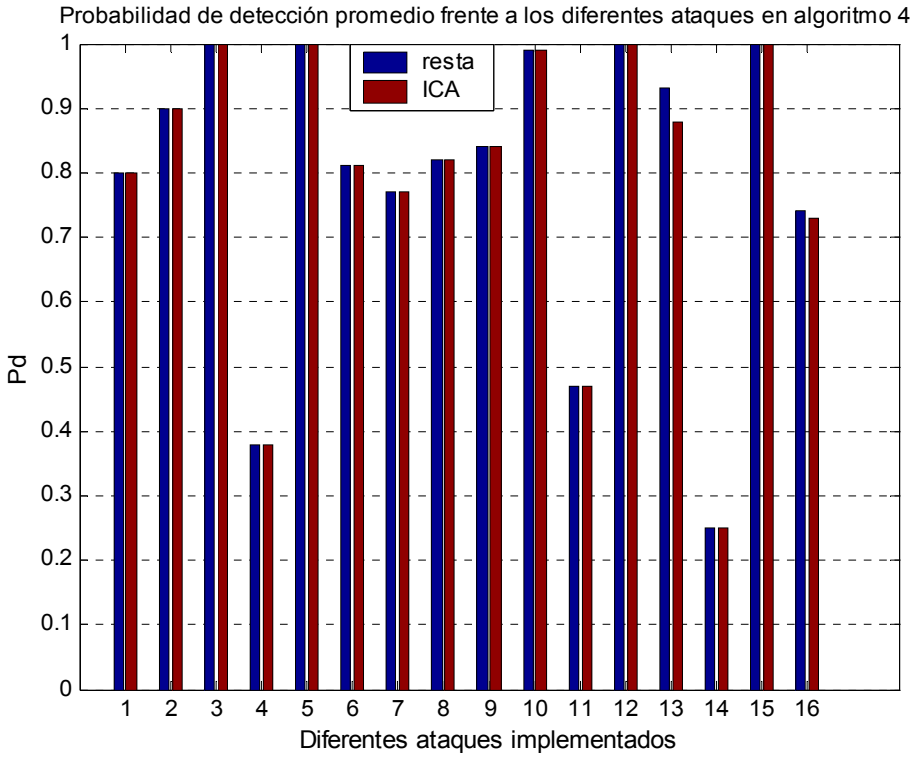


FIGURA 6.21. Características de detección correcta del algoritmo 4 frente a los diferentes ataques implementados, según los dos métodos de detección usados (resta e ICA).

6.4.3.2 Comparación de resultados promedio

Con el fin de poder comparar las características de robustez de los diferentes algoritmos y decidir cúal de ellos se comporta mejor, en general, mostramos una tabla con resultados globales, es decir, con el promedio total de los valores obtenidos para cada modalidad de ataque implementado.

PROBABILIDA DE DETECCIÓN PROMEDIO TOTAL	Resta	ICA
ALGORITMO 1	0.7360	0.7127
ALGORITMO 2	0.8169	0.8125
ALGORITMO 3	0.6300	0.6254
ALGORITMO 4	0.7938	0.7900

TABLA 6.14. Valores promedio globales de probabilidad de detección correcta para los diferentes algoritmos implementados y los dos métodos de detección usados (resta e ICA).

PROBABILIDA DE ERROR PROMEDIO TOTAL	Resta	ICA
ALGORITMO 1	0.0998	0.1102
ALGORITMO 2	0	0
ALGORITMO 3	0.1258	0.1337
ALGORITMO 4	0	0

TABLA 6.15. Valores promedio globales de probabilidad de error de bit de las marcas de agua estimadas, para los diferentes algoritmos implementados y los dos métodos de detección usados (resta e ICA).

En general, se observa que los algoritmos que hacen uso de técnicas de espectro expandido para la generación de la marca de agua presentan las mejores características de robustez y que la conjunción del empleo de los dominios DWT y DCT de la imagen para la selección de las componentes de la misma en las que insertar la *marca de agua binaria* no presenta mejoras con respecto al uso de la transformada DCT únicamente. Se marcan en cursiva los términos de marca de agua binaria porque en breve se verá que si esta marca fuera una secuencia continua aleatoria se pueden observar ciertos matices que hacen interesante el uso combinado de la trasformada DWT y detección mediante ICA. Remitimos al apartado 6.4.5 para ver detalles.

6.4.4 Eficiencia computacional

Para poder comparar la eficiencia en cuanto a tiempo de computación requerido por cada algoritmo, tanto para el proceso de inserción de la marca de agua en la imagen como para el de detección/extracción, se llevó a cabo el siguiente protocolo de prueba.

- **Proceso de sellado**. Se realizó el sellado de cien imágenes por cada uno de los algoritmos estudiados, las dimensiones de las imágenes usadas (las mismas de la figura 6.12) son las que se muestran en la tabla 6.9, que recordamos a continuación:
  - 206×345 píxeles
  - 305×200 píxeles
  - 297×200 píxeles
  - 287×200 píxeles
  - 150×200 píxeles

  Los tiempos promedio obtenidos muestran que el proceso de inserción de marca de agua más óptimo en cuanto a tiempo invertido en su realización es el asociado con el algoritmo 4, (aquel que insertaba la marca de agua en el dominio DCT de la subimagen aproximada resultado de aplicar la transformada Wavelet de nivel 2 a la imagen, usando, además técnicas de espectro expandido en la generación de la marca de agua), seguido del correspondiente al algoritmo3. El algoritmo 1, que realiza el marcado de la imagen en el dominio DCT de la misma, es el más lento con notable diferencia. Se ilustran los resultados en la tabla 6.16.

  Un estudio del tiempo requerido por cada etapa incluida en el proceso de sellado de los algoritmos 1 y 3, muestra que, aproximadamente, el 80% del tiempo total se invierte en el proceso de escaneado de los coeficientes de la imagen, lo que justifica los resultados obtenidos, pues el algoritmo 1 escanea los coeficientes DCT de la imagen completa.

- **Proceso de detección**. Al igual que en el caso anterior, se eligieron cien imágenes de las mismas características a las indicadas en el proceso de inserción y procesadas según los mismos ataques para

realizar el estudio de tiempos requeridos en la detección según las especificaciones de cada uno de los algoritmos, los resultados obtenidos se ilustran en la tabla 6.16.

ALGORITMO	Algoritmo1	Algoritmo2	Algoritmo3	Algoritmo4
$T_{MARCADO}$ (s)	5.8714	0.1542	0.2206	0.1461
$T_{DETECCIÓN/EXTRACCIÓN}$ (s)	8.2026	2.1018	1.4210	2.4805

TABLA 6.16. Tiempos promedio de sellado y recuperación de la marca de agua requeridos por cada uno de los algoritmos estudiados.

6.4.5 Comparación con algoritmo base en [Liu03].

Aunque en la técnica de watermarking propuesta en [Liu03] las secuencias correspondientes a las marcas de agua que se insertan en las imágenes son secuencias reales, para poder comparar resultados de probabilidades de error de bits presente en las diferentes estimaciones, nosotros en este documento las pasamos además a secuencias bipolares. Las características fundamentales de la técnica propuesta en [Liu03] pueden verse en mencionada referencia, pudiéndose comprobar que se trata de un método de sellado invisible de imágenes parecido a los implementados en el presente proyecto. Los parámetros de experimentación propuestos en dicha publicación son los indicados en la tabla 6.17.

	[Liu03]
Parámetros de inserción	Factor de escalado: $\alpha = 0.06$ Longitud marca de agua: n =1024bits (secuencia 1D)
Parámetros de detección	Umbral de detección: umbral_sim = 5

TABLA 6.17. Parámetros de inserción y de detección usados en los algoritmos de [Liu03].

Puesto que lo que emplean originalmente son secuencias pseudoaleatorias reales como marcas de agua, en principio sólo se podrían aportar resultados del proceso de detección, es decir, de la presencia/ausencia de una determinada marca de agua en la imagen bajo estudio (resultados que se aportan en [Liu03] y que nosotros hemos obtenido para los diferentes ataques implementados, véase tabla 6.18 y compárense con los aportados en la referencia), nosotros, además, hemos considerado la posibilidad de insertar en la

imagen marcas de agua en formato bipolar, con el objeto de poder aportar resultados relacionados con la probabilidad de error de bits presente en las diferentes estimaciones fruto del proceso de recuperación de la marca de agua, de modo que se puedan comparar resultados con los obtenidos en las técnicas implementadas (referidas como algoritmos 1,2 3 y 4). Indicamos además que en estas pruebas mostramos únicamente resultados de trabajar con imágenes de intensidad para ser fieles a la publicación referida.

En la siguiente tabla mostramos los resultados obtenidos para el estadístico de parecido promedio según la categoría de ataque implementado, para secuencias de marcas de agua en formato real y ambos métodos de detección utilizados, es decir, sustracción directa de la imagen original a la atacada y mediante ICA, (pueden compararse con [Liu03]). Si se utilizan marcas de agua en formato bipolar los resultados son, en general, similares a los obtenidos (algo más pequeños), aunque las diferencias entre los dos métodos de detección se hacen menos palpables. Si usamos secuencias bipolares podemos obtener el valor de la probabilidad de error de bits de las estimaciones tanto por resta como por ICA frente a los diversos ataques, valores que oscilan en el rango [0.27-0.45], donde las probabilidades de error de bits más reducidas se corresponden con aquellas estimas para las que el valor del estadístico es alto y viceversa.

ATAQUE	RESTA	ICA	ATAQUE	RESTA	ICA
Rotación negativa y escalado	10.56	11.27	**Filtro mediana**	14.78	18.83
Rotación positiva y escalado	11.33	11.31	**Ataque_FMLR**	18.42	19.98
Escalado	6.2820	12.4760	**Cambio relación de aspecto**	13.44	17.785
JPEG	28.7275	29.0592	**Filtrado gaussian 3×3**	23.81	23.89
Stirmark random bend	7.11	7.07	**Filtrado sharp 3×3**	15.27	15.13
Borrado simétrico/asimétrico de filas/columnas	9.1160	9.49	**Rotación negativa**	9.1833	9.0633
shearing	8.0150	8.4967	**Rotación positiva**	9.0833	9.0933
			Ruido gausiano (0,0.09)	17.84	19.25

TABLA 6.18. Valor del estadístico que mide la correlación entre las secuencias comparadas. En el algoritmo propuesto en [Liu03] con los parámetros de experimentación propuestos.

En la tabla 6.18 puede observarse que el valor del estadístico que mide el parecido es superior en el caso de detección por ICA que en la detección por resta, lo que se traduce en una mejor robustez en este mecanismo de detección en el sentido de que la posibilidad de que ocurran falsas alarmas deberá de ser más reducida, así como en el hecho de que, en general se observa una probabilidad de detección correcta superior a la obtenida mediante sustracción directa de la imagen original a la atacada, observándose situaciones límite en las que solo ocurre detección cuando extraemos la marca de agua mediante ICA.

No obstante los valores obtenidos para la probabilidad de error de bits son prácticamente iguales para ambos métodos, ello se debe a lo indicado más arriba relativo al hecho observado de que al cambiar las secuencias reales a formato bipolar, la distancia entre los valores obtenidos para el estadístico utilizado es mucho menor a la observada en la tabla 6.18. Por lo que los resultados por ambos métodos son más parecidos. Los valores obtenidos para las probabilidades de detección en función de cada ataque son del orden de las mostradas para el algoritmo 3, algo superiores, por lo que referimos al apartado 6.2.2.1 para el análisis de resultados, teniendo en cuenta que, en general los valores obtenidos para las probabilidades de error de bits son bastante peores como ya hemos señalado, (en el rango [0.27, 0.45]).

Notar que el valor del estadístico cuando las secuencias comparadas son muy parecidas es bastante alto. Se observa una relación directa entre la longitud de la secuencia de la marca de agua insertada y el valor que toma el estadístico cuando se compara la marca extraída con la insertada realmente, pues en el algoritmo 3 (que inserta en la imagen una secuencia de sesenta y cuatro bits) el valor máximo del estadístico, obtenido en el caso de comparar una marca de agua extraída de una imagen sin procesar con la realmente insertada es de, aproximadamente, ocho, tomando un valor superior a treinta y dos en el algoritmo propuesto por [Liu03] y que estudiamos en este apartado. En el apartado de anexos se incluye una versión desarrollada de la tabla 6.18.

6.5 Conclusiones

Tratar de decidir si existe algún algoritmo mejor que otro desde un punto de vista absoluto, puede resultar una tarea compleja y, en gran medida, carente de sentido, pues, en última aproximación, lo que decide si una determinada aplicación resulta adecuada depende de la finalidad o el propósito para el que haya sido creada. En el campo del sellado invisible de imágenes la efectividad

de una determinada técnica vendrá, entonces, impuesta por la aplicación para la que se destine en función de la que tendremos ciertos procesados y/o ataques con mayor probabilidad que otros. Aún así, se pueden enunciar ciertas propiedades apreciadas cuya característica repetitiva a largo de las pruebas realizadas nos induce a considerarlas como resultados concluyentes.

- En primer lugar, se observa que si la secuencia de marca de agua insertada en la imagen es real, en general, el valor que toma el estadístico que mide la correlación entre las señales extraídas y realmente insertadas es superior al que se obtiene cuando las secuencias que se añaden a la imagen presentan un formato bipolar. Además, en esta situación, puede observarse una clara superioridad del método ICA en recepción con respecto al de sustracción directa de la imagen original a la atacada, (fenómeno observado en [Liu03] y en el algoritmo 3 implementado, técnicas, ambas, de sellado invisible que añaden la marca de agua en el dominio DCT de la subbanda de mayor energía obtenida al aplicar la transformada Wavelet de nivel dos a la imagen).
  - Además, se observa que los valores de probabilidades de error de bits obtenidos frente a los diferentes ataques en el algoritmo 3 son bastante más pequeñas que las observadas en el método propuesto en [Liu03], ello se debe a la regla de selección de los coeficientes de la imagen candidatos a albergar la marca de agua, pues en [Liu03] se escogen los primeros coeficientes de la secuencia escaneada (con un offset inicial de treinta) mientras que en el algoritmo 3 se marcan los coeficientes de mayor valor (y por tanto los más significativos), que son más estables frente a ataques como compresión y filtrado paso bajo.
- De la comparación entre los resultados obtenidos en los algoritmos 1 y 3 también pueden extraerse ciertas consideraciones.
  - Por un lado, si bien los valores de detección correcta frente a los diferentes ataques son muy aproximados (algo mejores en el algoritmo 1, que únicamente hace uso de la transformada DCT de la imagen para seleccionar los coeficientes a marcar), en general, se observan valores para las probabilidades de error de bits superiores en el proceso de detección del algoritmo 3, lo que, en

cierto modo, puede explicarse por el empleo de dos transformaciones que añaden errores de redondeo, además del error de reconstrucción implícito en el uso de la transformada Wavelet discreta.

- Sin embargo, en el algoritmo 1, se observan mejores resultados en la detección mediante sustracción directa de la imagen original a la atacada, mientras que la adición de un nuevo dominio transformado, el de las Wavelet (que es lo que se hace en el algoritmo 3), añade una mejora considerable a la detección mediante ICA (con respecto al algoritmo 1), que aproxima los resultados obtenidos mediante ambos métodos hasta igualarlos. Lo que no significa que, en sentido general, el algoritmo 3 presente mejoras con respecto al primero (puesto que el promedio general de detección muestra lo contrario).

➢ Con respecto a los métodos que emplean técnicas de espectro expandido en la generación de la marca de agua a insertar (algoritmos 2 y 4) podemos indicar lo siguiente:

- El valor de la PSNR de las imágenes marcadas es más pequeño que el equivalente en los otros dos algoritmos, lo que indica una mayor degradación de las imágenes, aunque una medida subjetiva obtenida por observación directa de las mismas permite afirmar que dicha degradación es prácticamente imperceptible, véanse las figuras adjuntas de imágenes marcadas según los cuatro algoritmos para corroborar lo indicado (figuras 6.22, 6.23, 6.24 y 6.25).
- Se observa, también, que el proceso de recuperación de la marca de agua insertada en la imagen (en estos algoritmos) es perfecto en todas la pruebas realizadas (sin excepción), en el sentido de que en caso de que ocurra detección correcta de la marca de agua, la estima obtenida, tanto por resta como por ICA, coincide exactamente con la marca insertada, sin ningún bit erróneo.
- Se aprecia un aumento en los valores de probabilidad de detección correcta con respecto a los algoritmos 1 y 3.

FIGURA 6.22. Ejemplo de imágenes marcadas según el algoritmo 1.

FIGURA 6.23. Ejemplo de imágenes marcadas según el algoritmo 2.

FIGURA 6.24. Ejemplo de imágenes marcadas según el algoritmo 3.

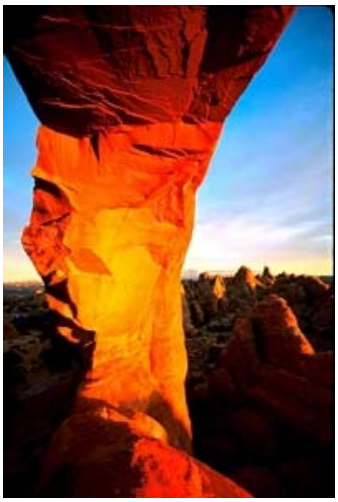


FIGURA 6.25. Ejemplo de imágenes marcadas según el algoritmo 4.

- No obstante, ciertos ataques muestran las debilidades más notorias de los algoritmos de sellado invisible de imágenes implementados en general, observándose la escasa robustez de los mismos frente a ataques que tratan de eliminar la sincronización en el receptor, como son, el rotado de la imagen por ángulos superiores al grado, el desplazamiento lineal general aplicado a la imagen, así como la

adición de desplazamientos locales arbitrarios unidos a un proceso de interpolación y adición de una pequeña cantidad de ruido a la imagen (ataque referido como stirmark en el documento).