INSTITUTO POLITÉCNICO NACIONAL

ESCUELA SUPERIOR DE INGENIERÍA MECÁNICA Y ELÉCTRICA

UNIDAD CULHUACAN

SECCIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

SISTEMA DE AUTENTICACIÓN DE IMÁGENES DIGITALES CON MARCA DE AGUA SEMI-FRÁGIL

T E S I S

QUE PARA OBTENER EL GRADO DE

DOCTORADO EN COMUNICACIONES Y ELECTRÓNICA

P R E S E N T A

M. EN C. CLARA CRUZ RAMOS

ASESORA

DRA. MARIKO NAKANO MIYATAKE

MÉXICO D.F. SEPTIEMBRE 2009

RESUMEN

En esta tesis se proponen dos algoritmos de marca de agua semi-frágil para la autenticación del contenido en imágenes digitales; ambos están basados en el contenido de los bloques de la imagen original, cabe mencionarse que no requieren de la imagen original para la extracción de la marca de agua, lo cual hace que sean algoritmos de extracción completamente a ciegas.

El primer algoritmo extrae una firma digital de la imagen huésped y la inserta en la misma como marca de agua, con la ventaja de que no se requiere almacenar la firma en un archivo por separado como generalmente se hace, esto evita que no se requiera ancho de banda adicional para el envió de este y así como también, se evita que al momento de transmitir el archivo que contiene la firma pueda ser alterado o extraviado. Para el proceso de inserción se utiliza el método de cuantificación vectorial en el dominio de la transformada wavelet discreta, lo cual lo hace robusto a ataques por compresión JPEG y ruido; pero sensible a alteraciones intencionales en la imagen como lo es el fotomontaje, con respecto a este ataque, cabe mencionar que el algoritmo propuesto detecta correctamente los bloques alterados en la imagen marcada. Generalmente en los sistemas cuya aplicación es solamente la autenticación de imágenes (como el propuesto) se limitan a determinar bloques erróneos, los cuales son considerados como ataques intencionales en la imagen marcada; es por esta razón que propusimos un algoritmo de verificación de bloques erróneos, el cual determina si estos son alterados por circunstancias intencionales o incidentales; en el caso de considerarse incidentales, estos bloques son eliminados de la imagen autenticada.

El segundo algoritmo propuesto además de autenticar la imagen, es capaz de recuperar la imagen original de los bloques alterados en la imagen marcada, este algoritmo divide la imagen original en dos regiones, la primera es la región de interés ROI (son regiones importantes que requieren protección contra modificaciones maliciosas y son seleccionadas por el propietario de la imagen) y la región de inserción ROE (es el resto de la imagen donde insertamos la marca de agua; y es una región diferente a la ROI). La generación de la marca de agua binaria consiste en extraer en su forma binaria el coeficiente DC y los 6 primeros coeficientes AC ordenados en zig-zag de cada bloque ROI transformado con la DCT, dicha marca es insertada en el bit menos significativo de los coeficientes DCT de frecuencia media de los bloques ROE seleccionados aleatoriamente por una llave secreta. Una vez determinados como alterados los bloques ROI, estos se reconstruyen con la marca extraída de la región ROE. El algoritmo es robusto a la compresión JPEG y además de detectar correctamente los bloques erróneos la reconstrucción de estos tiene una buena calidad visual.

Los algoritmos propuestos son comparados con algoritmos parecidos y en las mismas condiciones para mostrar que estos tienen un mejor desempeño que los propuestos por otros autores.

ABSTRACT

In this thesis, two image content authentication algorithms based on semi-fragile watermarking are proposed. Both algorithms are block-wise and content-based authentication; furthermore the watermark extraction process is performed using a blind scheme, where the original image is not required.

In the first algorithm, the digital signature is extracted from host image and it is embedded into the image as a watermark signal, the main advantage of this approach is that a digital signature (authentication code) is not required to transmit and/or save together with the image file, instead of an additional file composed of authentication code (digital signature), which could be lost or modified. The embedding process is performed using Vector Quantization method in the Discrete Wavelet Transform domain, reason why the proposed technique is robust to JPEG compression and noise but sensitive to intentional image tamper like photomontage, with regard to this attack it's important to mention that the proposed technique detects correctly the photomontage region of the images. Generally, almost all images authentication algorithms detect the error blocks in the watermarked tamper image as intentional attacks, so we propose an verification process of error blocks, which decide if it was altered intentionally or not. In the case of it was not intentional modified, the error block is removed from the authentication image.

The second image authentication scheme not only determines integrity of the image and detects regions suffered some modifications, but also it has recovery capacity of modified regions. The original image is divided into two

regions: Region of Interest ROI (are important regions that require protection against malicious modification; and the owner is the one who selects these regions), and Region of Embedding ROE (is the rest of the image where we can embed the watermark; obviously this region is different to ROI). ROI is selected manually by the owner's image, and ROE are regions of the host image different from ROI region. The binary watermark of each ROI's block is generated by the DC and 6 lowest AC components (in the zig-zag order) of DCT coefficients and they are embedded into the LSB of the corresponding DCT ROE's indicated by a mapping list previously generated using a secret key. Once that ROI blocks are detected as tampered, these are recovered replacing them by the watermark sequence extracted from ROE blocks. This algorithm is robust to JPEG compression and detects correctly the error blocks and the reconstructed image quality is good.

The proposed algorithms are compared with other authentication methods to show desirable performance of the proposed algorithms.

Agradecimientos

Mi más amplio agradecimiento a mi directora de tesis Dra. Mariko Nakano Miyatake y al Dr. Hector Manuel Perez Meana, cuyo invaluable y generoso apoyo, amistad y conocimientos hicieron posible la realización de esta tesis doctoral.

Quiero hacer patente mi agradecimiento a los Doctores, Volodymyr Ponomaryov, Hector Manuel Perez Meana Oleksiy Pogrebnyak y Francisco Javier García Ugalde, por las valiosas aportaciones que me hicieron para mejorar la presente investigación.

De todo corazón a mis padres Josefina y Fausto que siempre me enseñaron el bien y la perseverancia en el trabajo. A mis hermanas Eli, Mary e Inés por su muestra de solidaridad y apoyo. A mis suegros Florinda y Onésimo por su gran apoyo y comprensión.

Esta tesis doctoral pudo ser realizada gracias al apoyo de dos instituciones: el Instituto Politécnico Nacional, quien me abrigo en su programa doctoral y ha sido testigo de mi trayectoria escolar y laboral y al Consejo Nacional de Ciencia y Tecnología.

A Dios por poner en mi camino a todas las personas y circunstancias que me ayudaron a llegar a este momento de mi vida.

"Mejor luchar por algo que vivir por nada."
George S. Patton

A mis queridas hijas Amairani y Vania

Al gran amor de mi vida Rogelio

ÍNDICE

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

CAPÍTULO 1

INTRODUCCIÓN A LOS SISTEMAS DE AUTENTICACIÓN DE IMÁGENES DIGITALES CON MARCA DE AGUA SEMI-FRÁGIL

1.1. INTRODUCCIÓN

En 1954, Emil Hembrooke de la Corporación Muzac presento la patente "Identificación de sonidos y señales semejantes" [Frank, 1961] en el cual se describe un método para insertar imperceptiblemente un código de identificación dentro de la música con el propósito de conocer al propietario. La patente dice "La presente invención hace posible una identificación positiva del origen de una presentación musical y por ello constituye un efectivo significado de la prevención ante la piratería, por ejemplo puede ser semejante a una marca de agua en un papel". Entonces se dice que la marca de agua fue inventada. Desde entonces, un número de tecnologías de marca de agua han sido desarrolladas para una variedad de aplicaciones. El interés por insertar marcas de agua ha continuado a través de los siguientes 60 años. En la primera mitad de la década el interés por el tema se expande rápidamente y en la actualidad las memorias de conferencias enteras están dedicadas a este tema.

Este incremento en el interés fue motivado por la protección de derechos de autor que empezó a ser agudo con los avances en la tecnología computacional y por el desarrollo de la Web. Estas tecnologías permiten el perfecto copiado y distribución de material con derechos de autor casi a cualquier lugar del mundo y sin ningún costo. Para direccionar este asunto,

un número de grupos de la industria de la tecnología se establecieron, quizás los mejores grupos conocidos fueron el *Copy Protection Technical Working Group* (CPTWG) y el *Secure Digital Music Initiative* (SDMI) quienes inicialmente trabajaron con el almacenamiento del contenido del video digital en discos DVD y después con música digital.

Los sistemas de marca de agua insertan una señal, algunas veces llamada "señal insertada" o "marca de agua" dentro de otra señal llamada "señal huésped". La inserción debe hacerse de tal manera que la inserción de la señal no cause una seria degradación en la señal huésped. Al mismo tiempo, la inserción debe ser robusta a degradaciones comunes en la señal marcada, compuesta por la señal huésped y la marca de agua, dichas degradaciones en algunas aplicaciones resultan de ataques deliberados. Idealmente siempre que la señal huésped sobreviva a estas degradaciones, la marca de agua también sobrevive.

En adición a la facilidad de la duplicación, las señales digitales multimedia son también fáciles de alterar y manipular, la autenticación o detección de modificaciones en señales multimedia es otra aplicación de los métodos de marca de agua digital [Kundur, 1999]. En esta aplicación una firma digital se inserta como marca de agua, las marcas de agua son llamadas frágiles porque cambian cuando el compuesto de la señal es alterado significativamente, esto se traduce en que se detecta una falsificación. Alternativamente, se puede insertar una marca de agua robusta o una firma digital, por ejemplo dentro de la imagen de un mapa. Si el mapa es alterado, la marca de agua debe sobrevivir, pero no se debe percibir la marca de agua. En contraste con los métodos de autenticación tradicional, en ambos casos, la marca de agua robusta o frágil se inserta directamente dentro de la imagen huésped. Con esto no se requiere de un ancho de banda adicional y se pueden diseñar algoritmos de marca de agua que puedan autenticar

imágenes que fueron sometidas a ataques comunes como: cambio de formato o compresión por pérdida JPEG.

1.2. HIPÓTESIS

Algunas formulaciones acerca de sistemas de autenticación de imágenes generales fueron propuestas por Wu y Liu en [Wu, Liu, 1998] y por Lin y Chang en [Lin, Chang, 2000]. Ellos dicen que para que un sistema sea efectivo debe satisfacer los siguientes criterios:

1.- Susceptibilidad: el sistema bebe ser susceptible a manipulaciones maliciosas (por ejemplo modificaciones en el significado de la imagen) como recorte o alteración de la imagen en áreas especificas.

2.- Tolerancia: el sistema debe tolerar alguna pérdida de información (originada por algoritmos de compresión con pérdida) y a manipulaciones no maliciosas (generadas por ejemplo por proveedores multimedia o usuarios autorizados).

3.- Localización de regiones alteradas: el sistema debe localizar precisamente cualquier modificación maliciosa hecha a la imagen y verificar otras áreas como auténticas.

4.- Reconstrucción de regiones alteradas: el sistema puede necesitar la habilidad de reconstruir la imagen, inclusive si esta fue parcialmente alterada o destruida de tal manera que el usuario conozca cual fue el contenido original de las áreas manipuladas.

También se deben tomar en cuenta ciertas características como:

a) Almacenamiento: los datos de autenticación deben insertarse en la imagen como una marca de agua, en lugar de almacenarla en un archivo separado, como en el caso de una firma digital externa.

b) Modo de extracción: depende de la dependencia de la autenticación de los datos con la imagen, se requiere un modo de extracción a ciegas o no. Es claro que un sistema de extracción que no es a ciegas no tiene ningún sentido en un servicio de autenticación, ya que seria necesario conocer la imagen original.

c) Algoritmo asimétrico: Contrario a los servicios de seguridad clásicos como protección de derechos de autor, un sistema de autenticación requiere una marca de agua asimétrica, donde solo el autor de la imagen puede garantizar su autenticidad, pero cualquier usuario pueda ver el contenido de la imagen.

d) Visibilidad: la autenticación de datos debe ser invisible dentro de la observación normal. Es una cuestión de estar seguro que el impacto visual de la marca de agua es lo mas tenue posible para que la imagen marcada sea lo más parecida a la original.

e) Robustez y seguridad: no debe ser posible que un autenticador de datos sea falsificado o manipulado.

La primera clase de este tipo de métodos es un algoritmo de marcado de agua invertible [Fridrich, et al., 2001], en el sentido que, si la imagen se considera auténtica, la distorsión debida al proceso de marcado de agua puede ser eliminado para obtener la imagen original. Otra clase consiste en separar la imagen en dos zonas: una región de interés (ROI) la cual es la parte de la imagen usada para el diagnostico, donde la integridad de los datos debe ser estrictamente controlada; y una región de no interés (donde

la distorsión esta permitida), la cual se usa para insertar los datos de autenticación [Coatrieux et al., 2001].

1.3. OBJETIVO

Investigar, desarrollar y evaluar algoritmos de autenticación y auto recuperación de la información en imágenes digitales, basados en la inserción y extracción a ciegas de marcas de agua semi-frágiles imperceptibles, las cuales identifiquen las regiones alteradas intencionalmente y garanticen la integridad del contenido semántico de la imagen, aún cuando esta haya sido sometida a ataques intencionales y no intencionales.

1.3.1. OBJETIVOS PARTICULARES

Investigar, desarrollar y evaluar un algoritmo de autenticación y detección de regiones alteradas en imágenes digitales en donde la marca de agua semi-fragil insertada sea una firma digital extraída de la imagen original, la cual sea resistente a ataques no intencionales, pero susceptible a ataques intencionales, cuyo algoritmo de verificación pueda diferenciar entre ellos.

Investigar, desarrollar y evaluar un algoritmo de autenticación, detección y auto recuperación de regiones alteradas con una buena calidad visual en imágenes digitales, el cual este basado en marcas de agua semi-frágiles y sea resistente a ataques no intencionales, pero susceptible a ataques intencionales.

1.4. METAS

Investigar y analizar los algoritmos de autenticación de imágenes digitales existentes basados en marcas de agua.

Desarrollar e implementar métodos para extraer marcas de agua robustas del contenido de la imagen original.

Desarrollar e implementar algoritmos de inserción y extracción de la marca de agua en el dominio de la transformada.

Desarrollar e implementar algoritmos de verificación del contenido de la imagen marcada.

Desarrollar e implementar un algoritmo de auto recuperación del contenido protegido de la imagen original.

Analizar y evaluar la calidad en la imagen marcada y en la imagen recuperada de los algoritmos de autenticación y auto recuperación.

Evaluar la robustez de los algoritmos de autenticación y auto recuperación ante la compresión JPEG y adición de ruido.

Evaluar la sensibilidad de los algoritmos de autenticación y auto recuperación ante ataques de fotomontaje

Comparar y analizar los algoritmos propuestos en este trabajo de investigación con algoritmos similares propuestos por otros autores.

1.5. JUSTIFICACIÓN

Hasta ahora la mayoría de las publicaciones en el campo de las marcas de agua están direccionadas principalmente a la protección de los derechos de autor, otros servicios de seguridad, como la autenticación del contenido de

la imagen están todavía limitados y existen muchas cuestiones fundamentales abiertas.

En la comunidad de seguridad, un servicio de integridad es inequívocamente definido como el que asegura que el envió y recepción de datos es idéntico. Esta definición binaria puede ser aplicada a imágenes, aunque no esta estrictamente adaptada a este tipo de documento digital ya que los valores de sus pixeles podrían ser modificados pero no el contenido semántico de la imagen. En otras palabras el problema de la autenticación de imágenes concierne al contenido de la imagen, por ejemplo, cuando las modificaciones cambian el contenido o degradan visualmente la imagen. Con la finalidad de proveer un servicio de autenticación de imágenes es importante distinguir entre modificaciones maliciosas, las cuales cambian el contenido de la imagen original y modificaciones no maliciosas relacionadas al uso de la imagen, como conversión de formato, compresión ruido, etc.

Cabe mencionar que el enfoque de este trabajo es la autenticación de imágenes multimedia de propósito general, es interesante hacer notar que existen métodos dedicados a servicios de integridad muy específicos, como la autenticación de imágenes medicas o militares. En efecto estas imágenes no deben ser modificadas por ningún medio (incluido el marcado de agua) y por lo tanto requieren una definición estricta de seguridad.

Hoy en día las imágenes tienden a representar más información que el mismo texto, así como también es más fácil ignorar el contenido de información textual que el cuestionar sobre el origen y autenticidad de una fotografía. Antiguamente se asumía que la imagen de una cámara era auténtica o en otras palabras no podía ser falsa. Sin embargo, hoy en día es posible editar imágenes de manera muy sencilla y a muy bajo costo. Las imágenes resultantes pueden tener una alta calidad que las hacen parecer genuinas. En este contexto, es obvio que un servicio de autenticación de

imágenes no puede ser usado para verificar eventos, pero si puede ser usado para detectar una alteración posterior en la imagen. El uso de imágenes digitales en situaciones legales se vuelve cada día más y más cuestionable al mismo tiempo que las videocámaras de vigilancia son más comunes en ciudades y sitios públicos.

Con la finalidad de contrarrestar este problema en esta investigación se desarrollaron dos algoritmos de marca de agua digital semi-frágil para la autenticación, detección y auto recuperación de ataques intencionales y no intencionales en una imagen digital.

El primer algoritmo propone un sistema de autenticación y detección de regiones alteradas en los bloques de la imagen, el cual inserta una firma digital como marca de agua eliminando así la necesidad de generar un archivo adicional para la transmisión de la firma digital. Con la finalidad de incrementar la robustez de la marca de agua, esta se inserta en el dominio de la frecuencia utilizando la *Transformada Discreta Wavelet* (DWT) mediante un proceso de cuantificación controlado. Otro atributo a este algoritmo es que no se requiere conocer información de la imagen original para la extracción de la marca de agua, lo que incrementa la seguridad en el sistema. En lo subsecuente dicho algoritmo lo denominaremos como: Autenticación basado en Firma Digital como Marca de Agua **(AFDMA).**

El segundo es un algoritmo de marca de agua semi-frágil basado en bloques para la autenticación, detección y con la capacidad de recuperar la imagen original a partir de la imagen marcada y alterada en imágenes digitales. Antes de generar la secuencia de marca de agua, la imagen original se divide en dos regiones: región de interés (ROI) y región de inserción (ROE), la ROI es seleccionada manualmente por el propietario de la imagen y la ROE son todas las regiones de la imagen original fuera de la ROI. La generación de la marca de agua binaria consiste en extraer el coeficiente DC y los 6 primeros

coeficientes AC ordenados en zig-zag en forma binaria de cada bloque ROI transformado por la DCT, dicha marca es insertada en el bit menos significativo de los coeficientes DCT de frecuencia media de 6 bloques seleccionados aleatoriamente por medio de una llave secreta, la cual solo deberá conocer el propietario de la imagen. Para autenticar los bloques ROI se extrae la marca de agua de los de la ROI y de la ROE; al comparar ambas por medio de una operación XOR, si la diferencia es mayor a un umbral *Th*, entones consideramos que el bloque fue alterado y es reemplazado con el bloque reconstruido por la marca extraída de la región ROE. Este sistema no requiere la imagen original para poder recuperar la marca de agua simplemente se necesita de la llave secreta para conocer los bloques en donde se inserto la marca de agua. En lo subsecuente dicho algoritmo lo denominaremos: Autenticación y Auto-Recuperación basado en Marca de Agua **(AAMA).**

1.6. APORTACIONES

La inserción de la marca de agua en los algoritmos AFDMA AAMA se realiza en el dominio de la transformada debido a sus buenas características de adaptabilidad a la visión humana y su robustez a modificaciones no maliciosas relacionadas al uso de la imagen, como conversión de formato, compresión ruido, etc..

La principal aportación del algoritmo AFDMA es que además de detectar una modificación en la imagen, es capaz de diferenciar entre una modificación intencional de una no intencional, esto se logra con la propuesta de un algoritmo de verificación, el cual esta basado en el concepto de bloques erróneos aislados y bloques erróneos concentrados, ya que la mayoría de los algoritmos de autenticación propuestos por otros autores solo se limitan a determinar las regiones alteradas de la imagen, no importando el origen de estas.

Adicional al desarrollo del algoritmo AFDMA se hace una comparación de este con los algoritmo propuestos por Zhou et al. [Zhou et al., 2004] y Xie [Xie et al., 2007], los cuales propusieron sistemas basados en bloques para la autenticación de imágenes digitales utilizando marcas de agua semi-frágiles.

El algoritmo AAMA propone un novedoso y eficiente sistema de autenticación de imágenes digitales, el cual además de detectar las regiones modificadas intencionalmente puede recuperar la imagen original con una calidad visual superior a 30 dB´s y es robusto a compresión JPEG, es decir no detecta bloques erróneos ante factores de calidad mayores a 80 a diferencia de otros algoritmos en donde sus resultados de robustez ante la compresión JPEG se enfocan a la recuperación de los bloques erróneos.

Para autenticar y recuperar la imagen original en el algoritmo AAMA solamente se debe conocer la llave secreta utilizada en el proceso de inserción, lo que hace un algoritmo de extracción a ciegas, incrementando así la seguridad del sistema.

Se realizo una comparación entre el algoritmo AAMA y los algoritmos propuestos por Zhao et al. [Zhao et al., 2007] y Hassan et al. [Hassan et al., 2008] evaluados en las mismas condiciones. Los resultados muestran la superioridad en la calidad de la imagen recuperada así como una mayor robustez ante la compresión JPEG del algoritmo AAMA.

1.7. ORGANIZACIÓN DE LA TESIS

Esta tesis esta organizada en cinco capítulos, enfocados a la introducción de los sistemas de autenticación de imágenes digitales mediante marcas de agua semi-frágiles, desarrollo del sistema propuesto y resultados del mismo.

CAPITULO I. Introducción a los sistemas de autenticación de imágenes digitales usando marcas de agua simi-frágiles tomando en cuenta los aspectos de: descripción del problema, objetivo de los sistemas de autenticación de imágenes y la justificación de los mismos.

CAPITULO II. Marco teórico de los conceptos empleados en los sistemas de marca de agua, y términos relacionados con la implementación del sistema. Antecedentes de investigaciones previas al sistema propuesto; métodos empleados, alcances de robustez, capacidad de detección y en algunas aplicaciones capacidad de recuperación con las que evaluaron los sistemas, resaltando así la aportación del sistema propuesto y finalmente la descripción detallada de los algoritmos con los cuales se realizo la comparación de los algoritmos propuestos.

CAPITULO III. Desarrollo del sistema propuesto, el cual es una descripción detallada de las técnicas empleadas para el desarrollo e implementación del algoritmo propuesto.

CAPITULO IV. Resultados obtenidos.

CAPITULO IV. Conclusiones y futuras investigaciones.

Por último se presenta una sección de anexos, así como, los artículos publicados, los cuales reportan el avance del trabajo de investigación.

1.8. CONCLUSIONES

Debido a la facilidad de alteración y manipulación de las señales digitales multimedia la autenticación o detección de modificaciones en estas señales es una importante aplicación de los métodos de marca de agua digital. En esta aplicación una firma digital se inserta como marca de agua Las marcas

de agua son llamadas semi-frágiles porque cambian cuando el contenido de la señal es alterado significativamente esto se traduce en que se detecta una falsificación. En contraste con los métodos de autenticación tradicional, en ambos casos, la marca de agua robusta o frágil se inserta directamente dentro de la imagen huésped. Con esto no se requiere de un ancho de banda adicional y se pueden diseñar algoritmos de marca de agua que puedan autenticar imágenes que fueron sometidas a ataques comunes como: cambio de formato o compresión con pérdida por JPEG.

Los algoritmos propuestos cumplen con los principales criterios para que un sistema de autenticación sea efectivo ya que ambos son susceptibles a manipulaciones intencionales, toleran alguna pérdida de información, ya que son robustos a compresión JPEG, localizan precisamente cualquier modificación maliciosa hecha a la imagen y verifican otras áreas como auténticas, en el caso del algoritmo AAMA, este es capaz de reconstruir la imagen, inclusive si esta fue parcialmente alterada o destruida de tal manera que el usuario conozca cual fue el contenido original de las áreas manipuladas.

CAPÍTULO 2

ASPECTOS TEÓRICOS Y ANTECEDENTES DE LA AUTENTICACIÓN DE IMÁGENES

2.1. CÓMO SURGE Y QUÉ ES UNA MARCA DE AGUA DIGITAL

El gran crecimiento que ha tenido en los últimos tiempos la comercialización y distribución de información digital (textos, imágenes, audio y video) genera la necesidad de tener mecanismos que les aseguren a los propietarios de esta información que sus derechos y la integridad del contenido de los archivos serán protegidos.

La protección de los derechos de autor en la distribución de contenido (en el sentido de información con cierto valor) por medios más tradicionales como son documentos en papel, grabaciones analógicas o películas de celuloide, no implica un esfuerzo tan importante para los dueños, ya que si no se cuenta con el equipo apropiado, que generalmente es especializado y muy costoso, las copias obtenidas son de una calidad notoriamente menor a la del original. Como por ejemplo vemos el caso de la copia de documentos en papel; las fotocopiadoras tienen un uso muy generalizado, pero la calidad de la fotocopia obtenida generalmente es sensiblemente distinta a la del original. Para lograr una copia de buena calidad se necesitaría contar con acceso a una imprenta, y aún en ese caso reproducir exactamente el estilo (tamaño y tipo de letra) del texto y las imágenes del original no es una tarea sencilla.

La información multimedia digital ofrece varias ventajas para su comercialización sobre los medios tradicionales. Su duplicación es fácil y

barata: el equipo requerido no es muy costoso y es fácil de usar y a diferencia de lo que ocurre en el caso de los medios tradicionales, donde la calidad de la información obtenida se va degradando con cada copia, al copiar sucesivamente un archivo digital siempre se obtiene un archivo exactamente igual al original. Además, la distribución de información digital es más flexible y menos costosa, ya sea electrónicamente (por Internet) o físicamente (por ej. en CD-ROM).

Pero son estas mismas ventajas las que hacen que la tarea de protección del contenido sea mucho más difícil, puesto que es mucho más sencillo para una persona no autorizada obtener y distribuir copias perfectas de los archivos digitales originales. Es así que surge la tecnología de inserción de marcas de agua digitales, cuyo propósito es proveer una solución para el problema de la protección de estos datos. Una marca de agua digital puede compararse con una marca de agua tradicional. Las marcas de agua tradicionales son agregadas a algunos tipos de papel (el uso más común es en billetes) para brindar una prueba de autenticidad. Generalmente son textos o figuras que no son perceptibles a menos que el papel sea mirado a trasluz. Una marca de agua digital es una señal que se inserta en un archivo digital de manera de que pueda ser detectada por una computadora pero que no sea perceptible para el ojo u oído humanos. Esta señal generalmente contiene información relacionada con el contenido del archivo, como datos sobre sus dueños o creadores, restricciones de distribución u otra información adicional [Yeung, 1998].

2.2. CRIPTOGRAFÍA Y ESTEGANOGRAFÍA VS. MARCA DE AGUA DIGITAL

La encriptación es otra técnica muy usada cuando se quiere proteger información, pero cabe resaltar que, si bien tienen ciertas características similares, la criptografía y la inserción de marcas de agua son técnicas diferentes, que pueden considerarse complementarias.

El objetivo de la criptografía [Schneier, 1995], [Stinson, 1995] es proteger el contenido de los archivos durante su transmisión, modificando los datos originales de manera de que sólo puedan ser recuperados por personas autorizadas. Una vez que los datos son recibidos y desencriptados la protección del contenido se pierde y éste queda expuesto a diversos ataques. La inserción de marcas de agua complementa a la criptografía agregándole información adicional a los datos. El objetivo de la marca de agua es que esta siempre esté presente en los archivos de modo que nunca se pierda la protección que brinda; un archivo encriptado no puede ser utilizado de la misma forma que el original mientras que un archivo marcado digitalmente sí. El objetivo de esta protección no es evitar un uso indebido de los datos, sino poder detectarlo.

La inserción de marcas de agua está muy relacionada con la esteganografía, que es una técnica que se basa en disimular (esconder) la presencia de mensajes secretos [Johhnson and Jajodia , 1998], [Katzenbeisser and Petitcolas, 2000]. Si un mensaje encriptado es interceptado puede determinarse que se trata de información encriptada y se tratará de quebrar la encriptación. La idea principal de la esteganografía es esconder la información secreta de modo que si alguien intercepta el mensaje la existencia de información oculta no será revelada. La principal diferencia entre la marca de agua y la esteganografía es el uso que se le da a ambas técnicas.

La información que se oculta usando métodos esteganográficos generalmente no está relacionada con su "cubierta" (la señal donde está escondida) mientras que una marca de agua puede considerarse como un atributo del contenido de la señal donde es insertada, brindando información adicional sobre ésta. Además en la esteganografía no se le da mayor importancia a la robustez de la señal insertada, mientras que en la

mayoría de los casos en los que se usa marca de agua se considera fundamental que la marca no se pierda aún después de distintos procesamientos que pudieran hacérsele a los datos.

2.3. ACTORES INVOLUCRADOS EN LAS MARCAS DE AGUA

El proceso de utilización de marcas de agua digitales involucra a distintos actores, los que cumplen con distintos roles dentro del proceso. Estos son:

a) Dueño de la información: Es el que posee los derechos sobre el contenido de la información. Quiere asegurarse de que será recompensado y reconocido por su trabajo y de que éste será usado apropiadamente, según las condiciones que él determine.

Es el principal interesado en contar con técnicas de marca de agua, por lo que son sus necesidades las que determinan las propiedades que deben tener las marcas de agua y las opciones que debe brindar un software que implemente la inserción y detección de marcas de agua digitales.

b) Usuario del contenido: Es el que usa y se beneficia con el contenido de los archivos digitales. Generalmente un usuario adquiere una pieza de contenido digital para trabajar con ella de alguna manera, por lo que es de gran interés para éste que el archivo adquirido pueda usarse de una forma cómoda, sin necesidad de utilizar aplicaciones especiales o de realizar un procesamiento extra previo a su utilización. Cabe destacar que la mayoría de los usuarios están dispuestos a sacrificar algo de calidad de la información por una mayor comodidad en su uso [Craver, et. all, 1998]. Es por esto que es muy importante que la presencia de una marca de agua en una pieza de contenido no implique un uso o tratamiento especial de la información, como ocurriría si se usara encriptación para proteger los datos.

Lo ideal es que una pieza con una marca de agua se comporte igual que el original si es usada correctamente y respetando las condiciones de venta.

Un usuario puede convertirse en atacante cuando no respeta las condiciones de uso o distribución de la información que adquirió.

c) Atacante: Su objetivo es poder utilizar el contenido sin atenerse a las condiciones impuestas por los autores. Su principal interés es lograr eliminar o disminuir la efectividad de las marcas de agua insertadas por los dueños para proteger la autenticidad de los datos.

d) Distribuidor: Un último actor, que puede ser parte o no de este proceso, es el distribuidor de la información, que actúa como intermediario entre los dueños y los usuarios. Su tarea consiste en brindarle a los dueños del contenido digital un servicio de distribución de su producto que le proporcione ciertas garantías sobre el uso de sus archivos.

A los ojos de los usuarios el distribuidor se comportará como el dueño de la información pues es éste el que se encarga de hacerle llegar los archivos y el que debería encargarse de que estos estuvieran protegidos con marcas de agua de acuerdo a los requerimientos de los dueños. Por otro lado, para los dueños de la información los distribuidores también son posibles atacantes, por lo que éstos últimos deberán ofrecerles algún mecanismo que les permita verificar si han usado sus productos según las condiciones y requerimientos impuestos por los propios dueños.

2.4. APLICACIONES DE UNA MARCA DE AGUA

En esta sección se listan algunos escenarios en donde se aplican de manera eficiente una marca de agua [Memon and Wong, 1998], entre los que se encuentran los siguientes:

a) Verificación de Propiedad (Ownership Assertion): Esta aplicación es la más común de una marca de agua. En este caso la marca de agua insertada en el archivo es un gráfico, una frase o cualquier otro identificador que indique quién es su dueño. Si existe una disputa sobre quién es el propietario del contenido de un archivo, el verdadero dueño podrá demostrar que su marca de agua se encuentra en el archivo, probando así que éste le pertenece.

b) Detección de copia y distribución no autorizada: El autor inserta una marca distinta (que identifica al comprador) en cada copia que distribuye. Si una copia no autorizada es encontrada, se puede determinar el origen de la copia extrayendo la marca, la que indicará a quién se le entregó el archivo en una primera instancia.

c) Verificación de autenticidad e integridad de los datos: El contenido de un archivo puede ser usado con propósitos legales, en transacciones comerciales, aplicaciones médicas, etc. En estos contextos se quiere verificar el origen de los datos (quién los creó) y que éstos no hayan sido cambiados o manipulados. Cuando se chequean los datos marcados, se extrae la marca de agua y la integridad de los datos se verifica chequeando la integridad de la marca extraída (comparándola con la marca de agua insertada). Si la marca está cambiada significa que se le hicieron modificaciones al archivo.

d) Etiquetado: Simplemente se desea insertar una etiqueta que agregue información adicional acerca de los datos del archivo, con el uso de la marca

de agua esta etiqueta es más difícil de destruir o perder. Esta aplicación es muy importante en aplicaciones médicas, ya que previene equivocaciones.

e) Control de Uso: Este tipo de control puede realizarse cuando el contenido multimedia requiere de hardware especial para copiado o reproducción, por lo que no es una aplicación tan común. Se desea tener algún mecanismo para controlar la cantidad de veces o el tiempo que un producto es utilizado. Para esto se inserta una marca de agua digital que hace las veces de *flag* o contador y es modificada (o leída) por el hardware.

2.5. CLASIFICACIÓN DE UNA MARCA DE AGUA

Las marcas de agua digitales pueden clasificarse de diversas maneras según sus distintas propiedades. Para cada contexto de aplicación de una marca de agua se requerirán distintas características de las marcas de agua. Veremos entonces las características principales que diferencian a las marcas de agua y a las técnicas de inserción y detección.

2.5.1. MARCAS DE AGUA VISIBLES E INVISIBLES

Una primera clasificación de las marcas de agua indica si la marca modifica los datos de forma perceptible para los usuarios o no. Pueden insertarse marcas invisibles (también llamadas transparentes) o visibles.

a) Marcas invisibles: La inserción de una marca de agua implica necesariamente la modificación del archivo donde es insertada; lo que logran las técnicas que insertan marcas de agua invisibles es que estos cambios sean imperceptibles para los usuarios. Insertar información automáticamente de forma que no se deteriore la calidad perceptual de la imagen sólo es posible debido a que el sistema visual humano (SVH) no es

un detector perfecto y por lo tanto pueden aprovecharse sus limitaciones para lograr que ciertas modificaciones pasen desapercibidas.

El objetivo de la mayoría de las aplicaciones de marca de agua digital es marcar un archivo de forma tal que la utilidad que pudiera tener esa información para los usuarios no se vea afectada por el proceso de marcado. Es claro que en estos casos una marca de agua invisible se considerará más apropiada que una visible, puesto que con este tipo de marcas el usuario no percibe la diferencia entre la imagen marcada y la original.

Existen dos métodos para medir la calidad perceptual de una imagen, los cuales son llamados subjetivos y objetivos.

- Los métodos *subjetivos* involucran a un grupo de observadores los cuales determinan la diferencia entre la imagen original, la marcada y la recibida para determinar si estas han sufrido algún cambio o degradación. Este escenario determina la calidad visual de los datos usando métodos de estimulo simple o doble los cuales forzan a una elección. La recomendación ITU 500 provee normas para procesos de evaluación subjetiva de calidad en imágenes. Esta clasificación fue hecha típicamente en una escala de cinco puntos como en la ITU-R-BT.500 y la DSCQS (Double Stimulus Continuos Quality Scale) de este grupo de medidas se calcula la desviación estándar y la media, los cuales son el resultado numérico final que evalúa la calidad subjetiva de la imagen. Ya que los métodos para medir la calidad subjetiva requieren de la intervención humana, esta medida obviamente no es apropiada para un benchmarking automático.

- En cuanto a los métodos *objetivos* para la evaluación de la calidad visual en una imagen se encuentra la aproximación PSNR (the peak signal-to-noise ratio), el cual es bien conocido y ampliamente usado

para mediar la calidad, esta aproximación normaliza la medida de la varianza de la señal y calcula el logaritmo base 10 de esta relación, entonces se obtiene la relación señal a ruido SNR (signal-to-noise ratio). Si la normalización esta dentro del valor pico de la señal, se obtiene entonces el valor pico la relación señal a ruido (PSNR).

Aunque también es conocido que este criterio puede no concordar del todo bien con las clasificaciones subjetivas, el PSNR indica que la calidad de la imagen ha sido alterada en caso de haber filtrado la imagen o haber adicionado ruido aleatorio, ya que la diferencia entre los datos originales y los marcados es pequeña en general esperamos una correlación razonable entre las tasas iniciales de la calidad subjetiva de los datos marcados en la mayoría de los casos.

b) Marcas visibles: Las marcas de agua visibles son generalmente el logo de una empresa u otro tipo de imagen que indican quién es el propietario de los datos. Las propiedades que generalmente se piden de este tipo de marcas son:

- Ser obvias para cualquier persona con una visión normal o corregida (inclusive daltónicos)
- Ser lo suficientemente flexibles para hacerlas tan notorias como se quiera
- Tener características que formen una imagen por sí mismas
- Permitir que todas las partes de las imagen original se vean en la imagen marcada
- Deben ser muy difíciles de remover y falsificar

Lograr que una marca visible sea robusta es más difícil que para las marcas invisibles porque en este caso el atacante sabe en qué sectores de la imagen

se encuentra la marca y por lo tanto sabe dónde concentrar sus esfuerzos al intentar removerla.

Una marca de agua visible puede utilizarse para que los usuarios verifiquen rápidamente cuál es la fuente de los datos.

Otra aplicación para este tipo de marcas es brindar un preview de la información original. Supongamos que una persona tiene ciertas imágenes que desea comercializar. Una buena forma de publicitarlas es poniendo en Internet sus imágenes con marcas de agua muy visibles que permitan apreciar la calidad del original pero que hagan que la imagen marcada no pueda ser utilizada. De esta forma los posibles compradores pueden ver las características principales de las imágenes que van a comprar, pero de todas formas deberán recurrir a los dueños si desean utilizarlas.

2.5.2. MARCAS DE AGUA FRÁGILES Y ROBUSTAS

Las marcas de agua digitales también pueden dividirse en dos categorías según la resistencia de la marca a la manipulación del contenido de la información en este caso se dice que una marca de agua es frágil o robusta.

a) Marcas frágiles: Estas son marcas de agua que son corrompidas si el archivo en el que están embebidas sufre modificaciones. Generalmente se usan para la detección de ataques intencionales o no intencionales en las imágenes, que se basa en poder verificar que el contenido no haya sido cambiado o que el emisor sea quien dice ser. Una persona que va a usar un archivo con una marca de este tipo intentará extraer la marca de agua y la comparará con la marca de agua original. Si la marca no se encuentra o no coincide con el original se sabe que el archivo ha sufrido modificaciones.

Es importante que pequeños cambios en el contenido de la información sean detectados y es bueno que además la marca de agua ayude a localizar espacialmente las regiones en dónde se realizaron los cambios.

b) Marcas robustas: Generalmente se usan para lo que se llama protección de los derechos de autor, esta área de aplicación se basa principalmente en la identificación del dueño de la información y la identificación de los compradores del contenido marcado.

Para este tipo de aplicaciones es fundamental que la marca de agua resista distintos tipos de ataques intencionales o no intencionales. En la terminología de marca de agua un ataque es cualquier proceso que pueda afectar la detección de la marca de agua o la información que lleva la marca de agua. Obviamente las distorsiones deben ser limitadas, es decir, no deben ser excesivas, si no la imagen transformada podría ser inservible. En un ataque intencional, el objetivo del atacante es minimizar la probabilidad de detección de la modificación, mientras maximiza el impacto que produce su transformación en la imagen, esto lo hace sin conocer el valor de la llave secreta usado en el proceso de inserción de la marca de agua.

A continuación describimos algunos de los ataques más conocidos. Algunos de ellos pueden ser intencionales o no intencionales, dependiendo de su aplicación:

- *Ruido Aditivo:* Este se puede generar con ciertas aplicaciones, como el uso de convertidores analógico-digital y digital-analógico o por transmisores que generan errores. Sin embargo un atacante puede introducir ruido de manera intencional (esto es de manera imperceptible) con la máxima potencia inadvertida. Esto forza a un incremento en el umbral con el cual el detector de correlación trabaja.

- *Filtrado:* Un filtro pasa-bajas por ejemplo no introduce una degradación considerable en imágenes marcadas, pero puede afectar dramáticamente su desempeño, ya que la distribución del espectro de la marca de agua tiene muchos componentes espectrales de alta frecuencia.

- *Recorte:* Este es un ataque muy común, ya que el atacante esta interesado en pequeñas porciones del objeto marcado, tales como, partes de cierta pintura o cuadros de una secuencia de video. Con esto en mente, la marca de agua debe esparcirse por toda la imagen para que la marca de agua sobreviva a este tipo de ataque.

- *Compresión:* Este es generalmente un ataque no intencional, el cual se presenta constantemente en las aplicaciones multimedia. Prácticamente todas las imágenes, audio o videos que son distribuidos vía Internet son comprimidas. Si se requiere que la marca de agua sea resistente a diferentes niveles de compresión se recomienda que la inserción de la marca de agua se realice en el mismo dominio en que se realiza la compresión. Por ejemplo, las imágenes que se marcan en el dominio DCT son más robustas a la compresión JPEG que las que se marcan en el dominio espacial.

- *Rotación y Escalamiento:* Esta ha sido el gran reto científico y tecnológico de las marca de agua digitales, especialmente por su éxito en imágenes. La medida de correlación basada en la detección y extracción falla cuando la imagen marcada ha sido rotada o escalada debido a que la marca de agua y la versión generada localmente ya no comparten el mismo patrón espacial. Esto se podría lograr si se realizara un búsqueda exhaustiva de los diferentes ángulos de rotación y factores de escalamiento hasta que encontráramos una correlación pico, pero esto tendría un gran costo computacional muy

alto. En el caso de conocer la imagen original seria muy sencillo conocer la estimación de estos dos parámetros. Recientemente algunos autores han propuesto utilizar transformaciones invariantes a rotación y escalamiento (como la de Fourier-Mellin en [Herrigel, et. all, 1998]) pero esto reduce considerablemente la capacidad de ocultar información en la imagen.

- *Promedio estadístico:* Un ataque de esta tipo se genera cuando el atacante logra estimar la imagen marcada y la no marcada. Esto es muy peligroso si la marca de agua no depende sustancialmente de los datos de la imagen, ya que con diferentes imágenes marcadas es posible calcular el valor por medio de un promedio simple, es por esto que es importante generar marcas de agua usando una mascara perceptual.

- *Marca de agua multiple:* Se presenta cuando una imagen que ya ha sido marcada se vuelve a marcar por un atacante que posteriormente reclama la propiedad de la imagen. Una posible solución podría ser marcar con la hora en que se oculto la información y en presencia de una autoridad certificada.

- *Impresión y escaneado:* Se presenta cuando una imagen marcada llega a manos del atacante de forma impresa y este la escanea para distribuirla, pretendiendo decir que el es el propietario de la imagen.

- *Ataques de falsificación por parte de individuos mal intencionados:* Se presenta cuando un atacante reemplaza el producto original con segmentos de la imagen que no concuerdan con los datos originales

Lograr que la marca de agua sea robusta a todos los ataques es prácticamente imposible, por lo que se desarrollan distintas técnicas donde

cada una se centra en lograr la robustez de la marca frente a determinados ataques. La elección de la técnica a utilizar dependerá de la aplicación que se le quiera dar a la marca de agua.

2.5.3. DOMINIO DE INSERCIÓN

Por otro lado, las técnicas de marca de agua para imágenes también pueden dividirse según la forma en que modifican el archivo para insertar la marca de agua: pueden ser en el dominio espacial o en el dominio espectral/de transformadas.

a) Técnicas de dominio espacial: Las técnicas de dominio espacial son las primeras que se estudiaron e implementaron y se basan en esquemas relativamente simples. Este tipo de técnicas es el indicado cuando se quieren insertar marcas de agua visibles. Si se quieren utilizar para la inserción de marcas invisibles se debe tener en cuenta que se tendrá que sacrificar robustez en favor de lograr la invisibilidad de la marca pues, como ya mencionamos, ésta deberá ser insertada en los bits menos significativos de los pixels y por lo tanto será fácilmente removible..

La mayoría de estas técnicas modifican solamente los bits menos significativos del valor de los pixeles, obteniéndose de esta forma imágenes de alta calidad (la inserción de la marca de agua degrada muy poco la imagen original), por lo que son efectivas para la inserción de marcas invisibles.

Algunas desventajas de este método son:

- La marca es muy poco resistente a alteraciones comunes de la imagen puesto que los bits menos significativos son los que generalmente resultan modificados en el procesamiento comúnmente realizado sobre imágenes.

- Tienen "baja capacidad de bits", lo que significa que permiten insertar pocos bits de información, debido a que la inserción de cada bit de la marca requiere la modificación de muchos pixels de la imagen (para poder recuperar la marca aún cuando algunos bits hayan sido modificados por procesamientos posteriores).

b) Técnicas de dominio espectral: Las técnicas de dominio espectral no modifican directamente el valor de los pixeles de la imagen, sino que transforman las imágenes al dominio de frecuencias para luego insertar la marca en la señal obtenida en ese dominio. Las transformadas más comunes utilizadas en los algoritmos de marca de agua son:

- Transformada de Fourier **(FT)**,
- Transformada de Coseno Discreto **(DCT)**,
- Transformada Wavelet Discreta **(DWT)**,
- Transformada Slant **(SLT)**.

Al aplicar la transformada inversa a la señal modificada se obtiene la imagen marcada. Trabajando en el dominio espectral generalmente se logra una mayor robustez de la marca, pero los algoritmos utilizados son más complejos. Estas técnicas pueden insertar una gran cantidad de bits sin degradar notoriamente la calidad de la imagen, pero aquí hay un mayor compromiso entre invisibilidad y robustez puesto que la marca se aplica indiscriminadamente en el dominio espacial de la imagen.

Las técnicas de inserción de marca de agua en el dominio de frecuencias muchas veces se derivan por analogía con las comunicaciones de amplio espectro. En este tipo de comunicaciones uno transmite una señal de poco ancho de banda a través de un canal con un ancho de banda mucho mayor, de manera de que la energía de la señal trasmitida esté presente en todas las frecuencias y al mismo tiempo en ninguna sea tan fuerte como para ser

detectable. De la misma manera, el dominio de frecuencias de una imagen puede considerarse el canal de comunicación por el que se quiere trasmitir una señal (la marca de agua), que será distribuida en los distintos canales de frecuencia de la imagen.

Las transformadas utilizadas en las técnicas de inserción de los algoritmos de marca de agua, generalmente son conocidas y/o documentadas en la mayoría de la literatura especializada en el tema, cosa que no sucede con la Transformada Slant, debido a que esta transformada es utilizada por el algoritmo propuesto por Zhao et al. en [Zhao et al., 2007], el cual es comparado con el algoritmo AAMA es conveniente describirla a continuación.

2.5.3.1. TRANSFORMADA SLANT

Shibata y Enomoto han introducido transformadas ortogonales que contienen un vector base "slant" para datos de vectores de longitud 4 y 8 [Enomoto, Shibata, 1971]. El vector slant tiene una forma de onda diente de sierra discreto que decrece a pasos uniformes sobre su longitud, el cual es adecuado para una eficiente representación gradual de los cambios de brillo en una línea de imagen. El trabajo propuesto por los autores no da una indicación sobre la construcción de vectores de datos de gran tamaño así como también no utiliza algoritmos computacionales rápidos. Con el propósito de obtener un alto grado de compresión de la imagen usando técnicas de transformación, es necesario desarrollar transformaciones bidimensionales sobre bloques de al menos 16*16 elementos. Para tamaños de bloque grande, la complejidad computacional usualmente no es factible, a menos que se utilice un algoritmo rápido.

Tomando en consideración esos antecedentes se desarrollo la Transformada Slant **(SLT)** para codificación de imágenes cuya matriz posee las siguientes propiedades:

1. El conjunto de vectores básicos son ortonormales.
2. Un vector básico es constante.
3. Un vector básico es inclinado "Slant".
4. Propiedad secuencial.
5. Transformación de tamaño variable.
6. Algoritmo computacional rápido.
7. Alta compactación de energía.

2.5.3.1.1. TRANSFORMADA SLANT BIDIMENSIONAL

Sabemos que [*F*] es una matriz de *N*N* pixeles que representan valores de intensidad de los pixeles de una imagen y [f_i] es un vector de *N*1* que representa la i-esima columna de [*F*]. La transformada unidimensional de la i-esima línea de la imagen es

$$[\tilde{f}_i] = [S][f_i], \tag{2.1}$$

donde [*S*] es la matriz slant unitaria.

La transformada Slant bidimensional se lleva a cabo mediante transformaciones secuenciales de filas y columnas sobre [*F*]

$$[\tilde{F}] = [S][F][S]^T, \tag{2.2}$$

La transformación inversa para recuperar $[F]$ a partir de los componentes transformados $[\tilde{F}]$ se obtiene por

$$[F] = [S]^T [\tilde{F}][S]. \tag{2.3}$$

2.5.3.1.2. MATRIZ DE TRANSFORMACIÓN SLANT

La transformada Slant es un miembro de la clase de transformadas en las cuales las matrices son ortogonales, tiene una función constante para la primer fila y tiene una segunda fila la cual es una función lineal (slant) del índice de la columna. Las matrices son formadas mediante una construcción iterativa que exhibe las matrices como productos de matrices reducidas, las cuales hacen al algoritmo de transformación rápido.

La matriz de transformación Slant de orden dos consiste de una constante y un vector base "slant" dado por

$$S_2 = \frac{1}{2^{1/2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{2.4}$$

2.5.3.1.3. ALGORITMO COMPUTACIONAL RÁPIDO PARA LA TRANSFORMACIÓN SLANT

El algoritmo computacional rápido para la transformación slant esta basado en la matriz factorizada mostrada a continuación.

Generalizando para obtener la matriz slant de orden N ($n=2^n$, $n=1,2,3, \ldots$) en términos de la matriz slant de orden $N/2$ se da por la siguiente ecuación recursiva:

$$S_N = \frac{1}{2^{1/2}} \left[\begin{array}{cc|c|cc|c} 1 & 0 & & 1 & 0 & \\ a_N & b_N & 0 & -a_N & b_N & 0 \\ \hline 0 & & I_{(N/2)-2} & 0 & & I_{(N/2)-2} \\ \hline 0 & 1 & & 0 & -1 & \\ -b_N & a_N & 0 & b_N & a_N & 0 \\ \hline 0 & & I_{(N/2)-2} & 0 & & -I_{(N/2)-2} \end{array}\right] \left[\begin{array}{c|c} S_{N/2} & 0 \\ \hline & \\ 0 & S_{N/2} \end{array}\right], \tag{2.5}$$

donde

$$\begin{aligned} a_2 &= 1, \\ b_N &= [1+4(a_{N/2})^2]^{-1/2}, \\ a_N &= 2b_N a_{N/2}. \end{aligned} \tag{2.6}$$

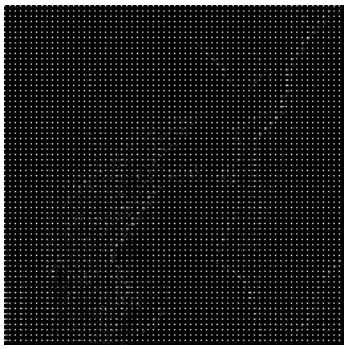
2.5.3.1.4. VENTAJAS DEL USO DE LA TRANSFORMACIÓN SLANT

Una de las ventajas del uso de la transformación Slant es que se puede implementar en varios sistemas de codificación de imágenes en escala de gris o a color.

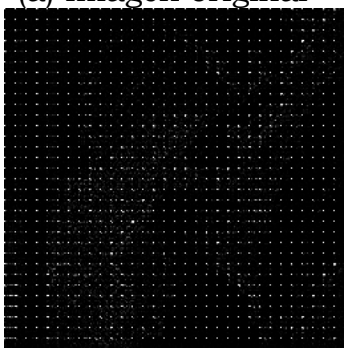
Otra importante cualidad de la transformación Slant es que se consigue una buena calidad de codificación con aproximadamente 1 a 2 bits/píxel para imágenes monocromáticas y 2 a 3 bits/píxel para imágenes a color.

La figura 2.1 muestra los coeficientes Slant de la imagen "Lena" transformada con una matriz de orden 4 (figura 2.1 (b)), de orden 8 (figura 2.1 (c)) y de orden 16 (figura 2.1 (d)), aplicando la transformada Slant Inversa en los tres casos se obtiene la imagen mostrada en la figura 2.1 (e) cuya calidad es exactamente igual a la imagen original.

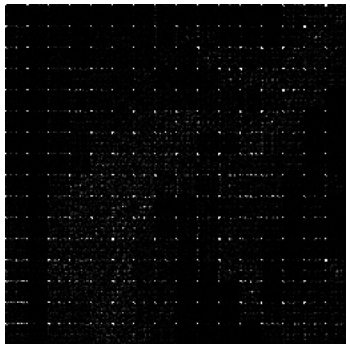
(a) Imagen original



(b) Coeficientes Slant orden 4



(c) Coeficientes Slant orden 8



(d) Coeficientes Slant orden 16

(e) Imagen recuperada

Figura 2.1. Coeficientes Slant empleando matrices de diferente orden

2.5.4. OTRAS CLASIFICACIONES

Otra forma de categorizar las técnicas de marca de agua es separar las que requieren de la imagen original en el proceso de detección de la marca de agua (non blind detection) y las que no utilizan la imagen original (blind detection).

Las técnicas que no son a ciegas tienen la ventaja de una mayor probabilidad de detectar la marca de agua en imágenes que han sido modificadas de varias maneras, es decir, son más robustas, pues pueden utilizar información secreta de la imagen original para la detección, pero este tipo de técnicas no puede combinarse con detectores automáticos o de búsqueda en Internet. En ciertos contextos la utilización de una técnica que no es a ciegas no es aplicable, como en el caso en que la marca de agua es utilizada para la autenticación de datos, donde la utilidad de la marca se basa en que una persona que no tiene acceso al archivo original, pueda asegurarse de todas maneras de que ha recibido una copia fiel de éste.

Las técnicas que usan la imagen original para generar la marca de agua se conocen como algoritmos adaptivos. Este tipo de técnicas puede aprovechar características de la imagen original para lograr la invisibilidad de la marca. Además la robustez también puede verse incrementada pues es más difícil para un atacante que tiene distintas imágenes marcadas con el mismo algoritmo determinar mediante análisis estadístico cuál es la marca de agua pues ésta será distinta para cada copia.

La utilización de la extracción a ciegas además de ser más conveniente en casos de autenticación del contenido de la imagen, incrementa la seguridad del sistema, ya que la imagen original permanece a resguardo del propietario de la imagen.

2.6. ETAPAS GENERALES DEL PROCESO DE MARCA DE AGUA EN IMÁGENES DIGITALES

La figura 2.2 muestra un esquema general de marca de agua para una imagen digital de tamaño *MxN*, de la cual pueden distinguirse tres etapas fundamentales: generación de la marca de agua, inserción de la marca y

extracción o detección de la misma, a continuación describiremos brevemente cada una de ellas.

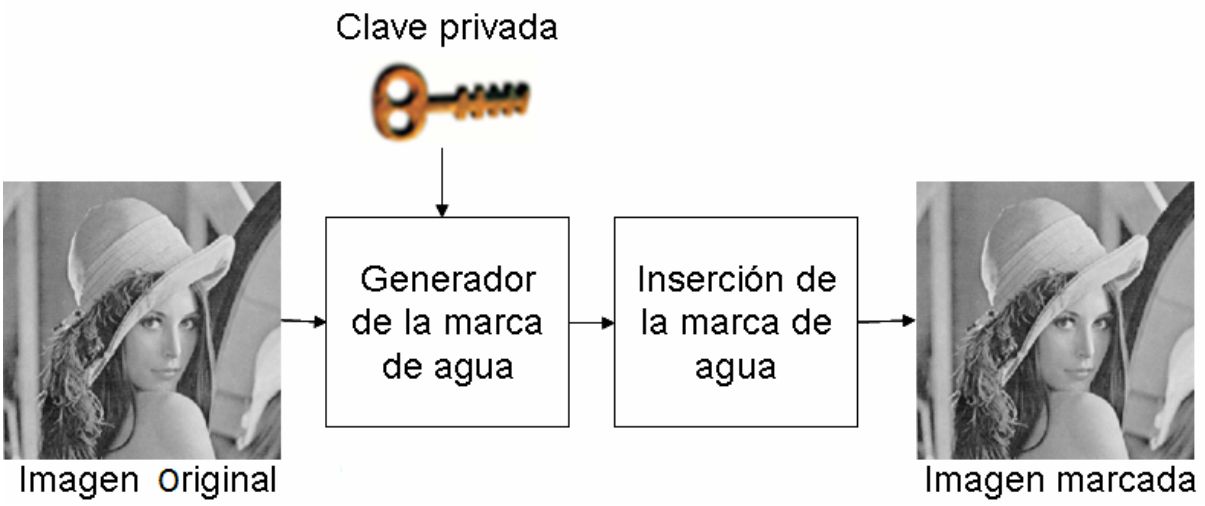


Figura 2.2. Esquema general del proceso de inserción de una marca de agua

2.6.1. GENERADOR DE LA MARCA DE AGUA

No todas las técnicas de marca de agua son capaces de insertar la misma cantidad o el mismo tipo de información, por lo que cada una presenta un grupo de condiciones que debe cumplir la marca de agua que se usará para marcar un archivo.

Las técnicas que insertan marcas frágiles o visibles son más flexibles con estas condiciones, pues no tienen en cuenta muchos aspectos de seguridad que limitan enormemente el tipo de marca que puede utilizarse. En este tipo de aplicaciones generalmente el dueño puede proporcionar la marca que será insertada, la cual comúnmente es un gráfico o una frase, solo deben atenderse ciertas restricciones con respecto al tamaño (por ej.: dimensiones de una imagen o largo de un texto) o al formato de la marca (por ej.: solamente imágenes representadas como mapas de bits, o textos que únicamente contengan caracteres ASCII).

Sin embargo, en las aplicaciones donde la robustez de la marca es un factor fundamental ésta debe tener una estructura mucho más específica para que su uso sea efectivo. En estos casos el dueño del archivo a marcar no proporciona la marca sino una clave privada K que se utilizará para generar de forma unívoca una marca de agua apropiada para la técnica de marcado. Es así que la marca de agua W puede verse como el resultado de aplicarle una función F a la clave privada K y, en el caso de algoritmos adaptativos, a la imagen original I.

En muchos casos F utiliza un generador de números pseudo-aleatorios que recibe como semilla la clave K. Si la técnica de extracción de la marca de agua es a ciegas, es decir, no utiliza la imagen original para la detección y se utiliza a I en la generación de la marca entonces F debería tomar en cuenta sólo propiedades robustas de la imagen, pues probablemente en la detección se utilice F con una versión modificada del original I.

$$W = F(I, K). \tag{2.7}$$

2.6.2. INSERCIÓN DE LA MARCA DE AGUA

En el proceso de inserción se obtiene una imagen marcada I_W a partir de la imagen original I y la marca de agua generada W. Este proceso puede denotarse como:

$$I_W = E(I, W), \tag{2.8}$$

donde E representa el algoritmo mediante el cual se inserta la marca W en la imagen I.

El método de inserción de la marca es distinto para cada técnica, ya que estas pueden diferir de acuerdo a la aplicación, al dominio en el cual se realizan los cambios a la imagen, e inclusive aunque las técnicas trabajen en el mismo dominio puede haber grandes diferencias en el proceso que realizan para insertar la información. Dentro de estas diferencias podemos mencionar la forma de elección de las partes de la imagen que serán modificadas. En algunos casos la marca se inserta una sola vez en un lugar determinado de la imagen (el ejemplo más claro es el caso de marcas visibles), en otros casos la misma información es insertada varias veces en toda la imagen (generalmente utilizado con marcas frágiles invisibles) y en otras aplicaciones se utiliza cierta información secreta para determinar dónde será insertada la marca (para lograr una mayor robustez). Esta información secreta puede ser una clave proporcionada por el usuario o información derivada de la propia imagen original.

2.6.3. EXTRACCIÓN O DETECCIÓN DE LA MARCA DE AGUA

En los algoritmos de marca de agua esta es la parte más importante pues es la que nos permite aseverar algo acerca de la información. Denotamos al procedimiento de detección como la función *D*, que puede involucrar la extracción de una marca de agua o simplemente la detección de una marca determinada.

Típicamente la **extracción** se utiliza para marcas de agua frágiles o semi-frágiles, donde la función *D* devuelve la señal que fue insertada en el archivo. Esta forma de detección se utiliza en las aplicaciones donde la marca se inserta para agregar información o en aquellas donde se quiere verificar la integridad de los datos. Como se muestra en la figura 2.3 en este caso se realiza un último paso que consiste en comparar la marca extraída con la marca que fue insertada en una primera instancia y donde los

cambios encontrados en la marca permiten determinar qué cambios sufrió la imagen marcada en cuestión.

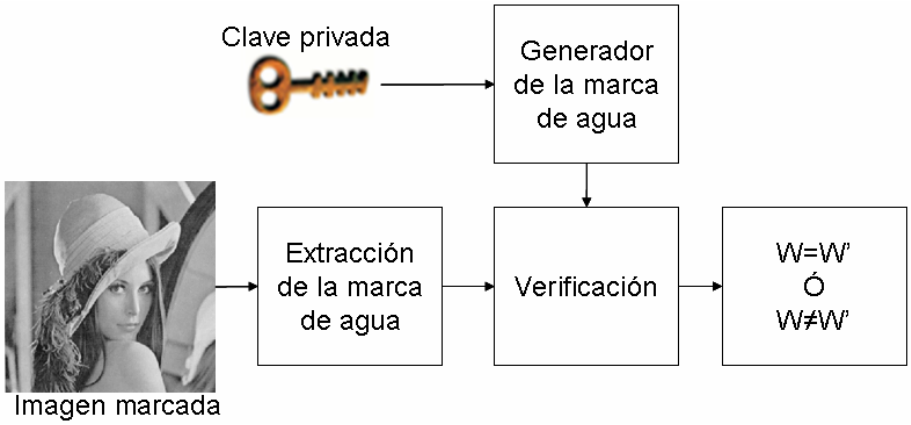


Figura 2.3. Esquema general del proceso de extracción y verificación de una marca de agua

$$D(I_W^{'}) = W^{'}. \tag{2.9}$$

En el caso de la **detección** la función D devuelve un valor de decisión (Sí/No) sobre la existencia de una marca de agua determinada en la imagen marcada en cuestión, preferentemente la detección se realiza sin el uso de la imagen original como se muestra en la figura 2.4. Cuando una imagen ha sido modificada de diversas maneras, la señal correspondiente a W aparecerá debilitada. Esto implica que muchas veces no puede determinarse con absoluta certeza la presencia o ausencia de W en el archivo es cuestión, por lo que D generalmente recibe un parámetro adicional T (umbral) que indica el valor mínimo para el cual la fuerza de la marca es considerada suficiente para asegurar que ésta marca fue insertada en la imagen.

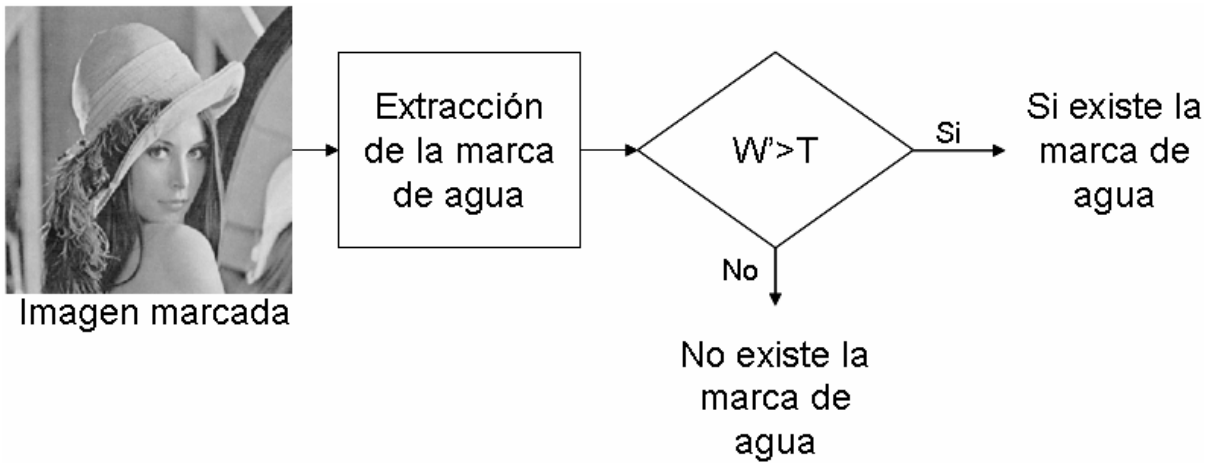


Figura 2.4. Esquema general del proceso de detección de una marca de agua

$$D(I_W^{'}, W, T) = \begin{cases} Si & si\ W > T \\ No & en\ otro\ caso \end{cases}. \qquad (2.10)$$

Debemos notar que en la señal de una imagen *Iω'* (Imagen marcada alterada) derivada de la imagen marcada *Iω*, la magnitud de *I* es mucho mayor que la de la señal insertada *W* o que la de la señal producida por las distorsiones que puede haber sufrido, puesto que de otra manera no se conservaría la fidelidad de la imagen. Esto significa que en un proceso de detección en el que se usa la imagen original se puede mejorar muchísimo la relación señal-ruido (donde se considera ruido a toda señal que no sea causada por la presencia de *W*) con simplemente restarle la señal de *I* a la imagen que se está probando. Esta es una de las principales ventajas de las técnicas de detección que no son a ciegas.

Después de que el detector D es aplicado se pueden generar los siguientes errores:

Error tipo I: Es cuando se detecta una marca de agua aunque esta no exista en los datos, también es conocido como falso positivo.

Error tipo II: Se presenta cuando no se detecta la marca de agua aunque esta este presente en los datos, también conocido como falso negativo.

Los errores anteriores ocurren con una probabilidad de falsa alarma (P_{fa}) y una probabilidad de falso negativo (P_{rej}) respectivamente. La figura 2.5 representa esquemáticamente la detección de estos errores cuando se realiza una prueba estadística basada en una distribución normal. Tenemos que $c=1-P_{fa}$, el cual representa la certeza de una detección positiva. La prueba de hipótesis puede ser usada para una estimación estadística segura y para la minimización en el error de detección [Papoulis, 1991]. Generalmente cuando un falso positivo es insignificante, es decir, $P_{fa}\rightarrow 0$ la probabilidad de falso negativo a la marca de agua incrementa $P_{rej}\rightarrow 1$ y viceversa.

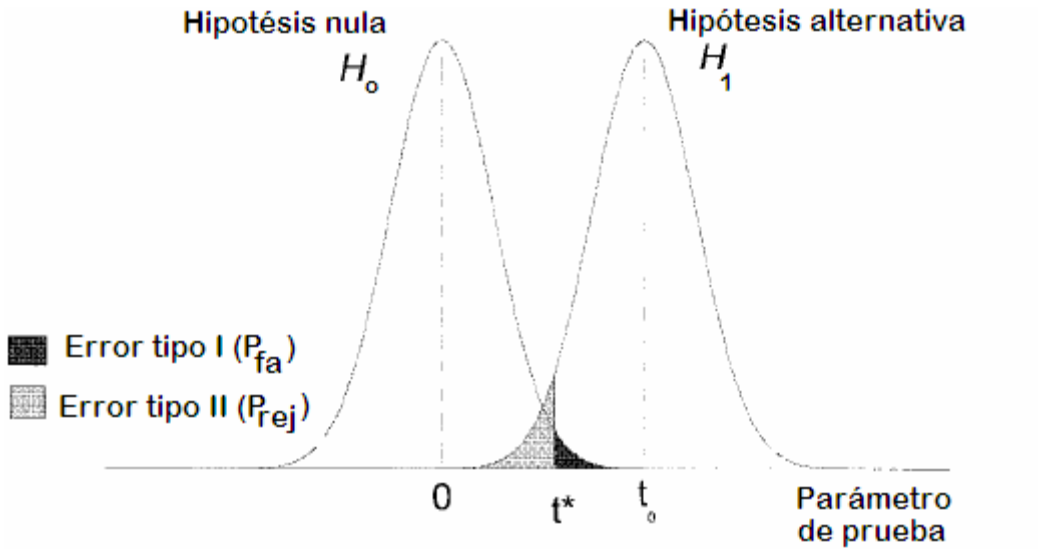


Figura 2.5. Prueba de hipótesis para la detección de marcas de agua basada en una distribución normal.

En muchos casos la detección se basa en la correlación entre la señal de marca de agua y el producto marcado [Tirkel, et. all, 1994], [Tirkel, et. all, 1998], [Swanson, et. all, 1998], [Cox, et. all, 1997]. También pueden

emplearse pruebas estadísticas en el proceso detección de la marca de agua [Pitas and Kaskalis, 1995], [Voyatzis and Pitas, 1998], [Bender, et. all, 1996]

2.7. ANTECEDENTES

En esta sección no pretendemos dar una completa y exhaustiva explicación de todos los métodos de autenticación de imágenes, cabe señalar que excluimos algoritmos que son utilizados para autenticación pero no abordan aspectos de marca de agua, como lo son los algoritmos basados en firma digital externa, dentro de los que se encuentran los sistemas criptográficos clásicos, por ejemplo MD-4, MD-5 (mensaje digerido), CRC-32 (código de redundancia cíclica de 32-bits), SHA-1 (algoritmo hash seguro) y algunos otros [SHA-1, 1995] [Bhattacharjee, 1998] [Lin, 1998a] [Lin, 1998b] [Wolfgang, 1996] [Wolfgang, 1999].

Los sistemas de autenticación de imágenes pueden ser clasificados de acuerdo a si se requiere una integridad estricta o se requiere la autenticación del contenido solamente, también depende del modo de almacenamiento de los datos de autenticación (por ejemplo marca de agua o firma digital externa).

Los métodos de marca de agua se clasifican en marca de agua frágil y marca de agua semi-frágil.

2.7.1. MÉTODOS DE AUTENTICACIÓN DE IMÁGENES BASADOS EN MARCA DE AGUA FRÁGIL

Muchos de los métodos que se han propuesto actualmente para la autenticación de imágenes están basados en una marca de agua frágil, la idea básica de estas técnicas es insertar una marca de agua especifica (generalmente independiente de los datos de la imagen [Yeung, 1997]) es por

eso que cualquier modificación o alteración en el contenido de una imagen altera la marca de agua. El proceso de autenticación consiste en detectar distorsiones en la marca de agua con la finalidad de localizar las regiones que fueron alteradas en la imagen. La mayor desventaja de este sistema es su dificultad para distinguir entre ataques maliciosos y no maliciosos, por ejemplo, todos los algoritmos basados en la marca de agua frágil consideran la compresión con pérdida en una imagen como una alteración intencional en la imagen.

Los principales métodos de inserción de la marca de agua frágil son los siguientes:

2.7.1.1. INSERCIÓN EN EL LSB

Una de las primeras técnicas usadas para la detección de modificaciones en la imagen se baso en la inserción de la suma de comprobación dentro de los bits menos significativos (LSB) de los datos de la imagen. El algoritmo propuesto por Walton [Walton, 1995], consiste en seleccionar un grupo de pixeles de manera pseudos-aleatoria de acuerdo a una llave secreta. El valor de la suma de comprobación se obtiene sumando los 7 bits más significativos (MSB) de los pixeles seleccionados, después el bit de la suma de comprobación se inserta en el LSB.

El proceso de chequeo consiste en comparar para cada bloque de la suma de comprobación determinado por el MSB de la imagen a evaluar con el valor original de la suma de comprobación recuperado del LSB.

Las ventajas principales de este método es que no producen cambios visibles en la imagen y provee una alta probabilidad de detección de alteraciones. Por ejemplo, si nosotros cambiamos solo dos pixeles de cualquier bloque, entonces la suma de comprobación se va a modificar

porque cada píxel p_j del bloque es multiplicado por un coeficiente diferente a_j. Además el paso aleatorio de los pixeles p_j y los coeficientes a_j son bloques dependientes, esto hace imposible cambiar o duplicar bloques enteros sin que la marca de agua detecte los cambios. Una de las desventajas de esta técnica es que es posible cambiar bloques homólogos (que son bloques de la misma posición) de dos imágenes autenticadas protegidas con la misma llave. Una solución simple para este tipo de ataque es hacer la marca de agua dependiente del contenido de la imagen. Esto se puede lograr usando el algoritmo de extracción de bit robusto propuesto por Fridrich [Fridrich, 1999].

En [Bravo, et al., 2008] se propone un esquema de marca de agua frágil el cual utiliza un logotipo para la autenticación y verificación de la integridad de la imagen. La seguridad del sistema propuesto esta basado en bloques y depende de un algoritmo de encripción de llave publica y una función hash. Los métodos de codificación y decodificación tienen la capacidad de detección incluso con la ausencia de los índices de la imagen y los logotipos originales. El detector auténtica automáticamente las imágenes de entrada y extrae los posibles logotipos e índices de la imagen, los cuales son usados para localizar las regiones alteradas.

2.7.1.2. AUTO INSERCIÓN

Fridrich y Goljan [Fridrich, 1999b] proponen un método original de auto inserción, una imagen dentro de si misma para proteger el contenido de la imagen. Este método permite que las regiones de la imagen que fueron alteradas con recorte o reemplazadas puedan ser parcialmente reparadas, el principio básico de este método es insertar una versión comprimida de la imagen dentro del LSB de sus pixeles. Como en todos lo métodos basados en la inserción de la marca de agua en el LSB, este tampoco introduce distorsiones visibles. El algoritmo consiste en dividir la imagen en bloques

de 8x8 pixeles. Se pone el LSB de cada píxel a cero y después se transforma cada bloque con la transformada DCT (Transformada Coseno Discreta). La matriz DCT es cuantizada con una matriz de cuantificación correspondiente a la compresión JPEG con un factor de calidad de 50. El resultado es codificado usando solo 64 bits y el código es insertado dentro el LSB de otro bloque. El bloque marcado debe estar lo suficientemente distante al bloque protegido para prevenir un deterioro simultáneo en la imagen y la recuperación de datos durante la alteración local de la imagen. La calidad de las regiones de recuperación de la imagen es algo peor que un 50% de calidad JPEG, pero suficiente para informar al usuario el contenido original de esas áreas. Los mismos autores proponen un método alternativo, el cual permite que la calidad de la imagen reconstruida sea mejorada. En esta variante se usan dos LSBs para insertar los coeficientes DCT cuantificados (se utilizan 128 bits en lugar de 64 bits). La mayor desventaja de este método es que la información insertada no es robusta. Si varias regiones distintas de la imagen han sido alteradas, la recuperación de los datos será errónea. Además ante modificaciones globales en la imagen marcada como: filtrado o compresión con pérdida, la mayoría de los datos reconstruidos serán erróneos debido a que los valores LSB son cambiados por esta clase de operaciones.

2.7.2. MÉTODOS DE AUTENTICACIÓN DE IMÁGENES BASADOS EN MARCA DE AGUA SEMI-FRÁGIL

Una marca de agua semi-frágil es otro tipo de autenticación de marca de agua. Las marcas de agua semi-frágiles son más robustas que las marcas de agua frágiles y menos sensibles a modificaciones como compresión JPEG. El objetivo de estos métodos es discriminar entre manipulaciones maliciosas como la adición o remoción de elementos significativos en la imagen y operaciones globales que preservan el contenido semántico de la imagen.

El uso de estos métodos esta principalmente justificado por el hecho de que las imágenes son generalmente transmitidas y almacenadas en forma comprimida. Además la mayoría de las aplicaciones de compresión con pérdida no afectan la integridad de la imagen dentro del significado de su interpretación.

2.7.2.1. MÉTODOS DE MARCA DE AGUA SEMI-FRÁGIL ROBUSTOS A COMPRESIÓN

Lin y Chang [Lin, Chang, 2000] propusieron un algoritmo de marca de agua semi-frágil que acepta compresión con pérdida JPEG y rechaza modificaciones maliciosas. Ellos resaltan y muestran dos propiedades invariantes de los coeficientes DCT con respecto a la compresión JPEG.

La primera propiedad muestra que si modificamos un coeficiente DCT con un entero múltiple al paso de cuantificación *Q'm*, el cual es mayor que el paso de cuantificación usado en la compresión JPEG, entonces estos coeficientes pueden ser reconstruidos después de la compresión JPEG.

El segundo es una relación invariante entre dos coeficientes homólogos en un par de bloques antes y después de la compresión. Porque todas las matrices de coeficientes DCT están divididas por la misma tabla de cuantificación en el proceso de compresión JPEG, la relación entre dos coeficientes DCT con la misma posición de coordenada de dos bloques no será cambiada después del proceso de cuantificación. La única excepción es que una estricta desigualdad puede favorecer igualdades simples dentro de la cuantificación.

El sistema de autenticación propuesto por Lin y Chang esta basado en esas dos propiedades, la primera es usada para insertar la firma y la otra es usada para generar los bits de autenticación. El proceso de inserción consiste en definir una relación de igualdad entre el LSB de los coeficientes

DCT preseleccionados y los bits de la firma. El proceso de autenticación consiste primero en extraer los bits de autenticación de las áreas marcadas de la imagen, posteriormente usamos estos para verificar si la relación de los coeficientes DCT en la firma se ajustan al criterio establecido. Si no es así esto significa que cualquiera de los bloques o posiblemente los dos bloques del par considerado fueron manipulados.

Los autores han propuesto algunas mejoras como la recuperación de bits, en donde la sobrecarga de bits es mayor. Por un lado ellos permiten una aproximación del bloque de la imagen original a ser reconstruido y por el otro ayudan a localizar las zonas precisas en donde las imágenes fueron desvanecidas (es decir, eliminan la ambigüedad de la identificación de los bloques alterados). La recuperación de bits es generada de un sub-muestreo y una versión comprimida de la imagen. Ellos insertan dentro de 4 bloques, el proceso de inserción de bits de recuperación es similar al de los bits de autenticación. La principal desventaja es que la tasa de compresión JPEG a la cual es robusto el sistema es aún baja.

En [Zhang, Huang, 2008] proponen un algoritmo de marca de agua robusto a transformaciones geométricas, utilizan la transformación de Radon para detectar e invertir las transformaciones geométricas y proponen un detector de características perceptual hash para extraer los puntos característicos más significativos de la imagen, finalmente comparan los puntos característicos de la imagen original con los de la imagen marcada generando así un valor de distancia Hausdorff modificado, el cual determina si la imagen es auténtica o no. Los resultados experimentales muestran que el algoritmo es robusto a ataques como compresión JPEG, operaciones comunes de procesamiento de imágenes y a transformaciones geométricas; también muestran que el sistema es sensible a cambios en el contenido de la imagen. Una desventaja a este algoritmo es que requiere la imagen original para poder realizar el proceso de autenticación.

2.7.2.2. MARCA DE AGUA BASADA EN BLOQUES

Las técnicas de marca de agua basadas en bloques consisten en dividir la imagen en bloques de n*n pixeles para posteriormente insertar una marca de agua "robusta" dentro de cada bloque. Para corroborar la integridad de una imagen, el autenticador prueba la presencia o ausencia de la marca en todos los bloques, si la marca de agua esta presente con una alta probabilidad en cada bloque, podemos afirmar que la imagen probada es auténtica.

En la técnica de marca de agua variable de dos dimensiones (VW2D) descrita por Wolfgang y Delp [Wolfgang, 1996] [Wolfgang, 1999] una marca de agua binaria *W(b)* es insertada en cada bloque *b* de una imagen *X*. Van Schyndel [Van, 1994], recomienda el uso de secuencias-m [Proakis, 1995] para generar la marca de agua, el uso de secuencias-m esta justificado por el hecho de que tienen excelentes propiedades de auto-correlación, tan buenas como una buena robustez en la adición de ruido. Para generar la marca de agua, se mapea una secuencia binaria de {0,1} a {-1, 1}, ordenada dentro de un bloque apropiado y después adicionado a los valores de los pixeles de la imagen. El proceso de verificación usado para probar si la imagen es auténtica consiste en calcular un puntaje estadístico *d* basado en una función de correlación cruzada espacial. Si *d* es menor a un valor de umbral definido por el usuario, el bloque es considerado como auténtico. Esta técnica es robusta a factores de calidad mayores a 85 para compresión JPEG y en sus resultados menciona que detecta correctamente las regiones alteradas en la imagen pero no las muestra.

Fridrich [Fridrich, 1998a] [Fridrich, 1998b] propone una técnica similar para prevenir movimientos no autorizados o distorsiones intencionales en la marca de agua, el autor recomienda hacer una marca de agua dependiente

de la imagen en la cual se va a insertar. La marca de agua binaria corresponde a una señal pseudo aleatoria generada por una llave secreta. Cada bloque es marcado usando una técnica de esparcimiento de espectro de Ó Ruanaidh [Ó Ruanaidh, 1997]. Los autores afirman que la marca de agua es bastante robusta con respecto a la adición de ruido y compresión JPEG moderada, además de que distingue entre manipulaciones maliciosas y cambios no maliciosos que son operaciones comunes de procesamiento de imagen.

El sistema propuesto por Zhou, et al. extrae una firma digital y utiliza un codificador de control de error BCH (7,4,1) [Zhou, et al., 2004]. El sistema detecta los bloques alterados erróneamente después de una compresión con factor de calidad menor a 85 y de una contaminación por ruido impulsivo con densidad mayor de 0.001, ruido gaussiano con una varianza mayor a 0.0001.

2.7.2.3. MARCA DE AGUA BASADA EN CARACTERÍSTICAS

La idea principal de este método [Dugelay, 1999] [Rey, 2000] consiste en extraer primero características de la imagen original e insertarla en ella en forma de marca de agua robusta e invisible, posteriormente para corroborar si una imagen ha sido alterada, simplemente comparamos sus características con la marca de agua que fue recobrada de la imagen original, si las características son idénticas, esto significa que la imagen no fue alterada, de otra forma las diferencias mostraran las áreas alteradas.

La selección de las características usadas de la imagen afectan directamente el tipo de alteraciones en la imagen que queramos detectar, adicionalmente estas características dependen del tipo de imagen en consideración (pinturas, imágenes satelitales, imágenes medicas, etc.). Las características seleccionadas son típicamente propiedades invariantes que se mantienen

ante alteraciones débiles en la imagen (compresión con pérdida) y débiles en manipulaciones maliciosas. Estas características pueden ser usadas para restaurar parcialmente las regiones alteradas de la imagen. Típicamente las características usadas para un sistema de autenticación son bordes, colores, gradiente, luminancia o combinaciones de estas características.

Este tipo de métodos imponen un número de restricciones, principalmente en términos de robustez y capacidad de almacenamiento de la firma, la robustez es requerida en orden en que la marca de agua permita pérdida. La exactitud de la detección y la cantidad de información insertada dentro de la imagen esta directamente relacionada, es necesario encontrar un buen compromiso para el tamaño de la firma para conseguirla robustez y exactitud en la detección de modificaciones.

Uno de los problemas que enfrenta este método es que la imagen es ligeramente modificada mientras se realiza la inserción de la marca de agua, es por eso que las características de la imagen y la marca de agua no son exactamente iguales, y existe un riesgo de una detección falsa positiva, el riesgo puede ser más o menos importante de acuerdo a la selección de las características.

En [Hu, Chen, 2007] se propone un algoritmo de marca de agua semi- frágil, el cual extrae características de la imagen de los componentes del dominio SVD (Singular Value Decomposition) para generar la marca de agua. En este método el proceso de generación e inserción de la marca de agua se lleva a cabo en la misma imagen y la autenticación e la imagen recibida no necesita información de la imagen o la marca de agua original, lo cual incrementa la seguridad de la marca de agua y previene la pérdida de la misma. De acuerdo a los resultados el sistema muestra robustez ante la compresión JPEG con factor de calidad mayor a 50, una densidad de ruido impulsivo de 0.001 para un factor σ de 256 y una varianza de ruido gaussiano de 0.0001

para el mismo factor σ. Una desventaja del sistema es que la detección de áreas alteradas no es tan precisa y para el proceso de detección se requiere un archivo de firma digital adicional.

En [Xie, et al., 2007] proponen un sistema de marca de agua semi-fragil para la autenticación de imágenes basado en la técnica propuesta en [Kundur, Hatzinakos, 1999]. En el nuevo esquema, la marca de agua es generada por una función de cuantificación en el dominio wavelet discreto de la imagen y se inserta cuantizando los coeficientes wavelet correspondientes. La imagen marcada es sin pérdidas ya que no presenta distorsiones evidentes con respecto a la imagen original. Esta técnica tiene la característica de que se puede insertar la marca de agua inclusive sin implementar la DWT inversa. Sus resultados muestran sensibilidad del algoritmo ante modificaciones en el contenido de la imagen pero no muestran la robustez ante ataques no intencionales.

En [Feng, Liu, 2008] proponen un sistema de autenticación de imágenes no supervisado basado en regiones y niveles llamado Abstracción de Contenido Estructural Bayesiano (BaSCA) por sus siglas en ingles, el cual es capaz de tolerar un ancho y dinámico rango de operaciones que no cambian el contenido de la imagen y es sensible a detectar operaciones reales que cambian el contenido de la imagen. Ellos modelaron el contenido estructural usando la red-estructurada de Markov Pixon de campo aleatorio, del cual derivamos el tamaño de la firma BaSCA. De acuerdo a los resultados muestran que el sistema propuesto es mejor que el propuesto en [Hu, Liao, 2003]. La principal desventaja de [Feng, Liu, 2008] y [Hu, Liao, 2003] es que la firma extraída de la imagen se debe almacenar en un archivo adicional a la imagen, requiriendo así un ancho de banda adicional para la transmisión de este.

2.7.3. MÉTODOS DE AUTENTICACIÓN Y RECUPERACIÓN DE IMÁGENES

Muchos esquemas de autenticación de imágenes solamente determinan la integridad del contenido de la imagen detectando las modificaciones que sufren ciertas regiones de la imagen, en algunas aplicaciones no basta solo con esa función, ya que se requiere conocer la imagen original. Actualmente en la literatura existen pocos sistemas que son capaces de recuperar la imagen original a partir de la imagen modificada [Zhao, et al., 2007] [Hassan, 2008] [Ho, et al., 2007] [Ho, 2007] [Lin, et al., 2004a] [Lin, et al., 2004b] [Tsai, Hu, 2005].

En los sistemas de autenticación y recuperación propuestos por [Lin, et al., 2004a] y [Lin, et al., 2004b], la imagen es dividida en subbloques de 8x8 y 4x4 pixeles respectivamente, el propietario de la imagen propone una llave secreta para generar una lista de mapeo entre los subbloques y así generar una imagen indistinguible basada en el método de mezcla caótica. En [Lin, et al., 2004a] los bits que componen a la secuencia de marca de agua son la salida de la función hash de los datos de los bloques de la imagen y un bit CRC (Cyclic Redundancy Check), mientras que en [Lin, et al., 2004b], los bits que componen a la secuencia de marca de agua son una versión reducida de un bloque de la imagen, un bit de autenticación y un bit de paridad. La principal desventaja de ambos esquemas es que la secuencia de marca de agua se inserta en los dos bits menos significativos del bloque correspondiente de la imagen, lo que hace que la marca de agua sea frágil, por lo que no es robusta a modificaciones no intencionales, como lo es compresión JPEG, contaminación por ruido, etc.

En [Tsai, Hu, 2005], los autores propusieron un esquema de autenticación capaz de recuperar el área modificada, en el cual la secuencia de marca de agua se genera en el dominio de la frecuencia y en el dominio espacial y se inserta usando una lista codificada SPIHT de los pixeles más significantes.

Este esquema usa su marca de agua espacial para realizar el proceso de recuperación, aunque la robustez ante ataques intencionales y no intencionales no es mostrada por los autores.

En [Ho, et al., 2007] y [Ho, 2007] se proponen dos algoritmos para la autenticación y recuperación automática basados en marcas de agua semi-fragiles, ambos algoritmos usan la transformada Seno Pinned (PST). En [Ho, et al., 2007] los autores se enfocan en la autenticación y recuperación de imágenes relacionadas con la caligrafía china y en [Ho, 2007] se enfocan en imágenes relacionadas con escenas del crimen, durante el proceso de inserción de la marca de agua ambos algoritmos dividen la imagen original en bloques traslapados de 10x10 pixeles, la marca de agua es un patrón binario pseudo-aleatorio de longitud L y es insertada en la banda de frecuencia media de los coeficientes de la transformada seno del campo Pinned de cada bloque. Los autores mencionan que cuando el algoritmo detecta una alteración en la imagen, esta puede ser recuperada usando el método de proyección reportado en [Fridrich, Goljan, 1999] y en [Zhu, et al., 2005], con respecto a la robustez de la marca de agua, los autores no muestra resultados.

En [Zhao, et al., 2007] se propone un algoritmo de marca de agua digital semi-fragil basado en la transformada Slant (SLT), el cual es capaz de detectar modificaciones y recuperar la imagen original. La imagen se divide en bloques de 8x8 y los bits de la marca de agua se conforman de una secuencia pseudo-aleatoria generada por una llave secreta y posteriormente se inserta en la banda de frecuencia media de los coeficientes de la SLT. Después de que se inserta la marca de agua, los coeficientes de frecuencia de cada bloque de la imagen marcada son extraídos usando la transformada Slant Inversa (ISLT). Para realizar el proceso de recuperación, la imagen original se divide en sub bloques, los cuales son transformados con la SLT, comprimidos y almacenados descartando los coeficientes de alta frecuencia.

Los bits de la imagen comprimida reemplazan a los LSB de la imagen marcada, lo que hace que sea una marca de agua frágil, por lo tanto no es robusta a ataques no intencionales.

En [Hassan, 2008] se propone un esquema de marca de agua hibrida basada en bloques, el cual utiliza una esquema de marca de agua robusto para la auto-corrección de la imagen original y un esquema de marca de agua frágil para una autenticación sensible. La imagen original se divide en bloques de 8x8 pixeles, antes de insertar la marca de agua frágil en el LSB de cada bloque, estos son divididos en 4 sub bloques para generar así la imagen original aproximada, esto se logra conservando los coeficientes DC de los sub bloques. Estos valores son escalados e insertados aleatoriamente en 4 localidades de los coeficientes de frecuencia media de un bloque seleccionado aleatoriamente. Para verificar la autenticidad de una imagen marcada, primero se prueba la marca de agua frágil que esta almacenada en los LSB de la misma, una vez que se determina que bloque fue alterado para reconstruir una aproximación de este, los 4 DC´s son redondeados, saturados e interpolados usando una interpolación bilinear/cúbica/spline de dos dimensiones para reproducir un bloque de 4x4. Este esquema puede detectar y recuperar la imagen original después de una modificación intencional pero es frágil ante ataques no intencionales.

2.8. ALGORITMOS COMPARATIVOS

En esta sección describiremos los algoritmos propuestos Zhou et al. en [Zhou et al., 2004], Xie et al. en [Xie et al., 2007] los cuales son comparados posteriormente con el algoritmo AFDMA, tambien se describen los algoritmos propuestos por Zhao et al. en [Zhao et al., 2007] y Hassan et al. en [Hassan et al., 2008], los cuales son comparados con el algoritmo AAMA.

2.8.1. ALGORITMO PROPUESTO POR ZHOU [ZHOU ET AL., 2004]

El autor propone usar una imagen en escala de grises de tamaño N*N pixeles como imagen original. El dominio utilizado para la extracción de la firma en el método propuesto es el dominio espacial y el método usado para insertar la marca de agua es el dominio de la frecuencia para el cual utilizan la transformada wavelet.

Para generar la firma digital primero, la imagen original es dividida en bloques no traslapados de 16*16. Asumiendo que existen n bloques después de la división y que k bits de la firma son extraídos de cada bloque, entonces puede definirse a un vector b_i de longitud k como la firma del bloque i; así el tamaño total de la firma es $k*n$ bits, posteriormente se calcula el nivel promedio de gris de cada bloque, el cual se codifica para generar la firma digital que posteriormente será insertada. Debido a que al insertar la marca de agua dentro de la imagen los valores promedio de los bloques cambian surge el problema de que la firma extraída no sea la misma que la firma original insertada aún cuando esta no haya sido ilegalmente manipulada; por esta razón se hace necesario el uso de un código corrector de error para solventar este problema. Los autores utilizan el código BCH (7,4,1).

El método propuesto para la inserción de la firma digital se resume en la figura 2.6 y el sistema de autenticación propuesto por los autores se muestra en la figura 2.7.

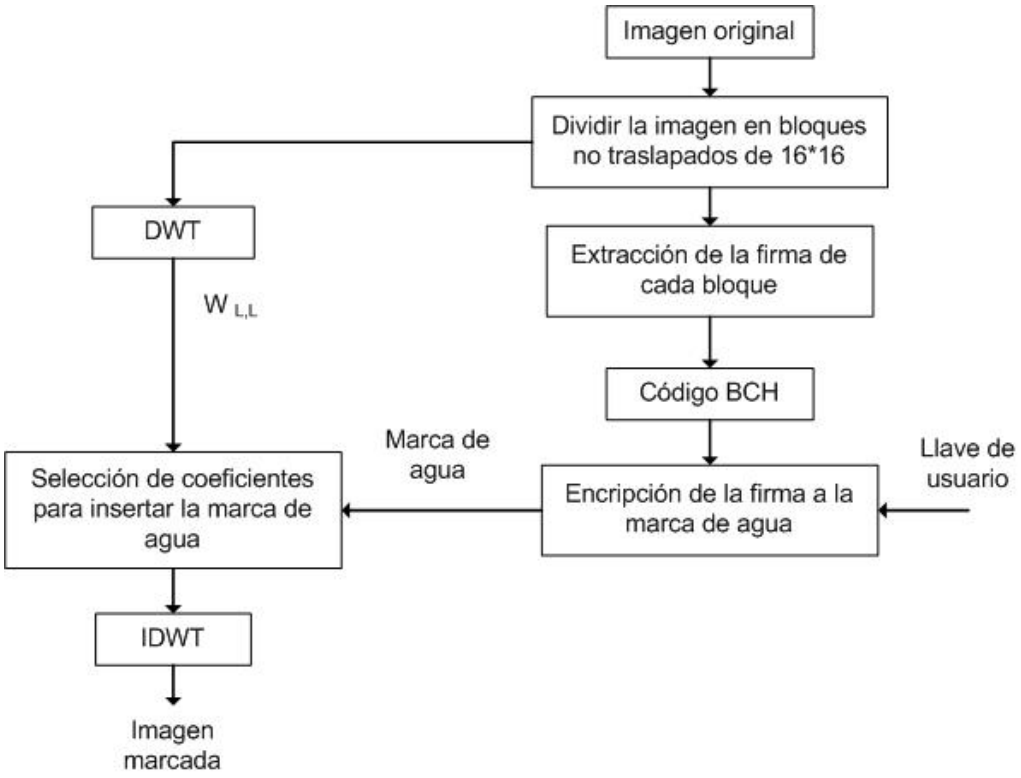


Figura 2.6. Sistema de inserción de la marca de agua por [Zhou et al., 2004]

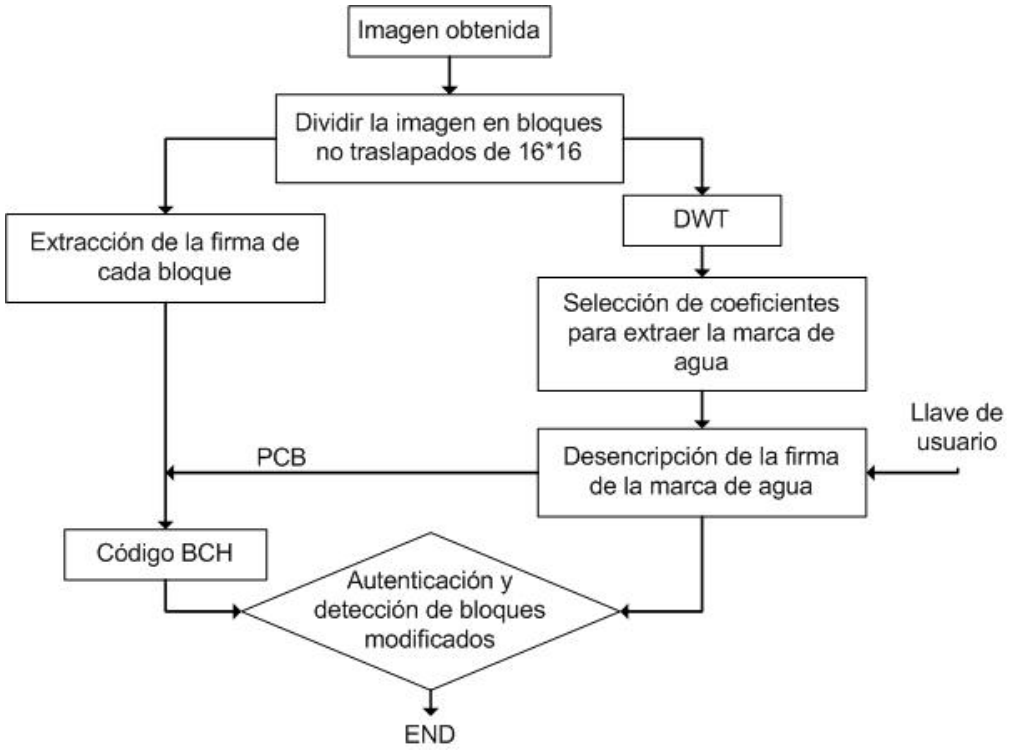


Figura 2.7. Sistema de autenticación por [Zhou et al., 2004]

2.8.2. Algoritmo propuesto por Xie [Xie et al., 2007]

Xie et al. proponen un método de marca de agua semi-frágil para autenticación de imágenes. Fundamentan su algoritmo con el hecho de que cuando se inserta una marca de agua en los coeficientes wavelet o cuando se aplica la transformada wavelet inversa se genera un error numérico, el cual podría interpretarse como una alteración en la imagen original.

Generalmente una marca de agua binaria se genera usando un ID proporcionado por el usuario o por una secuencia de bits aleatoria, la cual al insertarla en los coeficientes wavelet se generaría un cambio en los mismos; los autores proponen generar una marca de agua que no haga ningún cambio en los coeficientes. Es por eso que simplemente generan la marca de agua mediante una función de cuantificación en el dominio de la transformada wavelet discreta (DWT) *Q(f)* para un Δ dado expresado en la ecuación 2.11.

$$Q(f)=\begin{cases}0 & r\le f<(r+1)\Delta, \quad r=0\pm 2\pm 4,...\\ 1 & r\le f<(r+1)\Delta, \quad r=1\pm 3\pm 5,...\end{cases} \tag{2.11}$$

Con lo que insertan una marca de agua dentro de los coeficientes wavelet $f(m.n)$ sin aplicar la DWT inversa ya que no hicieron nada, excepto generar la marca de agua. La imagen marcada es exactamente a la imagen original siempre que esta no sea alterada. De hecho podemos decir que la marca de agua generada por Q(f) es insertada, dado que se puede extraer si la imagen marcada no ha sido alterada.

Posteriormente los autores definen una secuencia de parámetros de cuantificación $\{\Delta_1, \Delta_2, ... \Delta_n\}$, la cual satisface que $\Delta_1 < \Delta_2 < .... \Delta_n$, entonces se pueden generar simultáneamente múltiples marcas de agua en la misma imagen mediante la definición de varios parámetros de cuantificación sin

distorsionar la imagen original, lo cual provee más información para la autenticación de la misma, lo cual mejora la habilidad para autenticar la imagen.

2.8.3. ALGORITMO PROPUESTO POR ZHAO [ZHAO ET AL., 2007]

Zhao et al. proponen un método de marca de agua digital semi-frágil basado en la transformada Slant (SLT) para la autenticación de imágenes y auto corrección. Parten de una imagen original de la cual se toman los primeros 7 bits y después se sub-divide en bloques de 8*8. Los bits de la marca de agua se insertan en la región de frecuencia media de cada bloque después de aplicar la SLT en la imagen original. La imagen original es comprimida y posteriormente insertada en los LSB's de la imagen marcada, lo cual servirá para su posterior auto corrección. Con respecto a la autenticación, las regiones alteradas de la imagen marcada pueden ser detectadas y localizadas mediante la comparación de la marca de agua extraída y la marca de agua insertada. Una vez localizadas las regiones alteradas estas son auto corregidas mediante la extracción de los LSB's de la imagen marcada. La figura 2.8 muestra un diagrama a bloques del sistema de inserción de la marca de agua propuesto por los autores.

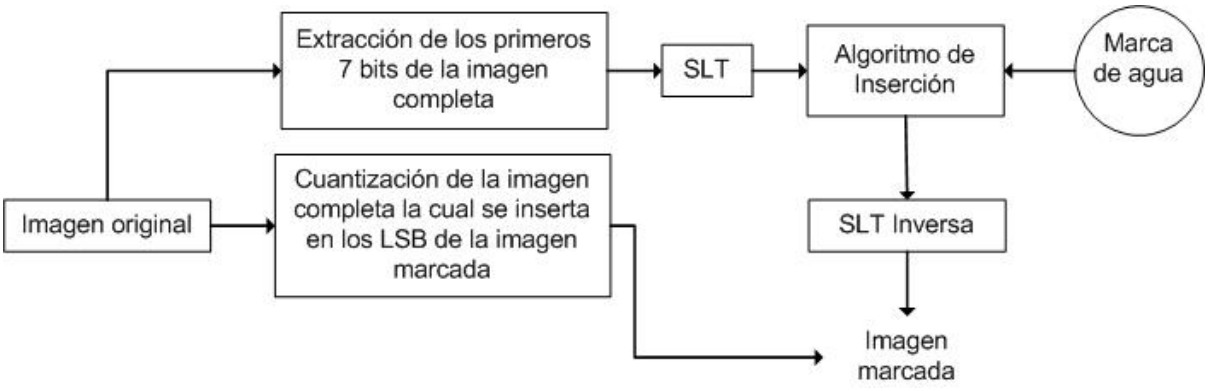


Figura 2.8. Inserción de la marca de agua por [Zhao et al., 2007]

La figura 2.9 muestra un diagrama a bloques del algoritmo de autenticación y recuperación de la imagen marcada propuesto por los autores.

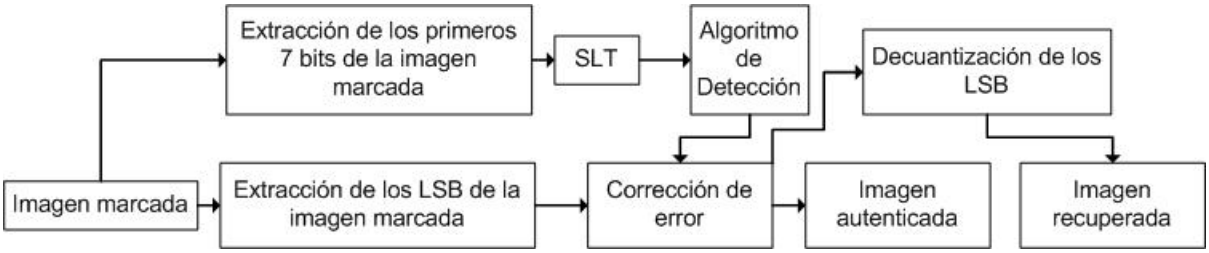


Figura 2.9. Autenticación y recuperación de la imagen marcada por [Zhao et al., 2007]

2.8.4. ALGORITMO PROPUESTO POR HASSAN [HASSAN ET AL., 2008]

Hassan et al. propusieron un método de autenticación hibrido de imágenes a ciegas basado en bloques con auto inserción de una marca de agua robusta para auto corrección y una marca de agua frágil capaz de detectar alteraciones locales.

Partiendo de una imagen original X de tamaño N*M la cual es dividida en bloques de n*n obteniéndose M_b bloques por fila y N_b bloques por columna. Propusieron insertar una versión aproximada de X dentro de la misma imagen, la cual se obtiene excluyendo los LSB's en cada bloque de n*n (n=8) el cual a su vez se subdivide en 4 sub-bloques y la versión aproximada del bloque se obtiene almacenando únicamente los primeros n_a=1 coeficientes de la DCT ordenados en zig-zag para cada sub-bloque. Estos valores son escalados por un factor *sf* y aleatoriamente insertados en 4 localidades de los coeficientes de frecuencia media de la DCT. Las figuras 2.10 y 2.11 muestran los procesos de inserción y detección de la marca de agua en los bloques de la imagen, así como la recuperación del bloque original usando solamente la imagen marcada.

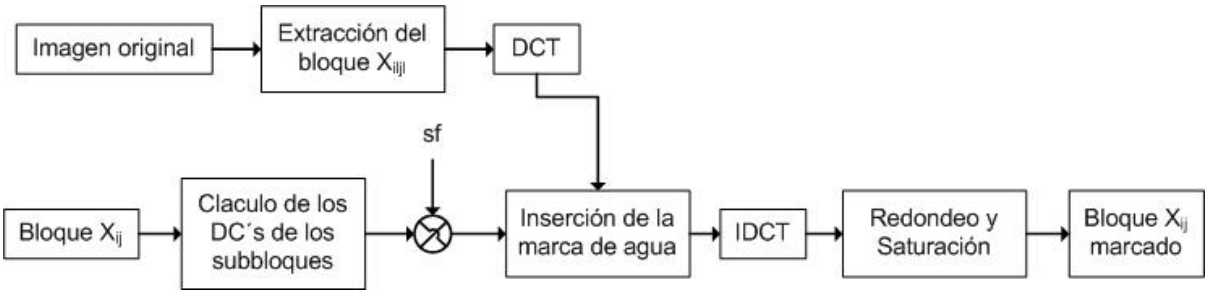


Figura 2.10. Inserción de la marca de agua en un bloque por [Hassan et al., 2008]

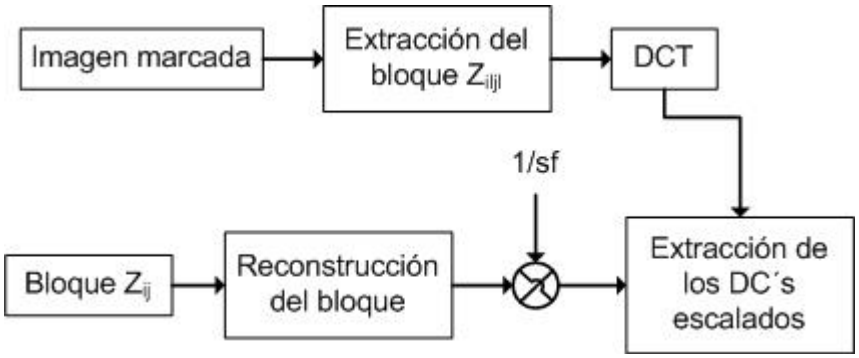


Figura 2.11. Detección y recuperación del bloque original por [Hassan et al., 2008]

2.9. CONCLUSIONES

En este capitulo se investigaron los conceptos relacionados con las marcas de agua, de los cuales se determino que el algoritmo propuesto en esta tesis debe diseñarse con una marca de agua invisible para que no se altere el contenido de la imagen, la marca de agua debe ser semi-frágil para que sea más robusta a ataques no intencionales, también se determino que la inserción de esta debe ser en el dominio espectral, ya que se puede insertar una mayor cantidad de bits sin degradar la calidad en la imagen.

En este capitulo también se investigaron y analizaron los diferentes métodos de autenticación de imágenes con marcas de agua frágiles y semi-frágiles, destacando que al utilizarse marcas de agua semi-frágiles, estas son más robustas a alteraciones comunes del procesamiento de imágenes como compresión JPEG y adición de ruido. Por otro lado se analizaron y compararon métodos de autenticación de imágenes basados en marca de agua semi-frágil, los cuales se dividen principalmente en marcas de agua basadas en bloques y basadas en características, las primeras tienen la ventaja de que se inserta una marca de agua robusta en cada bloque y el problema que enfrenta la segunda es que la imagen es ligeramente modificada mientras se realiza la inserción de la marca de agua, lo que conlleva a que las características de la imagen y la marca de agua no sean exactamente iguales, existiendo así un riesgo de una detección falsa positiva, este riesgo puede ser más o menos importante de acuerdo a la selección de las características

Finalmente se investigaron y analizaron diferentes métodos de autenticación y recuperación de imágenes, los cuales se implementaron utilizando diferentes transformadas (DCT, SLT y PST), pero desafortunadamente los resultados reportados no mencionan la robustez de los algoritmos ante ataques no intencionales.

CAPÍTULO 3

DESARROLLO DE LOS SISTEMAS PROPUESTOS

3.1. ALGORITMO DE AUTENTICACIÓN CON FIRMA DIGITAL COMO MARCA DE AGUA (AFDMA).

Este algoritmo utiliza una firma digital como marca de agua para determinar los bloques alterados de la imagen marcada y modificada, a este método lo llamaremos Autenticación basado en Firma Digital como Marca de Agua (AFDMA).

El algoritmo AFDMA desarrolla un sistema de marca de agua digital semi-frágil, el cual es robusto a ataques intencionales y no intencionales para la autenticación de una imagen digital. El procedimiento se basa en extraer una firma digital robusta de la imagen huésped y posteriormente insertarla dentro de la misma como marca de agua invisible eliminando así la necesidad de generar un archivo adicional para la transmisión de la firma digital, como generalmente se hace. La inserción de la marca de agua se realiza en el dominio de la transformada discreta wavelet con el objetivo de hacer mas robusta la inserción y que la marca de agua sea menos perceptible, una característica importante de este algoritmo es que no se requiere la imagen original para la extracción de la marca de agua, lo que incrementa la seguridad en el sistema. Una aportación importante a este algoritmo es la propuesta de un proceso de verificación, el cual determina que si hay 3 ó más bloques erróneos consecutivos, la imagen digital ha sido modificada intencionalmente y por consecuencia es necesario considerar la imagen marcada como alterada intencionalmente, en caso contrario se

considera que la imagen marcada ha sufrido una modificación no intencional.

La figura 3.1 (a) muestra un diagrama a bloques general del proceso de generación e inserción de la marca de agua en donde podemos observar que la imagen huésped es dividida en bloques, en donde cada uno de ellos es procesado de manera independiente para extraer su firma digital correspondiente y posteriormente el mismo bloque es sometido al proceso de inserción de la firma, generando así el bloque marcado; el conjunto de todos los bloques marcados conforman la imagen marcada.

La figura 3.1 (b) muestra el diagrama a bloques general del proceso de extracción de la marca de agua de un bloque marcado, en donde podemos observar que requerimos solamente de la imagen marcada para poder extraer la marca de agua y la firma digital, cabe mencionarse que al insertar y extraer la firma digital de cada bloque por separado podemos detectar exactamente que bloque se ha modificado de la imagen marcada. Una vez que contamos con todos los bloques erróneos de la imagen marcada aplicamos el proceso de verificación; el cual nos indicara si la modificación o alteración en la imagen marcada se bebe a una modificación intencional o no.

3.1.1. GENERACIÓN DE LA FIRMA DIGITAL

El algoritmo que genera la firma digital de la imagen huésped lo hace por medio de la extracción de bits en los bloques de la imagen huésped con el propósito de que todos los bloques que sean similares, ya sea que estén marcados, desmarcados o atacados, produzcan casi la misma secuencia de bits de longitud N [Fridrich, 1999a].

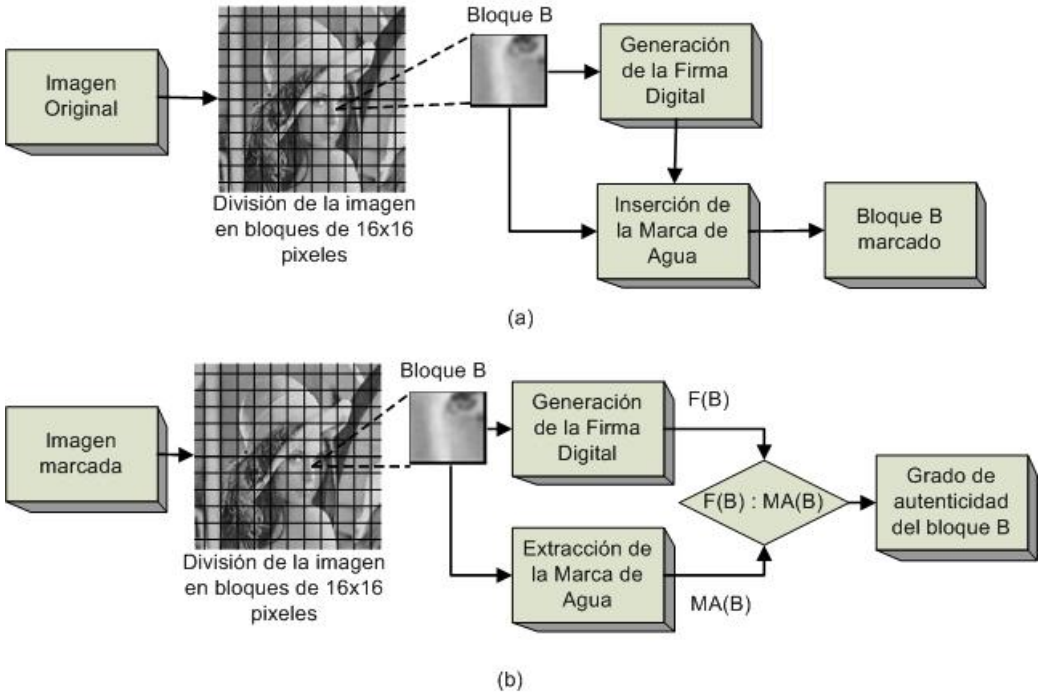


Figura 3.1. (a) Diagrama general de inserción del sistema AFDMA; (b) Diagrama general de extracción del sistema AFDMA

Para llevar a cabo la generación de la firma digital, partimos de una imagen original K de tamaño $X*Y$ y se llevan a cabo los siguientes pasos:

1. La imagen original se divide en bloques de 16*16 píxeles
2. Cada bloque B_k es usado de manera independiente para la extracción de la firma digital.
3. Utilizamos una llave secreta, la cual solamente es conocida por el propietario de la imagen K para generar N matrices aleatorias, las cuales están uniformemente distribuidas en un intervalo de [0,1] donde $1 \leq N \leq 16$.
4. Aplicamos repetidamente un filtro pasa bajas *FPB* a cada matriz para obtener N patrones suavizados.
5. Todos los patrones son hechos DC-libres sustrayendo el promedio de cada patrón.
6. Considerando el bloque y el patrón como vectores, el bloque de la imagen B_k es proyectado con cada patrón P_i, $1 \leq i \leq N$.

7. Su valor absoluto es comparado con un umbral *Th* para obtener *N* bits b_i, representado en la ecuación (3.1).

$$\begin{aligned} si \quad & |B * P_i| < Th \quad b_i = 0 \\ si \quad & |B * P_i| \geq Th \quad b_i = 1 \end{aligned} \quad . \tag{3.1}$$

El proceso descrito anteriormente se muestra gráficamente en la figura 3.2, la cual muestra el proceso de generación de la firma digital.

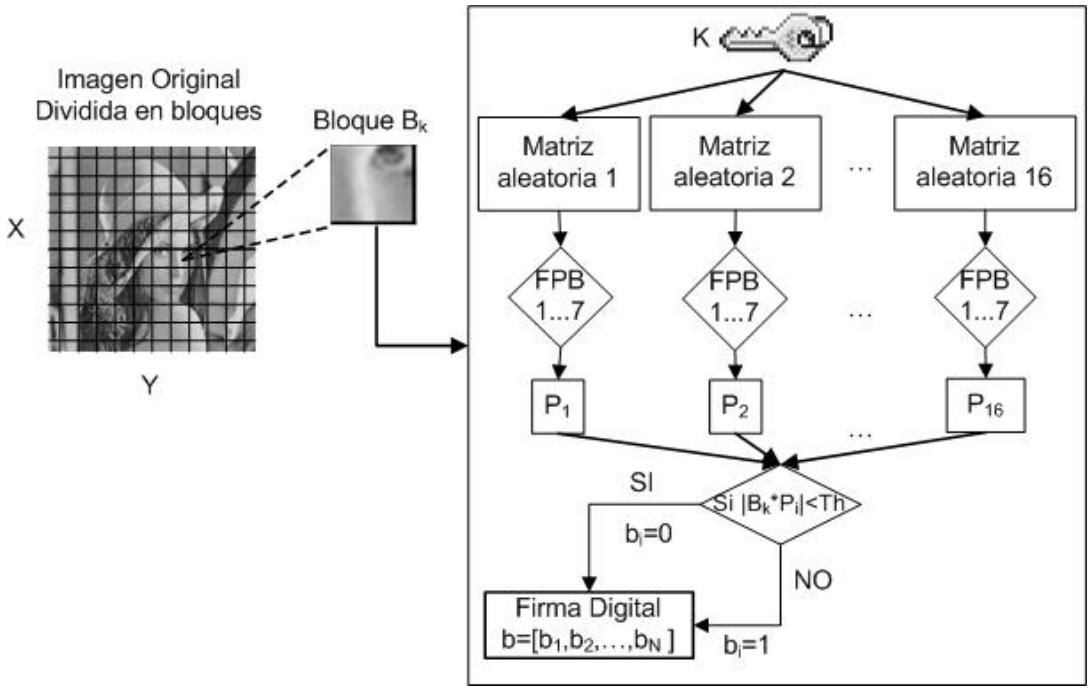


Figura 3.2. Proceso de generación de la firma digital del sistema AFDMA

Ya que los patrones P_i tienen media cero, las proyecciones no dependen de los valores promedio de gris del bloque sino que dependen de las mismas variaciones del bloque. La distribución de las proyecciones dependen de la imagen y se debe ajustar de acuerdo a que aproximadamente la mitad de los bits b_i son ceros y la otra mitad unos, esto garantiza un mayor contenido de información en los *N* bits extraídos. La selección adaptiva del umbral es importante para aquellas operaciones en las imágenes que cambian

significativamente la distribución de proyecciones, como lo es el ajuste de contraste.

3.1.2. INSERCIÓN DE LA MARCA DE AGUA

Una de las características en la inserción de la marca de agua es que esta sea imperceptible al sistema visual humano y robusto ante ataques comunes de procesamiento de imágenes como: compresión JPEG y adición de ruido, es por eso que la marca de agua la insertamos en los componentes de baja frecuencia de la imagen usando un proceso de cuantificación controlado, los datos son extraídos usando solamente el paso de cuantificación y la amplitud promedio de los componentes de baja frecuencia sin tener acceso a la imagen original.

Los pasos para llevar a cabo la inserción de la marca de agua son los siguientes:

1. Partimos del bloque B_k utilizado para la extracción de la firma digital, donde la firma digital funge como la marca de agua binaria y la representamos como $w_k=[b_1, b_2,\ldots,b_N]$ del bloque B_k.
2. El bloque B_k se descompone en sub-bandas usando la transformada wavelet bidimensional (DWT) generando así una sub-banda de información LL_1 y tres sub-bandas de detalles LH_1, HL_1 y HH_1.
3. La sub-banda LL_1 se subdivide en bloques de b_x*b_y, es decir 2*2 pixeles B_{kj}.
4. Calculamos el promedio M_j de los coeficientes wavelet de B_{kj}.
5. Dado un paso de cuantificación Q=5 llamado también intensidad de la inserción, se cuantifica el valor promedio de cada bloque por medio de la ecuación (3.2).

$$q = round\left[\frac{M_j}{Q}\right]. \tag{3.2}$$

6. Como siguiente paso calculamos la diferencia entre q y $fix\left[M_j/Q\right]$, ecuación (3.3).

$$dif = abs\left(q - fix\left(\frac{M_j}{Q}\right)\right). \tag{3.3}$$

7. Modificamos el valor de cuantificación q de acuerdo al valor de w_j, q y dif de la siguiente manera:

- Si w_j=0 y q es un número impar ó Si w_k=1 y q es un número par entonces obtenemos q' con la ecuación (3.4)

$$q' = \begin{cases} q+1 & para \quad dif = 0 \\ q-1 & para \quad dif = 1 \end{cases}. \tag{3.4}$$

8. Una vez encontrado q' calculamos $M_j'=q'*Q$ y posteriormente adicionamos la diferencia entre M_j' y M_j representada por δM_j a los coeficientes wavelet de B_{kj}. Debido a que los elementos de LL_1 varia por la adición de δM_j debemos calcular y guardar el promedio LM' de LL_1'.
9. Finalmente aplicamos la transformada wavelet inversa para obtener el bloque marcado, la imagen marcada I' la construimos usando todos los bloques marcados.

Podemos representar en forma de diagrama de bloques el proceso de inserción de la marca de agua que fue descrito anteriormente por medio de la figura 3.3.

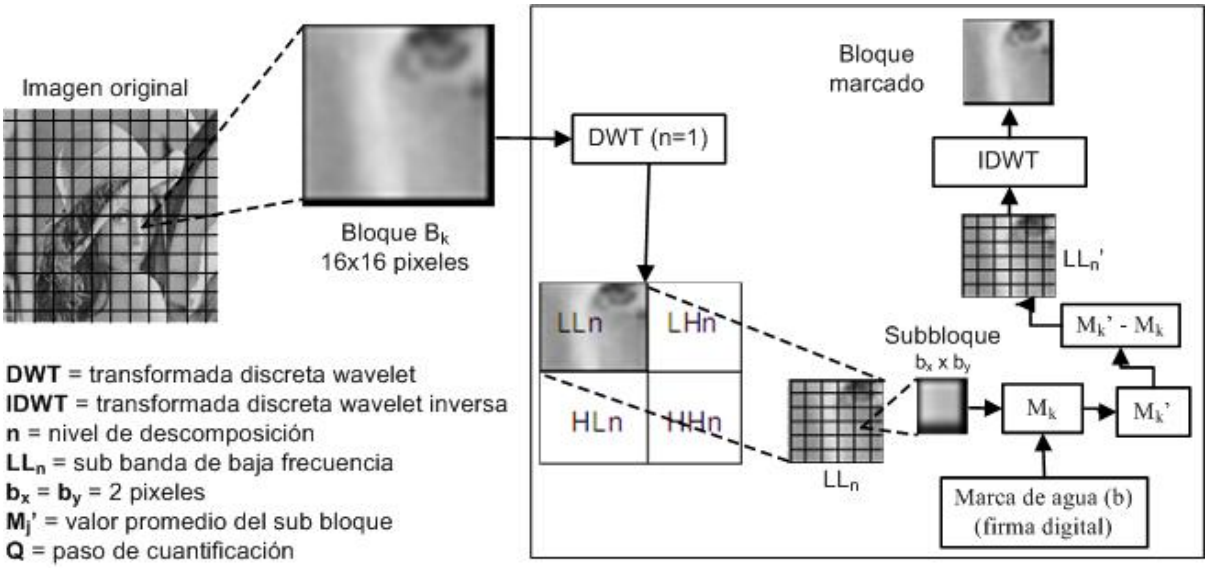


Figura 3.3. Algoritmo de inserción de la marca de agua del sistema AFDMA

3.1.3. EXTRACCIÓN DE LA MARCA DE AGUA

Para extraer la marca de agua, es necesario conocer el nivel de descomposición de la DWT; es decir el valor de *n*, el tamaño de los bloques de sub-banda LL'_n $(b_x * b_y)$, el paso de cuantificación *Q* y *LM'*. El algoritmo de extracción propuesto esta basado en el algoritmo de cuantificación propuesto en [Inoue, et al. 2000], los procesos del algoritmo de extracción de la marca de agua se describen a continuación:

1. Se divide la imagen marcada y distorsionada recibida en bloques de 16*16 pixeles (B'_k).
2. Se descompone B'_k usando la DWT hasta el nivel *n*.
3. La sub-banda LL'_n se dividen en sub-bloques de tamaño $b_x * b_y$ (B'_{kj}).
4. Se calcula el valor promedio M'_j de cada sub-bloque B'_{kj}.

5. Se calcula la diferencia entre M_j' y M_j'' que corresponde a Δm
6. Calculamos el valor de cuantificación S usando la ecuación (3.5)

$$S_j = round\left[\frac{M_k^{''} - \Delta m}{Q}\right]. \tag{3.5}$$

7. Finalmente el bit del j-estimo elemento de marca de agua se extrae por la ecuación (3.6)

$$w_j = \mathrm{mod}(S_j, 2). \tag{3.6}$$

La razón por la que utilizamos el valor promedio de la sub-banda LL'_n es que cuando la imagen no recibe ningún ataque, el valor promedio de sub-banda de baja frecuencia no cambia su valor, sin embargo si la imagen marcada recibe ataques, este valor cambia ligeramente.

3.1.4. AUTENTICACIÓN DE LA IMAGEN RECIBIDA

Una vez extraídas las secuencias de la firma digital ($\tilde{b}_k$) y de la marca de agua ($\tilde{W}_k$) de un mismo bloque de la imagen marcada recibida, se determina un umbral (Th_v) para decidir mediante una operación XOR si el bloque es erróneo o no, expresado en la ecuación (3.7)

$$Si\begin{cases} \sum \tilde{W}_k \otimes \tilde{b}_k < Th_v & el\ bloque\ es\ autentico \\ \sum \tilde{W}_k \otimes \tilde{b}_k \geq Th_v & el\ bloque\ es\ \mathrm{mod}\,ificado \end{cases}. \tag{3.7}$$

El umbral Th_v fue determinado a prueba y error, para lo cual el Th_v seleccionado fue 4; es decir, ya que la longitud de la firma digital es de 16

bits, si al menos 12 de estos bits son correctos se considera que el bloque es auténtico. Pero en algunas ocasiones observamos que se registraban bloques erróneos en regiones que no fueron alteradas; es por eso que propusimos es siguiente proceso de verificación.

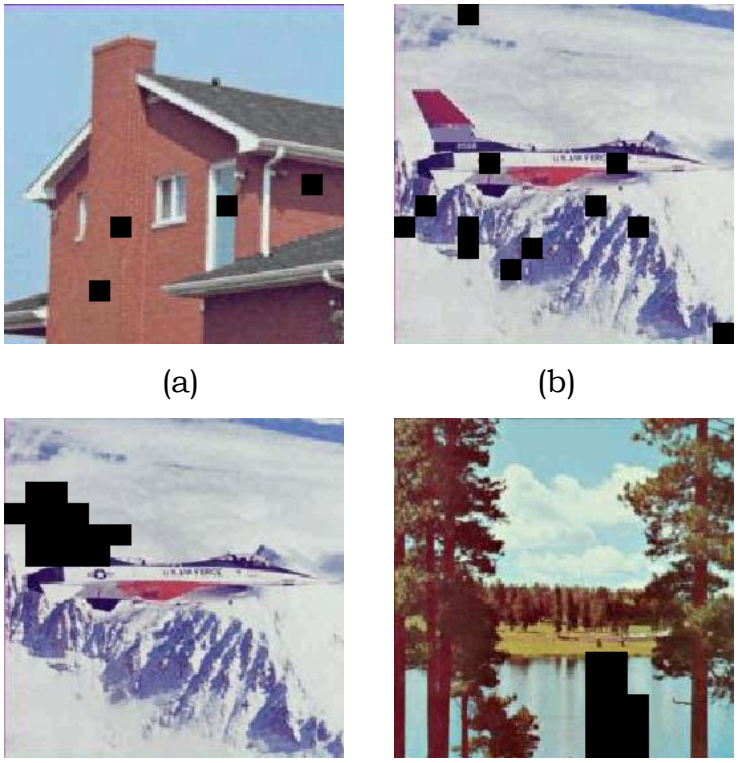
3.1.4.1 PROCESO DE VERIFICACIÓN PROPUESTO

De acuerdo a las pruebas de autenticación realizadas a 200 imágenes, se llego a la siguiente conclusión: cuando se presentan bloques erróneos en regiones no alteradas, estos bloques se presentan de forma aislada; es decir bloques erróneos no concentrados, como se muestra en las figuras 3.4. (a,b). En el caso de imágenes modificadas intencionalmente los bloques erróneos detectados se encuentran en forma concentrada como se muestra en las figuras 3.4. (c,d), por lo tanto cuando se detectan bloques erróneos aislados estamos hablando de un ataque no intencional y en el caso contrario nos referimos a ataques intencionales.

Para establecer un criterio que determine si la modificación ocurrida en un bloque es de tipo intencional o del tipo no intencional observamos de la figura 3.4. que si hay más de tres bloques erróneos consecutivos la región de la imagen fue modificada intencionalmente; en caso contrario la modificación fue no intencional. Aplicando el concepto de conectividad entre los 8 vecinos de los bloques erróneos se pueden determinar las regiones modificadas intencionalmente de las que no lo son. Este criterio se representa matemáticamente con la ecuación (3.8)

$$\text{Re}\,gión\begin{cases} autentica & si \quad \tilde{B} \le 3\, consecutivos \\ no\ autentica & si \quad \tilde{B} > 3\, consecutivos \end{cases}, \tag{3.8}$$

donde $\tilde{B}$ representa un bloque erróneo, por lo tanto si hay más de tres bloques erróneos consecutivos la región ha sido modificada intencionalmente.



(a) (b)

(c) (d)

Figura 3.4. (a,b) Presentación de bloques erróneos aislados en regiones no alteradas; (c,d) bloques erróneos concentrados en regiones alteradas

3.2. ALGORITMO DE AUTENTICACIÓN Y AUTO-RECUPERACIÓN CON MARCA DE AGUA (AAMA).

Este algoritmo de autenticación de imágenes digitales es capaz de detectar y recuperar los bloques de la imagen original a partir de la imagen marcada y alterada. A este método lo llamaremos en lo subsecuente como: Autenticación y Auto-Recuperación basado en Marca de Agua (AAMA).

El algoritmo AAMA desarrolla un sistema de autenticación con marca de agua digital semi-frágil basado en bloques, el cual, además de detectar correctamente las regiones alteradas de la imagen marcada es capaz de recuperar la información de la imagen original de ciertas regiones de interés seleccionadas previamente por el propietario. El algoritmo propuesto es robusto a ataques no intencionales propios del procesamiento de imágenes como lo es la compresión JPEG y la adición de ruido. El procedimiento se basa dividir la imagen original en dos regiones: región de interés (ROI) y región de inserción (ROE), los bloques ROI son seleccionados manualmente por el propietario de la imagen y los bloques ROE son todos aquellos bloques restantes de la imagen diferentes a los bloques ROI, en los cuales inserta la marca de agua. La marca de agua es una secuencia binaria generada de cada bloque ROI, compuesta por el coeficiente DC y los 6 primeros coeficientes AC de baja frecuencia ordenados en zig-zag de los coeficientes DCT, posteriormente la marca de agua se inserta en los bits LSB de los bloques DCT correspondientes indicados por una lista generada con una llave secreta, la cual solo debe conocer el propietario de la imagen. Para autenticar los bloques ROI se extrae la marca de agua de los de la ROI y la ROE si la diferencia es mayor a un umbral *Th*, entones consideramos que el bloque fue alterado y es reemplazado con el bloque reconstruido por la marca extraída de la región ROE. El sistema propuesto solo requiere la llave secreta para extraer y autenticar la marca de agua sin necesidad de conocer la imagen original, lo que aumenta la seguridad en el sistema. Cabe destacarse que la literatura reporta muy poca investigación en este campo, las técnicas desarrolladas en este algoritmo no se habían propuesto por otro autor.

3.2.1. GENERACIÓN DE LA MARCA DE AGUA

La figura 3.5. ilustra el proceso para generar la marca de agua, el cual se puede describir de la siguiente manera:

1. Se modifican los niveles de gris de la imagen original en el rango [-127 a 128] con el propósito de reducir el valor del coeficiente DC ya que con esto aseguramos que podremos representar a este con un máximo de 8 bits.
2. En muchas aplicaciones solamente algunas regiones u objetos en la imagen tienen mayor importancia para el propietario. Basándonos en este principio este sistema divide la imagen original en 2 regiones para mejorar la capacidad de detección y/o localización de ataques, así como mejorar la calidad en la imagen reconstruida. En lo subsiguiente nos referiremos a las 2 regiones en la imagen como:
    a. Región de Interés (ROI); son regiones importantes en la imagen que requieren una mayor protección contra modificaciones maliciosas, estas regiones son seleccionadas manualmente por el propietario de la imagen

       La programación del algoritmo para generar las regiones ROI es como sigue:

        i. Por cada píxel de la imagen seleccionado manualmente por el propietario de la imagen, automáticamente el sistema un bloque de 24*24 pixeles, el cual esta compuesto por los 8 pixeles vecinos a el y por los 8 bloques vecinos de 8*8, por lo tanto, un bloque ROI puede ser de tamaño 24*24 o mayor y pueden estar concentrados en una región o esparcidos por varias regiones de la imagen, lo cual depende directamente de los requerimientos del propietario, es importante mencionar que si existen bloques ROI duplicados el sistema los elimina automáticamente.

b. Región de Inserción (ROE): esta conformada por el resto de la imagen que no forma parte de la ROI y es en esta región en donde se inserta la marca de agua.

3. La región de interés ROI es dividida en bloques de 8x8 pixeles no traslapados X_{ij}.
4. A cada bloque X_{ij}.
    a. Calculamos la Transformada Coseno Discreto bidimensional.
    b. Representamos los coeficientes DCT en su forma binaria
        i. Representamos el coeficiente DC con 10 bits más 1 bit mas que represente su signo. Para demostrar que 10 bits son suficientes para representar al coeficiente DC aplicamos la DCT a un bloque de 8*8 donde todos sus valores fueron 128 (ya que es el máximo valor que puede tomar un píxel en la imagen), el valor del coeficiente DC fue 1016 cuya representación en binario es 1111111000 (10 bits).
        ii. Usamos 8+1 bits para representar cada uno de los 6 coeficientes AC de baja frecuencia ordenados en zig-zag.
    c. El conjunto de los bits generados en (b) forman la marca de agua W_{ij} cuya longitud es de 66 bits, donde:
        i. 11 bits corresponden al coeficiente DC
        ii. 54 bits corresponden a los 6 coeficientes AC (9 bits * 6 coeficientes).
        iii. 1 bit adicional para poder dividir e insertar la marca de agua en 6 partes iguales.

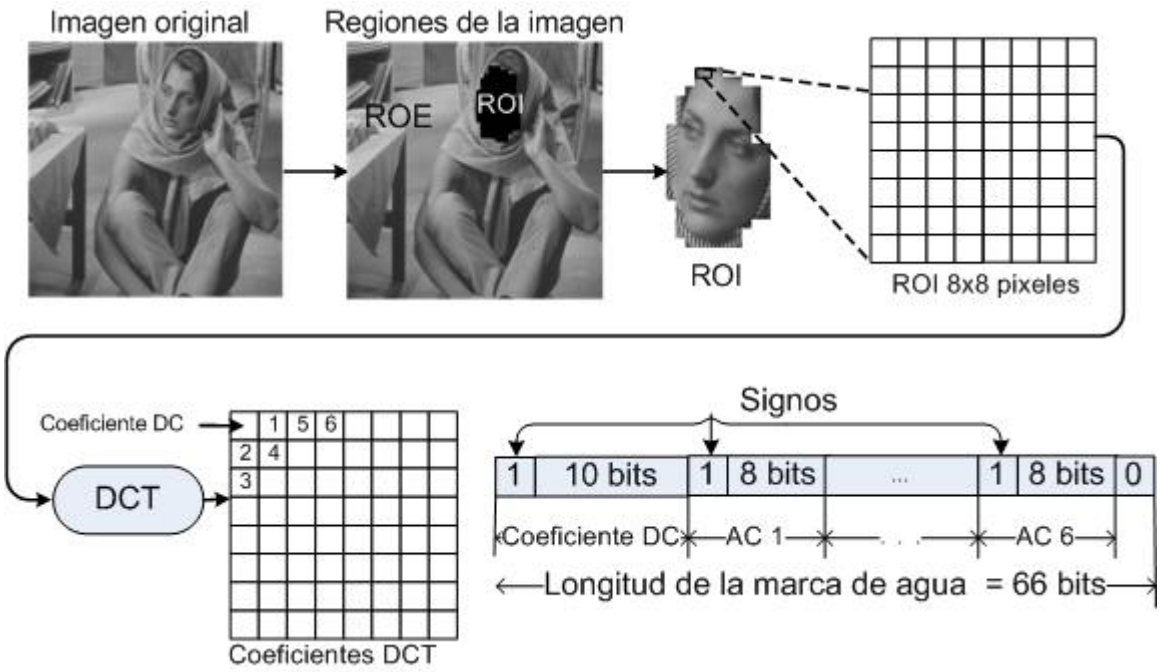


Figura 3.5. Generación de la marca de agua para el algoritmo AAMA

3.2.2. INSERCIÓN DE LA MARCA DE AGUA

El proceso de inserción se realiza en el dominio de la transformada DCT con la finalidad de hacerla robusta ante ataques no intencionales. Los siguientes pasos describen el proceso de inserción:

1. Usando una llave secreta *K*, la cual es propuesta por el propietario de la imagen, se construye una lista de mapeo de los bloques ROE en donde se debe insertar la marca de agua de manera aleatoria.
2. Usando esta lista de mapeo se seleccionan 6 bloques ROE para insertar la marca de agua generada por un bloque ROI X_{ij}.
3. Cada uno de los 6 bloques ROE es:
    a. Transformado usando la DCT bidimensional
    b. Cuantificado usando una matriz de cuantificación Q=70, la cual corresponde a una calidad de compresión JPEG de 70%, nosotros seleccionamos este factor de calidad, ya que probamos con factores de calidad menores para poder obtener una mayor

robustez del sistema ante ataques JPEG pero observamos que ante factores de calidad menores la calidad perceptual de la imagen marcada disminuía. La ecuación (3.9) muestra la cuantificación

$$\tilde{C}(u,v) = \left\lfloor \frac{C(u,v)}{Q(u,v)} \right\rfloor, \tag{3.9}$$

donde $C(u,v)$ y $\tilde{C}(u,v)$ son los (u,v)-esimos coeficientes DCT y los coeficientes DCT transformados respectivamente del bloque, $\lfloor x \rfloor$ es el entero más cercano a cero del valor de x.

c. Cada 11 bits de la marca de agua (W_{ij}/6) es insertado en el LSB de 11 coeficientes DCT de la banda de frecuencia media.
d. Los bloques DCT marcados son multiplicados por Q.
e. Transformados por la DCT Inversa para obtener los bloques marcados.

4. Concatenando todos los bloques marcados, obtenemos la imagen marcada.

La figura 3.6 presenta un diagrama a bloques en donde podemos observar gráficamente el proceso de inserción de la marca de agua en una imagen huésped.

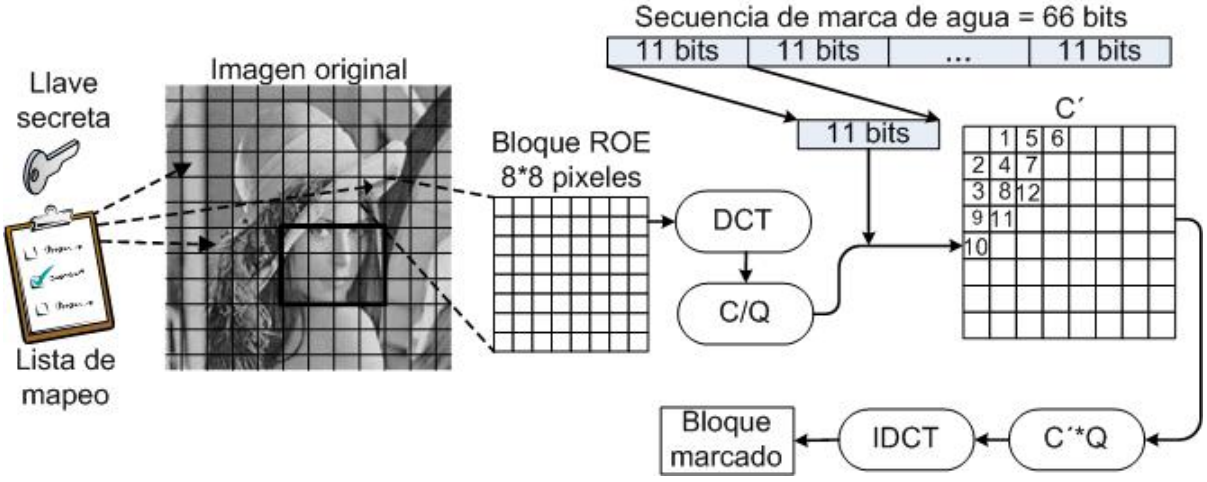


Figura 3.6. Proceso de inserción de la marca de agua para el algoritmo AAMA

3.2.3. EXTRACCIÓN DE LA MARCA DE AGUA

Para la extracción de la marca de agua requerimos de la imagen marcada, la llave secreta del usuario y conocer la región de interés que propuso el usuario

1. Utilizamos la llave secreta *K* para construir la lista de mapeo para conocer cuales son los bloques marcados en donde se inserto la marca de agua.
2. Cada uno de estos bloques es:
    a. Transformado usando la DCT bidimensional
    b. Cuantizados por una matriz de cuantificación Q=70
3. Extraemos la marca de agua del LSB de los primeros 12 coeficientes AC de los coeficientes DCT cuantizados
4. Este proceso se repite para todos los bloques marcados generando así los 67 bits de la secuencia de marca de agua W_{ext},

3.2.4. AUTENTICACIÓN DE LA IMAGEN

El proceso de autenticación determina si el contenido de la imagen de la imagen marcada recibida ha sido modificado. Para autenticar la imagen se deben comparar dos marcas de agua:

1. La primer marca de agua W_{ROIext} es extraída de los bloques ROI como se describió en la sección 3.2.1 (1-4).
2. La segunda marca de agua W_{ROEext} se extrae de los bloques ROE usando la llave secreta *K* usada en el proceso de inserción, con la cual generamos la lista de mapeo que nos indicará en que bloques esta insertada la marca de agua.
3. Cada uno e los 6 bloques ROE seleccionados es:
    a. Transformado usando la DCT bidimensional.

b. Cuantizados por la matriz de cuantificación Q.
c. Se extraen 11 bits del LSB de los 11 coeficientes AC de frecuencia media.

4. Posteriormente por medio de la operación XOR W_{ROIext} y W_{ROEext}, son comparados; usando un umbral *Th* se determinará la autenticidad del bloque, dicha autenticidad esta dada por la ecuación (3.10)

$$\begin{aligned} &Si \sum XOR(W_{ROIext}, W_{ROEext}) < Th \; el \quad bloque \quad es \quad autentico \\ &Si \sum XOR(W_{ROIext}, W_{ROEext}) \geq Th \, el \quad bloque \quad es \quad \mathrm{mod}\, ificado' \end{aligned} \tag{3.10}$$

La figura 3.7. muestra un esquema a bloques del proceso de extracción y verificación de la marca de agua.

3.2.5. AUTO-RECUPERACIÓN DE LA LOS BLOQUES ALTERADOS

Una vez que el algoritmo determina que un bloque ROI fue alterado maliciosamente, la localización del bloque alterado es marcado con un bloque negro, se procede a recuperar la imagen original, el desarrollo de este proceso se lista a continuación:

1. Se elimina el último bit de la marca de agua extraída W_{ROEext}, ya que como se menciono en la sección 3.2.1. este se adiciono para completar los 66 bits de la misma
2. La marca de agua W_{ROEext} (65 bits) es dividida como sigue:
   a. 11 bits para representar el componente DC con su signo.
   b. Los restantes 54 bits son divididos en 9 bits para cada uno de los 6 coeficientes AC con sus respectivos signos.
3. Convertir el coeficiente DC y los 6 coeficientes AC es su forma decimal.
4. Reemplazarlos en su pocisión original (zig-zag) y el resto de los coeficientes son representados por ceros.
5. Calcular la transformada DCT inversa

6. Reemplazar el bloque ROI modificado por el bloque reconstruido.

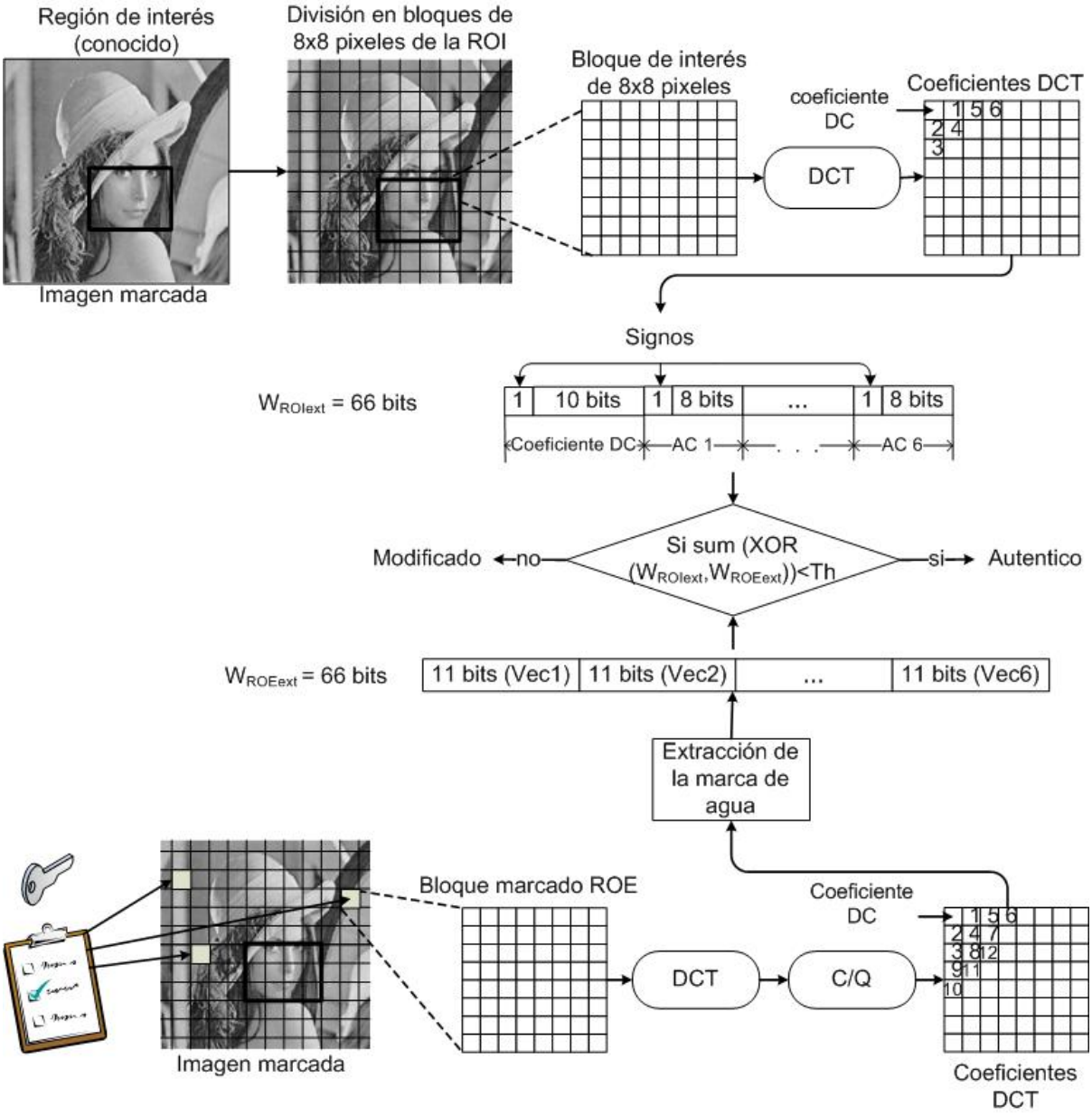


Figura 3.7. Proceso de extracción y verificación de la marca de agua para el algoritmo AAMA

3.3. CONCLUSIONES

Se propusieron dos algoritmos de autenticación de imágenes con marcas de agua y los identificamos como:

1. Algoritmo de Autenticación con Firma Digital como Marca de Agua (AFDMA) y
2. Algoritmo de Autenticación y Auto-Recuperación con Marca de Agua (AAMA).

El algoritmo AFDMA se compone de las siguientes etapas:

- Generación de la firma digital a partir de los bloques de la imagen original, inserción de la marca de agua en la imagen huésped, extracción de la marca de agua de la imagen marcada, autenticación del contenido de la imagen e identificación de ataques intencionales de los no intencionales.

El algoritmo AAMA se compone de las siguientes etapas:

- Generación de la marca de agua a partir de los bloques ROI de la imagen original, inserción de la marca de agua en los bloques ROE de la imagen huésped, extracción de la marca de agua de la imagen marcada, autenticación del contenido de la imagen y auto-recuperación de la imagen original de los bloques atacados.

CAPÍTULO 4

RESULTADOS EXPERIMENTALES Y COMPARACIONES

En este capítulo se muestran los resultados que se obtuvieron en los dos algoritmos propuestos y descritos en el capítulo 3, estos sistemas se probaron utilizando imágenes en escala de gris y a color de tamaños y texturas diversos.

Las pruebas realizadas a los algoritmos están conducidas a tres experimentos principales:

1. Prueba de imperceptibilidad de la marca de agua
2. Autenticación del contenido de la imagen recibida y en su caso auto-recuperación de la imagen original
3. Robustez de la marca de agua ante ataques no intencionales

4.1. RESULTADOS DEL ALGORITMO AFDMA

Para la implementación de este algoritmo se utilizaron imágenes a color y en escala de gris de tamaño 512x512 y 256x256 pixeles como imágenes huésped, cada píxel es representado por 24 bits en el espacio RGB y 8 bits en escala de gris, la marca de agua es una secuencia binaria con una longitud de 16 bits. La programación y las pruebas realizadas a este algoritmo se realizaron en Matlab 7.0.

Inicialmente el algoritmo se probó en imágenes en escala de grises y al observar que este respondía eficientemente, se prosiguió a implementarlo en

imágenes a color ya que actualmente este tipo de imágenes tienen mayor difusión en la Internet con excelentes resultados.

Los valores mostrados en la tabla 4.1 son los utilizados para las pruebas en el algoritmo. Cabe mencionarse que los valores de los parámetros Q y Th_V fueron determinados en base a prueba y error durante la fase de prueba del algoritmo.

Tabla 4.1. Valores de los parámetros usados en el algoritmo AFDMA

Parámetros	*Descripción*	*Valor*
No. de imágenes probadas	256 niveles de gris (8 bits/pixel)	100
	Color (24 bits/pixel)	100
N	Número de bits de la firma digital por bloque	16
Th	Valor de umbral usado en la ecuación (3.1)	adaptivo
(b_x,b_y)	Tamaño de los sub-bloques	(2,2)
Q	Paso de cuantificación usado en la ecuación (3.2)	5
Th_v	Valor de umbral usado en la ecuación(3.8)	4

4.1.1. IMPERCEPTIBILIDAD DE LA MARCA DE AGUA

Ya que uno de los objetivos de los sistemas de marca de agua es que está sea imperceptible al sistema visual humano (SVH) calculamos la relación señal a ruido pico (PSNR) para evaluar la imperceptibilidad, esto lo logramos evaluando la calidad de la imagen marcada con respecto a la de la imagen original de acuerdo a la ecuación (4.1) [Chen, et al., 2001]

$$PSNR_{dB} = 10\log_{10}\frac{255^2}{\delta_q^2}, \tag{4.1}$$

donde δ_q^2 es el error cuadrático medio entre la imagen original y la imagen marcada.

De acuerdo a los resultados obtenidos el valor de PSNR en las imágenes en escala de grises marcadas es en promedio 45 dB´s y el PSNR en las imágenes a color marcadas es de 50 dB´s en promedio, lo que garantiza la imperceptibilidad de la marca de agua ante el sistema visual humano sin importar la textura de la imagen original.

4.1.1.1. LONGITUD DE LA MARCA DE AGUA

La marca de agua utilizada es una secuencia binaria formada de 16 bits por cada bloque de la imagen huésped, por lo tanto, la longitud total de la marca de agua depende directamente del tamaño de la imagen huésped, la tabla 4.2 muestra la longitud de la marca de agua dependiendo del tamaño de la imagen huésped.

Tabla 4.2. Longitud de la secuencia de marca de agua en el algoritmo AFDMA.

Tamaño de la imagen huésped (pixeles)	*Numero de bloques de 16*16 pixeles en la imagen huésped*	*Longitud total de la marca de agua (bits)*
512x512	32x32=1024 bloques	16384
256x256	16x16=256 bloques	4096
128x128	8x8=64 bloques	1024

La figura 4.1 muestra la imperceptibilidad de la marca de agua, es decir, las imágenes marcadas de “Bárbara” “Lena” y “chiles” en realidad no muestran

una diferencia en cuanto al contenido de la imagen original, las figuras 4.1 (c, f e i) son las imágenes de error, es decir, la diferencia multiplicada por 1000 entre la imagen original y la imagen marcada.

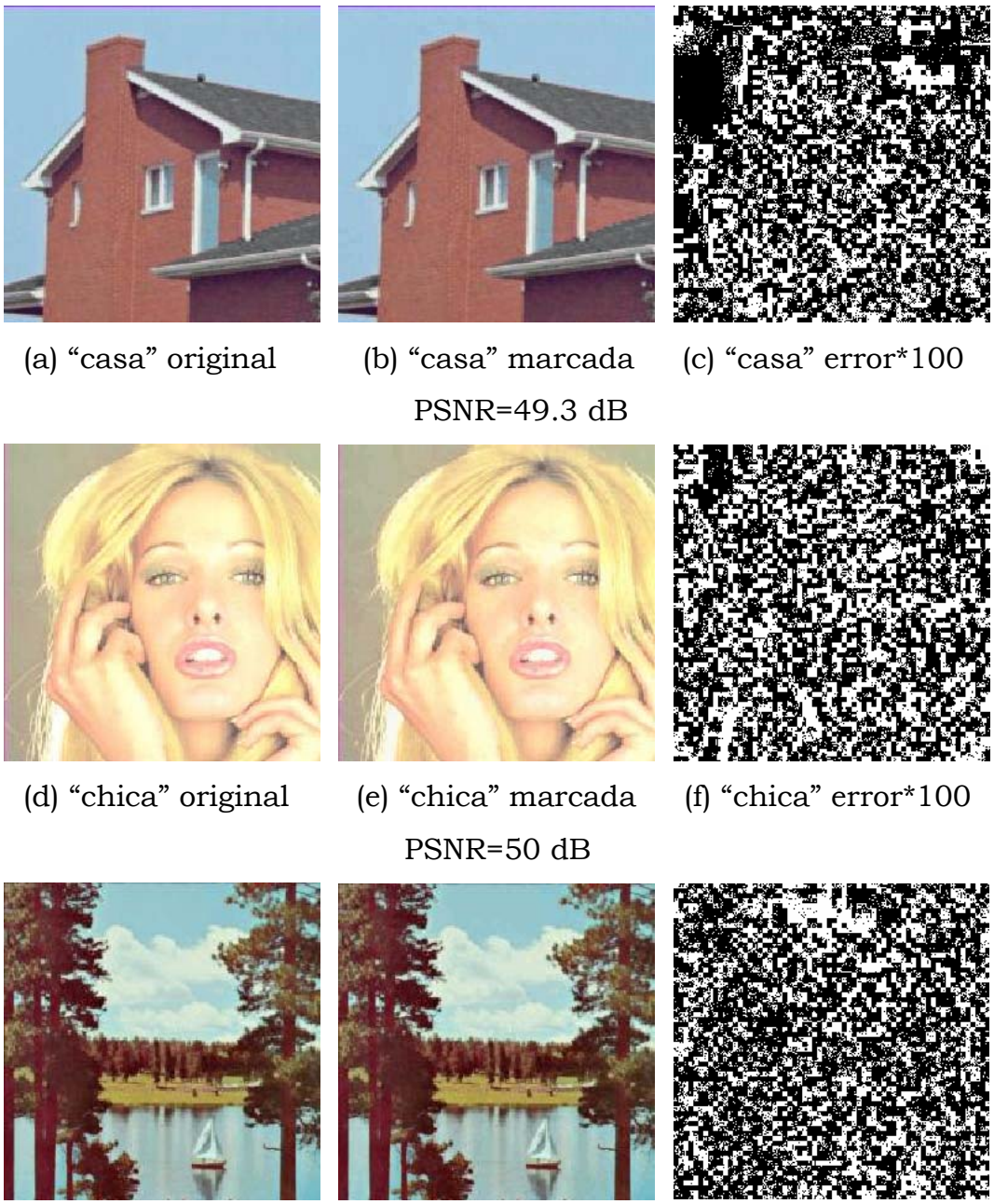
(a) "Bárbara" original (b) "Bárbara" marcada (c) "Bárbara" error*1000 PSNR=45 dB

(d) "Lena" original (e) "Lena" marcada (f) "Lena" error*1000 PSNR=44.5 dB

(g) "chiles" original (h) "chiles" marcada (i) "chiles" error*1000 PSNR=45.3 dB

Figura 4.1. Imperceptibilidad de la marca de agua en imágenes en escala de grises.

La figura 4.2 muestra las imágenes originales a color de “casa”, “chica” y “barco” con sus correspondientes imágenes marcadas, en donde podemos observar que la marca de agua no se percibe en la imagen marcada.



(a) “casa” original (b) “casa” marcada PSNR=49.3 dB (c) “casa” error*100

(d) “chica” original (e) “chica” marcada PSNR=50 dB (f) “chica” error*100

(g) “barco” original (h) “barco” marcada PSNR=49.5 dB (i) “barco” error*100

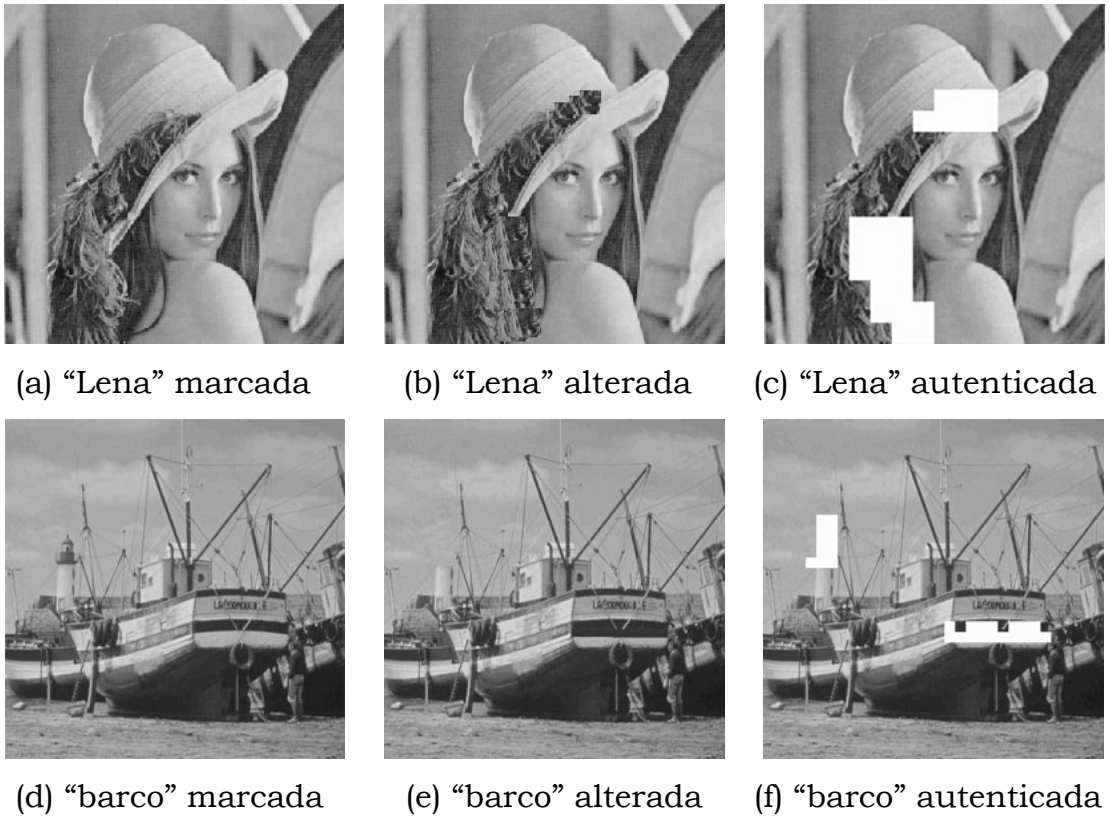
Figura 4.2. Imperceptibilidad de la marca de agua en imágenes a color.

De la figura 4.2 también podemos observar que el valor de PSNR no varia mucho entre las imágenes marcadas, debido a que la longitud de la marca de agua esta en función al tamaño de la imagen. Las figuras de error, es decir la diferencia entre la imagen original marcada multiplicada por 100 son mostradas en las figuras 4.2 (c, f e i) con la finalidad de mostrar la degradación que sufre la imagen original al ser marcada.

4.1.2. CAPACIDAD DE DETECCIÓN Y VERIFICACIÓN DE REGIONES ALTERADAS

En esta sección mostramos que el sistema propuesto es capaz de detectar exactamente los bloques alterados en la imagen marcada, como se menciono en el capítulo 2 las modificaciones que cambian el contenido de la imagen son considerados como modificaciones maliciosas. Para esta prueba se modificaron directamente ciertas regiones de la imagen marcada, las cuales solo sería posible detectarlas si se tuviera la imagen original, el software utilizando para este propósito fue: Photoshop.

La figura 4.3 muestra los resultados del proceso de autenticación; el inciso (a) muestra la imagen "Lena" original, el (b) muestra la imagen "Lena" marcada y alterada; el inciso (c) muestra los bloques alterados de la imagen representados con bloques blancos, con lo cual demostramos la localización exacta de la modificación hecha a la imagen, si esta imagen es sometida al proceso de verificación propuesto y descrito en la sección 3.1.4.1 la imagen resultante será la misma, porque determinará que las alteraciones en la imagen fueron intencionales ya que los bloques erróneos no están aislados. Un proceso similar es aplicado en la figura 4.4 inciso (a) al (d), la diferencia con el caso de la imagen "Lena" es que la imagen autenticada de "chica" (4.4 (d)) presenta bloques erróneos aislados (bloques negros), los cuales son identificados como modificaciones no intencionales por el proceso de verificación; la imagen resultante del proceso de autenticación se muestra en la figura 4.4 (d).



(a) "Lena" marcada (b) "Lena" alterada (c) "Lena" autenticada

(d) "barco" marcada (e) "barco" alterada (f) "barco" autenticada

Figura 4.3. Capacidad de detección y verificación de regiones alteradas en imágenes en escala de grises.

4.1.3. ROBUSTEZ DE LA MARCA DE AGUA A ATAQUES NO INTENCIONALES

4.1.3.1. ROBUSTEZ A LA COMPRESION JPEG

Uno de los procesos más comunes para la transmisión de imágenes digitales es la compresión JPEG, ya que se requiere reducir el tamaño del archivo pero sin que se altere el contenido de la misma, otra razón por la cual se comprime la imagen es que se reduce el ancho de banda necesario para la transmisión de la misma.

(a) “chica” marcada (b) “chica” marcada alterada (c) “chica” autenticada

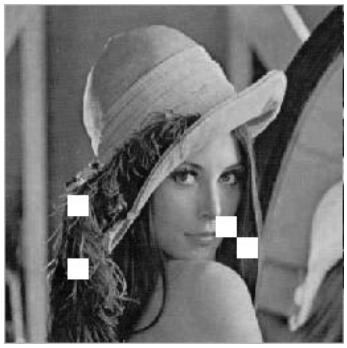
(d) “chica” verificada

(e) “lago” marcada (f) “lago” marcada alterada (g) “lago” autenticada

Figura 4.4. Capacidad de detección y verificación de regiones alteradas en imágenes a color.

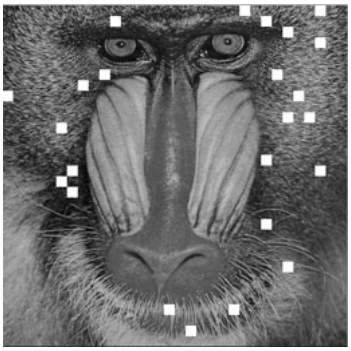
Para comprobar la robustez del sistema ante ataques de compresión sin pérdida JPEG, la cual es considerada como un ataque no intencional, las imágenes marcadas se sometieron a diferentes tasas de compresión.

La figuras 4.5 inciso (a) y (c) muestran los bloques erróneos (cuadros blancos) detectados por el sistema de autenticación después de aplicar compresión JPEG a imágenes marcadas en escala de grises con tasas de compresión de 1.29 bits/píxel y 2.24 bits/pixel respectivamente, las figuras 4.5 inciso (b) y (d) muestran las imágenes resultantes del proceso de verificación, en donde podemos ver que son exactamente iguales a la imagen marcada y comprimida, ya que elimina los bloques erróneos aislados considerándolas como no alterada.

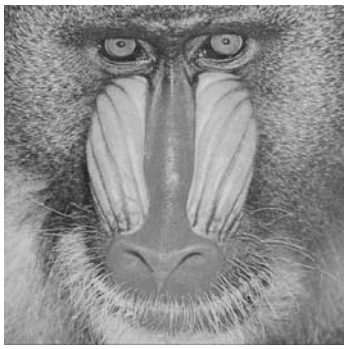


(a) "Lena" autenticada JPEG Q=75

(b) "Lena" verificada JPEG Q=75



(c) "mandril" autenticada JPEG Q=80

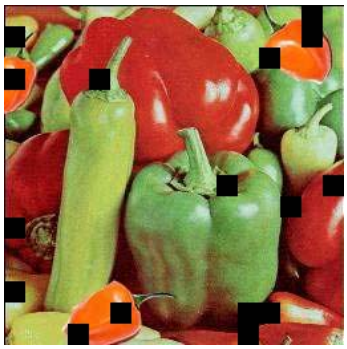


(d) "mandril" verificada JPEG Q=80

Figura 4.5. Robustez ante compresión JPEG en imágenes en escala de grises

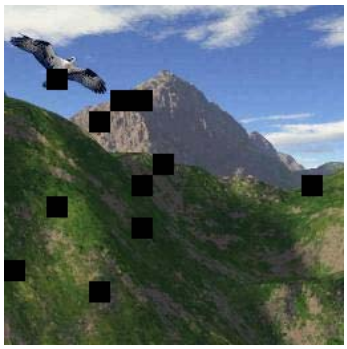
La figura 4.6 muestra los bloques erróneos (cuadros negros) detectados por el sistema en imágenes a color con tasas de compresión de 0.58bits/pixel y 0.49 bits/píxel respectivamente en donde de acuerdo a la posición de los bloques erróneos estas imágenes se consideran no alteradas.

La tabla 4.3. muestra la tasa de compresión máxima a la cual las imágenes son determinadas por el sistema como auténticas en donde podemos observar la buena respuesta del sistema ante la compresión JPEG ya que la reducción del tamaño de la imagen marcada en escala de grises en promedio se puede reducir 6.6 veces, debido a que en lugar de utilizar 8 bits/pixel en promedio requerimos 1.34 bits/pixel en el caso de imágenes a color observamos que para codificar un píxel se requieren en promedio 0.46 bits, es decir las imágenes marcadas se reducen en promedio 7.54 veces.



(a) “chiles” autenticada JPEG Q=65

(b) “chiles” verificada JPEG Q=65



(c) “montaña” autenticada JPEG Q=75

(d) “montaña” verificada JPEG Q=75

Figura 4.6. Robustez ante compresión JPEG en imágenes a color

Tabla 4.3. Tasa de compresión JPEG máxima soportada por el algoritmo AFDMA.

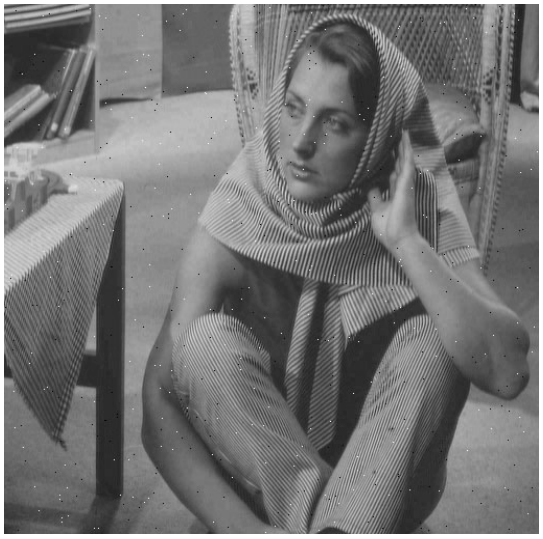
Imagen en escala de grises	*Factor de calidad*	*Tasa de compresión (bits/pixel)*	*Imagen a color*	*Factor de calidad*	*Tasa de compresión (bits/pixel)*
Bárbara	70	1.10	*avión*	70	0.45
barco	80	1.20	*casa*	70	0.37
bridge	75	1.79	*chica*	75	0.41
cámara	80	1.26	*chiles*	65	0.58
chiles	75	0.97	*lago*	70	0.58
goldhill	70	1.11	*Lena*	65	0.45
Lena	75	1.29	*montaña*	75	0.49
mandril	80	2.24	*personas*	70	0.37
pájaro	80	0.93			

4.1.3.2. ROBUSTEZ AL RUIDO

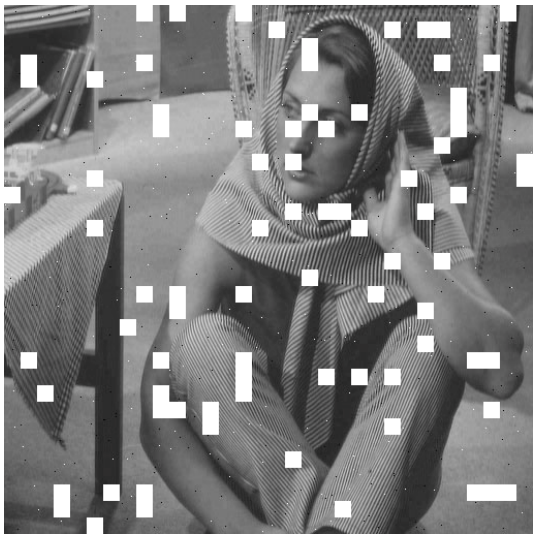
Para simular el ruido que se adiciona en la transmisión de una imagen en el canal de comunicaciones, las imágenes marcadas fueron modificadas con diferentes niveles de ruido impulsivo y ruido gaussiano. La tabla 4.4 muestra los valores máximos de densidad de ruido "impulsivo" en imágenes en escala de grises y a color, para los cuales el sistema determina que la imagen marcada y alterada con ruido aditivo es auténtica, de acuerdo a estos resultados podemos destacar la eficiencia del sistema ante ataques de ruido "impulsivo", ya que en promedio soporta una densidad de ruido igual a 0.002.

La figura 4.7 (a) muestra la degradación que sufre la imagen en escala de grises "Bárbara" cuyo PSNR es de 32 dB entre la imagen marcada y la imagen marcada contaminada por ruido impulsivo con una densidad de 0.002, la figura 4.7 (b) muestra la imagen resultante del proceso de verificación, en donde podemos observar que los bloques detectados como erróneos (bloques blancos) se presentan en forma aislada, por lo tanto, esta imagen se considera como auténtica.

La figura 4.7 inciso (c) y (d) muestran hasta que valor de densidad es robusto el algoritmo en imágenes a color y que degradación sufre la imagen marcada y contaminada por ruido impulsivo expresado en decibeles.

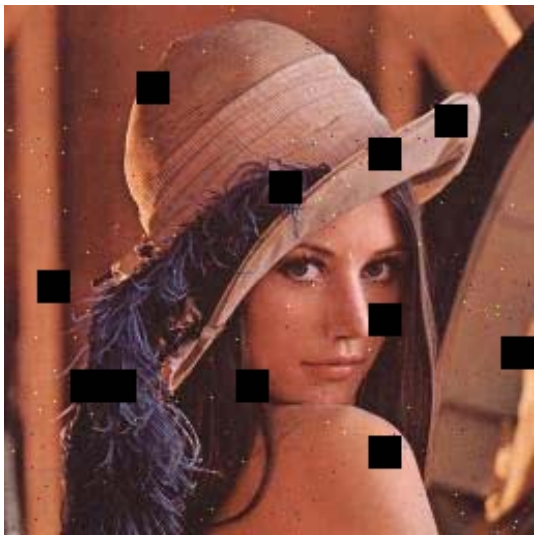


(a) "Bárbara" den. impulsivo=0.002
PSNR=32 dB



(b) "Bárbara" autenticada

(c) "Lena" den. impulsivo=0.0025
PSNR=30 dB



(d) "Lena" autenticada

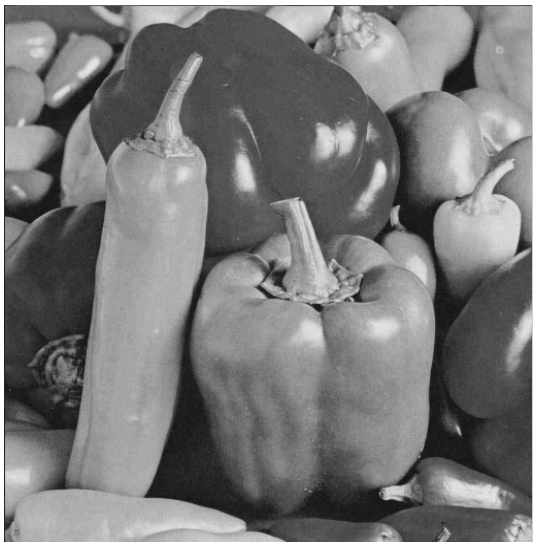
Figura 4.7. Robustez al ruido impulsivo.

De los resultados arrojados por el sistema en la adición de ruido gaussiano aplicado en imágenes en escala de grises y a color mostrados en la tabla 4.4 podemos observar que en promedio el sistema considera una imagen con ruido gaussiano como auténtica si la varianza de este es menor a 0.0001.

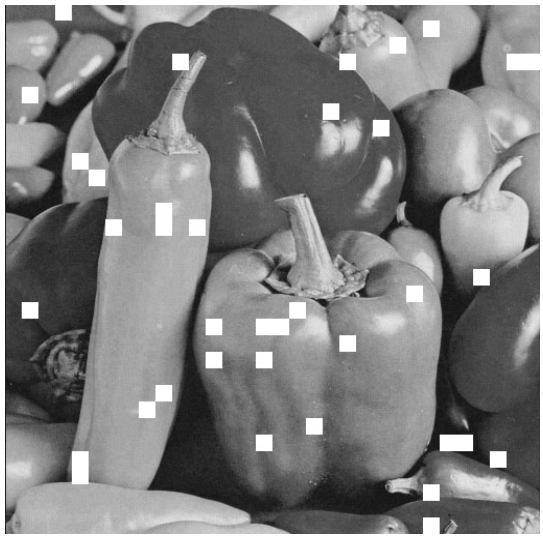
Tabla 4.4. Prueba de resistencia a la adición de ruido impulsivo y gaussiano del algoritmo AFDMA.

Imagen en escala de grises	*Densidad de ruido impulsivo*	*Varianza de ruido gaussiano*	*Imagen en color*	*Densidad de ruido impulsivo*	*Varianza de ruido gaussiano*
Bárbara	0.002	0.00011	*avión*	0.0016	0.00011
barco	0.002	0.0001	*casa*	0.0016	0.00031
bridge	0.002	0.00014	*chica*	0.0015	0.00027
cámara	0.002	0.00011	*chiles*	0.0024	0.00027
chiles	0.002	0.00011	*lago*	0.0018	0.00033
goldhill	0.002	0.00011	*Lena*	0.0025	0.00027
Lena	0.002	0.00014	*montaña*	0.0015	0.00031
mandril	0.002	0.00011	*personas*	0.0015	0.00027
pájaro	0.0009	0.00014			

La figura 4.8 (a) muestra la imagen de “chiles” en escala de grises contaminada con ruido gaussiano con una varianza de 0.00011, mientras que la figura 4.8 (b) muestra la ubicación de los bloques erróneos localizados por el sistema de autenticación (bloques blancos). Las figuras 4.8 (c) y (d) muestran los resultados de ruido gaussiano en imágenes a color.

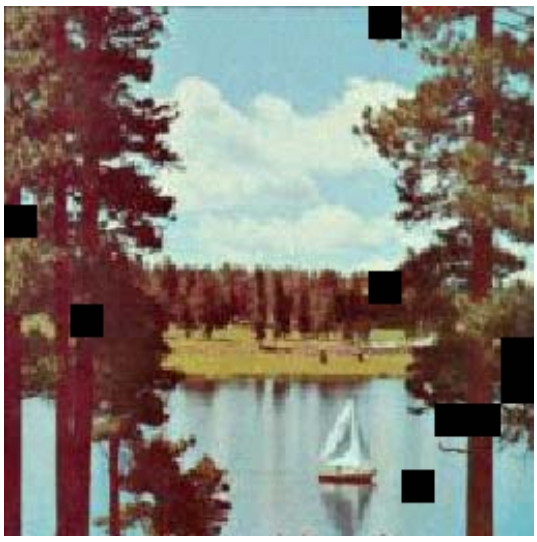


(a) “chiles” var. gaussiano=0.00011



(b) “chiles” autenticada

(c) “lago” var. gaussiano=0.00033



(d) “chiles” autenticada

Figura 4.8. Robustez al ruido gaussiano.

4.1.4. COMPARACIONES DEL ALGORITMO AFDMA CON OTROS MÉTODOS

En esta sección realizaremos la comparación entre el algoritmo AFDMA con los algoritmos propuestos por Zhou et al. en [Zhou et al., 2004] y Xie et al. en [Xie et al., 2007], los cuales fueron descritos en el capítulo 2. Dichas comparaciones están basadas en tres aspectos principales:

1. Imperceptibilidad de la marca de agua
2. Localización de los bloques alterados
3. Robustez a la compresión JPEG

4.1.4.1. Imperceptibilidad de la marca de agua

Para medir la imperceptibilidad de la marca de agua utilizamos la relación señal a ruido entre la imagen original y la imagen marcada. El valor de PSNR promedio de 100 imágenes probadas en los tres algoritmos fue:

- Imperceptibilidad de la marca de agua para el algoritmo *AFDMA* es 45 dB.
- Imperceptibilidad de la marca de agua para el algoritmo propuesto por Zhou es 50 dB.
- Imperceptibilidad de la marca de agua para el algoritmo propuesto por Xie es 39.7 dB.

Durante el proceso de discretización de los valores de los pixeles de la imagen original en la fase de inserción de la marca de agua es común que estos sufran una variación, la cual podría reflejarse como una modificación en la imagen marcada durante el proceso de verificación e identificar a una imagen no alterada como alterada; lo que se conoce como falso positivo. La tabla 4.5 muestra la probabilidad de errores falso positivo en las imágenes marcadas sin ninguna alteración en los tres algoritmos.

En la tabla 4.5 podemos observar que el error producido al momento de insertar la marca de agua en la imagen original es muy pequeño; por consecuencia el error falso positivo también lo es, en este caso no podemos medir el error falso negativo, ya que no modificamos la imagen marcada.

Tabla 4.5. Comparación de la probabilidad de error falso positivo en imágenes marcadas sin alteraciones.

	Algorítmo AFDMA	*Algoritmo de Zhou*	*Algoritmo de Xie*
Falso positivo (Pf_p)	1%	2%	1%

4.1.4.2. Localización de los bloques alterados

Para llevar a cabo esta prueba, las imágenes marcadas de los tres algoritmos fueron alteradas usando Photoshop. La tabla 4.6 muestra los resultados obtenidos en términos de probabilidad de error falso positivo (detecta bloques erróneos en donde no se altero la imagen) y error falso negativo (no detecta errores en donde si los hay) para la detección de bloques erróneos en la imagen.

Tabla 4.6. Comparación de la probabilidad de error falso positivo y falso negativo en imágenes marcadas y alteradas.

	Algoritmo AFDMA	*Algoritmo de Zhou*	*Algoritmo de Xie*
Falso positivo (Pf_p)	1%	5%	1%
Falso negativo (Pf_n)	10%	66%	70%

Debido a que la probabilidad de error falso negativo es más perjudicial para cualquier sistema de autenticación podemos decir que el algoritmo menos eficiente de acuerdo a la comparación mostrada en la tabla 4.6 es el propuesto por Xie.

4.1.4.3. Tolerancia a la compresión JPEG

Para llevar a cabo esta prueba, sometimos a las 100 imágenes marcadas por los tres algoritmos a compresión JPEG con una variación del factor de calidad de 100 a 60.

La figura 4.9 muestra una gráfica de la variación de la probabilidad de error falso positivo *(Pf_p)* para imágenes marcadas comprimidas con factores de calidad 60, 65, 70, 75, 80, 85, 90, 95 y 100. En la gráfica la línea representada por asteriscos es la respuesta del sistema *AFDMA,* la línea que tiene cuadrados es la respuesta del sistema propuesto por Zhou y finalmente la línea representada por círculos es la respuesta del sistema propuesto por Xie, en donde podemos observar que el algoritmo propuesto por Xie es robusto a la compresión JPEG para un factor de calidad mayor a 60, en el caso del algoritmo *AFDMA* es robusto para factores de calidad mayor a 70, pero en el caso del algoritmo propuesto por Zhou es robusto para factores de calidad mayor a 95.

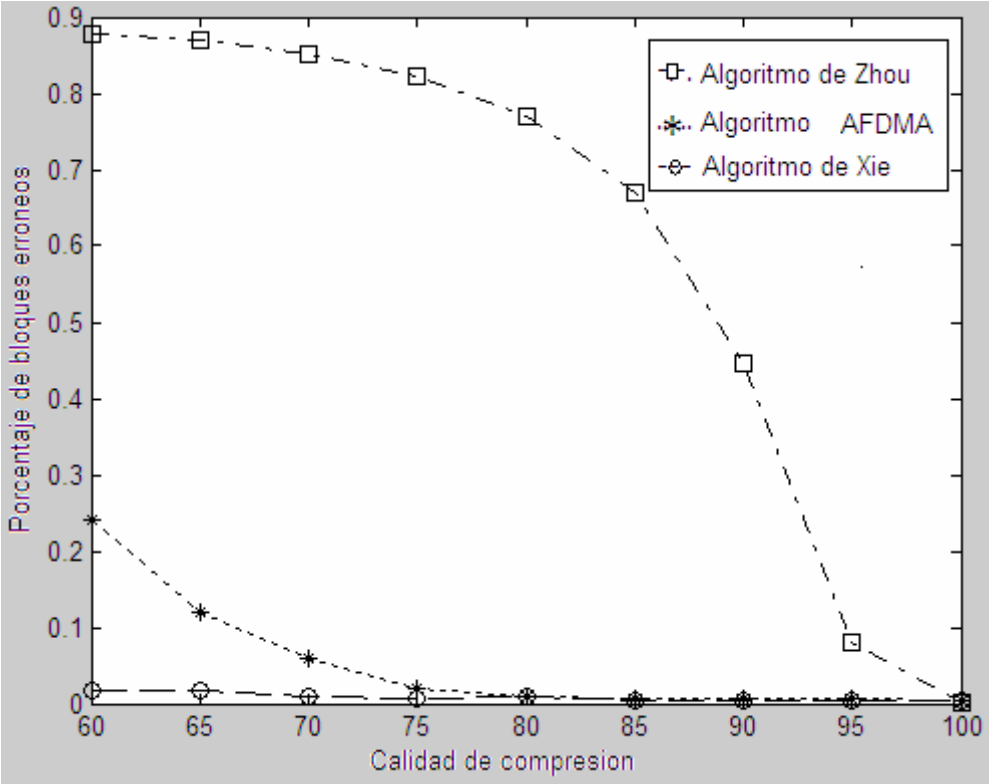


Figura 4.9. Comparación de la robustez a la compresión JPEG de los 3 algoritmos

4.2. RESULTADOS DEL ALGORITMO AAMA

Para la implementación de este algoritmo se utilizaron imágenes en escala de grises de tamaño 512x512 y 256x256 pixeles como imágenes huésped, cada píxel es representado por 8 bits, la marca de agua es una secuencia binaria cuya longitud es de 66 bits. La programación y las pruebas realizadas a este algoritmo se realizaron en Matlab 7.0.

Los valores mostrados en la tabla 4.7. son los utilizados para las pruebas en el algoritmo.

Tabla 4.7. Valores de los parámetros usados en el algoritmo AAMA

Parámetros	*Descripción*	*Valor*
No. de imágenes probadas	256 niveles de gris (8 bits/pixel)	100
W	Secuencia de marca de agua por bloque	66 bits
Q	Factor de calidad usado en la ecuación (3.10)	70
Th	Valor de umbral usado en la ecuación (3.11)	13

4.2.1. DEMOSTRACIÓN MATEMÁTICA DEL VALOR DE UMBRAL TH EN EL PROCESO DE AUTENTICACIÓN

A continuación se muestra la demostración matemática del valor del *Th* óptimo usado en el sistema de autenticación de los bloques de la Región de Interés (ROI) de la imagen marcada, descrito en la sección 3.2.4

Generalmente el problema de autenticación de imágenes puede ser considerado como un problema de prueba de hipótesis, las cuales son del tipo:

Hipótesis 1: el bloque ROI es auténtico
Hipótesis 2: el bloque ROI no es auténtico

Debido a que la función de densidad de probabilidad (pdf) de la suma XOR expresada en la ecuación 4.2 de la hipótesis 1 tiene distribución gaussiana con media=2.2681 y varianza=2.3369.

$$\begin{aligned} &Si \sum XOR(W_{ROIext}, W_{ROEext}) < Th \quad el\ bloque\ es\ autentico \\ &Si \sum XOR(W_{ROIext}, W_{ROEext}) \geq Th \quad el\ bloque\ es\ \text{mod}ificado \end{aligned} \tag{4.2}$$

Para obtener la probabilidad de error se debe utilizar la función de distribución acumulativa complementaria

$$Q(x) = \int_{x}^{\alpha} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{t^2}{2}\right\} dt \,. \tag{4.3}$$

Pero debido a que la función de distribución de probabilidad (pdf) gaussiana no es normal, es decir $N\left(\mu, \sigma^2\right)$ se debe hacer un cambio de variable

$$F(x) = \int_{x}^{\alpha} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt \,, \quad \text{si} \quad u = \frac{t-\mu}{\sigma} \,, \quad \text{y} \quad dt = \sigma du \,, \tag{4.4}$$

entonces

$$F(x) = \int_{x}^{\alpha} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{= \frac{1}{2}\left(\frac{(x-\mu)}{\sigma}\right)^2\right\} dt \,, \tag{4.5}$$

$$F(x) = \int_{\frac{x-\mu}{\sigma}}^{\alpha} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{u^2}{2}\right\} \sigma du \,, \tag{4.6}$$

por lo tanto

$$F(x) = \int_{\frac{x-\mu}{\sigma}}^{\alpha} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du \,, \tag{4.7}$$

lo que implica que F(x) tiene N(0,1). Basándonos en la ecuación (4.7) y haciendo un cambio de variable

$$z^2 = \frac{u^2}{2} \,, \tag{4.8}$$

entonces

$$z = \frac{u}{\sqrt{2}} \quad \text{y} \quad du = \sqrt{2}dz \tag{4.9}$$

la ecuación (4.7) se puede reescribir como

$$Q(x) = \int_{\frac{x-\mu}{\sqrt{2}\sigma}}^{\alpha} \frac{1}{\sqrt{\pi}} \exp\{-z^2\} dz \,, \tag{4.10}$$

$$Q(x) = \frac{1}{2} erfc\left(\frac{x-\mu}{\sqrt{2}\sigma}\right), \tag{4.11}$$

donde

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\alpha} \exp(-y^2) dy \,. \tag{4.12}$$

Para una probabilidad de error falso positivo aceptable se requiere de un umbral que satisfaga $P_{fp} \leq 10^{-12}$ por lo tanto se requiere obtener el valor de x a partir de que $Q(x) = 10^{-12}$, de la ecuación (4.11)

$$\frac{1}{2} erfc\left(\frac{x-\mu}{\sqrt{2}\sigma}\right) = 10^{-12} \,, \tag{4.13}$$

dado que

$$erfc(x) = 1 - erf(x) \,, \tag{4.14}$$

entonces

$$\frac{1}{2}\left[1-erf\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right]=10^{-12}, \tag{4.15}$$

finalmente despejando x de la ecuación (4.15)

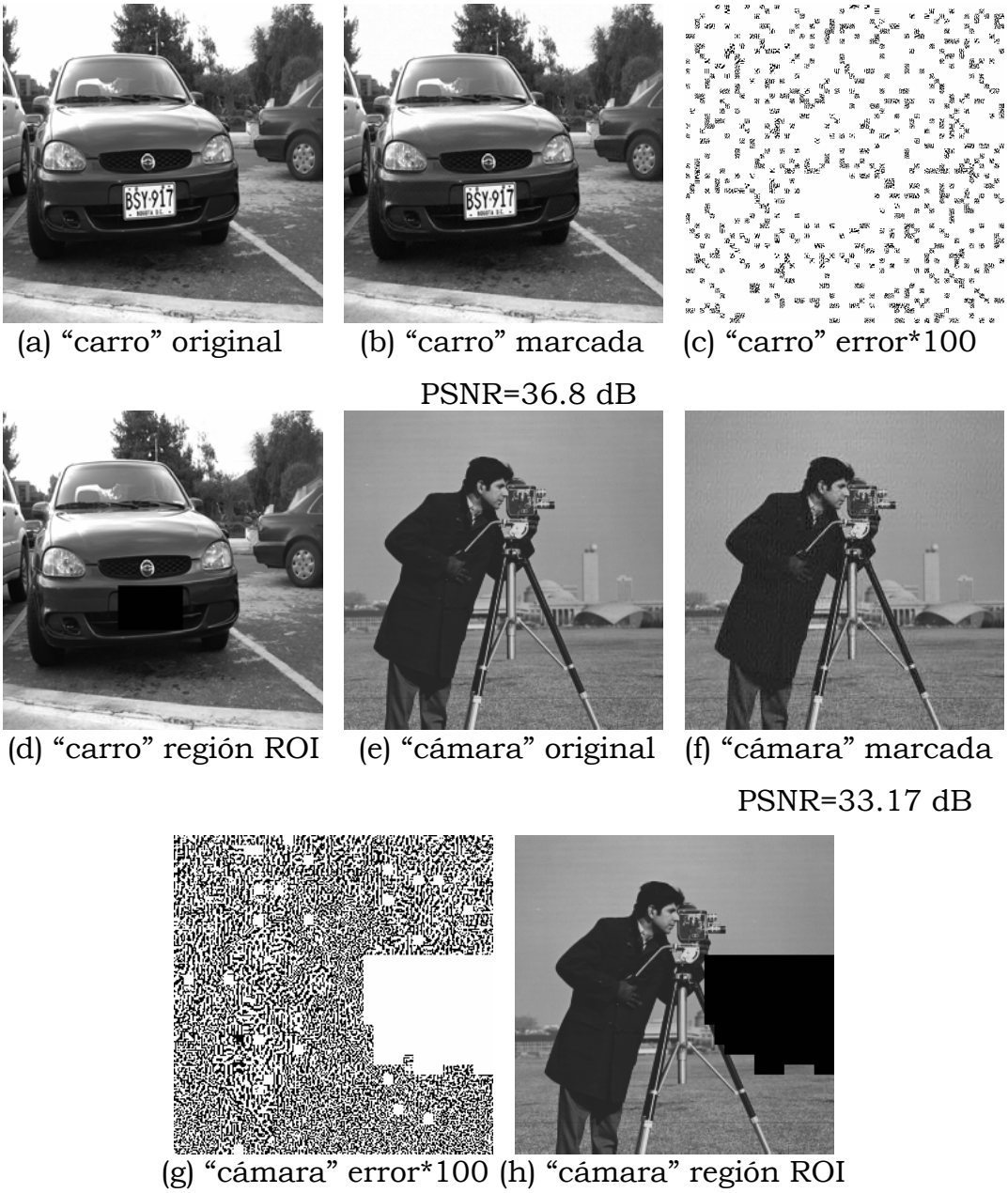
$$\frac{x-\mu}{\sqrt{2}\sigma}=erf^{-1}\left(1-2\left(10^{-12}\right)\right), \tag{4.16}$$

$$x=\sqrt{2}\sigma erf^{-1}\left(1-2\left(10^{-12}\right)\right)+\mu\,. \tag{4.17}$$

Para este caso en específico obtenemos que $Th = x \approx 13$.

4.2.2. IMPERCEPTIBILIDAD DE LA MARCA DE AGUA

Las imágenes originales “carro” y “cámara” se muestran en la figura 4.10 (a) y en la figura 4.10 (e), las figuras 4.10 (b) y 4.10 (f) muestran las imágenes marcadas correspondientes cuyos valores de PSNR entre la imagen original y la imagen marcada son 36.8 dB y 33.17 dB respectivamente estos valores dependen directamente de la longitud de la marca de agua y esta a su vez esta en función de la región de interés (ROI) representadas en las figuras 4.10 (d) y 4.10 (h), las figuras 4.10 (c) y 4.10 (g) muestran las imágenes de error las cuales son la diferencia entre la imagen original y la imagen marcada multiplicada por 100 en donde podemos observar que la marca de agua modifica solamente la región ROE.



(a) "carro" original (b) "carro" marcada (c) "carro" error*100

PSNR=36.8 dB

(d) "carro" región ROI (e) "cámara" original (f) "cámara" marcada

PSNR=33.17 dB

(g) "cámara" error*100 (h) "cámara" región ROI

Figura 4.10. Imperceptibilidad de la marca de agua.

Estos resultados indican que la distorsión en la imagen marcada cuado se lleva a cabo el proceso de inserción de la marca de agua no es visible por el sistema visual humano.

4.2.2.1. LONGITUD DE LA MARCA DE AGUA

La longitud de la marca de agua en la imagen entera depende directamente del número de bloques ROI seleccionados por el propietario de la imagen, el cual esta limitado debido a que por cada bloque ROI de 8*8 pixeles extraemos una secuencia de marca e agua de 66 bits y necesitamos 6 bloques ROE para insertar 11 bits por cada bloque ROE como se describió en la sección 3.2.2. Para conocer el máximo número de bloques ROI que puede tener una imagen se calcula mediante la ecuación (4.18)

$$No.\max bloques\ ROI = \frac{No.\ \ total\ \ de\ \ bloques\ \ de\ \ 8*8\ \ en\ \ la\ \ imagen}{7}. \qquad (4.18)$$

Por ejemplo para una imagen de 256*256 pixeles el máximo número de bloques ROI son 146. Las figuras 4.10. (c) y 4.10. (f) muestran el ejemplo de unas posibles ROI que seleccionaría el usuario en la imagen de “carro” y “cámara”, las cuales están representadas por bloques negros.

La tabla 4.8 muestra los valores de PSNR promedio en imágenes de tamaño de 512*512 y 256*256 en escala de grises con diferentes características para diferentes porcentajes de bloques ROI seleccionados por el usuario.

La figura 4.11 muestra una gráfica del comportamiento del PSNR promedio de imágenes marcadas ante la variación del número de bloques ROI para imágenes de tamaños 512*512 representado por ‘*’ y 256*256 representado por ‘.’ En donde podemos ver que si la cantidad de bloques ROI aumenta el la calidad en la imagen marcada disminuye.

Tabla 4.8. PSNR de las imágenes marcadas con diferentes porcentajes de bloques ROI.

Tamaño de la imagen				
512x512			256x256	
Porcentaje	No de bloques ROI	PSNR promedio	No de bloques ROI	PSNR promedio
100	585	33.71	146	32.75
90	526	34.21	131	33.35
80	468	34.71	116	34.08
70	409	35.31	102	34.51
60	351	35.98	87	35.44
50	292	36.82	73	35.98
40	234	37.78	58	37.17
30	175	39.14	43	39.17
20	117	40.8	29	40.5
10	58	44.17	14	45.5

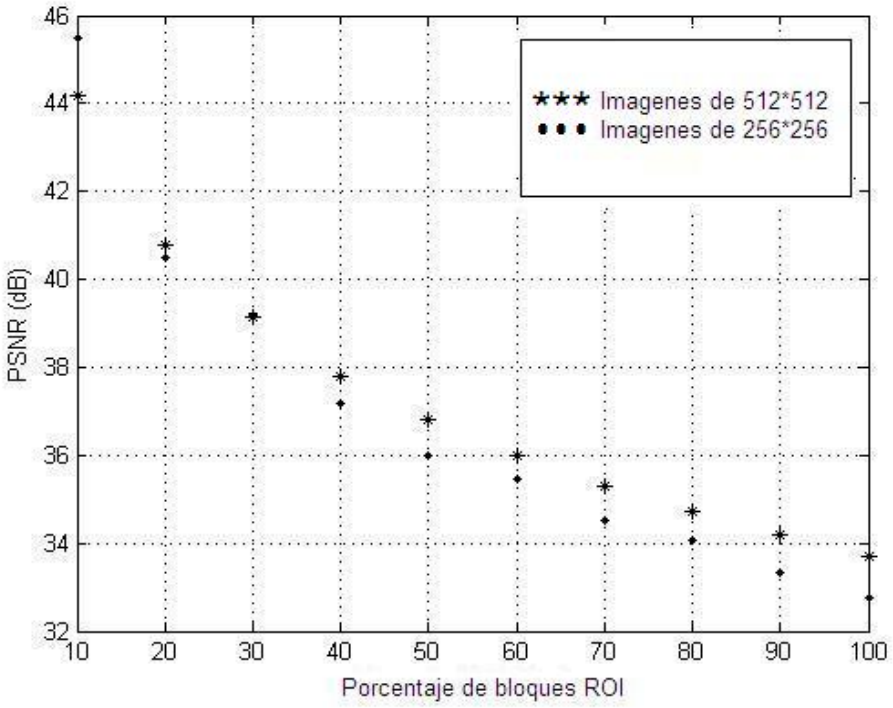


Figura 4.11. Calidad en imágenes marcadas.

Debido a que la longitud de la marca de agua esta en función de la cantidad de bloques ROI en la imagen, si esta se incrementa, la calidad en la imagen disminuye. La tabla 4.9 muestra algunos ejemplos de la calidad en la imagen marcada de tamaño 512*512.

Tabla 4.9. Ejemplo de la variación del PSNR ante la variación en la longitud de la marca de agua en el algoritmo AAMA.

Numero de bloques ROI	*Longitud de la marca de agua (bits)*	*PSNR (dB)*
117	117x66=7722	39.85
182	182x66=12012	37.99
441	441x66=29106	33.27
453	453x66=29898	33.17

4.2.3. CAPACIDAD DE DETECCIÓN DE REGIONES ALTERADAS Y AUTO-RECUPERACIÓN

Para evaluar la eficiencia del sistema de autenticación propuesto las imágenes marcadas fueron alteradas usando Photoshop. Las figura 4.12 (b) y 4.12 (f) muestran las imágenes alteradas, en la primera el número 7 de la imagen “carro” fue sustituida por el número 9; en la segunda, la torre de la imagen “cámara” fue borrada. Las figuras 4.12 (c) y 4.12 (g) muestran por medio de bloques negros la localización de los bloques erróneos en la imagen marcada, posteriormente al haber localizado los bloques erróneos se aplica el algoritmo de auto-recuperación cuyo resultado se muestra en las figuras 4.12 (d) y 4.12 (h).


a) "carro" marcada


(b) "carro" alterada


(c) "carro" autenticada


(d)"carro" Auto-recuperada


(e) "cámara" marcada


(f) "cámara" alterada


(g) "cámara" región ROI


(h) "cámara" Auto-recuperada

Figura 4.12. Capacidad de detección y auto-recuperación de regiones alteradas.

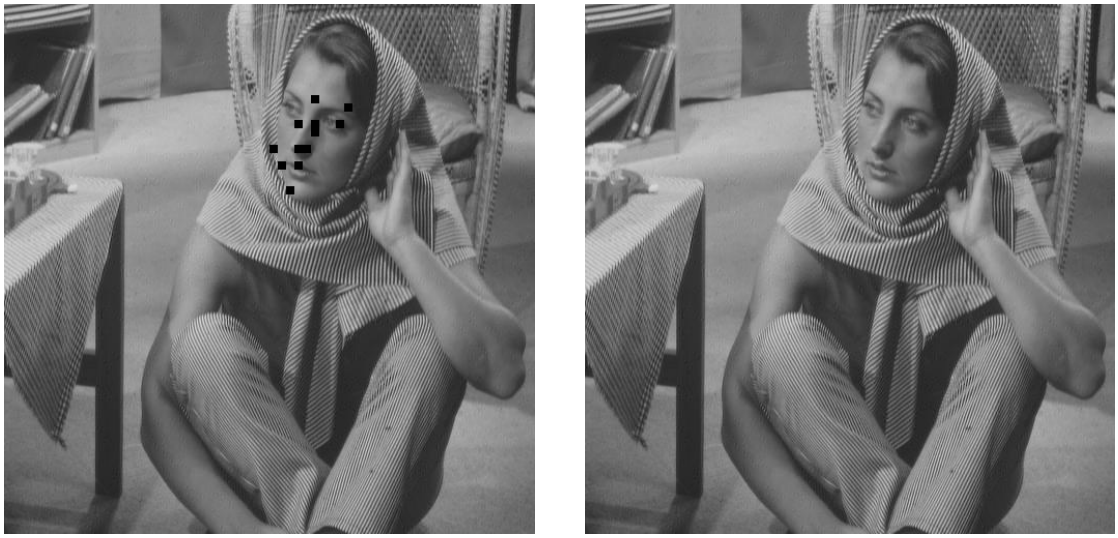
4.2.4. ROBUSTEZ DE LA MARCA DE AGUA A ATAQUES NO INTENCIONALES

4.2.4.1. ROBUSTEZ A LA COMPRESION JPEG

Las imágenes probadas pueden resistir a la compresión JPEG con un factor de calidad variable de 100 a 80, en otras palabras, los datos ocultos (marca

de agua) se pueden extraer sin ningún error cuando las imágenes marcadas han sigo comprimidas en este rango.

En el caso donde las imágenes marcadas son comprimidas con un factor de calidad dentro del rango (79 a 70) los bloques detectados como erróneos pueden ser auto-recuperados. La figura 4.13 muestra un ejemplo para una imagen comprimida con un factor de calidad de 75.



(a) "Bárbara" autenticada comprimida Q=75

(b)"Bárbara" auto-recuperada

Figura 4.13. Auto-recuperación en imágenes comprimidas con JPEG

4.2.4.2. ROBUSTEZ AL RUIDO

Para probar la robustez del la marca de agua a ataques de ruido, adicionamos ruido impulsivo y gaussiano a las imágenes marcadas, de las cuales observamos lo siguiente:

- El sistema propuesto es robusto al ruido impulsivo adicionado a las imágenes marcadas con una densidad menor a 0.0005, para poder

determinar cuanto se degrada la imagen marcada al aplicar esta densidad de ruido se calculo el PSNR entre la imagen marcada y la imagen marcada y ruidosa el cual es 38.6 dB.

- Para que el algoritmo sea robusto al ruido gaussiano las imágenes marcadas deben ser contaminadas con una varianza menor a 0.00009 cuyo valor de PSNR entre la imagen marcada y la imagen marcada contaminada es 40.4 dB.

4.2.5. COMPARACIONES DEL ALGORITMO AAMA CON OTROS MÉTODOS

Ya que el algoritmo propuesto es un sistema de marca de agua semi-frágil basado en bloques en cual puede localizar y recuperar a su estado original la imagen alterada realizamos una comparación con los algoritmos propuestos por Zhao [Zhao et al., 2007] y Hassan [Hassan et al., 2008] descritos en el capitulo 2. Esta comparación esta enfocada en tres aspectos principales:

1. Calidad en la imagen recuperada
2. Localización y auto-recuperación de los bloques alterados
3. Robustez a la compresión JPEG

Para que la comparación sea en las mismas condiciones en los tres algoritmos, adaptamos el algoritmo de Zhao en regiones ROI y ROE y retomamos el concepto del algoritmo propuesto por Hassan acerca de la región ROI. Antes de mostrar los resultados, es importante mencionar que:

- La imagen original en los tres algoritmos es dividida en bloques de 8*8 pixeles.
- La imagen original en los tres algoritmos es dividida en regiones ROI y ROE.
- La longitud de la marca de agua de nuestro algoritmo es 66 bits.

- La longitud de la marca de agua frágil del algoritmo propuesto Zhao es 20 bits y el de la marca de agua robusta es 64 bits.
- La longitud de la marca de agua frágil del algoritmo propuesto Hassan es 20 bits y el de la marca de agua robusta es 4 números decimales.

4.2.5.1. CALIDAD DE LA IMAGEN RECUPERADA

En esta comparación mostraremos la calidad de la imagen recuperada usando a "Bárbara" como la imagen original de tamaño 512*512 pixeles mostrada en la figura 4.14 (a), la cual es la misma para los tres algoritmos. La figura 4.14 (b) muestra los bloques ROI representados por bloques negros en la imagen (compuesto por 169 bloques) y los bloques ROE representados por el resto de la imagen (compuesto por 3927 bloques), los cuales fueron propuestos por nosotros como posibles bloques ROI. Las figuras 4.14 del inciso (c) al (e) muestra las tres imágenes marcadas cuyo valor de PSNR es 39.85 dB (algoritmo AAMA), 42.18 dB (algoritmo de Zhao) y 44.19 dB (algoritmo de Hassan). Los bloques ROI de las tres imágenes marcadas fueron modificados completamente (atacadas) como se muestra en la figura 4.14 (f). Finalmente las figura 4.14 del inciso (g) al (i) muestran las imágenes recuperadas con un valor de PSNR igual a 42.9 dB (algoritmo AAMA), 40.37 dB (algoritmo de Zhao) y 39.15 dB (algoritmo de Hassan).

De acuerdo a estos resultados es importante puntualizar que el valor de PSNR de la imagen recuperada por nuestro esquema es mayor que el de los otros autores; es decir, la imagen recuperada por el algoritmo AAMA se degrada 2.5 dB menos que la de Zhao y 3.5 dB menos que la de Hassan.

4.2.5.2. CAPACIDAD DE LOCALIZACIÓN Y AUTO-RECUPERACIÓN

En esta sección pretendemos probar cual algoritmo detecta mejor las alteraciones en una imagen marcada. La figura 4.15 (a) muestra la imagen

original, la figura 4.15 (b) muestra las regiones ROI y ROE de la imagen huésped. Para probar los algoritmos, intencionalmente reemplazamos el número 7 de la placa del carro por el número 9 en las 3 imágenes marcadas por los 3 diferentes algoritmos como se muestra en la figura 4.15 (c). Las figuras 4.15 (d) al (f) muestran los bloques modificados detectados por los sistemas (bloques negros). Como podemos observar el proceso de detección del sistema AAMA es más exacto, ya que no detecta bloques erróneos fuera de la región modificada como lo hacen los otros dos algoritmos. Finalmente en las figuras 4.15 (g) al (i) se pueden observar las imágenes recuperadas cuyos valores de PSNR calculado entre la imagen marcada sin modificar y la imagen recuperada son:

- 43 dB para el algoritmo AAMA
- 39.41 dB para el algoritmo propuesto por Zhao y
- 37 dB para el algoritmo propuesto por Hassan

De acuerdo a estos resultados podemos decir que el algoritmo propuesto en esta investigación (AAMA) tiene una mayor calidad en las imágenes recuperadas.

4. 2.5.3. ROBUSTEZ A LA COMPRESIÓN JPEG

Las imágenes marcadas fueron codificadas usando la compresión JPEG estándar para probar la robustez de los tres algoritmos variando el factor de calidad de 100 a 75.

La tabla 4.10 muestra una comparación de la robustez del algoritmo AAMA con los algoritmos propuestos por Zhao y Hassan en donde podemos observar que el más robusto es el algoritmo AAMA. En la tabla 4.10 el símbolo (A) representa que la imagen marcada y alterada es detectada como auténtica, es decir no se detectan bloques erróneos en el proceso de

autenticación. En esta tabla podemos observar que el algoritmo AAMA detecta a la imagen marcada y comprimida como auténtica hasta con un factor de calidad mayor a 80, a diferencia de los algoritmos propuestos por Zhao y Hassan en donde estos detectan la imagen como auténtica con factores de calidad mayores a 95 debido a que insertan una marca de agua frágil en los LSB de la imagen huésped, lo que la hace poco robusta a la compresión JPEG.

Como podemos observar en la tabla 4.10 algunos datos están expresados en decibeles, debido a que estos representan la calidad de los bloques ROI recuperados con respecto a los bloques ROI de la imagen marcada; esto quiere decir la imagen marcada y comprimida con esas tasas de compresión no son detectadas como auténticas, sino que los bloques erróneos localizados son auto-recuperados.

Tabla 4.10. Comparación de robustez a la compresión JPEG del algoritmo AAMA con el algoritmo de Zhao y Hassan.

QF	*100*	*95*	*90*	*85*	*80*	*75*
Tasa de compression (bpp)	4.66	2.53	1.95	1.57	1.37	1.20
Algoritmo AAMA	(A)	(A)	(A)	(A)	(A)	45.04 dB
Algoritmo Zhao	(A)	(A)	30 dB	28.34 dB	21.47 dB	21.99 dB
Algoritmo Hassan	(A)	29 dB	27.8 dB	24.78 dB	20 dB	18.9 dB

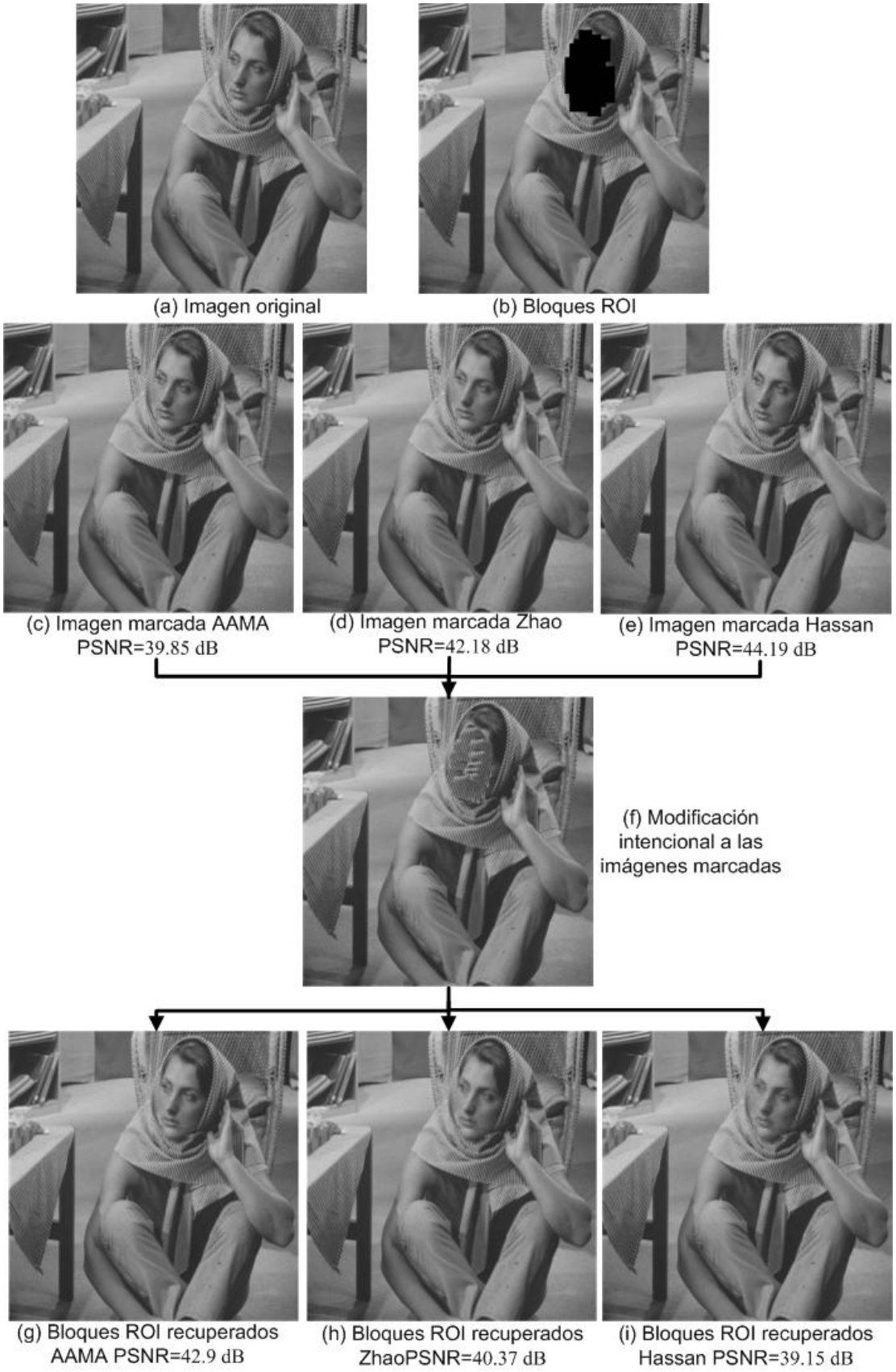


Figura 4.14. Comparación de la calidad en la auto-recuperación

a) “carro” marcada

(b) “carro” Región ROI

(c) “carro” alterada

(d) “carro” detección AAMA

(e) “carro” detección Zhao

(f) “carro” detección Hassan

(g) “carro” recuperación AAMA PSNR=43 dB

(h) “carro” recuperación Zhao PSNR=39.41dB

(i) “carro” recuperación Hassan PSNR=37 dB

Figura 4.15. Comparación de la capacidad de detección y auto-recuperación

4.2.5.4. RESUMEN DE LAS COMPRACIONES DE LOS ALGORTIMOS AFDMA Y AAMA

Como ya se menciono en la sección 4.1.4 el algoritmo AFDMA se comparó con el algoritmo propuesto por Zhou [Zhou, et al., 2004] y el algoritmo propuesto por Xie [Xie, et al., 2007], finalmente queremos mostrar por medio de la tabla 4.11 un resumen de los resultados de estas comparaciones, en donde podemos puntualizar que aunque la calidad en la imagen marcada es mayor en el algoritmo propuesto por Zhou, en el algoritmo AFDMA y en el propuesto por Xie la calidad en la imagen marcada es aceptable, ya que no se distingue la marca de agua; Otra ventaja que tiene el algoritmo AFDMA es que la probabilidad de error falso negativo es mucho menor a la de los otros dos algoritmos. En cuanto a la robustez a la compresión aunque el algoritmo propuesto por Xie es robusto para un factor de calidad mayor a 60 la robustez que presenta el algoritmo AFDMA (75) es muy aceptable, ya que para valores de FC menores a 70, la imagen marcada se degrada y pierde su valor comercial.

Tabla 4.11. Comparación final entre AFDMA, el algoritmo propuesto por Zhou y el algoritmo propuesto por Xie.

	Algoritmo AFDMA	*Algoritmo propuesto por Zhou*	*Algoritmo propuesto por Xie*
Imperceptibilidad de la marca de agua	45 dB	50 dB	39.7 dB
Localización de los bloques alterados	$Pf_{n=}$ 10%	$Pf_{n=}$ 66%	$Pf_{n=}$ 70%
Robustez a la compresión JPEG	FC>75	FC>95	FC>60

En la tabla 4.12 resumimos los resultados de la comparación hecha entre el algoritmo AAMA y los algoritmos propuestos por Zhao [Zhao, et al., 2007] y Hassan [Hassan, et al., 2008] descritos en el capitulo 2; cabe mencionarse que debido a que el valor PSNR depende de la cantidad de bloques ROI seleccionados, estos resultados se obtuvieron usando imágenes de 512*512 con una cantidad de bloques ROI igual a 169. Aunque los tres algoritmos detectan y localizan correctamente los bloques erróneos en las imágenes marcadas y alteradas, en cuanto a la calidad de la imagen reconstruida, el algoritmo AAMA es mejor, permitiendo así identificar plenamente cual fue el contenido de la imagen alterado. Al hacer la comparación de los algoritmos a la compresión JPEG el algoritmo AAMA tiene la ventaja de que es robusto a factores da calidad mayores a 80, el cual es un factor de calidad aceptable, aunque los otros dos algoritmos tienen la capacidad de recuperar los bloques detectados como alterados, tendrían que pasar todavía por el proceso de auto recuperación.

Tabla 4.12. Comparación final entre AAMA, el algoritmo propuesto por Zhao y el algoritmo propuesto por Hassan.

	Algoritmo AAMA	*Algoritmo propuesto por Zhao*	*Algoritmo propuesto por Hassan*
Calidad en la imagen recuperada	42.9 dB	40.37 dB	39.15 dB
Localización y auto-recuperación de los bloques alterados	Los tres algoritmos localizaron correctamente los bloques erróneos, pero en cuanto a la calidad en la recuperación fue mayor en el algoritmo AAMA.		
Robustez a la compresión JPEG	FC>80	FC>90	FC>95

4.3. CONCLUSIONES

Las pruebas realizadas a los algoritmos propuestos en esta tesis (AFDMA y AAMA) estuvieron conducidas a evaluar la imperceptibilidad de la marca de agua, la cual fue mayor a 30 dB en ambos algoritmos; a evaluar la capacidad de los algoritmos para autenticar el contenido de la imagen recibida cuyo resultado fue satisfactorio ya que los dos algoritmos propuestos detectan correctamente las regiones alteradas. En el caso del algoritmo AAMA la calidad de la imagen recuperada es superior a los 30 dB aunque se inserte la marca de agua correspondiente al 100% de los bloques ROI. Otro análisis que se hizo fue la robustez de los algoritmos propuestos cuando la imagen marcada se somete a compresión JPEG a diferentes factores de calidad, el resultado de este análisis mostró que el algoritmo AFDMA es robusto a valores de FC mayores a 70 y el algoritmo AAMA es robusto a valores de FC mayores a 80.

Las comparaciones que se realizaron en este capitulo reflejan las ventajas de los algoritmos propuestos con respecto a algoritmos similares en términos de probabilidad de error falso negativo, calidad en la imagen recuperada y compresión JPEG, cabe mencionar que dichas comparaciones se realizaron en condiciones similares.

CAPÍTULO 5

CONCLUSIONES GENERALES Y TRABAJO FUTURO

5.1. CONCLUSIONES GENERALES

De acuerdo al estado del arte realizado en esta tesis es difícil afirmar que algoritmo es el más adecuado y que asegure un servicio eficiente de integridad en imágenes digitales, es decir, no existe una única técnica que pueda ser utilizada en todos los casos, sino que existen distintos tipos de técnicas de marca de agua cuyas propiedades varían de acuerdo al contexto de aplicación para el que fueron desarrolladas.

En esta tesis se propusieron dos algoritmos de autenticación y localización de regiones alteradas en imágenes digitales con marcas de agua semi-frágiles, cuya extracción de la marca de agua es a ciegas, dichos algoritmos los identificamos como: 1) Autenticación basado en Firma Digital como Marca de Agua (AFDMA) y 2) Autenticación y Auto-Recuperación basado en Marca de Agua (AAMA).

En este trabajo de investigación podemos concluir que el haber extraído la marca de agua del contenido de la imagen en el dominio de la transformada, por un lado nos permitió incrementar la robustez de los métodos ante ataques maliciosos como el fotomontaje y por el otro pudimos recuperar la imagen que fue alterada. También podemos decir que esta robustez fue especificada por el coeficiente de cuantificación en la inserción de la marca de agua en los algoritmos AFDMA y AAMA.

De las pruebas realizadas a los algoritmos propuestos en esta tesis se puede concluir lo siguiente:

La imperceptibilidad de la marca de agua en el algoritmo AFDMA tuvo un valor PSNR promedio de 47 dB´s. Además, este algoritmo detecta correctamente las modificaciones intencionales que sufre la imagen marcada, ya que tuvo una probabilidad de error falso negativo del 10% y fue capaz de diferenciar entre una modificación intencional de una no intencional, ya que el algoritmo de verificación propuesto en bloques aislados y concentrados, elimina adecuadamente las modificaciones no intencionales.

Al realizar la comparación del algoritmo AFDMA con los algoritmos propuestos por Xie y por Zhou podemos concluir que aunque la calidad en la imagen marcada es mayor en el algoritmo propuesto por Zhou; la calidad en el algoritmo AFDMA y en el de Xie es aceptable, ya que no se distingue la marca de agua; Otra ventaja que tiene el algoritmo AFDMA con respecto a los otros dos algoritmos es que la probabilidad de error falso negativo es mucho menor. En cuanto a la robustez a la compresión aunque el algoritmo propuesto por Xie es robusto para un factor de calidad mayor a 60 la robustez que presenta el algoritmo AFDMA (75) es muy aceptable, ya que para valores de FC menores a 70, la imagen marcada se degrada y pierde su valor comercial.

La imperceptibilidad de la marca de agua es aceptable aún cuando se inserta el máximo número de bloques ROI en la imagen original, así mismo la calidad en la imagen recuperada es buena, ya que se puede distinguir la información original, bastando solo con conocer la llave secreta utilizada en el proceso de inserción, incrementando así la seguridad del sistema. Otro atributo al novedoso sistema de autenticación y auto recuperación propuesto es su robustez a la compresión JPEG, ya que detecta a la imagen

marcada y comprimida hasta un factor de calidad de 80 como autentica. Por lo cual, podemos concluir que el sistema de auto recuperación propuesto trabaja eficientemente.

Una vez realizada la comparación entre el algoritmo AAMA y los algoritmos propuestos por Zhao y Hassan, podemos concluir que aunque los tres algoritmos detectan y localizan correctamente los bloques erróneos en las imágenes marcadas y alteradas, en cuanto a la calidad de la imagen reconstruida, el algoritmo AAMA es mejor, permitiendo así identificar plenamente cual fue el contenido de la imagen alterado. En cuanto a la compresión JPEG el algoritmo AAMA tiene la ventaja de que es robusto a factores da calidad mayores a 80, el cual es un factor de calidad aceptable, aunque los otros dos algoritmos tienen la capacidad de recuperar los bloques detectados como alterados, tendrían que pasar todavía por el proceso de auto recuperación.

5.2. TRABAJO FUTURO

Como ya mencionamos anteriormente el campo de investigación aún es muy amplio y crecerá aún más conforme las aplicaciones en el campo del procesamiento de imágenes sean más sofisticadas, por lo cual de este trabajo de investigación surgen de manera natural como extensiones del mismo un número de direcciones de investigación:

1) Mientras los algoritmos basados en el dominio DCT sean robustos a compresión con pérdida JPEG, estos permanecerán pero se requiere que también sean robustos a la compresión basada la DWT.
2) Los algoritmos de autenticación y recuperación que usan la DCT deben ser resistentes a cualquier cambio geométrico.

3) Se deben probar algoritmos de compresión de imágenes más compactos que nos ayuden a recuperar la imagen con una mayor calidad utilizando una marca de agua más pequeña.
4) Los algoritmos de auto recuperación deben cubrir la mayor parte de la imagen para poder autenticar y recuperar cualquier región de la misma.

REFERENCIAS

[Arce et al., 1999] Gonzalo Arce, Charles G. Boncelet, Richard F. Graveman and Lisa M. Marvel. *"Applications of Information Hiding"*. In Proc. of Third Annual Federated Laboratory Symposium on Advanced Telecommunications and Information Distribution Research Program, pp. 423-427, February 1999.

[Bender, et. all, 1996] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, *"Techniques for Data Hiding,"* IBM Syst. J., vol. 35, no. 3/4, pp. 313–335, 1996.

[Bhattacharjee, 1998] S. Bhattacharjee and M. Kutter, *"Compression Tolerant Image Authentication"*, in Proc. 5th IEEE International Conference on Image Processing (ICIP '98), pp. 435–439, Chicago, USA, October 1998.

[Bravo, et al., 2008] S. Bravo, L. Gan, A. K. Nandi and A. F. Aburdene, *"Fragile Logo Watermarking for Public Authentication"*, ICASSP, pp. 1669-1672, 2008.

[Coatrieux et al, 2001] G. Coatrieux, B. Sankur, and H. Maitre, *"Strict Integrity Control of Biomedical Images,"* in Security and Watermarking of Multimedia Contents III, vol. 4314 of SPIE Proceedings, San Jose, Calif, USA, January 2001.

[Cox, et. all, 1997] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, *"Secure Spread Spectrum Watermarking for Multimedia," IEEE Trans. Image Processing,* vol. 6, pp. 1673–1687, Dec. 1997.

[Chen, et al. 2001] Chen T., Wang J. and Zhou Y., *"Combined Digital Signature and Digital Watermark Scheme for Image Authentication"*, IEEE, (2001) 78-82, 2001.

[Craver, et. all, 1998] S. Craver, B. Yeo, M. Yeung, *"Technical Trials and Legal Tribulations"*, Communications of the ACM, vol. 41, no.7, pp. 45-54, Jul. 1998.

[Dugelay, 1999] J.-L. Dugelay and S. Roche, *"Process for Marking a Multimedia Document, Such an Image, by Generating a Mark,"* Pending patent EP 99480075.3, EURECOM 11/12 EP, July 1999.

[Enomoto, Shibata, 1971] H. Enomoto and K. Shibata, *"Orthogonal Transform Coding System for Television Signals"* IEEE Trans. Electromagn. Compa., vol. EMC-13 pp. 11-17 Aug. 1971.

[Feng, Liu, 2008] W. Feng and Zhi-Quiang Liu, *"Region-Level Image Authentication Using Bayesian Structural Content Abstraction"*, IEEE Transactions on Image Processing, vol. 17, No. 12, pp. 2413-2424, December 2008.

[Frank, 1961] Emil Frank Hembrooke, *"Identification of sound and like signals"*, United States Patent, 3,004,104, 1961.

[Fridrich, et al., 2001] J. Fridrich, M. Goljan, and R. Du, *"Invertible Authentication,"* in Proc. SPIE Conf. Security and Watermarking of Multimedia Contents III, vol. 4314, pp. 197–208, San Jose, Calif, USA, January 2001.

[Fridrich, 1998a] J. Fridrich, *"Image Watermarking for Tamper Detection,"* in Proc. IEEE International Conference on Image Processing, vol. 2, pp. 404–408, Chicago, Ill, USA, October 1998.

[Fridrich, 1998b] J. Fridrich, *"Methods for Detecting Changes in Digital Images,"* in Proc. IEEE International Conference on Image Processing, Chicago, Ill, USA, October 1998.

[Fridrich, 1999a] J. Fridrich, *"Robust Bit Extraction from Images"*, Proceedings of IEEE International Conference on Multimedia Computing and Systems (ICMCS'99), Vol. 2, pp. 536-540, 1999.

[Fridrich, 1999b] J. Fridrich and M. Goljan, *"Protection of Digital Images Using Self Embedding,"* in Symposium on Content Security and Data Hiding in Digital Media, New Jersey Institute of Technology, Newark, NJ, USA, May 1999.

[Hassan et al., 2008] Y. M. Y. Hassan and A. M. Hassan, *"Tamper Detection with Self-Correction Hybrid Spatial-DCT Domains Image Authentication Technique"*, Communication Systems Software and Middleware and Workshops COMSWARE, pp. 608-613, 2008.

[Hernández, et. all, 1998] J. R. Hernández, F. Pérez-González, and J. M. Rodríguez, *"Coding and Synchronization: A boost and a Bottleneck for the Development of image Watermarking,"* in Proc. of the COST #254 workshop on Intelligent Communications, (L'Aquila, Italia), pp. 77–82, SSGRR, June 1998.

[Herrigel, et. all, 1998] A. Herrigel, J. O'Ruanaidh, H. Petersen, S. Pererira, and T. Pun, *"Secure Copyright Protection Techniques for Digital Images,"* in *Information Hiding* (D. Aucsmith, ed.), vol. 1525 of *Lecture Notes in Computer Science*, (Berlin), pp. 169–190, Springer-Verlag, 1998.

[Ho, et al., 2007] A. T. S. Ho, X. Zhu and L. H. Tang, "*Digital Watermarking Authentication and Restoration for Chinese Calligraphy Images*", Proc. Of the 2007 15th Intl. Conf. on Digital Signal Processing (DSP 2007), pp. 483 – 486, 2007.

[Ho, 2007] A.T.S. Ho, "*Semi-fragile Watermarking and Authentication for Law Enforcement Applications*" Innovative Computing, Information and Control, 2007. ICICIC '07. Second International Conference on 5-7 Sept. 2007, pp. 286 – 286, 2007.

[Hu, Chen, 2007] Y.-P. Hu, Z.-G. Chen, *"An SVD-Based Watermarking Method for Image Authentication"*, Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007, pp. 1723-1728, 2007.

[Inoue, et al. 2000] H. Inoue, A. Miyazaki and T. Katsura, *"A Digital Watermark for Images Using the Wavelet Transform"*, Integrated Computer – Aided Engineering, Vol. 7, No. 2, pp. 105-115, 2000.

[Johhnson and Jajodia , 1998] JOHHNSON Neil F. and JAJODIA Jushil. Steganography: Seeing the Unseen [en linea]. IEEE Computer, February 1998 [citado 1 septiembre 2009]. Disponible World Wide Web: <http://www.jjtc.com/pub/r2026.pdf>

[Katzenbeisser and Petitcolas, 2000] Stefan Katzenbeisser and Fabien A.P. Petitcolas (Editores) Artech House, *"Information Hiding Techniques for Steganography and Digital Watermarking"*, Ene. 2000, ISBN: 1580530354

[kundur, 1999] D. Kundur and D. Hatzinakos, *"Digital Watermarking for Telltale Tamper Proofing and Authentication"*, Proceedings of the IEEE, 87(7); pp. 1167-1180, July 1999

[Lin, Chang, 2000] C.-Y. Lin and S.-F. Chang, *"Semi-fragile Watermarking for Authenticating JPEG Visual Content,"* in Proc. SPIE International Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, San Jose, Calif, USA, January 2000.

[Lin, 1998a] C.-Y. Lin and S.-F. Chang, *"Generating Robust Digital Signature for Image/Video Authentication,"* in Proc. Multimedia and Security Workshop at ACM Multimedia '98, Bristol, UK, September 1998.

[Lin, 1998b] C.-Y. Lin and S.-F. Chang, *"A robust Image Authentication Method Surviving JPEG Lossy Compression,"* in Proc. SPIE Storage and Retrieval of Image/Video Database, vol. 3312, San Jose, Calif, USA, pp. 296–307, January 1998.

[Lin, et al., 2004a] P. -L. Lin, P. -W. Huang, A. -W. Peng, *"A Fragile Watermarking Scheme for Image Authentication with Localization and Recovery*", Proc. of the IEEE sixth Int. Symp. on Multimedia Software Engineering, 2004, pp.146-153.

[Lin, et al., 2004b] P. -L. Lin, C. -K. Hsieh, P. -W. Huang, "*Hierarchical Watermarking Scheme for Image Authentication and Recovery*", IEEE Int. Conf. on Multimedia and Expo, 2004, pp. 963-966.

[Lu, Liao, 2003] C. Lu and H. Liao, *"Structural Digital Signature for Image Authentication: An Incidental Distortion Resistant Scheme," IEEE Trans. Multimedia*, vol. 5, no. 2, pp. 161–173, Jun. 2003.

[Memon and Wong, 1998] N. Memon , P. W. Wong, *"Protecting Digital Media Content"*, Communications of the ACM, Jul. 1998, vol. 41, no.7, pp. 35-43

[Ó Ruanaidh, 1997] J. J. K. Ó Ruanaidh and T. Pun, *"Rotation, Scale and Translation Invariant Digital Image Watermarking,"* in Proc.

IEEE International Conference on Image Processing, vol. 1, Santa Barbara, Calif, USA, pp. 536–539,October 1997.

[Papoulis, 1991] A. Papoulis, "*Probability and Statistics".* Englewood Cliffs, NJ: Prentice-Hall, 1991.

[Pérez, Hernández,] Fernando Pérez-González and Juan R. Hernández, *"A Tutorial on Digital Watermarking",* Work partially funded by CICYT under project TIC-96-0500-C10-10

[Peticolas et al. 1999] F. A. P. Peticolas, R. J. Anderson and M. G. Kuhn, *"Information Hiding a Survey",* Proceedings of the IEEE, 87(7): pp. 1062-1078, July 1999.

Pitas and Kaskalis, 1995] I. Pitas and T. H. Kaskalis, *"Applying Signatures on Digital Images,"* in *Proc. 1995 IEEE Workshop Nonlinear Signal and Image Processing,* North Marmaras, Greece, June 20–22, 1995, pp. 460–463.

[Proakis, 1995] J. G. Proakis, *"Digital Communications"*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.

[Rey, 2000] C. Rey and J.-L. Dugelay, *"Blind Detection of Malicious Alterations on Still Images Using Robust Watermarks,"* in Secure Images and Image Authentication Colloquium, IEE Electronics & Communications, London, UK, 2000.

[Schneier, 1995] Bruce Schneier, *"Applied Cryptography: Protocols, Algorithms and Source Code in C"*, John Wiley & Sons, Oct. 1995, ISBN: 0471117099

[SHA-1, 1995] Secure Hash Standard [en linea]. U.S Department of Commerce/National Institute of Standards and Technology: Federal information processing standards publication, April 1995 [citado 1 septiembre 2009]. Disponible World Wide Web: <http://www.itl.nist.gov/fipspubs/fip180-1.htm>.

[Stinson, 1995] Douglas R. Stinson, "Cryptography: Theory and Practice (Discrete Mathematics and its Applications", CRC Press, Mar. 1995, ISBN: 0849385210

[Swanson, et. all, 1998] M. D. Swanson, B. Zhu, A. H. Tewfik, and L. Boney, *"Robust Audio Watermarking Using Perceptual Masking," Signal Processing,* vol. 66, no. 3, pp. 337–355, 1998.

[Tirkel, et. all, 1994] A. Z. Tirkel, R. G. Schyndel, and C. F. Osborne, *"A Digital Watermark,"* in *Proc. ICIP'94,* 1994, vol. II, pp. 86–90.

[Tirkel, et. all, 1998] A. Z. Tirkel, C. F. Osborne, and T. E. Hall, *"Image and Watermark Registration,"* Signal Processing, vol. 66, no. 3, pp. 373–383, 1998.

[Tsai, Hu, 2005] P. Tsai, Y.-C. Hu, "*A Watermarking-Based Authentication with Malicious Detection and Recovery*", Int. Conf. of Information, Communication and Signal Processing, 2005, pp. 865-869.

[Van, 1994] R. G. Van Schyndel, A. Z. Tirkel, and C. F. Osborne, *"A Digital Watermark,"* in Proc. IEEE International Conference on Image Processing, vol. 2, Austin, Texas, USA, pp. 86–90, November 1994.

[Voyatzis and Pitas, 1998] G. Voyatzis and I. Pitas, *"Digital Image Watermarking Using Mixing Systems,"* Comput. Graphics, vol. 22, no. 3, pp. 405–416, 1998.

[Walton, 1995] S. Walton, *"Information Authentication for a Slippery New Age,"* Dr. Dobbs Journal, vol. 20, no. 4, pp. 18–26, 1995.

[Wolfgang, 1999] R. B. Wolfgang and E. J. Delp, *"Fragile Watermarking Using the VW2D Watermark,"* in Security and Watermarking of Multimedia Contents, vol. 3657 of SPIE Proceedings, San Jose, Calif, USA, p.40–51, January 1999.

[Wolfgang, 1996] R. B. Wolfgang and E. J. Delp, *"A watermark for Digital Images,"* in Proc. 1996 IEEE International Conference on Image Processing, vol. 3, Lausanne, Switzerland, pp. 219–222, September 1996.

[Wu, Liu, 1998] M. Wu and B. Liu, *"Watermarking for Image Authentication,"* in Proc. IEEE International Conference on Image Processing, vol. 2, Chicago, Ill, USA, pp. 437–441, October 1998.

[Xie et al., 2007] Xie R., Wu K., LI Ch. and Zhu S.: *"An Improved Semi-fragile Digital Watermarking Scheme for Image Authentication"*, IEEE, (2007)

[Yeung, 1997] M. M. Yeung and F.Mintzer, *"An Invisible Watermarking Technique for Image Verification,"* in Proc. IEEE International Conference on Image Processing, vol. 2, Santa Barbara, Calif, USA, pp. 680–683, October 1997.

[Yeung, 1998] Minerva M. Yeung, *"Digital Watermarking"*, Communications of the ACM, Jul. 1998, vol. 41, no.7, pp.31-33

[Zhao et al., 2007] X. Zhao, A.T.S. Ho, H. Treharne, V. Pankajakshan, C. Culnane, W. Jiang, *"A Novel Semi-Fragile Image Watermarking, Authentication and Self-Restoration Technique Using the Slant Transform"*, Intelligent Information Hiding and Multimedia Signal Processing, 2007. IIHMSP 2007. Third International Conference on Volume 1, pp. 283 – 286, 26-28 Nov. 2007.

[Zhang, Huang, 2008] Han-ling Zhang, Sheng Huang, *"A Novel Image Authentication Robust to Geometric Transformations"*, Congress on Image and Signal Processing, IEEE Computer Society, pp. 654-658, 2008.

[Zhou et al, 2004] Zhou X., Duan X. and Wang D., *"A Semi-Fragile Watermark Scheme for Image Authentication"*, Proc. of Int. Conf. Multimedia Modeling Conference, 2004.

APÉNDICE A

GLOSARIO

AAMA	Autenticación y Auto-Recuperación basado en Marca de Agua
AC	Valores superpuestos o análogos de la DCT
AFDMA	Autenticación basado en Firma Digital como Marca de Agua
ASCII	Código Estadounidense Estándar para el Intercambio de Información
BaSCA	Abstracción de Contenido Estructural Bayesiano
BCH	Código Corrector de Errores (Acrónimo de las iniciales de sus inventores)
CD-ROM	Disco Compacto de Solo Lectura
CPTWG	Grupo de Trabajo Técnico en la Protección de Copiado
CRC	Código de Redundancia Cíclica
dB	Decibel
DC	Valor promedio de los coeficientes DCT
DCT	Transformada Coseno Discreto
DSCQS	Escala de Calidad Continua de Doble Estímulo
DVD	Disco Versátil Digital
DWT	Transformada Discreta Wavelet

FPB	Filtro Pasa Bajas
FT	Transformada de Fourier
IDWT	Transformada Discreta Wavelet Inversa
ISLT	Transformada Slant Inversa
ITU	Unión de Telecomunicaciones Internacional
JPEG	Grupo de Expertos en Fotografía
LSB	Bit Menos Significativo
MD	Algoritmo de Resumen de Mensaje
MSB	Bit Más Significativo
PDF	Función de Distribución de Probabilidad
$\mathbf{P_{fa}}$	Probabilidad de Falsa Alarma
$\mathbf{P_{rej}}$	Probabilidad de Falso Negativo
PSNR	Relación Señal a Ruido Pico
PST	Transformada Seno Pinned
RGB	Sistema de señal de video que utiliza las señales Rojo, Verde y Azul por separado.
ROE	Región de Inserción
ROI	Región de Interés
SDMI	Iniciativa de Música Digital Segura
SHA	Algoritmo Hash Seguro
SLT	Transformada Slant
SNR	Relación Señal a Ruido

SPIHT	Algoritmo de Árboles Jerárquicos en Partición por Grupos
SVD	Descomposición de Valor Singular
SVH	Sistema Visual Humano
VW2D	Marca de Agua de Dos Dimensiones
XOR	O exclusiva

APÉNDICE B

IMÁGENES UTILIZADAS

A continuación se muestran algunas imágenes en escala de grises a 8 bits de tamaños 512x512, 256x256 y 128x128, las cuales fueron utilizadas en los experimentos de los algoritmos ADFMA y AAMA.


Barbara


Barco

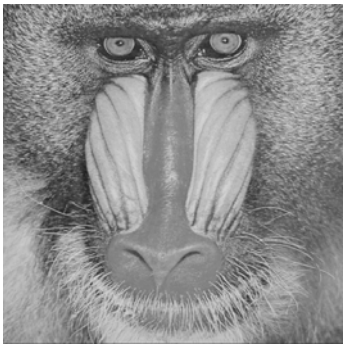

Camera


Bridge


Chiles


Goldhill


Lena


Mandrill


Pajaro

Placa


Avión 1

Avión 2


Desayuno 1


Gato


Tren 1


Tren 2


Tren 3


Tren 4


Tren 5


Tren 6


Tren 7

Paisaje 1

Animales 1


Paisaje 2


Paisaje 3


Gente 1


Tren 8


Gente 2


Paisaje 4


Paisaje 5

Paisaje 6


Paisaje 7

Desayuno 2


Flores 1

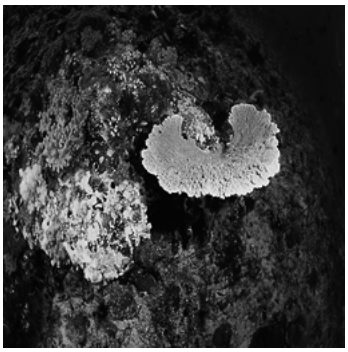

Flores 2


Animales 2


Animales 3

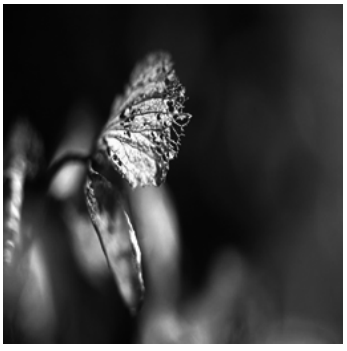

Gente 3


Gente 4


Flores 3


Paisaje 8


Paisaje 9


Animales 4


Paisaje 10


Paisaje 11


Paisaje 12


Gente 5


Paisaje 13


Paisaje 14


Paisaje 15


Paisaje 16


Paisaje 17


Gente 6


Tren 9


Paisaje 18


Flores 4


Paisaje 19

A continuación se muestran algunas imágenes a color a 24 bits de tamaños 512x512, 256x256 y 128x128, las cuales fueron utilizadas en los experimentos del algoritmo ADFMA.

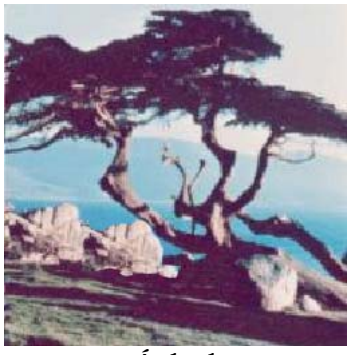

Avión


Casa


Chica

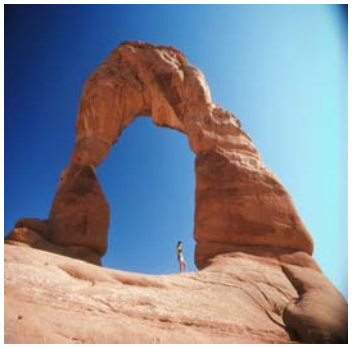

Lago


Montaña

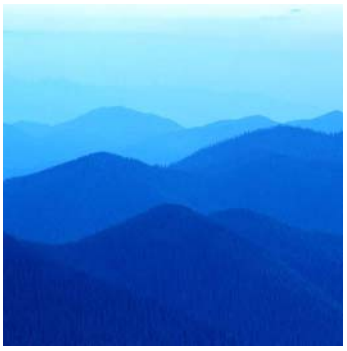

Árbol


Gente


Chiles


Paisaje 1


Paisaje 2


Paisaje 3


Pez


Paisaje 4


Paisaje 5


Paisaje 6


Flores 1


Paisaje 7

Flores 2

Flores 3

Flores 4

Paisaje 8


Telaraña


Paisaje 9


Paisaje 10


Paisaje 11


Paisaje 12


Paisaje 13


Pintura 1


Pintura 2


Paisaje 14


Paisaje 15


Pintura 3


Pintura 4

Paisaje 16

Paisaje 17

Paisaje 18

Paisaje 19

Paisaje 20

Paisaje 21

Pintura 5


Pintura 6


Pintura 7


Pintura 8


Pintura 9


Pintura 10

APÉNDICE C

CÓDIGO FUENTE

A continuación se muestra el código fuente de los algoritmos de inserción y extracción de la marca de agua, del algoritmo AFDMA, los cuales fueron realizados en Matlab 7.0

```
%
%              Inserción de la Marca de agua
%
%     Este programa realiza siguientes operaciones
%
%     1. Obtener la firma digital por bloque de la imagen original

clear all;
close all;

ruta_res='C:\Documents and Settings\Clarita\My Documents\Clara\
doctorado\programas\imagen_gris\';
[X,Map,nom_arch_res]=get_image(ruta_res);

figure(1);
imshow(X,Map);

[T1,T2]=size(X);
TB=16; %    TB: Tamaño de bloque

NB1=T1/TB;
NB2=T2/TB;
Bloque=zeros(TB,TB,NB1*NB2);
Bloque_w=zeros(TB,TB,NB1*NB2);

Q=9;
N=7; % longitud de código

%%%%% obtención de la firma digital %%%%%%

llave2=123;
[M,Bloque]=ext_bit_robust(X,TB,N,llave2); %M:  firma digital codificada
```

```
%     2. Insertar la marca de agua por bloque en la imagen original

Y=zeros(TB,TB);
key=llave;
for num_bloque=1:NB1*NB2
   Y=Bloque(1:TB,1:TB,num_bloque);
   [WLL,WHL,WLH,WHH]=dwtper2(double(Y),'db2');  % Transformada DWT
                                                    periódica
   for i=1:N  % N=7: longitud de marca de agua para cada bloque
      Coef(i)=WLL(key(i));
      Coef_w=coef_marca(M(num_bloque,i),Coef(i),Q);
      WLL(key(i))=Coef_w;
   end
      YW=idwtper2(WLL,WHL,WLH,WHH,'db2');  % Transformada inversa de
                                              la DWT periódica
      Bloque_w(1:TB,1:TB,num_bloque)=YW;
end
num_bloque=1;
for i=1:NB1
   for j=1:NB2
      Xw((i-1)*TB+1:i*TB,(j-1)*TB+1:j*TB)=Bloque_w(1:TB,1:TB,num_bloque);
      num_bloque=num_bloque+1;
   end
end

Xw=uint8(Xw);
figure(2);
imshow(Xw,Map);

R=PSNR(X,uint8(Xw));
fprintf(' PSNR= %f \n',R);  %  fidelidad de la imagen marcada respecto a la
                               imagen original
imwrite(Xw,Map,nom_arch_res);
```

Función que extrae los bits robustos

function [SB,B]=ext_bit_robust(X,BS,N,K);

```
[T1,T2]=size(X);
Num_bloque=1;
for i=1:fix(T1/BS)
  for j=1:fix(T2/BS)
    B(:,:,Num_bloque)=X((i-1)*BS+1:i*BS,(j-1)*BS+1:j*BS);
    Num_bloque=Num_bloque+1;
```

```
   end
end
rand('seed',K);
for n=1:N
   P(:,:,n)=rand(BS,BS);
end
OP=P;

%%%%%%%% aplicar filtro pasa baja repetidamente %%%%%%%%%%

for rep=1:7   %%%%%%% número de repetición
  for n=1:N
     P(:,:,n)=filtro_pb(P(:,:,n),1);
  end
end

%%%%%%% Generar DC-free %%%%%%%%

for n=1:N
   P(:,:,n)=P(:,:,n)-sum(sum(P(:,:,n)))/(BS*BS);
end
Num_bloque=Num_bloque-1;
k=0;
for bn=1:Num_bloque
   BB=double(B(:,:,bn));
   for n=1:N
      PP=double(P(:,:,n));
      proj(bn,n)=abs(sum(sum(BB.*PP)));
   end
end

Th=median(proj(:));

for bn=1:Num_bloque
   SB(bn,:)=proj(bn,:)>Th;
end
```

Función que cuantifica los valores de los coeficientes de acuerdo al valor de la marca de agua

```
function Cw=coef_marca(Mb,C,Q)


if mod(C,Q)<Q/2
```

```
   cw0=C-mod(C,Q);
else
   cw0=C+(Q-mod(C,Q));
end

[val,pos]=min([abs(C-(cw0-Q/2)),abs(C-(cw0+Q/2))]);

switch pos
case 1, cw1=cw0-Q/2;
case 2, cw1=cw0+Q/2;
end

if Mb==0
   Cw=cw0;
else
   Cw=cw1;
end

%
%     Extracción de la Marca de agua y autenticación de la imagen por
        bloque
%
%     Este programa realiza siguientes operaciones
%
%     1. Extracción de marca de agua por bloque de la imagen
        marcada


clear all;
close all;

%%%%% Obtener la firma digital de la imagen bajo análisis

ruta_res='C:\Documents and Settings\Clarita\My Documents\ Clara\
doctorado\programas\Robust_bit\';
nom_im=input('Teclea nombre de imagen: ','s');
nom_imw=strcat(ruta_res,nom_im);
[X,Map]=imread(nom_imw,'bmp');

[T1,T2]=size(X);
TB=16;

NB1=T1/TB;
NB2=T2/TB;
Bloque=zeros(TB,TB,NB1*NB2);
```

```
Bloque_w=zeros(TB,TB,NB1*NB2);
llave2=123;
Q=9;
N=7; % longitud de codigo
[M,Bloque]=ext_bit_robust(X,TB,N,llave2);
key=llave;
for num_bloque=1:NB1*NB2
   Y=Bloque(1:TB,1:TB,num_bloque);
   [WLL,WHL,WLH,WHH]=dwtper2(double(Y),'db2');  % Transformada de
                                                          DWT periódica
   for i=1:N  % N=7: longitud de marca de agua para cada bloque
      Coef_w=WLL(key(i));
      M_ext(num_bloque,i)=ext_bit(Coef_w,Q);
   end
end

%     2. Autenticación

err_count=0;
err_bloque=[];
for num_bloque=1:NB1*NB2
  if sum(M(num_bloque,:)==M_ext(num_bloque,:))<6  % error: cuando
                                                  menos un bit de diferencia
     err_bloque=[err_bloque,num_bloque];
     err_count=err_count+1;
  end
end

fprintf(' Numero de bloques con error=%d \n',err_count);
fprintf(' bloques erroneos:');
fprintf(' %d ',err_bloque);
fprintf(' \n');

Z=X;

for i=1:length(err_bloque)
  b_err=err_bloque(i);
  pos_ren=fix(b_err/NB1)+1;
  pos_col=mod(b_err,NB2);
  if pos_col==0
     pos_col=NB2;
     end
 Z(TB*(pos_ren-1)+1:TB*pos_ren,TB*(pos_col-1)+1:TB*pos_col)=ones(TB);
end
figure(2);
```

```
imshow(Z,Map);
```

Función que extrae los bits de marca de agua del bloque marcado

function M=ext_bit(C,Q)

```
Cq1=Q*round(C/Q);
Cq2=Cq1+Q/2;
Cq3=Cq1-Q/2;
if abs(C-Cq1)<abs(C-Cq2) & abs(C-Cq1)<abs(C-Cq3)
   M=0;
   else
   M=1;
end
```

A continuación se muestra el código fuente de los algoritmos de inserción y extracción de la marca de agua, del algoritmo AAMA, los cuales fueron realizados en Matlab 7.0

```
%
%     Inserción de la Marca de agua que contiene el coeficiente
       DC y los 6 primeros coeficientes AC con sus respectivos signos
%
%    Este programa realiza siguientes operaciones
%
%    1. Obtener  la Región de Interés (ROI)

clear all;
close all;

ruta='C:\Documents and Settings\Clarita\My Documents\ Clara\
doctorado\programas\imagen_gris\';
ruta2='C:\Documents and Settings\Clarita\My Documents\ Clara\
doctorado\programas\Corrección\PROGRAMA\imagenes\';
ruta3='C:\Documents and Settings\Clarita\My Documents\Clara\
doctorado\programas\Corrección\PROGRAMA\coeficientes\';

Logo_arch='bridge.bmp';

nom_arch=strcat(ruta,Logo_arch);
nom_arch2=strcat(ruta2,Logo_arch(1:(length(Logo_arch)-4)));
nom_arch3=strcat(ruta3,Logo_arch(1:(length(Logo_arch)-4)));

[Xori,Map]=imread(nom_arch,'bmp');
X=double(Xori)-128;
```

```
figure(1);
imshow(Xori,Map);

X2=X;

[T1,T2]=size(X);

fprintf('      Indicar región de interés usando mouse (botton izquierdo) \n');
fprintf('      Finalización es con el botón derecho de mouse \n');

i=1;
while 1
   [cp(i),rp(i),botton]=ginput(1);
   if botton==3
      break;
   end
   if cp(i)>=1 & cp(i)<=T1 & rp(i)>=1 & rp(i) <=T2
      i=i+1;
   end

end

L=length(rp);
rp=rp(1:L-1);
cp=cp(1:L-1);
L=L-1;
rp=round(rp);
cp=round(cp);

L=length(rp);

%%%% obtener numero de bloque partir de (rp,cp);
ind=1;
for k=1:L
   [Num_bloque,pos_i,pos_j]=get_block_num(rp(k),cp(k),T1,T2);
   [V,num_vecino]=vecinos8(pos_i,pos_j,T1,T2);
   B_relevante(ind,:)=[Num_bloque,pos_i,pos_j];
   B_relevante(ind+1:ind+num_vecino,:)=V;
   ind=ind+num_vecino+1;
end

%%%% Desplegar bloques para confirmación %%%%%

%%%% Eliminar duplicación %%%%%
B_rel2=unique(B_relevante,'rows');
```

```
Y=Xori;
L2=size(B_rel2,1);
B_negro=ones(8,8);
for k=1:L2
   ip=B_rel2(k,2);
   jp=B_rel2(k,3);
   Y((ip-1)*8+1:ip*8,(jp-1)*8+1:jp*8)=B_negro;
end

%     2. Extraer la marca de Agua  de la Región de Interés

D=unique((B_relevante(:,1))');
s=size(D',1);
t=1;
T=zeros(s*6,12);
[Z]=chusyutu_6coef_tot(X); %Z contiene la marca de agua en binario de
                                    todos los bloques de la imagen
for i=1:s
    k=D(i);
    for j=1:6
    T(t,1:11)=Z(k,12*(j-1)+1:11*j); % T contiene la marca de agua en binario
                                        de la región de interés pero dividid en 6
                                         bloques de 11 bits cada uno
    t=t+1;
    end
end

%     3. Inserta la marca de Agua  de la Región de Inserción (ROE)

key=input('Key No : ');
V=ransu(X,B_rel2,key); %% Ordena aleatoriamente los bloques fuera de la
                                    región de interés
k=0;
for i=1:s*6
    k=k+1;
    R=ryousika(X,V(i)); % Aplica la DCT y cuantiza al bloque en donde se va a
                                insertar la marca de agua
    R=dainyu(k,R,T); % Inserta la marca de agua a los coeficientes
                            cuantizados de frecuencia media
    R2=gyaku_ryousika(R); %%% Decuantiza los coeficientes marcados
    t=V(i);

    %%% Sustituye los bloques marcados por los que no estaban marcados
```

```
   if T1==512
      X2(8*(fix((t-1)/64)+1)-7:8*(fix((t-1)/64)+1),8*(mod(t-1,64))+1:8*(mod(t-
1,64))+8)=R2;
   else
      X2(8*(fix((t-1)/32)+1)-7:8*(fix((t-1)/32)+1),8*(mod(t-1,32))+1:8*(mod(t-
1,32))+8)=R2;
   end
end
X2=X2+128;
```

Función que extrae las posiciones de los bloques de la región ROI

function [Num_bloque,pos_i,pos_j]=get_block_num(rp,cp,T1,T2)

```
BS=8;
NB1=fix(T1/BS);
NB2=fix(T2/BS);

pos_i=fix(rp/BS)+1;
pos_j=fix(cp/BS)+1;

if pos_i>NB1
   pos_i=NB1;
end
if pos_j>NB2
   pos_j=NB2;
end

Num_bloque=(pos_i-1)*NB2+pos_j;
```

Función que localiza los bloques vecinos de 8x8 del bloque ROI seleccionado

function [V,Num_vecinos]=vecinos8(pos_i,pos_j,T1,T2)

```
BS=8;
BN1=fix(T1/BS);
BN2=fix(T2/BS);

if pos_i<2

   if pos_j<2
      Num_vecinos=3;
      Vecinos=[pos_i,pos_j+1;pos_i+1,pos_j;pos_i+1,pos_j+1];
```

```
  elseif pos_j>BN2-1
    Num_vecinos=3;
    Vecinos=[pos_i,pos_j-1;pos_i+1,pos_j;pos_i+1,pos_j-1];
  else
    Num_vecinos=5;
    Vecinos=[pos_i,pos_j-1;pos_i,pos_j+1;pos_i+1,pos_j-
1;pos_i+1,pos_j;pos_i+1,pos_j+1];
  end

elseif pos_i>BN1-1

  if pos_j<2
    Num_vecinos=3;
    Vecinos=[pos_i,pos_j+1;pos_i-1,pos_j;pos_i-1,pos_j+1];

  elseif pos_j>BN2-1
    Num_vecinos=3;
    Vecinos=[pos_i,pos_j-1;pos_i-1,pos_j;pos_i-1,pos_j-1];
  else
    Num_vecinos=5;
    Vecinos=[pos_i,pos_j-1;pos_i,pos_j+1;pos_i-1,pos_j-1;pos_i-1,pos_j;pos_i-
1,pos_j+1];
  end

else  % pos_i>=2 & pos_i <=T1-1

  if pos_j<2
    Num_vecinos=5;
    Vecinos=[pos_i-1,pos_j;pos_i-
1,pos_j+1;pos_i,pos_j+1;pos_i+1,pos_j;pos_i+1,pos_j+1];

  elseif pos_j>BN2-1
    Num_vecinos=5;
    Vecinos=[pos_i-1,pos_j;pos_i-1,pos_j-1;pos_i,pos_j-
1;pos_i+1,pos_j;pos_i+1,pos_j-1];

  else
    Num_vecinos=8;
    Vecinos=[pos_i-1,pos_j-1;pos_i-1,pos_j;pos_i-1,pos_j+1;pos_i,pos_j-
1;pos_i,pos_j+1;pos_i+1,pos_j-1;pos_i+1,pos_j;pos_i+1,pos_j+1];
  end
end

for i=1:Num_vecinos
  V(i,2:3)=Vecinos(i,:);
```

```
   V(i,1)=(V(i,2)-1)*fix(T1/BS)+V(i,3);
end
```

Función que extrae la marca de agua de los bloques ROI

function [Z]=chusyutu_6coef_tot(X)

```
[S1,S2]=size(X);
BS1=8;
BS2=8;
B1=fix(S1/BS1);
B2=fix(S2/BS2);
Block=zeros(B1*B2,BS1,BS2);
count=1;
for i=1:B1
   for j=1:B2
   Block(count,1:BS1,1:BS2)=X((i-1)*BS1+1:i*BS1,(j-1)*BS2+1:j*BS2);
      count=count+1;
   end
end
T1=fix(BS1/2);
T2=fix(BS2/2);
Block_length=size(Block,1);
for k=1:Block_length
   B=reshape(Block(k,:,:),BS1,BS2);
   Num_coef=6;
   H(k,:)=zeros(1,72);
   coef_B=dct2(B);
   coef_DC=coef_B(1,1);

coef_choise=[coef_DC,coef_B(1,2),coef_B(2,1),coef_B(3,1),coef_B(2,2),coef_B(1
,3),coef_B(1,4)];
   sig_coef_choise=sign(coef_choise);
   count2=0;
   r=length(sig_coef_choise);
   for m=1:r
      if sig_coef_choise(m)==-1
         sig=0;
      else
         sig=1;
      end
      if m==1
         binary_sec_coefDC=double(dec2bin(abs(coef_choise(m)),12))-48;
         binary_sec=[sig,binary_sec_coefDC];
         H(k,1:13)=binary_sec;
```

```
        end
        if m>1
            binary_sec=double(dec2bin(abs(coef_choise(m)),8))-48;
            d=length(binary_sec);
            if d>8
                binary_sec=binary_sec(1:8);
            end
            binary_sec=[sig,binary_sec];
            H(k,13+(((count2-1)*9)+1):13+(count2*9))=binary_sec;
        end
        count2=count2+1;
    end
     H(k,68:72)=0;
end
Z=H;
```

Función que selecciona y cuantifica el bloque ROE

function R=ryousika(X,t)

```
[G,H]=size(X);

% Selecciona el bloque en donde se va a insertar la marca de agua
if G==512
A=X(8*(fix((t-1)/64)+1)-7:8*(fix((t-1)/64)+1),8*(mod(t-1,64))+1:8*(mod(t-
1,64))+8);
else
    A=X(8*(fix((t-1)/32)+1)-7:8*(fix((t-1)/32)+1),8*(mod(t-1,32))+1:8*(mod(t-
1,32))+8);
end
QF=70;
Q=Q_matriz(QF); % Genera la matriz de cuantificacion
B=dct2(A);      % Aplica la transformada DCT
R=round(B./Q);  % Cuantifica los coeficientes DCT
```

Función que inserta la marca binaria en los bits menos significativos de los coeficientes DCT cuantizados de frecuencia media

function R=dainyu(k,R,T)

```
T(k,1);
if R(2,3)<0
 binary_sec=double(dec2bin((abs(R(2,3))),8))-48;
 binary_sec(8)=T(k,1);
 Pix_value=bin2dec_func(binary_sec(1:8));
```

```
 R(2,3)=-Pix_value;
else
 binary_sec=double(dec2bin((abs(R(2,3))),8))-48;
 binary_sec(8)=T(k,1);
 Pix_value=bin2dec_func(binary_sec(1:8));
 R(2,3)=Pix_value;
end
T(k,2);
if R(3,2)<0
 binary_sec=double(dec2bin((abs(R(3,2))),8))-48;
 binary_sec(8)=T(k,2);
 Pix_value=bin2dec_func(binary_sec(1:8));
 R(3,2)=-Pix_value;
else
 binary_sec=double(dec2bin((abs(R(3,2))),8))-48;
 binary_sec(8)=T(k,2);
 Pix_value=bin2dec_func(binary_sec(1:8));
 R(3,2)=Pix_value;
end

%%% Se repite hasta T(k,11)
 R;
```

Función que decuantiza los coeficientes DCT marcados

function R2=gyaku_ryousika(R)

```
QF=70;
Q=Q_matriz(QF);
A=R.*Q;
R2=round(idct2(A));

%
%     Extracción de la marca e agua, Autenticación y Recuperación de
      la imagen
%
%    Este programa realiza siguientes operaciones
%
%    1. Extracción de la marca de agua

%  leer imagen

clear all;
close all;
clc;
```

```
%%%%%%%%% Obtener marca de agua de la imagen bajo analisis

ruta2='C:\Documents and Settings\Clarita\My Documents\Clara\
doctorado\programas\Corrección\PROGRAMA\imagenes\';
nom_im=input('Teclea nombre de imagen: ','s');
nom_imw=strcat(ruta2,nom_im);
[Xori,Map]=imread(nom_imw,'bmp');
figure(1);
imshow(Xori,Map);
X=double(Xori)-128;
X2=Xori;
X3=X;
[T1,T2]=size(X);

L=length(rp);
ind=1;
for k=1:L
   [Num_bloque,pos_i,pos_j]=get_block_num(rp(k),cp(k),T1,T2);
   [V,num_vecino]=vecinos8(pos_i,pos_j,T1,T2);
   B_relevante(ind,:)=[Num_bloque,pos_i,pos_j];
   B_relevante(ind+1:ind+num_vecino,:)=V;
   ind=ind+num_vecino+1;
end

%%%% Desplegar bloques para confirmación %%%%%

%%%% Eliminar duplicación %%%%%%
B_rel2=unique(B_relevante,'rows');

Y=Xori;
L2=size(B_rel2,1);
B_negro=ones(8,8);
for k=1:L2
   ip=B_rel2(k,2);
   jp=B_rel2(k,3);
   Y((ip-1)*8+1:ip*8,(jp-1)*8+1:jp*8)=B_negro;
end

figure(2);
imshow(Y,Map);

%% Extracción  de la marca de la región de interés

D=unique((B_relevante(:,1))');
s=size(D',1);
```

```
t=1;
T_interes=zeros(s*6,12);
[Z]=chusyutu_6coef_tot(X);
for i=1:s
   k=D(i);
   for j=1:6
   T_interes(t,1:12)=Z(k,12*(j-1)+1:12*j);
   t=t+1;
   end
end
```

%% Extracción de la marca de los bloques marcados

```
key=input('Key No : ');
V=ransu(X,B_rel2,key); %% Ordena aleatoriamente los bloques fuera de la
region de interes
D=unique((B_rel2(:,1))');
s=size(D',1);
T=zeros(s*6,12);
k=0;
for i=1:s*6%*2  % *2 porque se inserta 2 veces
   k=k+1;
   R=ryousika(X,V(i));
   T=toridashi(k,R,T);
end
T_ext=T(1:s*6,:)
T_dif=xor(T_interes,T_ext)
S=reshape(T_ext',1,s*6*11);
S2=reshape(T_dif',1,s*6*11);
p=0;
U=zeros(1,7);
Block=zeros(8,8);
pru=ones(24,24);
```

% ***2. Autenticación de los bloques***

```
for h=1:s
   err=0;
   p=p+1;
   t=D(h);
   A=S((p-1)*72+1:p*72);
   B=S2((p-1)*72+1:p*72);
   G=sum(B);
   if G>18
      err=1;
```

```
   end

%     3. Recuperación de los bloques erróneos

% Convertir de binario a decimal  los coeficientes
    i=1;
   for a=1:7
      if a==1
         A2=A(1:13);
        U(a)=bin2dec_func(A2(2:13));
        if A2(1)==0
           U(a)=U(a)*(-1);
        end
      end
        if a>1
           A2=A(1,13+(9*(a-2)+1):13+(9*(a-1)));
           U(a)=bin2dec_func([A2(2:9)]);
           if A2(1)==0
              U(a)=U(a)*(-1);
           end
        end
   end
   Block(1,1)=U(1);
   Block(1,2)=U(2);
   Block(2,1)=U(3);
   Block(3,1)=U(4);
   Block(2,2)=U(5);
   Block(1,3)=U(6);
   Block(1,4)=U(7);
   % Calculo de la DCT inversa
   Block2=idct2(Block);
   t=D(h);
   if T1==512
      X2(8*(fix((t-1)/64)+1)-7:8*(fix((t-1)/64)+1),8*(mod(t-1,64))+1:8*(mod(t-
1,64))+8)=Block2+128;
   else
      X2(8*(fix((t-1)/32)+1)-7:8*(fix((t-1)/32)+1),8*(mod(t-1,32))+1:8*(mod(t-
1,32))+8)=Block2+128;
   end
   if err==1
      Block3=ones(8,8);
      if T1==512
         X3(8*(fix((t-1)/64)+1)-7:8*(fix((t-1)/64)+1),8*(mod(t-
1,64))+1:8*(mod(t-1,64))+8)=Block3;
      else
```

```
            X3(8*(fix((t-1)/32)+1)-7:8*(fix((t-1)/32)+1),8*(mod(t-
1,32))+1:8*(mod(t-1,32))+8)=Block3;
        end
        if T1==512
        X2(8*(fix((t-1)/64)+1)-7:8*(fix((t-1)/64)+1),8*(mod(t-1,64))+1:8*(mod(t-
1,64))+8)=Block2+128;
    else
        X2(8*(fix((t-1)/32)+1)-7:8*(fix((t-1)/32)+1),8*(mod(t-1,32))+1:8*(mod(t-
1,32))+8)=Block2+128;
    end
    end
end
X3=X3+128;
```

APÉNDICE D

PUBLICACIONES

Artículos en revistas internacionales

1. V. Hernández-Guzman; C. Cruz-Ramos; M. Nakano-Miyatake; H. Pérez –Meana, "*Algoritmo de Marca de Agua Basado en la DWT para Patrones Visualmente Reconocibles*", IEEE Latinoamérica indexada a INSPEC y SCOPUS, www.ieee.org/transactions-r9, Vol. 4, issue 4, pp. 257-267, junio, 2006.

2. C. Cruz-Ramos, R. Reyes-Reyes, J. Mendoza-Noriega, M. Nakano-Miyatake, H.Pérez-Meana, "*A Novel Verification System Process to Image Content Authentication Systems Based on Semi-Fragile Watermarking*", Journal and Telecommunication and Radio Engeneering, indexada a SCOPUS, Vol. 67, issue 19, pp. 1777-1790, 2008.

Capítulo en libro

1. C. Cruz-Ramos, R. Reyes-Reyes, M. Nakano-Miyatake, H.Pérez-Meana, "*Image Authentication Scheme Based on Self-embedding Watermarking*", aceptado para su publicación en el Springer LNCS 5856 Volume Editor(s): Prof. Eduardo Bayro-Corrochano, Prof. Jan Olor Eklundh.

Artículos en memorias de Congresos Internacionales

1. C. Cruz-Ramos, R. Reyes-Reyes, M. Nakano-Miyatake, H.Pérez-Meana, "*Robust Image Watermarking System Based in Digital Signature*", 6th Mexican International Conference on Artificial Intelligence, Workshop in Computer Security, Aguascalientes, México, 5-9, Noviembre, 2007

2. C. Cruz-Ramos, R. Reyes-Reyes, M. Nakano-Miyatake, H.Pérez-Meana, "*Image Content Authentication System Based on Semi-Fragile Watermarking*", 51st IEEE International Midwest Symposium on Circuits Systems, Knoxville, Tenessee, 10-13 August, pp. 306-309, 2008.

3. J. Mendoza-Noriega, C. Cruz-Ramos, M. Nakano-Miyatake, H.Pérez-Meana, "*Content Authentication Schemes for Digital Images*", 5th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE 2008), Mexico City, México, 12-14 November, pp. 292-297, 2008.

Artículos en memorias de Congresos Nacionales

1. C. Cruz-Ramos, J. A. Martínez-Ñonthe, R. Reyes-Reyes, M. Nakano-Miyatake, H.Pérez-Meana, "*Extracción e Inserción de Características Robustas a Compresión JPEG en Imágenes Digitales Mediante Marcas de Agua*", Congreso de Instrumentación SOMI XXII, Monterrey, Nuevo León, México, 30 de Septiembre-4 de Octubre, 2007.

2. J. Mendoza-Noriega, C. Cruz-Ramos, J. A. Martínez-Ñonthe, M. Nakano-Miyatake, "*Firma Digital Estructural para Autenticación de Imágenes*", Reunión de otoño, ROC&C´2007, Acapulco, Gro.,25-30 Noviembre, 2007.