Algoritmo Esteganográfico Adaptivo Basado en Texturas

Por

Dulce Rosario Herrera Moro

Tesina

Sometida como requisito parcial

para obtener el grado de

Licenciada en Ciencias Computacionales

en la

Benemérita Universidad Autónoma de Puebla

Octubre 2007

Puebla, Puebla

Supervisada por:

Dra. Claudia Feregrino Uribe, INAOE

M.C. José Alfonso Garcés Báez, BUAP

Dedicatoria

A mis Padres:

Yrene Moro Torres
Eliazar Herrera González

Agradecimietos

A mis asesores Dra. Claudia Feregrino Uribe y M.C. José Alfonso Garcés Báez mi más sincero agradecimiento por su apoyo, comentarios y consejos a lo largo de la elaboración de este trabajo.

A mis hermanos Martha Neri y Eleazar Herrera Moro por su cariño incondicional.

A mi compañero y amigo Angel Valentin Bautista Durán por su amor y paciencia.

A la Benemérita Universidad Autónoma de Puebla (Buap) por la educación que me brindó.

A mis compañeros de licenciatura por brindarme su amistad.

Abstract

In this work we present a steganographic algorithm that allows an image process to identify the regions in a cover image where it is less probable to detect a hidden message using visual attacks. The regions selected are those whose texture is not homogeneous. This is because those kind of regions are originally noisy and is too difficult to detect extra information. Since the pixels where a message is hidden depend directly on the features of the cover image the steganographic system becomes adaptive. Tests were carried out with different kind of gray scale images and the results are compared against other approaches.

Resumen

En el presente trabajo de investigación se presenta un algoritmo estaganográfico que utiliza un preproceso en la imagen cubierta para localizar las regiones en las que es menos probable detectar el mensaje oculto mediante un ataque visual. Las regiones seleccionadas son aquellas que presentan texturas no homogéneas pues en este tipo de regiones son por naturaleza ruidosas y es más difícil detectar la información oculta. Dado que los pixeles que son modificados durante el proceso de insercción dependen directamente de las catacterísticas de la imagen el algoritmo propuesto se considera adaptivo. El algortimo propuesto se comparó con los algoritmos no adaptivos: LSB y LSBSpread, así como con los algoritmos adaptivos: ConDith y ConDithSpread presentados en [1]. Los experimentos realizados en diferentes imágenes en escala de grises y demuestran que el mensaje oculto no se detecta con el algoritmo propuesto.

Índice general

Capítulo 1

Introducción

La esteganografía es una técnica usada por civilizaciones antiguas como la romana y la griega para enviar mensajes de forma segura. Básicamente, esta técnica consiste en ocultar un mensaje en un medio que no levante sospechas. A diferencia de la criptografía en donde el mensaje es modificado de tal manera que sea ilegible para una persona no autorizada, el principal interés en esteganografía no es el contenido del mensaje si no de mantener oculta la existencia del mensaje más que su contenido.

En la actualidad, las técnicas de estaganografía moderna utilizan a los medios digitales tales como video, audio, imágenes y texto que por sus características permiten insertar información extra usando la información redundante sin alterar la calidad del medio. Además de la redundancia de información los métodos esteganográficos utilizan los medios digitales para cubrir mensajes por que son de fácil almacenamiento, son fáciles de manipular y sobretodo por que pueden ser transmitidos vía internet a cualquier parte del mundo y por lo tanto mensaje oculto pude llegar a destinos remotos rápidamente.

La esteganografía al ocultar la información presenta una ventaja sobre la criptografía pues al enviar un mensaje codificado en un medio como Internet difícilmente se le permitirá llegar a su destino aun si no se puede descifrar su contenido Mas aun la esteganografía puede utilizar criptografía para incrementar la seguridad del mensaje y de

compresión de datos para incrementar la cantidad de información que puede ocultarse en un medio.

En el presente trabajo se presentan las ventajas de tener algoritmos esteganográficos que se adapten a las características del medio con el fin de disminuir la posibilidad de detección del mensaje. El medio usado en este trabajo son imágenes digitales en escala de grises. En este trabajo se presenta un algoritmo que se adapta a las características de la imagen basado en texturas y se compara con dos algoritmos propuestos en [1]. Los resultados que se obtuvieron demuestran que el algoritmo propuesto incrementa la capacidad de inserción en imágenes que a pesar no tener alto contraste en los niveles de gris presentan texturas en las que se puede ocultar información conservando la imperceptibilidad a ataques visuales que se logra con los algoritmos de [1].

1.1. Objetivo General

- Diseñar e implementar un algoritmo esteganográfico para imágenes en escala de grises que se adapte a las características de la imagen para evitar que el mensaje oculto sea detectado mediante un ataque visual.

1.2. Objetivos Específicos

- Mostar la ventaja de los algoritmos esteganográgicos adaptivos sobre los no adaptivo.

- Establecer un criterio de insercción que además de permitir al algortimo adaptarse a las características de la imagen obtenga mayor capacidad de insercción que los algoritmos propuestos en [1].

1.3. Estructura del trabajo

El trabajo esta organizado de la siguiente manera: En el capítulo 1 se presenta una revisión del marco teórico de la esteganografía, en el capítulo 2 se presenta la definición de textura en imágenes digitales así como diferentes métodos que se utilizan para su análisis. La definición de textura se utilizó para definir el criterio de insercción en el algoritmo propuesto . En capítulo tres se describen cada uno de los algoritmos que se presentan. En el capitulo 5 se presentan los resultados que se obtenidos por los algoritmos adaptivos y se comparan con los no adaptivos . Finalmente en el capitulo 6 se presentan conclusiones y trabajos futuros.

Capítulo 2

Esteganografía

Esteganografia proviene de las palabras griegas $estegos$ que significa cubierta y $graphic$ que significa escritura, por lo que esteganografía significa literalmente escritura cubierta[3]. Su objetivo principal es esconder el mensaje oculto en un medio de manera que no pueda ser detectado.

En la siguiente sección se presenta un resumen de cómo han ido evolucionando las técnicas esteganográficas a lo largo de la historia hasta nuestros días.

2.1. Historia y evolución de la esteganografía

La esteganografía a pesar de lo que se piensa es una técnica antigua utilizada principalmente como mecanismo de comunicación segura en tiempos de guerra, uno de los primeros documentos en los que se registra el uso de estas técnicas es en las historias de Herodoto.

En una de sus historias, Herodoto describe la manera en que el general persa Harpagus mandaba mensajes ocultos en el abdomen de liebres muertas que el mensajero transportaba disfrazado de cazador. En otra historia, relata cómo Histaieus rapó a uno de sus esclavos de confianza para tatuar en su cuero cabelludo instrucciones a sus aliados para comenzar la revolución en contra de los Persas y los pueblos mediterráneos.

En la cultura griega se utilizaban tablillas de madera cubiertas de cera en las que escribían con estoletes de madera, metal o marfil. Esta técnica de escritura representó un gran avance en su tiempo, pues a diferencia de otros métodos se podía borrar lo escrito con una espátula denominada pinake. Herodoto cuenta cómo Demeratus utilizó estas tablillas para avisarles a los griegos las intenciones de invasión del emperador persa Xerxes el Grande. La técnica de Demeratus consistía en remover la cera de la tabla para grabar el mensaje directamente sobre la madera que después se recubría nuevamente con cera para dar la impresión de que las tablas no habían sido utilizadas. Esta técnica funcionó muy bien al principio, pero una mujer llamada Gorgo removió la cera pues sospechaba que ocultaba algo, de esta manera, Gorgo se convirtió en la primera mujer criptoanalista [6]FALTA.

Kahn habla de una práctica que se volvió común en China que consistía en insertar códigos de ideogramas en lugares predeterminados de un mensaje. Esta práctica es parecida al sistema de rejillas usado durante la edad media en Europa. El sistema de rejillas consistía en el uso de un patrón hecho de papel o de madera que al ser colocado sobre un texto aparentemente inocuo revelaba las posiciones de las letras de un mensaje oculto [6].

Un método que por su sencillez se ha venido usando a lo largo de los siglos para ocultar información consiste en oculta los caracteres de un mensaje entremezclándolos a determinada distancia con los de otro texto que aparentemente es inofensivo.

Un ejemplo de lo anterior lo podemos encontrar en el siguiente mensaje enviado por un espía Alemán durante la Segunda Guerra Mundial.

textitA**p**parently n**e**utral's p**r**otest i**s** t**h**oroughly d**i**scounted a**n**d i**g**nored. I**s**man h**a**rd hit. B**l**ockade i**s**sue a**f**fects p**r**etext f**o**r e**m**bargo o**n** b**y** p**r**oducts, e**j**ecting s**u**ets a**n**d v**e**getable o**i**ls.

Al extraer la segunda letra de cada palabra del mensaje se descubre el mensaje oculto:

Pershing sails from NY June 1.

En el ejemplo anterior se observa lo fácil que resulta esconder información en textos

escritos sin levantar sospecha. El truco para asegurar que el mensaje no sea detectado es que éste debe estar inmerso en una gran cantidad de información.

Las continuas mejoras a los lentes, cámaras fotográficas y películas permitieron reducir cada vez más las fotografías. Durante la segunda guerra mundial el avance alcanzado en esa área permitió desarrollar a los alemanes su tecnología de espionaje llamada micropuntos. Los micropuntos son imágenes tan pequeñas como puntos, de manera que se podían usar como signos de puntuación en un texto o se ocultaban detrás del sello postal. Este mecanismo permitía la transmisión de una gran cantidad de información y debido a su tamaño pasaba desapercibido sin necesidad de esconderlo.

En 1983 Simmons en [7] formula $El problema de los prisioneros$ en el que plantea el siguiente escenario: Alice y Bob están encarcelados y desean realizar un plan de escape, se les ha permitido tener comunicación escrita pero cada mensaje que se envían es revisado por el guardián Willie. Si Alice y Bob decidieran encriptar sus mensajes para evitar que Willie los leyera y descubriera sus planes, el guardián a pesar de no saber exactamente lo que dicen los mensajes sospecharía las intenciones de los prisioneros y los frustraría confinándolos incomunicados. La solución al problema de Alice y Bob es encontrar la manera de ocultar la información referente a su plan de manera que Willie al revisar los mensajes solo vea un texto totalmente inofensivo y lo deje pasar.

En la actualidad los métodos esteganográficos utilizan los distintos medios digitales tales como audio, video, texto e imágenes para insertar información y enviarla vía Internet. Además de enviar mensajes ocultos los métodos esteganográficos se han utilizado para proteger la propiedad de los medios digitales. Esto se hace insertando en el medio información acerca del autor o proveedor.

En la siguiente sección se mostrará el esquema general de los sistemas esteganográficos así como las características que deben cumplir.

2.2. Características de los procesos esteganográficos

En la Figura 2.1 se presenta el esquema general de un proceso esteganográfico.

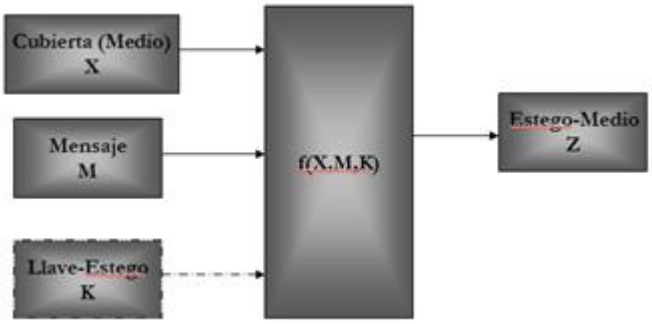


Figura 2.1: Esquema general de un proceso esteganográfico

En la Figura 2.1, el mensaje M es la información que se desea ocultar, esta puede ser: texto, imagen, video o cualquier otra información que pueda ser representada por una cadena de bits.

La cubierta o host X es el medio en el que se oculta la información.

Los procesos esteganográficos pueden usar una $stego-keyK$ cuya función es similar a la llave que se utiliza en criptografía. La llave esteganográfica se utiliza para controlar el proceso de inserción del mensaje seleccionando los píxeles, coeficientes, etc. que se modificarán al insertar el mensaje. El propósito de la llave es asegurar que sólo los receptores que la conoz can podrán extraer el mensaje de la estego-imagen [4]

La función de inserción f recibe como parámetros la cubierta X, el mensaje My opcionalmente la llave K, como resultado la función devuelve el medio con información oculta Z denominado estego-medio. La función f depende primordialmente de la estructura del medio cubierta y consta de dos fases: en la primera se seleccionan los bits o coeficientes en los que se insertará la información, en la segunda fase se inserta la información en los lugares seleccionados de acuerdo a un determinado método.

El estego-medio debe ser muy similar al medio original de manera que el usuario común no perciba una disminución en la calidad del medio con información oculta y más aun, se debe buscar que el mensaje no pueda ser detectado por un alguna persona que maliciosamente busque indicios que delaten la presencia del mensaje oculto. La mayoría de las técnicas esteganográficas descomponen la información que se desea

ocultar en pequeñas porciones de manera que los cambios sean demasiado pequeños como para percibirlos. Sin embargo, existen otras técnicas en las que se utilizan las características del medio para insertar la información.

En general los algoritmos esteganográficos deben garantizar lo siguiente:

- La información oculta en un medio no debe percibirse visualmente.
- Los usuarios ordinarios no deben percibir alguna ambigüedad en la claridad del medio que oculta información.
- La información oculta sólo debe poder ser recuperada por el creador y usuarios autorizados.

Sin embargo, existen tres características propias de las técnicas esteganográficas que se contraponen entre sí, por lo que para cada técnica que se desarrolle de debe establecer un compromiso entre ellas de acuerdo a la aplicación. En la Figura 2.2 se muestra la relación que guardan estas características.

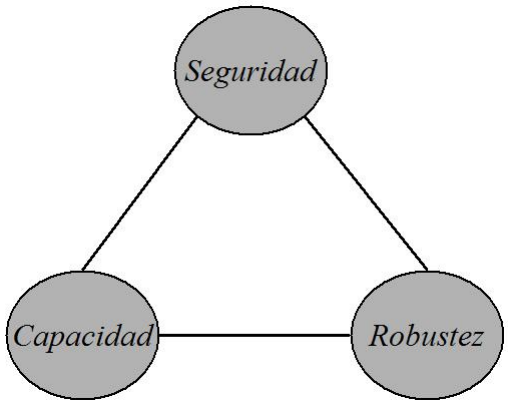


Figura 2.2: Características de los métodos esteganográficos

- **Capacidad:** Que se refiere a la cantidad de información que se puede insertar en un medio conservando la calidad del medio y la imperceptibilidad de la información oculta.

- **Seguridad:** Está directamente relacionada a la imperceptibilidad de la información oculta. La mayoría de las técnicas esteganográficas para lograr la imperceptibilidad emplean las limitaciones del Sistema de Audición Humana o HAS por sus siglas en ingles o las de el Sistema de Visión Humana ,HVS según sea el caso.

- **Robustez:** Se refiere a la cantidad de modificaciones que el medio soporta antes de que se destruya el mensaje oculto en él.

Como ya se mencionó se debe establecer un compromiso entre las características dependiendo del tipo de aplicación con la que se va a trabajar. De acuerdo a lo anterior se distinguen dos líneas generales:

- *Protección contra detección.* El objetivo en este esquema, como su nombre lo indica, es evitar que un mensaje sea descubierto por lo que en este tipo de técnicas se debe establecer prioridad a la seguridad. La segunda característica a la que se le da importancia es a la capacidad y por último a la robustez. Estas técnicas se utilizan principalmente para comunicación segura de mensajes. El problema del prisionero es un ejemplo que ilustra la protección contra detección.

- *Protección contra eliminación.* El propósito en este esquema es evitar que se elimine la información insertada en el medio. Este esquema se utiliza principalmente para proteger los derechos autor. Existen dos tipos de técnicas que sobresalen en este esquema: Watermarking o Marcas de agua y Fingerprinting o Firmas digitales. La diferencia entre estas dos técnicas es el tipo de información que se inserta en el documento: en las marcas de agua la información pueden ser sonidos, imágenes, códigos de barras, etc. que identifiquen al autor o proveedor. En Fingerprinting la información que se inserta es un número de serie que permite identificar de manera única a cada una de las copias que se hagan del documento.Figura 2.3.

Figura 2.3: Ramas de aplicación de la esteganografía

2.3. Principales Métodos

En esta sección se realiza un revisión de los principales métodos esteganográficos utilizados en diferentes medios.

2.3.1. Ocultación de información en las capas del modelo OSI

En 1984, la Organización Internacional de Estandarización (ISO) desarrolló un modelo llamado OSI, Open Systems Interconection o Interconexión de sistemas abiertos. En el modelo OSI el propósito de cada capa es proveer los servicios para la siguiente capa superior de manera que le sean transparentes los detalles de implementación. Las capas son abstraídas de tal manera que cada capa cree que se está comunicando con la capa homónima en otra computadora, cuando realmente cada capa se comunica sólo con las capas adyacentes de las misma computadora. Figura 2.4.

Las capas del modelo OSI presentan lugares que en esteganografía se han utilizado para ocultar información.

En la ***capa física*** la transmisión de información se hace a través de hardware especial que utiliza señales de control y temporización. La información en esta capa se

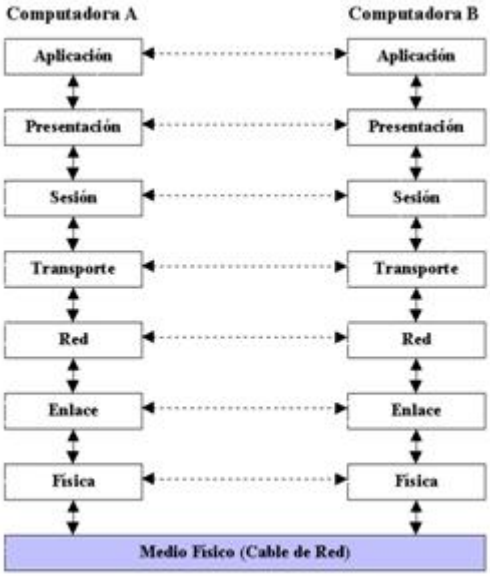


Figura 2.4: Esquema del modelo OSI

puede ocultar manipulando la tensión de las líneas de comunicación. Es decir, al modificar los 5 volts que representan a un ”1”binario de manera que si perciben 5.5. volts el receptor leerá ”1”del mensaje oculto mientras que 5 volts representarían ”0”. Un aspecto importante a resaltar es que la capa física sólo se puede usar para la transmisión de información oculta cuando no se hace uso de un repetidor, brige, router, etc. entre emisor y receptor.

En la ***capa de enlace*** se puede modificar el sistema de detección de errores.

En la ***capa de red*** agrega información suplementaria como cabeceras para direccionar y posibles errores de control en la que se agrega información adicional del mensaje. En la figura 2.5 se muestra la organización de la cabecera de los paquetes IPv4 resaltando los bits sin uso específico que se utilizan para almacenar información extra. La implementación para Internet de la **capa de transporte** es el *Transmission Control Protocol (TCP)*. Cada segmento TCP comienza con un formato de fijo de 20-bytes de cabecera. En la figura 2.6 se muestran los bytes 13 y 14 de la cabecera, como se puede observar existen 6 bits disponibles que se pueden usar para insertar información.

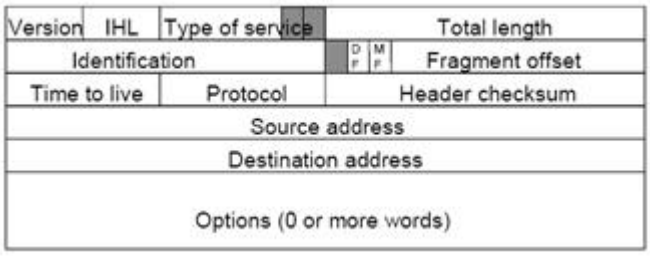


Figura 2.5: Cabecera de IPv4

Figura 2.6: Bytes 13 y 14 de la cabecera TCP

2.4. Ocultación de información en documentos de texto

Ocultar información en documentos de textos es históricamente el método más importante de la esteganografía. Es sólo cuando aparece el Internet que disminuye su importancia. Sin embargo, existe una necesidad de proteger las publicaciones que se distribuyen electrónicamente vía CD-ROMs, en línea, bases de datos, librerías electrónicas etc. El mayor problema con este tipo de distribuciones es que las copias digitales de textos son idénticas a la original. Para que la publicación electrónica tenga éxito se debe garantizar que los beneficios no disminuirán debido a las pérdidas producidas por las copias ilegales. Las técnicas usadas para proteger los documentos insertan información en las copias de un documento, dicha marca puede ser la misma o diferente para cada copia.

Las técnicas para marcar documentos alteran el formato o las características de los caracteres. El propósito es que las modificaciones que se hagan no sean visibles pero si decodificables. Los documentos están constituidos por dos partes en donde se pueden aplicar las modificaciones. La primera es la imagen de un documento que lo describe en términos de un arreglo de píxeles. La segunda representación es el archivo de formato

del documento que describe el contenido del documento así como estilo. De el archivo de formato se genera la imagen que se le presenta al lector.

El esquema general de las técnicas usadas para marcar un documento requieren de un codificador y un decodificador. El documento original se preprocesa, el codificador recibe como entrada un documento virtual en el que se harán las modificaciones, Figura 2.7. El decodificador recibe un archivo como entrada y presenta como salida la marca

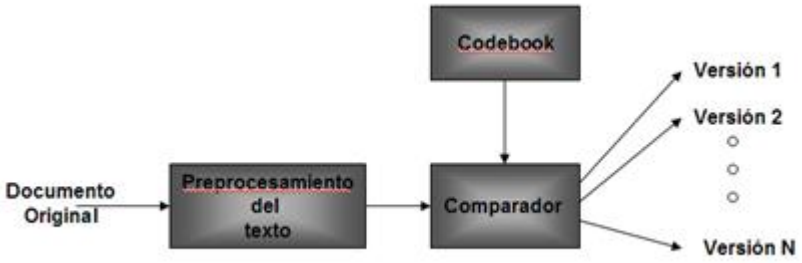


Figura 2.7: Esquema general para el marcado de documentos.

insertada en él. Figura 2.8 Existen tres técnicas principales para ocultar información en

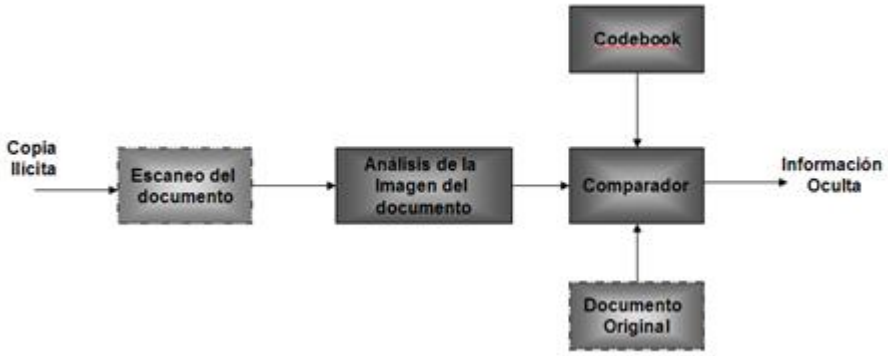


Figura 2.8: Esquema de decodificación de documentos marcados.

documentos:

- **Codificación por cambio de línea:** En esta técnica el codificador mueve las líneas hacia arriba o hacia abajo, de acuerdo a la marca que se desea insertar.

El decodificador mide la distancia entre dos líneas adyacentes. Esto se puede hacer de dos manera: medir la distancia entre la base de dos líneas o midiendo la distancia entre sus centroides.

- **Codificación por cambio de palabra:** Este esquema altera el documento moviendo horizontalmente la posición de las palabras. Por lo que, los espacios entre dos palabras adyacentes deben ser diferentes utilizando este esquema. Sin embargo generalmente los espacios entre palabras varia para acomodar los espacios en blanco de documentos justificados por los que el decodificador de este esquema requiere la imagen original.

- **Codificación por características:** En este esquema, se examinan las características del documento para seleccionar algunas que serán modificadas, tales características pueden ser por ejemplo, la longitud las líneas verticales de las letras b, d, h, k, etc.. También se pueden cambiar las alturas de carácter de una fuente dada. Existen técnicas en las cuales se cambian las palabras del texto por sinónimos, en este esquema, generalmente hay dos pares de sinónimos y el uso de uno u otro es equivalente a insertar un ”0.º un ”1”.

2.4.1. Ocultación de información en archivos de audio

La representación de la información en los archivos de sonido se realiza generalmente mediante el registro consecutivo de las muestras que componen el sonido. El sonido puede representarse en forma de bits por lo que el método esteganográfico LSB (Least Significative Bit) se puede usar fácilmente. El método LSB consiste en reemplazar los bits menos significativos del medio original por los bits de la información que se desea ocultar.

Los métodos que insertan información en archivos de audio utilizan las propiedades del Sistema de Audición Humana o SAH. El oído humano percibe un rango de frecuencias entre 20 Hz y 20 KHz sin embargo es más sensible a las frecuencias

comprendidas entre 2000 y 5000 Hz, decreciendo la sensibilidad para el resto de frecuencias. Además se sabe que la sensibilidad auditiva es menor en la zona de frecuencias bajas y que aunque el SAH percibe amplio rango dinámico de frecuencias, no distingue pequeñas variaciones entre los rangos. Por lo que existen algunas distorsiones que son ignoradas por el SAH. En el sonido ocurre una cosa muy curiosa: da prácticamente igual que la calidad sea alta o baja, pues en calidades altas la cantidad de bits usados para representar el sonido es muy alta, y un cambio ínfimo no influye en el resultado; mientras que en los de calidad baja, aunque la modificación sea más llamativa, al tratarse de un sonido de baja calidad, cualquier modificación puede pasar inadvertida como ruido de fondo.

Otra característica del sistema de audición es el efecto conocido como enmascaramiento.

Las características mencionadas anteriormente y otras se utilizan en los procesos de compresión de audio como Mp3 para eliminar elementos que no son percibidos por el SAH y que por lo tanto su presencia es redundante en los archivos de audio. Sin embargo, a pesar de los procesos de compresión de audio, existe información redundante en los archivos que se puede utilizar en procesos esteganográficos. El considerar las características del SAH en los métodos esteganográficos asegura que al alterar el archivo de audio se haga de tal manera que no se perciba.

2.4.2. Ocultación de información en imágenes

En los últimos años, las imágenes se han convertido en el medio más usado en la esteganografía. Existen diferentes formatos en los archivos de imágenes muchos de ellos diseñados para aplicaciones específicas por lo que existen diferentes algoritmos esteganográficos. En esta sección se presentan algunas de las técnicas más comunes para ocultar información en imágenes digitales.

Para una computadora una imagen es una colección de números que representan diferentes intensidades de luz en diferentes áreas de la imagen. La representación numérica está dada en forma de matriz y a los puntos individuales se les denomina píxeles. El número de bits que se utiliza para representar el color en un píxel es llamado profundidad. Las imágenes comúnmente se almacenan usando 24 bits que se dividen en tres bytes que constituye en canal rojo, verde y azul en una imagen. Estas imágenes si se guardaran tal cual, ocuparían demasiado espacio de almacenamiento. Por lo que se utilizan técnicas de compresión para facilitar el almacenamiento y manejo de las imágenes.

Existen dos tipos de compresión: *Compresión sin pérdida*, en la que se conserva la estructura original de la imagen. Ejemplos de este formato de imagen son las imágenes GIF. El otro tipo es la *compresión con pérdida*, en este tipo de compresión no se conserva la estructura original de la imagen pues se eliminan los detalles de la imagen que no percibe el ojo humano por lo que se obtienen archivos mucho más pequeños que usando la compresión sin pérdida. Un ejemplo de formato de que utiliza este tipo de compresión es el JPEG (Joint Photographic Experts Group).

Anteriormente los métodos esteganográficos precisaban que se debía utilizar compresión sin pérdida pues de lo contrario, la información insertada se eliminaba. Sin embargo dado que en la actualidad se utiliza con frecuencia el formato JPEG para la transmisión vía Internet, se han creado métodos esteganográficos cuyas características les permiten recuperar el mensaje incluso después de una compresión con pérdida. Estos métodos generalmente insertan el mensaje en los coeficientes de la transformada discreta de coseno que se utiliza en el proceso de compresión JPEG.

Las técnicas esteganográficas en imágenes se pueden dividir en dos grupos:

- *Técnicas que trabajan en el dominio espacial:* Estas técnicas insertan el mensaje directamente en la intensidad de los píxeles. Este tipo de técnicas

proporcionan una mayor capacidad insertando y manipulando la información en forma de ruido sobre la imagen.

- *Técnicas que trabajan en el dominio de la frecuencia:* En este tipo técnicas primero transforman la imagen y después insertan la información en áreas significativas de la imagen. Este tipo de técnicas son independientes del formato de la imagen por lo que el mensaje insertado puede recuperarse entre la conversión entre una compresión sin pérdida a una compresión con pérdida.

2.5. Estegoanálisis

El estegoanálisis se refiere a un conjunto de técnicas que se utilizan para atacar a un sistema esteganográfico, el éxito del ataque se obtiene al detectar, destruir, extraer o modificar un mensaje oculto. En comunicación segura, el éxito de una ataque consiste en detectar y probar la existencia de algún tipo de información oculta mas no recuperar el mensaje. Para los propósitos de la piratería, el ataque debe retirar o modificar la marca del autor sin que el medio se degrade significativamente. Por lo tanto, la complejidad de las técnicas esteganográficas, así como el significado de éxito depende de la aplicación.

Los tipos de ataques a sistemas esteganográficos pueden ser [5].

- *Pasivos:* si el atacante sólo es capaz de interceptar y analizar los datos.
- *Activos:* cuando el atacante además puede manipular los datos, estos ataques consisten principalmente en aplicar a la imagen algún tipo de función que modifique su estructura o los valores de intensidad de los píxeles con el fin de destruir un posible mensaje.

Por el tipo de información que utilizan para realizar un ataque, los métodos de estegoanálisis se pueden clasificar de la siguiente manera: beginitemize

- *Estego-solo*, si el atacante sólo dispone del estego-medio.
- *Ataques por repetición de cubierta:* se da cuando el creador de los esteganogramas ha usado la misma cubierta para ocultar información en más de una ocasión.
- *Ataque por cubierta conocida*, en este ataque se intercepta el esteganograma y se conoce la imagen original que se usó como cubierta.
- *Estego elegido*, en este ataque se tiene acceso al estego-medio pero además conoce el método esteganográfico empleado.
- *Mensaje elegido*, en él se genera el estego-medio de un mensaje conocido para encontrar las firmas que le permitirán detectar a otros.

En la actualidad existen un gran número de técnicas de estegoanálisis tanto para la detección de mensajes como para modificar o retirar las marcas de agua que protegen los derechos de autor, sin embargo, dado que en este trabajo se presentan técnicas de estegoanálisis para la comunicación segura se utilizó un ataque visual cuyo propósito es detectar el mensaje oculto. A continuación se describe el ataque visual.

2.5.1. Ataques visuales

Si un mensaje se oculta en el bit menos significativo de los píxeles de una imagen un usuario normal no apreciará la diferencia, sin embargo si el usuario intencionalmente busca un mensaje en el plano del bit menos significativo de los píxeles de una imagen es probable que aprecie el mensaje oculto como ruido en la imagen e incluso pueda llegar a calcular su longitud. El ataque visual consiste en extraer los bits menos significativos de una imagen para formar una imagen binaria de la siguiente manera: si el bit LSB en la posición (x, y) de la imagen es 1 entonces el píxel correspondiente de la imagen binaria tomará el valor más alto de la escala de grises en otro caso tomara el valor 0. En la figura 2.9 se muestra un ejemplo de un ataque visual a) es la imagen original, en b) se presenta el plano LSB de la imagen sin mensaje oculto y c) muestra el plano

de las misma imagen con mensaje oculto. Como se puede apreciar en el ejemplo de la figura 2.9 el mensaje oculto se aprecia principalmente en las regiones homogéneas de la imagen.

a) b) c)

Figura 2.9: Ataque visual. a) imagen original. b) Plano LSB de la imagen sin mensaje. c) Plano LSB de la imagen con mensaje oculto.

Como podemos apreciar en la figura 1.8 las regiones en las que se detecta con mayor facilidad el mensaje oculto en la imagen son aquellas en las que los pixeles tienen el mismo color, en este caso el fondo blanco y la torre que es de color negro en su mayoría. Ahora notemos que si el mensaje fuera insertado en la parte inferior de la imagen, en donde no existe una tonalidad predominante, seria más difícil determinar la existencia de un mensaje. En este trabajo se presenta un algoritmo esteganográfico que inserta la información en aéreas de la imagen en donde es difícil detectar el mensaje oculto mediante un ataque visual. El algoritmo utiliza las regiones ricas en textura. Por lo que en el siguiente capitulo se presenta el concepto de textura empleado imágenes digitales y se describen brevemente los principales métodos que se utilizan para su caracterización.

Capítulo 3

Textura

Una de las etapas de la visión por computadora consiste en segmentar la imagen en elementos que sean de interés para la aplicación. La segmentación de una imagen basada en texturas resulta de gran utilidad para el análisis de imágenes aéreas, biomédicas y sísmicas, así como en la automatización de procesos industriales.

En esteganografía el análisis de texturas se puede utilizar como criterio para seleccionar las regiones de la imagen que se utilizará para insertar información. En el presente trabajo se presenta un algoritmo esteganográfico que inserta información en aquellas regiones de la imagen que contienen texturas no homogéneas.

La razón de utilizar regiones con texturas no homogéneas para la inserción de información es debe a que este tipo de regiones son por naturaleza ruidosas y no es posible detectar mediante un ataque visual el ruido que delata la presencia de un mensaje oculto en regiones homogéneas. Así, el algoritmo evita las regiones homogéneas de la imagen y reduce la probabilidad de detección del mensaje bajo un ataque visual.

Los humanos tenemos intuitivamente la noción de textura y podemos fácilmente distinguir las diferentes tipos de texturas presentes en los objetos de una escena. Sin embargo, no existe una definición de textura que nos indique con base en qué hacemos esta discriminación.

Al no contar con una definición exacta de lo que es la textura la segmentación

automática de una imagen con base en la textura de los objetos que la conforman se convierte en una tarea complicada puesto que a la computadora se le tienen que proporcionar datos específicos que le permitan distinguir una textura de otra.

En las siguientes secciones se introduce y explica el concepto de textura que han utilizado algunos autores para imágenes digitales y se describen brevemente los tipos de métodos que se utilizan para su análisis.

3.1. Definición de Textura en imágenes Digitales

Una de las definiciones de textura que se maneja para imágenes digitales es:

Textura es la repetición de un patrón o más patrones sobre una región.

Cada elemento de un patrón tiene características propias como: tamaño, forma, color y orientación. Para describir diferentes texturas, los patrones se repiten a lo largo de una región variando las características en cada tipo de textura.

En la figura 3.1 se muestra un ejemplo de cómo se pueden caracterizar diferentes texturas en imágenes digitales. Las texturas mostradas en los incisos a) y b) la textura se puede caracterizar por la geometría de un patrón, en c) y d) la textura se puede caracterizar por la variación en los niveles de gris y en e) por repetición de un objeto, que este caso la piedra.

3.2. Análisis de texturas en imágenes digitales

En la figura 3.2 se muestra la clasificación general de los métodos usados para caracterizar los diferentes tipos de texturas que aparecen en una imagen digital.

3.2.1. Métodos espectrales

En los métodos espectrales se utiliza la transformada de fourier para analizar el espectro de la imagen para detectar los patrones periódicos de las texturas. Los picos

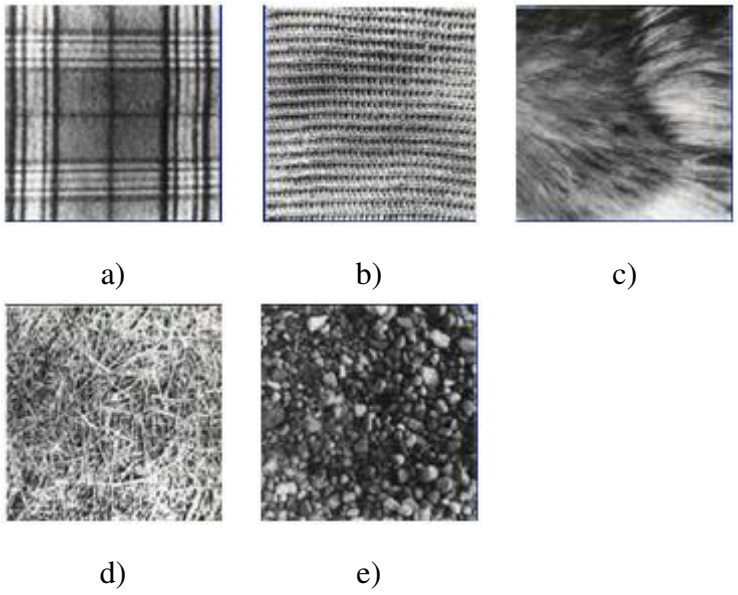


Figura 3.1: Caracterización de texturas a), b) por la geometría del patrón. c), d) por variación en los niveles de gris, e) por repetición de un objeto (piedra)

prominentes del espectro proporcionan la dirección principal de los patrones y los picos en el plano de frecuencia indica el período espacial de los patrones.

Para simplificar la detección de texturas se emplean coordenadas polares. Sea $S(r, \theta)$ la función del espectro donde θ representa dirección y r frecuencia. Para cada θ se puede obtener la función unidireccional $S_\theta(r)$ que describe el comportamiento del espectro a lo largo de una dirección radial desde el origen figura 3.3-b). De la misma forma para cada r se obtiene la función $Sr(\theta)$ que describe el comportamiento a lo largo de un

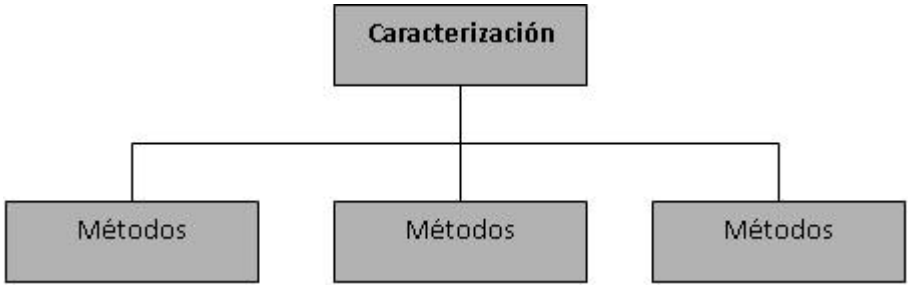
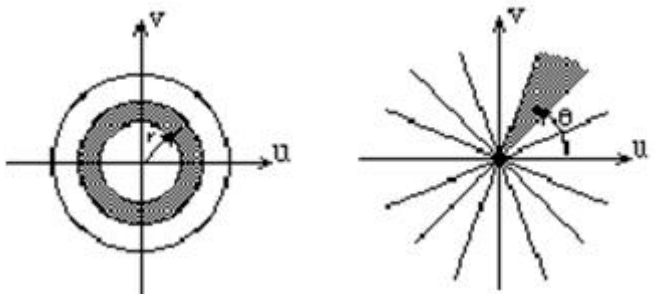


Figura 3.2: Clasificación general de los métodos para la caracterización de texturas

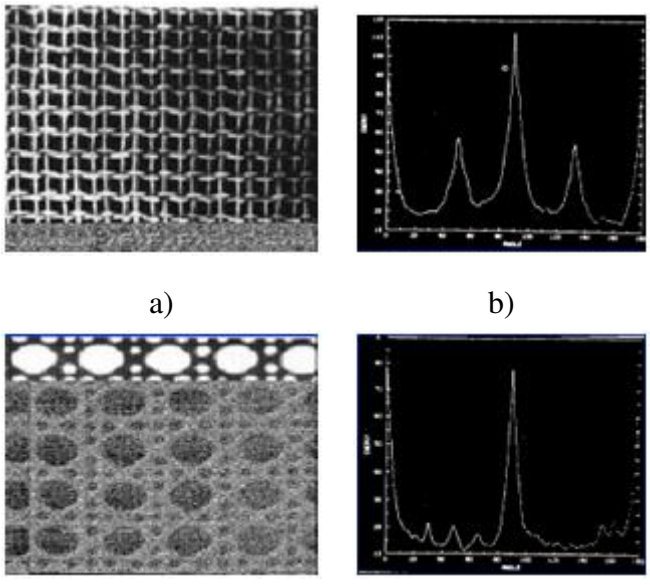
círculo centrado en el origen 3.3-b).



a) b)

Figura 3.3: Gráficas de las funciones a) $S(r)$, b)$S(\theta)$

Variando los parámetros r y θ las funciones describen la energía espectral de la textura. Los descriptores de las funciones $S_\theta(r)$ y $Sr(\theta)$ son usados para cuantizar su comportamiento. En la figura 3.4 se ilustra el empleo de $S(\theta)$ para diferenciar entre dos formas de textura.



a) b)

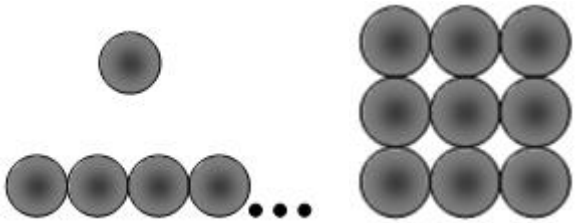
c) d)

Figura 3.4: Ejemplo del uso del espectro de Furier para describir textura en una imagen

3.2.2. Métodos estructurales

El objetivo en este tipo de métodos es encontrar las características y distribución de una primitiva que permita describir los patrones que conforman a la textura. Comúnmente se utilizan gramáticas para describir a través de reglas como se forman los patrones a partir de un número pequeño de símbolos. Para representar de manera realista las variaciones que se presentan en texturas naturales los métodos agregan probabilidades a diferentes reglas.

Ejemplo: supóngase que tenemos una regla de la forma $S \rightarrow aS$. Si a representa un círculo y se asigna el significado de *círculo a la derecha* a una cadena de la aaa, esta regla permite generar el patrón que se muestra en la figura 3.5 a). Si agregamos nuevas reglas a este esquema $S \rightarrow bA$, $A \rightarrow cA$, $A \rightarrow c$, $A \rightarrow b$, $S \rightarrow a$, donde b representa la presencia de un círculo abajo y c la de un circulo a la izquierda podemos formar la cadena $aaabccbaa$ que representa el patrón de que se muestra en la figura 3.5 b). Además de círculos se le pueden agregar a estas reglas, nuevos símbolos para crear patrones más complejos.



a) b)

Figura 3.5: Ejemplo de uso de gramáticas para describir textura.

3.2.3. Métodos estadísticos

Los métodos estadísticos analizan la distribución del histograma mediante métricas estadísticas para caracterizar las texturas presentes en la imagen. Por ejemplo, el valor de intensidad de los píxeles que pertenecen a una región con la misma textura proba-

blemente sea muy similar por lo que la textura se pude identificar por el promedio.

Además del promedio que resulta insuficiente para la mayoría de las imágenes se pueden usar otras medidas como la desviación estándar que se utiliza para distinguir entre los píxeles que conforman una textura y los que son parte del fondo.

El promedio y la desviación estándar son conocidos como momentos estadísticos cuya ecuación general se muestra en la ecuación (3.1). El promedio está relacionado al primer momento, y la desviación depende del segundo momento. Existen otros momentos que también pueden ser usados para caracterizar la textura de las regiones. En la ecuación (3.1) se presenta la fórmula general estos momentos.

$$Mn = \frac{\sum (x - x^{-})^{n}}{N} \tag{3.1}$$

Donde N es el número de píxeles, n es el orden del momento, X es la media y x.

Cuando se analiza una textura considerando únicamente la distribución en los niveles de gris, las medidas estadísticas no proporcionan información de la relación espacial que guardan entre sí los píxeles pertenecientes a la misma textura.

Existen técnicas que además de considerar la distribución del histograma considera las relaciones espaciales que guardan los píxeles con un valor de intensidad similar. Una de estas técnicas es el llamado análisis co-ocurrente en el que se crea una matriz de co-ocurrencia que contiene el número de píxeles con intensidades similares que están separados por d posiciones en una dirección dada. Normalmente se usan cuatro direcciones: horizontal, vertical, y dos diagonales.

La creación de esta matriz se ilustra con el siguiente ejemplo. Considérese la imagen con tres niveles de niveles de gris de la figura 3.6 a). Considérese la dirección diagonal derecha y una distancia $d = 1$. El primer paso es crear una matriz A de kxk cuyos elementos a_{ij} son el número de veces que aparecen píxeles con nivel de gris z_i en la dirección establecida y a la distancia d de los píxeles con nivel de gris z_j , con $1 <= i$, $j <= k$. En la imagen de la figura 3.6 a) tenemos tres niveles de gris por lo que tenemos $z_1 = 0$, $z_2 = 1$, $z_3 = 2$. de la se obtiene la siguiente matriz de la figura 3.6 b)

. La matriz de co-ocurrencia se obtiene dividiendo A entre el número total de píxeles considerados figura 3.6 c). Como la matriz de co-ocurrencia depende de la dirección

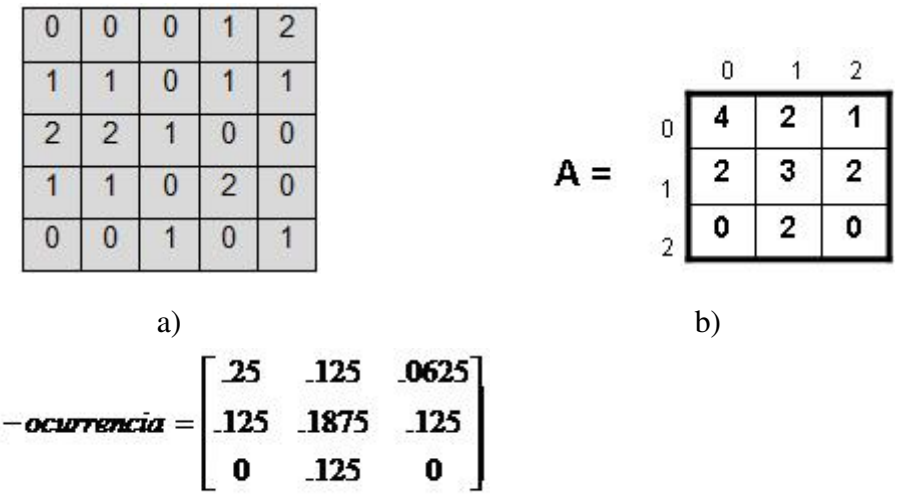


Figura 3.6: Ejemplo de una matriz de co-ocurrencia para una imagen con 3 niveles de gris.

y la distancia a considerar se tendrá una matriz de co-ocurrencia por la variación de estos parámetros. Otro parámetro que se considera es el número de niveles de gris que existenten en la imagen pues determina el tamaño de la matriz de co-ocurrencia. Para evitar que la matriz de co-ocurrencia sea demasiado grande se recomienda cuantificar las intensidades a pocas bandas en el nivel de gris. Por último, para obtener información acerca de la textura, la matriz de co-ocurrencia es analizada por medio de descriptores. En la tabla 2.1 se muestran tres los descriptores más comunes.

En el presente trabajo se utiliza la definición de textura de los métodos estadísticos presentada en la sección 2.2.3 para insertar información en las regiones con texturas no homogéneas en la imagen. Se escogió esta definición de textura puesto que no nos interesa saber qué patrones componen las diferentes texturas de la imagen como lo proponen los métodos estructurales. En nuestro caso solo queremos determinar si una

región contiene o no una textura homogénea, sin embargo, si utilizáramos algún método espectral para nos podría decir si en una imagen contiene o no texturas homogéneas pero no nos daría la posición en coordenadas espaciales de las regiones de interés.

Puesto que la definición de textura que dan los métodos estadísticos esta directamente relacionada con los niveles de grises presentes en una región una región con textura no homogénea podría definirse como aquella en donde se encuentran varias tonalidades intercaladas.

En el siguiente capítulo se presenta el algoritmo ConText que utiliza este concepto para determinar los pixeles de una imagen en los que se puede insertar información reduciendo la probabilidad de detección ante el ataque visual presentado en la sección.

Capítulo 4

Algoritmos estaganográficos adaptivos

En la esteganografía moderna, el medio más usado como cubierta son las imágenes digitales. En la figura 4.1 se muestran los elementos que conforman el esquema general de los algoritmos esteganográfico que utilizan imágenes digitales como medio para esconder información. A la imagen original se le denomina imagen cubierta o cover image, la estego llave se utiliza para controlar la posición del mensaje en la imagen cubierta, generalmente la estego llave es la semilla de una función de números pseudo aleatorios, la función de inserción es la manera en la que se modifica la imagen para contener el mensaje, como resultado de la función de inserción se obtiene la imagen modificada por el mensaje a la que se le denomina stego imagen.

La función de inserción puede modificar la imagen en el dominio espacial o de las frecuencias. Uno de los métodos más utilizados para modificar la stego imagen es el llamado LBS (Least Significant Bit), esto debido principalmente a su sencillez.

LSB modifica el valor del bit menos significativo de los píxeles cuando se trabaja en el dominio espacial, o valor del bit menos significativo de los coeficientes que representan a la imagen en el dominio de una determinada función.

El principal problema del esquema general mostrado en la figura 4.1 . es que al no considerar las características de la imagen cubierta, el mensaje insertado en la imagen puede provocar artefactos en la imagen que pueden revelar su existencia. Como

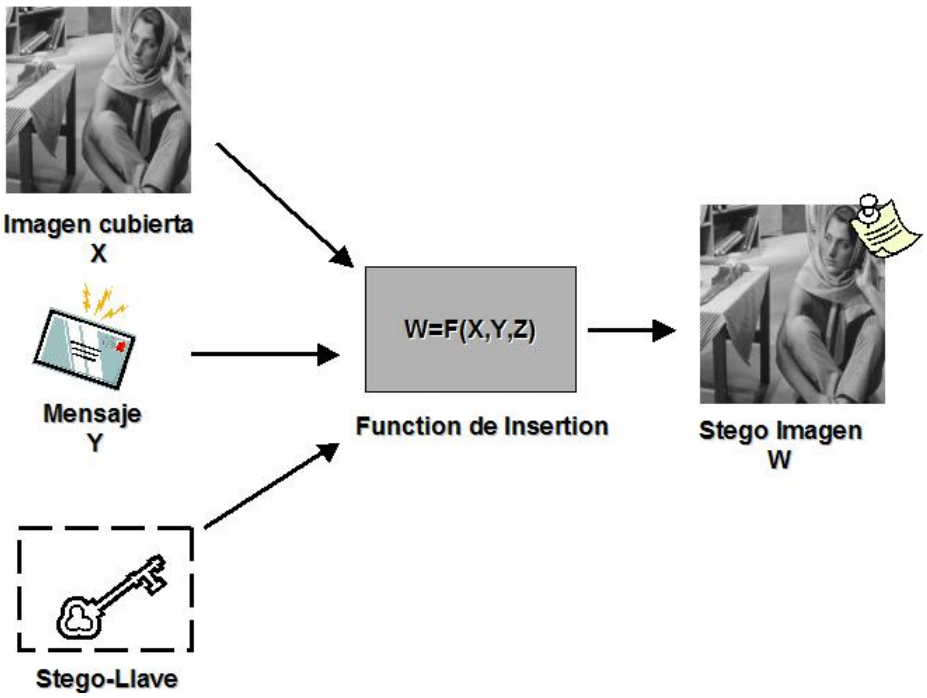


Figura 4.1: Esquema general de esteganografía usando imágenes digitales.

vimos en la sección 1.4 los ataques visuales pueden revelar la existencia de mensajes ocultos en forma de ruido en la imagen, esto se presenta principalmente en la regiones homogéneas.

Para solucionar este problema se han propuesto esquemas de esteganografía que se adaptan a las características de la imagen cubierta. En la figura 4.2 se muestra el esquema general de los algoritmos esteganográficos adaptivos.

Como se aprecia en la figura 4.2 en el esquema adaptivo se agrega un proceso para la extracción de características de la imagen misma que se utilizan para controlar la función de inserción.

Existen diferentes maneras en las que se pueden utilizar las características de una imagen lograr la adaptabilidad en un método esteganográfico. En [1] se reconocen tres aspectos de la función de inserción que pueden ser modificados para lograr la adaptabilidad en los algoritmos.

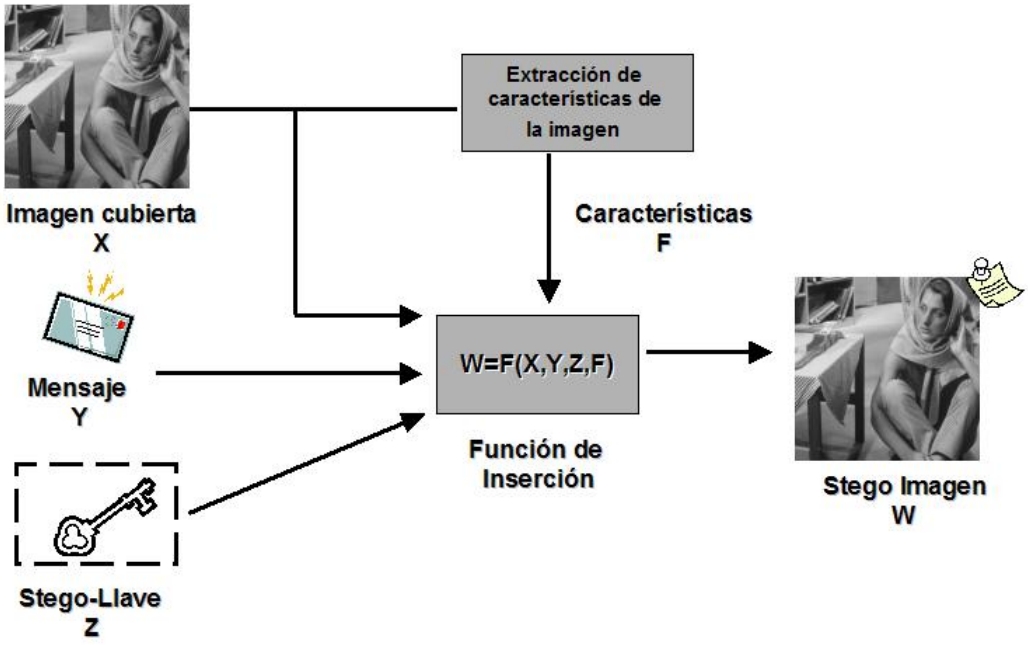


Figura 4.2: Esquema general de los métodos esteganográficos adaptivos.

- ***La selección de los píxeles que contendrán al mensaje.*** Es decir, la adaptabilidad de logra en la manera de seleccionar los píxeles a modificar de acuerdo a las características de la imagen. De esta manera, el mensaje se inserta en las regiones de la imagen en las que es menos probable la su detección.

- ***La representación de los bits del mensaje.*** Aquí el mensaje se representa de tal manera que afecte lo menos posible las características originales de la imagen.

- ***La manipulación de la imagen cubierta durante la inserción.*** Las características de la imagen se utilizan para determinar la mejor método para modificar los píxeles de la imagen con el mensaje. Por ejemplo en [] de acuerdo a las características estadísticas de una región se escoge entre tres métodos de inserción.

Para determinar las posiciones en las que se inserta el mensaje de o el método de inserción adecuado para el tipo de imagen. Dado que la inserción del mensaje es diferente en cada imagen se logran métodos esteganográficos adaptable en la figura 4.2

se muestra un esquema general de los métodos esteganográficos adaptivos.

Los algoritmos que se describen en el presente trabajo, siguen el esquema de esteganografía adaptiva logrando la adaptabilidad por la forma en que escogen los píxeles en los que se inserta el mensaje. Cada uno de los algoritmos procesa la imagen cubierta con la finalidad de crear una imagen binaria que se utiliza como plantilla para marcar los píxeles que se utilizan para la inserción del mensaje. Los píxeles seleccionados por los algoritmos pertenecen a regiones de la imagen que por naturaleza son ruidosas por lo que es difícil detectar un mensaje oculto mediante un ataque visual. Estas regiones pueden ser bordes, texturas no homogéneas, regiones con alto contraste.

En las siguientes secciones se describen a detalle cada algoritmo. Los primeros dos algoritmos fueron propuestos en [3] bajo los nombres de *ConDith* y *ConDithSpread*. El tercer algoritmo lo hemos denominado *Context*.

4.1. ConDith

Este algoritmo selecciona los píxeles de áreas de la imagen en donde existe un alto contraste local. El método que utiliza para el procesamiento de la imagen cubierta es una modificación del proceso Dithering.

Dithering es un método de reducción de color utilizado principalmente en los monitores monocromáticos y en las impresoras. El objetivo es simular la presencia de más colores de los que realmente se tienen disponibles. En el caso de las imágenes en escala de grises el proceso Dithering puede representar todas las tonalidades usando solo los colores blanco y negro. Existen diferentes métodos de Dithering para la representación de imágenes en escala de grises sin embargo la idea básica es procesar los píxeles de la imagen original mediante algún método después el resultado se compara con un umbral que si es sobrepasado se le asigna al correspondiente píxel en la imagen Dithering el valor blanco, en caso contrario le será asignado el negro.

Condith al igual que el proceso Dithering obtiene una imagen binaria a partir de una en escala de grises, la modificación que se hace al proceso Dithering esta en la manera

en la que se escoge el umbral. En Context lo que se desea es resaltar aquellos píxeles en los que exista alto contraste con respecto a sus vecinos. El criterio que se utiliza es el siguiente.

Píxel[y,x] es usable, si dither[y,x]=0

$$dither[x,y] = \begin{cases} 0; & diff > Cim \\ 255: & otro \end{cases} \tag{4.1}$$

En donde *diff* es la diferencia entre el píxel analizado y sus en las posiciones $(y, x+1)$, $(y+1, x-1)$, $(y+1, x)$ y $(y+1, x+1)$. *Cmin* es un valor de contraste mínimo. De esta manera se escogen aquellos píxeles en donde el contraste local es mayor, estos píxeles son principalmente bordes.

En la imagen binaria creada por ConDith los píxeles marcados con color negro indican los píxeles de la imagen original que se deben usar para insertar el mensaje. Para insertar el mensaje se utiliza el método LSB (Least Significant Bit) que consiste en insertar los bits del mensaje en los bits menos significativos de los píxeles de la imagen. En la figura 4.3 se muestra como se insertan los bits de una mensaje utilizando ConDith.

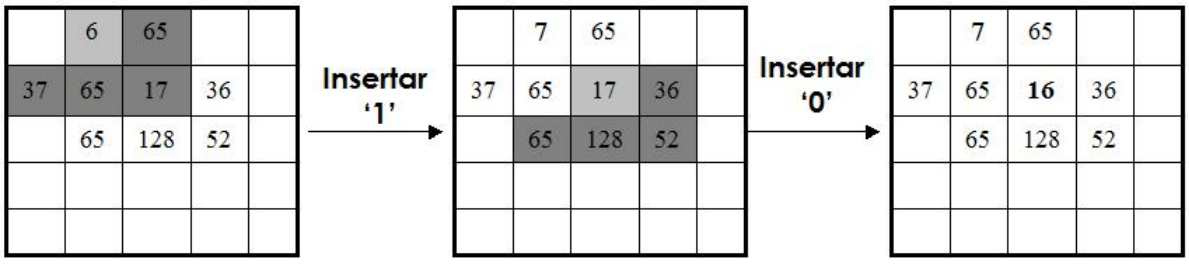


Figura 4.3: Ejemplo del uso de Condith.

4.2. ConDithSpread

Es una modificación de ConDith en la que se pretende aumentar la aleatoriedad de los píxeles seleccionados al insertar una condición extra a la presentada en ContDith Así tenemos:

Píxel[x,y] es usable, si dither [x,y]= 0

$$dither[x,y] = \begin{cases} 0; & diff > Cim \wedge rnd[x,y] < (n+1)/2 \\ 255: & otro \end{cases} \tag{4.2}$$

En donde *rnd* es un número seudoaleatorio en el intervalo de 0 a n. Al aumentar la condición $rnd[x,y] < (n+1)/2$ se disminuyen a aproximadamente a la mitad los píxeles seleccionados. Sin embargo la semilla del generador del número aleatorio se puede usar como llave esteganográfica. Esto aumenta la seguridad pues con la llave la posición del mensaje no sólo depende de las características de la imagen sino además del numero seleccionado como semilla.

4.3. ConText

Siguiendo el esquema de los algoritmos anteriores Context crea una imagen binaria que se utiliza para indicar las posiciones de los píxeles que se pueden modificar en la imagen cubierta. Sin embargo, en este algoritmo el criterio de selección de píxeles se basa en la detección de regiones en donde existen texturas no homogéneas.

Como se vio en el capítulo dos a pesar de no haber una definición única de lo que es la textura en imágenes digitales siguiendo los métodos estadísticos la textura se pude definir en términos de la distribución en los niveles que existen en una región así se puede decir que una región contiene textura no homogénea si existe una diversidad significativa en los niveles de gris. Así el algoritmo busca regiones en las que localmente existe diversas tonalidades.

El algoritmo Context se presenta en la tabla 4.1.

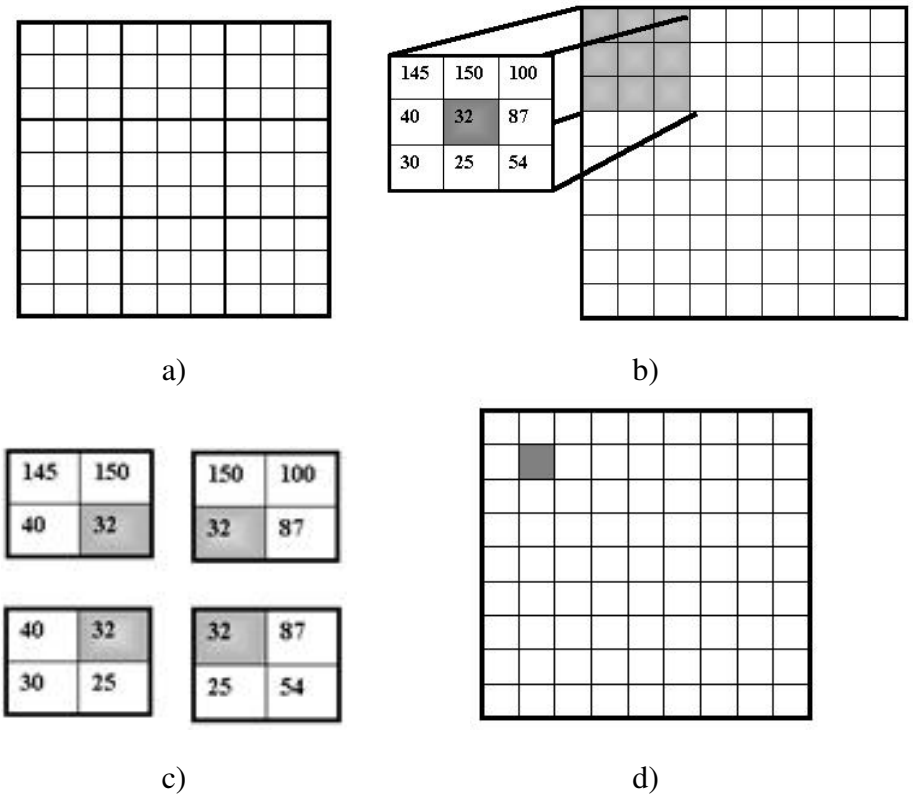


Figura 4.4: Ejemplo de un píxel seleccionado por Context para la inserción

El algoritmo se puede modificar de la Context de la misma forma que ConDith-Spread para aumentar la seguridad del algoritmo.

En general en los tres algoritmos se pueden modificar para emplear un método diferente a LSB para incrementar su robustez. Sin embargo, en este trabajo solo se mostrarán los resultados obtenidos usando LSB en la fase de inserción.

En la figura 3 se muestra un ejemplo de los píxeles que se seleccionan para la inserción del mensaje usando los tres algoritmos anteriores. En ella se presenta un claro ejemplo de cómo ConText selecciona más pixeles para la inserción en las imágenes en las que existe una diversidad de tonalidades que el algoritmo ConDith.

En el siguiente capítulo se presentan los resultados obtenidos después de aplicar el ataque visual descrito en la sección 1.4.1. El objetivo de las pruebas realizadas es mostrar es presentar la ventaja que presentan los algoritmos adaptivos aquí descriptos

frente a los esquemas tradicionales de LSB.

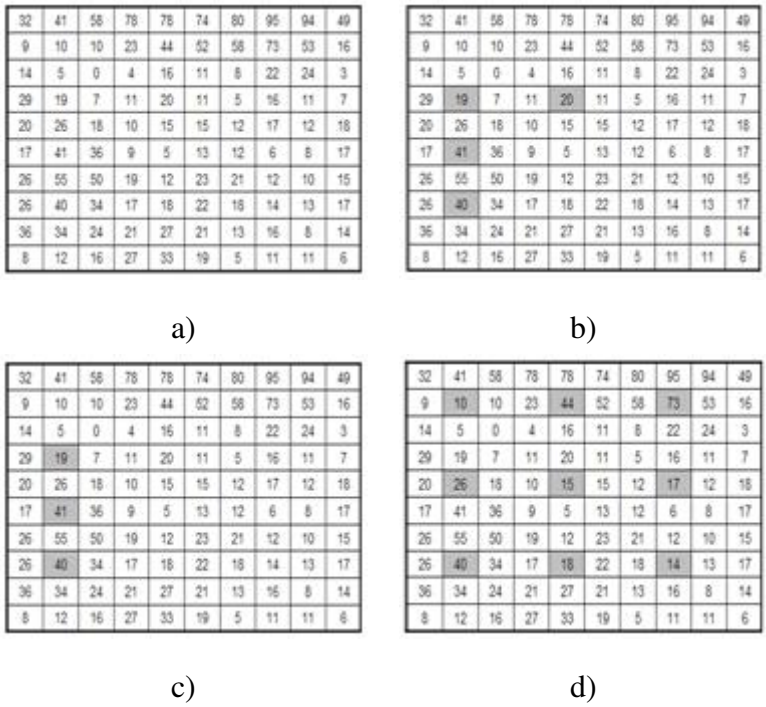


Figura 4.5: Resultados de los criterios de selección. a)Imagen original. b) ConDith. c) ConDithSpread e) ConText

Cuadro 4.1: **Algoritmo Context**

Entrada:

- Imagen en escala de grises (imagen cubierta), mensaje de texto.

Salida:

- Imagen en escala de grises modificada por los bits del mensaje (Setgo-imagen).

1. Se divide la imagen cubierta en bloques no traslapados de $3x3$.
2. Cada bloque generado se divide en cuatro sub-bloques de tamaño $2x2$ de manera que cada uno contengan al píxel central.
3. Se dice que un sub-bloque es *bueno* si contiene por lo menos tres niveles de gris distintos.
4. El píxel central de un bloque se elige para insertar el mensaje si los cuatro sub-bloques son *buenos*.
5. Se inserta el mensaje en la imagen cubierta empleando LSB en los píxeles seleccionados.

Capítulo 5

Resultados

El objetivo de este cápitulo es mostrar las diferencias entre los algoritmos adaptivos presentados en el cápitulo 3 y los algotimos tradicionales no adaptivos *LSB* y una de sus variantes denominado *LSBSpread* cuando se enfrentan a un ataque visual.Los algoritmos fueron implementados usando Matlab 7.0 al igual que los experimentos realizados.

Como primera sección en este cápitulo se muestra un ejemplo de los algortimos tradicionales no adaptivos LSB para imágenes usados durante los experimentos

5.1. Algortimos no adaptivos

El algoritmo *LSB* inserta la información de manera consecutiva en cada uno de los píxeles de la imagen. En la figura 5.1 se presenta un ejemplo del algoritmo *LSB* tradicional.

LSBSpread es una modificación del algoritmo LSB en el que la información no se inserta de manera consecutiva en los pixeles de la imagen cubierta, sino a distancias variables determinadas por un generador de números semi-aleatorios.En nuestro caso las distancias que separan a los pixeles están en el rango de 1 a 5 pixeles.

El objetivo del algoritmo *LSBSpread* es dispersar el mensaje en la imagen de tal forma que no sea tan fácil determinar las posiciones de los pixeles modificados. Para ello,

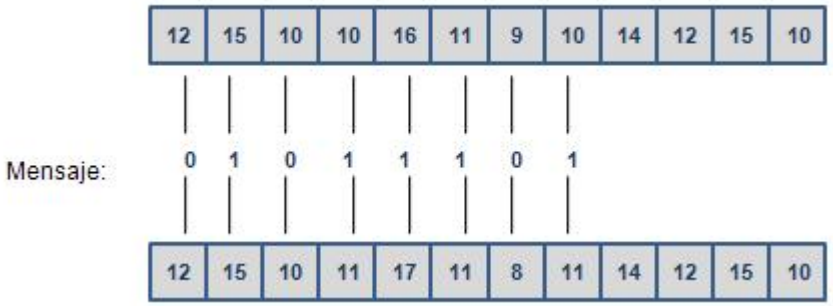


Figura 5.1: Ejemplo de funcionamiento del algoritmo *LSB* tradicional.

la semilla utilizada por el generador de números semi-aleatorios se utiliza como stego-llave que incrimenta la seguridad del sistema. En la figura 5.2 se ilustra el algoritmo *LSBSpread*.

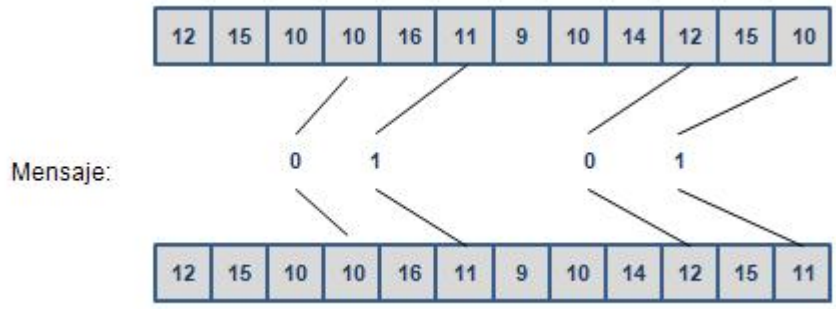
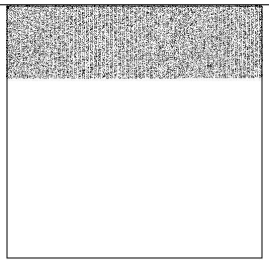
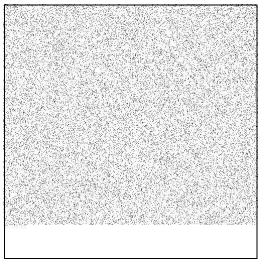


Figura 5.2: Ejemplo de funcionamiento del algoritmo *LSBSpread*

Como se puede observar en las figuras fig:Ejn y fig:Ejns, al no insertar la información en pixeles continuos de la imagen, el algoritmo *LSBSpread* inserta una menor cantidad de información en la misma área.

En la tabla 5.1 se presentan los resultados obtenidos al aplicar los algoritmos *LSB* y *LSBSpread* para insertar un archivo de texto de 10k en la imagen en escala de grises de *Lena*. En la primera columna de la tabla se muestra la imagen original de *Lena* usada como cubierta. En la tercera columna se muestran las imágenes stegos obtenidas. En la

última columna se presentan las imágenes que se obtienen al restar a la imagen stego la imagen original, la imagen resultante muestran los pixeles modificados.

Imagen Cubierta	Método	Stego Imagen	Diferencia entre original y stego
	LSB		
	LSBSpread		

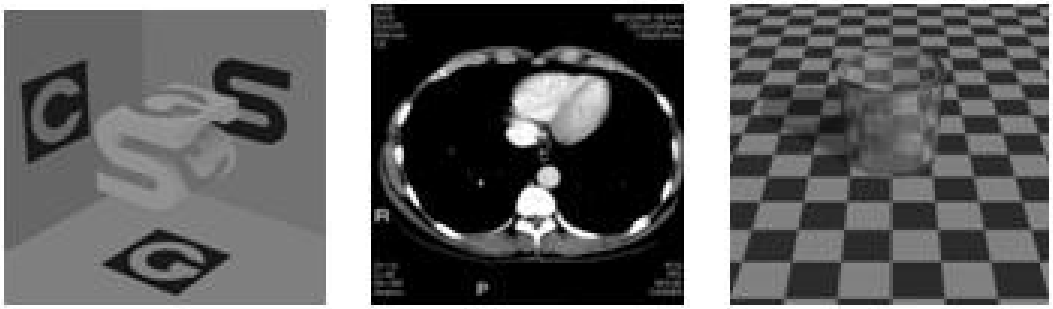
Cuadro 5.1: Diferencias entre LSB y LSBSpread.

5.2. Experimentos

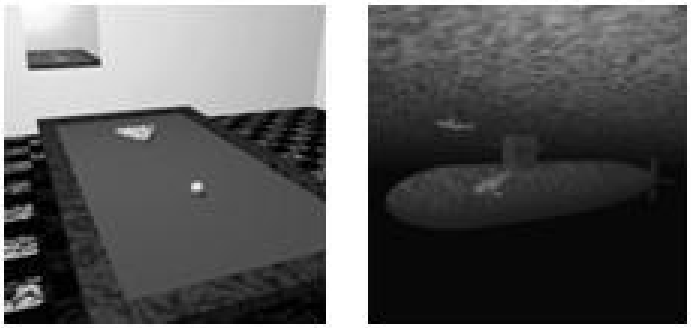
En ésta sección se describen los experimentos realizados así como los resultados obtenidos.

Durante los experimentos se utilizaron imágenes en escala de grises de $512x512$ pixeles y la información insertada fue un archivo de texto de 10k del cual fueron insertados en las imágenes tantos bit como el algotimo lo permitia. En la figura 5.3 se presentan las imágenes utilizadas durante los experimento, estas imágenes se caracterizan por tener amplias regiones homogéneas en las que es fácil detectar un mensaje oculto mediante un ataque visual.

Los experimentos tienen el objetivo de mostrar el comportamiento de los algoritmos en dos aspectos:



a) b) c)



d) e)

Figura 5.3: Imágenes de prueba

- *Imperceptibilidad ante ataques visuales:*Es decir que después de un ataque visual(descrito en la sección 2.5.1) no se pueda percibir la presencia de información extra presente en la imagen.

- *Capacidad:* Es decir la cantidad de información que permite almacenar cada algoritmo.

Para mostrar el comportamiento de los algoritmos en estos aspectos se realizaron 2 experimentos.

5.2.1. Experimento 1

El objetivo del primer experimento es mostar el comportamiento de los algoritmos no adaptivo *LSB* y *LSB-Spread* y los algoritmos adaptivos *ConDith*, *ConDith-Spread*

y *Context*. El experimento consistió en insertar un archivo de texto de 2.5K usando las diferentes imágenes de prueba una vez insertada la información se aplica un ataque visual con la finalidad de detectar la presencia del mensaje oculto en la estegoimagen.

En las figuras 5.4 a 5.5 se presentan los resultados obtenidos durante este experimento. En cada una de las figuras se muestran 6 imágenes. La primer imagen muestran en el resultado de aplicar el ataque visual a la imagen original sin mensaje oculto, las siguientes imagenes corresponden a las obtenidas despúes de aplicar un ataque visual en imágenes en las que se habia insertado un archivo de texto usando los algoritmos LSB, LSB-Spread, ConDith, ConDith-Spread y Context respectivamente.

En la figura 5.4 se muestran los resultados del experimento aplicado en la figura *CSG*. En esta imagen se presentan grandes regiones homogeneas que se conservan después de un ataque visual. Como se observa, cuando la información se inserta utilizando los agoritmos no adaptivos estas regiones no se evitan y el mensaje se revela en forma de ruido. En cambio, cuando se utilizan los algoritmos adaptivos se evitan estas regiones y el mesaje oculto no se detecta.

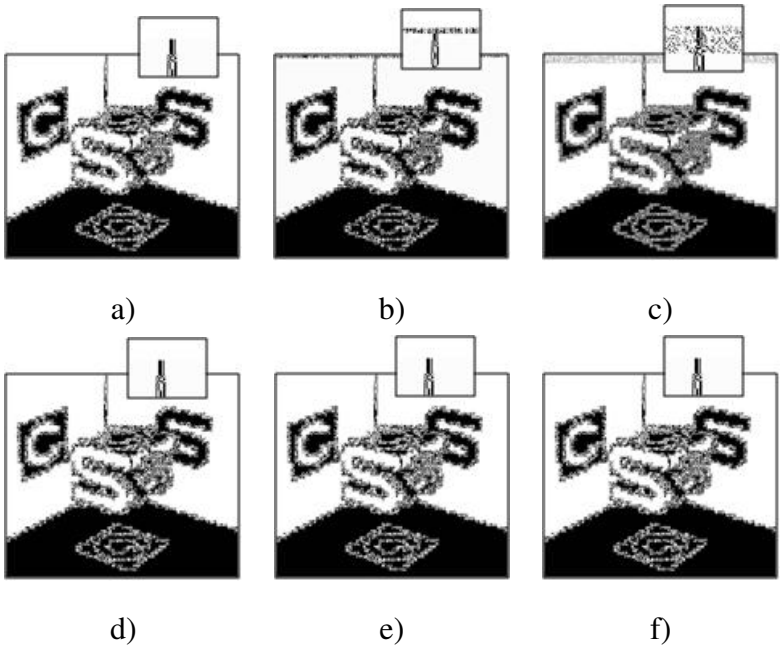
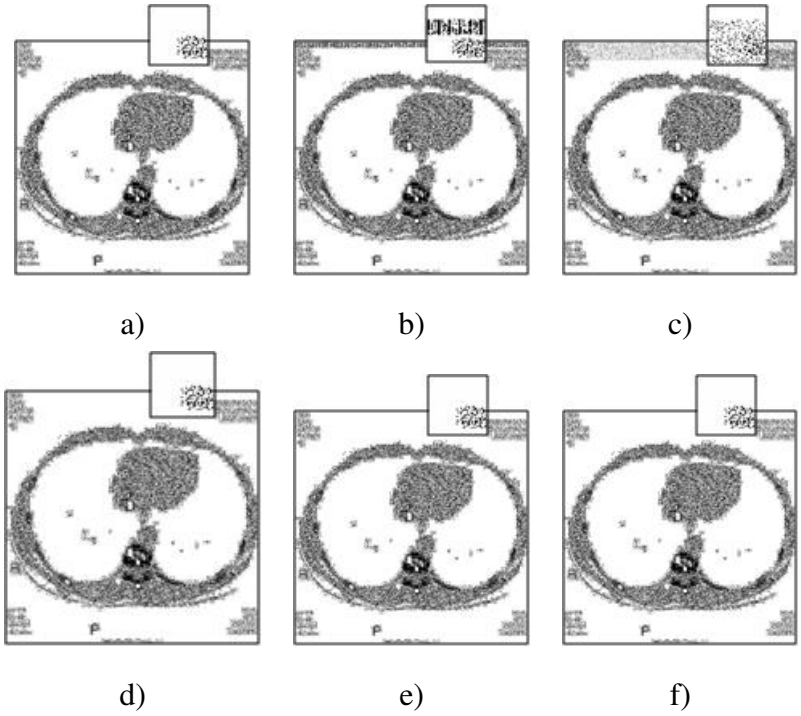


Figura 5.4: Resultados de un ataque visual a la imagen Csg

En la figura 5.5 se presentan los resultados obtenidos del experimento en la imagen medica *Ultra*.



a) b) c)

d) e) f)

Figura 5.5: Resultados de un ataque visual a la imagen Ultra

En la figura fig:revi2 se presentan los resultados obtenidos del experimento en la imagen medica *Pool*. La imagen resultante del ataque visual a la imagen *Pool* a diferencia de las imágenes obtenidas de *Csg y Ultra* no presenta regiones homógeneas grandes. Sin embargo, siguen un patrón en el que es posible detectar un mensaje oculto cuando se utilizan algoritmos no adaptivos.

5.2.2. Experimento 2

En este experimento no fue insertada información en las imágenes. El experimento consiste en aplicar el criterio de selección de pixeles sobre las imágenes de prueb y mostrar y contar los pixeles que son considerados para la insección.

Aqui no se aplicaron los algoritmos LSB y LSBSpread. Dado que la capacidad de

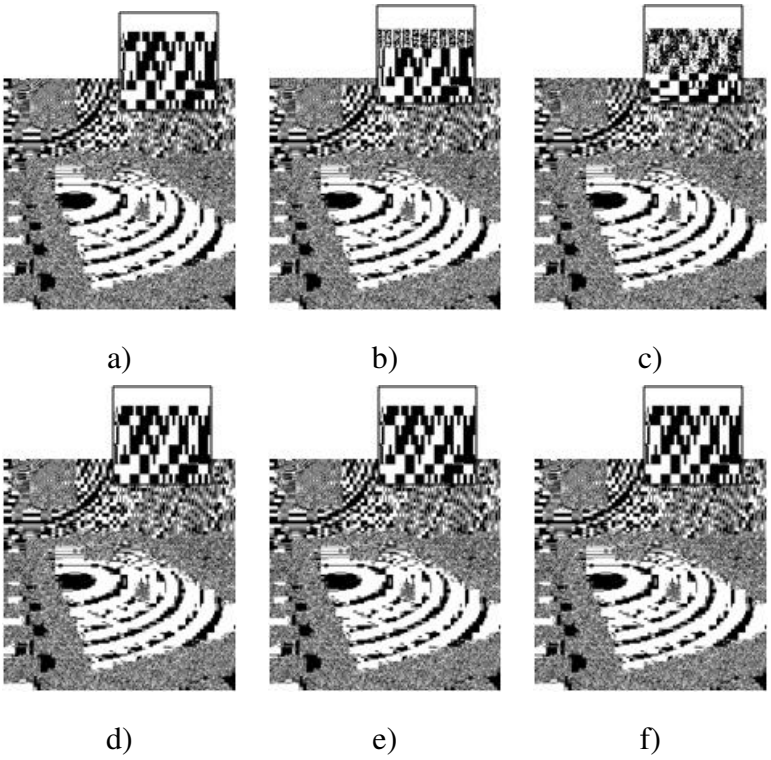


Figura 5.6: Resultados de un ataque visual a la imagen Pool

los algoritmos tradicionales de LSB depende directamente del tamaño de la imagen y no de su contenido. Así, los algoritmos tradicionales de LSB pueden almacenar informació en cada uno de los los píxeles de la imagen. En el caso de caso de LSBSpread la capacidad del algoritmo ademas depende de la distancia máxima entre pixeles a la que se puede insertar información. Así, en el caso de las imágenes de prueba que son de 512x512 pixeles se pueden insertar información en 262144 píxeles. Miestras que con el algoritmo LSBSpread se pueden insertar en el peor de los casos, es decir que el generador de números aleatorios diera en todos los casos una distancia de 5 pixeles, se tendrían 52428 píxeles para insertar información .

En los algoritmos adaptivos la cantidad de información que se puede almacenar depende directamente del contenido de la imagen y no se puede determinar a priori la cantidad de información que puede ser insertada pues además de cumplir la condición se necesita que una vez insertada la información se cumpla de nueva cuenta la condición

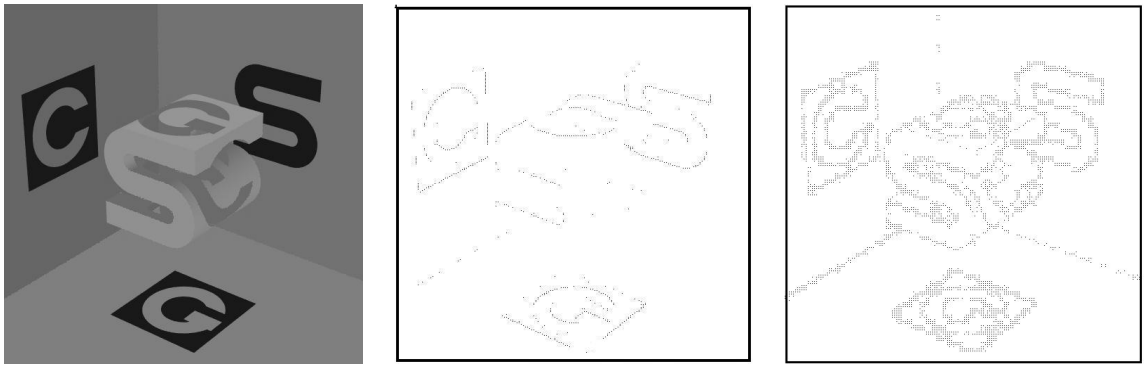
para poder recuperarla .

En la tabla 5.2 se presenta la capacidad alcanzada por los algoritmos adaptivos alcanzada en las imágenes de prueba. la capacidad se mide por el número de pixeles que de acuerdo con el criterio de los diferentes algortimos se pueden utilizar durante el proceso de inserción. Como se puede observar en la tabla, en todos los casos el algoritmo propuesto Context obtiene una mayor capacidad de insercción.

		Métodos		
		ConDithSpread	ConDith	ConText
Imágenes	Csg	323	654	3740
	Pool	742	1516	8693
	Ultra	1057	2220	7164
	Glass	1082	2148	7988
	Subs	1488	3100	13669

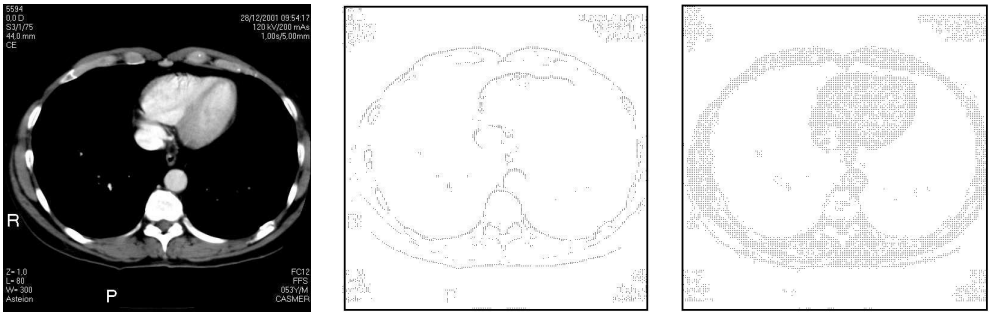
Cuadro 5.2: Capacidad obtenida por los algoritmos adaptivos en las imágenes de prueba

En las figuras 5.7 a 5.11 se muestran los pixeles que fueron seleccionados para la inserción de los algoritmos *ConDith* y *ConText*. En donde se puede observar que ninguno de los aldoritmos seleccionan pixeles en las regiones homógeneas. Sin embargo, mientras que ConDith se concentra pixeles principalmente en los bordes, ConText aprovecha regiones con variedad en texturas que le permiten obtener mayor capacidad de insercción.



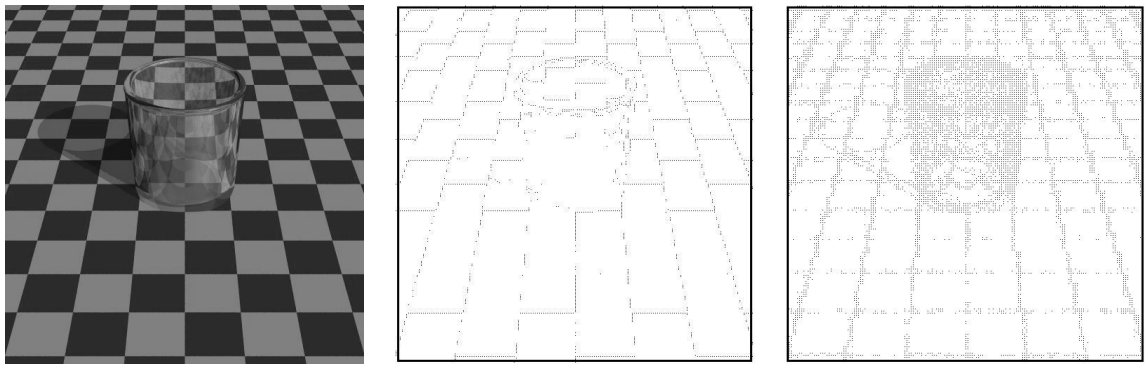
a) b) c)

Figura 5.7: Píxeles seleccionados para la insercción usando ConDith y Context en la imagen Csg



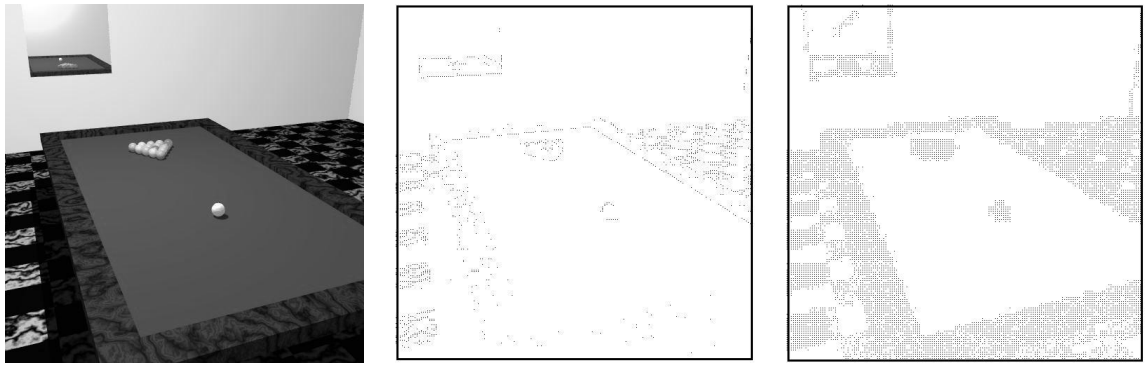
a) b) c)

Figura 5.8: Píxeles seleccionados para la insercción usando ConDith y Context en la imagen Pool



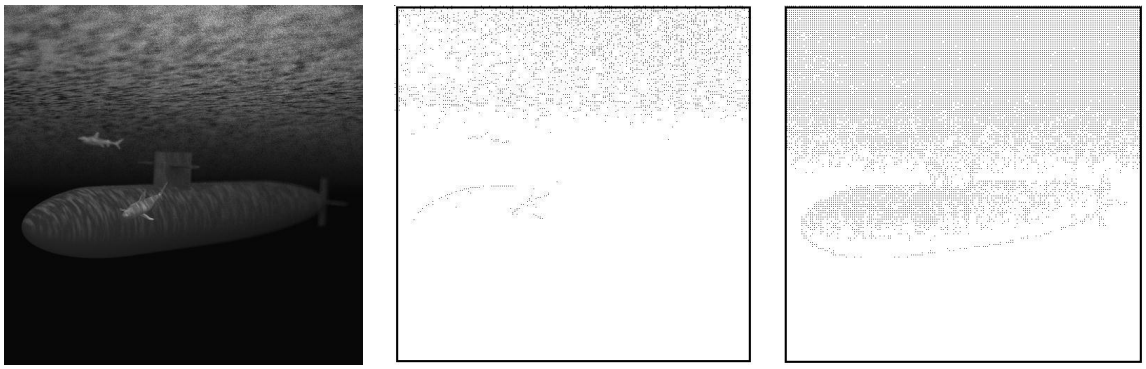
a) b) c)

Figura 5.9: Píxeles seleccionados para la insercción usando ConDith y Context en la imagen Ultra



a) b) c)

Figura 5.10: Píxeles seleccionados para la insercción usando ConDith y Context en la imagen Glass



a) b) c)

Figura 5.11: Píxeles seleccionados para la insercción usando ConDith y Context en la imagen Subs

Capítulo 6

Conclusiones y trabajos futuros

Se presentó un algoritmo esteganográfico llamado Context que se adapta a las características de la imagen cubierta al seleccionar para la inserción del mensaje aquellos píxeles que pertenecían a regiones de con textura homogénea. Como resultado del presente trabajo de invertigación se publico el trabajo denominado *Adaptive Steganography based on textures* en el 17th International Conference on Electronics, Communications and Computers (CONIELECOMP'07) [2].

El algoritmo Context fue comparado con los algoritmos adaptivos *ConDith* y *ConDithSpread* presentados en [1] que sirvieron como base para el desarrollo del presente trabajo. La principal contribución de [1] es la selección de píxeles con base a las características de la imagen cubierta por lo que las posiciones del mensaje cambian de acuerdo al tipo de imagen lo que hace que los algoritmos sean adaptivos. Esto reduce la probabilidad de detección al insertar la menor cantidad de ruido.

Como podimos apreciar en los resultados, los algoritmos adaptivos son imperceptibles ante ataques visuales pues tanto *ConDith* como *ConText* evitan las regiones homogéneas. Sin embargo, *ConDith* se enfoca en los contornos de la imagen mientras que *ConText* en las texturas de la imagen, logrando mayor capacidad.

Indudablemente, el presente trabajo es suceptible de mejoras que podrían ser investigadas en un futuro cercano. Algunos de los trabajos futuros pueden ser:

- Ampliar el algoritmo para que maneje imágenes a color.

- Una de las mayores desventajan que presentan tanto los algoritmos propuestos en [1] como el algoritmo propuesto,es que una ves seleccionados los pixeles el proceso de insercción propiamente es el mismo que el utilizados por los algoritmos LSB que es suceptible a cambios en la imagen. Uno de los trabajos futuros es encontrar un proceso de insercción en el que no se pierda la información cuando se manipule la imagen.

- Como ya se mencionó en los algoritmos adaptivos no se puede determinar apriori la cantidad de información que puede ser insertada en la imagen y esta es mucho menor que la capacidad obtenida por los algoritmos tradicionales, sin embargo, se podría combinar la esteganografí con un proceso de compresión para lograr mejores resultados.

Índice de figuras

Índice de cuadros

Bibliografía

[1] A. Schneidewind A. Franz. Adaptive steganography based on dithering. *Proceedings of the Workshop on. Multimedia and Security, Magdeburg, Germany*, pages 56–62, 2004.

[2] C. Feregrino-Uribe D.R. Herrera-Moro, R. Rodriguez-Colin. Adaptive steganography based on textures. In *conielecomp, p. 34, 17th International Conference on Electronics, Communications and Computers (CONIELECOMP'07),*, 2007.

[3] R. Du J. Frifrich. Secure steganographic methods for palette images. *Information Hiding, 3rd nt. Workshop*, pages 47–60, 1999.

[4] E.T. Lin and E.J. Delp. A review of data hiding in digital images. *in Proceedings of the Image Processing, Image Quality, Image Capture Systems Conference, PICS '99*, pages pp. 274–278., 1999.

[5] P. Honeyman N. Provos. Hiden and seek: An introduction to steganography. *IEEE Security & Privacy Magazine*, 1:32–44, 2003.

[6] F.A.P Petitcolas R.J Anderson. On the limits of steganography. *Selected Areas in Communications, IEEE Journal*, vol 16:p.p. 474–481, 1998.

[7] G.J. Simmons. The prisoners' problem and the subliminal channel. *CRYPTO*, pages 411–431, 1983.