DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA ELÉCTRONICA Y TELECOMUNICACIONES

TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERA EN ELECTRÓNICA Y TELECOMUNICACIONES

TEMA: DESARROLLO Y ANÁLISIS DE UNA TÉCNICA ESTEGANOGRÁFICA EN ZONAS RUIDOSAS DE LA IMAGEN MEDIANTE TRANSFORMACIONES DE COLOR REVERSIBLES

AUTOR: ONOFRE CONCHA, GABRIELA ESTEFANÍA

DIRECTOR: ING. ACOSTA BUENAÑO, FREDDY

SANGOLQUÍ

2016

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA ELECTRÓNICA Y TELECOMUNCACIONES

CERTIFICACIÓN

Certifico que el trabajo de titulación, "DESARROLLO Y ANÁLISIS DE UNA TÉCNICA ESTEGANOGRÁFICA EN ZONAS RUIDOSAS DE LA IMAGEN MEDIANTE TRANSFORMACIONES DE COLOR REVERSIBLES" realizado por la señorita GABRIELA ESTEFANÍA ONOFRE CONCHA, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar a la señorita GABRIELA ESTEFANÍA ONOFRE CONCHA para que lo sustente públicamente.

Sangolquí, 03 de octubre del 2016

ING. FREDDY ROBERTO ACOSTA BUENAÑO

DIRECTOR

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

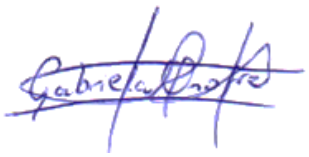
CARRERA DE INGENIERÍA ELECTRÓNICA Y TELECOMUNCACIONES

AUTORÍA DE RESPONSABILIDAD

Yo, GABRIELA ESTEFANÍA ONOFRE CONCHA, con cédula de identidad N° 172262770-8, declaro que este trabajo de titulación "DESARROLLO Y ANÁLISIS DE UNA TÉCNICA ESTEGANOGRÁFICA EN ZONAS RUIDOSAS DE LA IMAGEN MEDIANTE TRANSFORMACIONES DE COLOR REVERSIBLES" ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, 03 de octubre del 2016



GABRIELA ESTEFANÍA ONOFRE CONCHA

C.C. 1722627708

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA ELECTRÓNICA Y TELECOMUNCACIONES

AUTORIZACIÓN

Yo, GABRIELA ESTEFANÍA ONOFRE CONCHA, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación "DESARROLLO Y ANÁLISIS DE UNA TÉCNICA ESTEGANOGRÁFICA EN ZONAS RUIDOSAS DE LA IMAGEN MEDIANTE TRANSFORMACIONES DE COLOR REVERSIBLES" cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Sangolquí, 03 de octubre del 2016

GABRIELA ESTEFANÍA ONOFRE CONCHA

C.C. 1722627708

DEDICATORIA

El presente trabajo es dedicado especialmente a mi familia, que es un ejemplo de superación y perseverancia y sin ellos no hubiera podido cumplir esta meta tan importante en mi carrera profesional.

A mis padres Franklin y Amparito que han sabido comprenderme, apoyarme incondicionalmente y brindarme un hogar lleno de amor donde he encontrado la fuerza necesaria para seguir adelante y cumplir mis metas y sueños.

A mi hermano Danny, quien mas que un hermano ha sido un amigo con el que siempre puedo contar, que me ha acompañado en los buenos y malos momentos, un ejemplo a seguir impulsándome a no rendirme y ser mejor cada día.

AGRADECIMIENTO

Agradezco a mis padres Franklin y Amparo y a mi hermano Danny, por todo el sacrificio y amor incondicional que me han brindado a lo largo de mi vida, por sus enseñanzas y su ejemplo de perseverancia a pesar de los obstáculos que se presentan. Gracias a ellos y por ellos he cumplido con mis metas académicas y he logrado culminar esta difícil etapa.

A mi novio Julio Enríquez que ha estado a mi lado en los momentos buenos y momentos difíciles apoyándome incondicionalmente y brindándome su amor, su tiempo, su comprensión y ayuda para que cumpla mis objetivos.

A mis amigas del colegio con las que he logrado formar una sólida amistad que ha sobrevivido y se ha fortalecido a través de los años. A los grandes amigos que he encontrado durante la vida universitaria con los que he vivido experiencias inolvidables, viajes y noches de desvelo para cumplir con las actividades académicas; gracias por compartir conmigo sus conocimientos y acompañarme en este arduo trayecto.

Finalmente, un agradecimiento especial al Ing. Freddy Acosta, mi director de proyecto que ha sido un ejemplo de persona y de docente, me ha inculcado sus conocimientos, me ha dado su tiempo y guiado durante la realización de este proyecto. Y al Ing. Julio Larco por su colaboración y consejos a lo largo de este trabajo.

ÍNDICE DE CONTENIDOS

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

RESUMEN

La esteganografía es una ciencia que consiste en ocultar información, presente desde tiempos inmemorables, con el objetivo de enviar un mensaje secreto que solo puede ser leído por su destinatario. Con el avance de la tecnología, se han desarrollado nuevas técnicas esteganográficas en medios portadores digitales como imágenes y a la par han ido evolucionando métodos de estegoanálisis para descubrir la información oculta e impedir la comunicación. Por este motivo, este trabajo busca crear una técnica esteganográfica en imágenes, robusta ante ataques estadísticos, mediante la combinación de dos métodos: uno basado en transformaciones de color reversibles que modifican la imagen secreta a transmitir, para la obtención de un mosaico que luzca similar a una imagen portadora seleccionada previamente; y un método que busca las zonas más aptas (texturas y bordes) en el mosaico creado, para ocultar la información relevante requerida y recuperar la imagen secreta. Estas zonas son de naturaleza ruidosa, por lo que representan un reto para los estegoanalistas al momento de extraer características para entrenar clasificadores encargados de definir si una imagen tiene información embebida o no. Para la detección de texturas y bordes en la imagen, se utiliza operadores diferenciales y filtros, y se obtiene un mapa donde se incrusta la información usando la técnica esteganográfica *LSB matching*. Por último, se realizaron mediciones de calidad de la imagen y se evaluó la efectividad del método propuesto mediante el análisis de histogramas y una comparación de la indetectabilidad estadística respecto al uso del sistema *LSB matching* en toda la imagen con la herramienta StegExpose.

Palabras claves

- **ESTEGANOGRAFÍA**
- **ESTEGOANÁLISIS ESTADÍSTICO**
- **ZONAS RUIDOSAS**
- **TRANSFORMACIONES DE COLOR**
- **INDETECTABILIDAD ESTADÍSTICA**

ABSTRACT

Steganography is a science which consists in hiding information, present during the History, in order to send a secret message that can only be read by its recipient. With the advancement of technology, it has developed new steganographic techniques in digital media carriers such as images and at the same time it has evolved steganalysis methods to discover hidden information and prevent communication. For this reason, this work seeks to create a steganographic technique in images, robust against statistical attacks by combining two methods: one based on transformations of reversible color which change the secret image to be transmitted, to obtain a mosaic that looks like a previously selected carrier image; and a method which seeks the most suitable areas (textures and borders) in the mosaic created to hide relevant required information and recover the secret image. These areas are noisy nature, so they represent a challenge for estegoanalysts when extract features to train classifiers with the purpose of defining whether an image has embedded information or not. Differential operators and filters are used for detection of textures and edges in the image, and there is obtained a map where information is embedded using the steganographic technique *LSB matching*. Finally, measurements of image quality were performed and the effectiveness of the proposed method was evaluated by histogram analysis and comparison of statistical undetectability with regard to the use of *LSB matching* system throughout the image with *StegExpose* tool.

Keywords:

• STEGANOGRAPHY

• STATISTICAL STEGANALYSIS

• NOISY AREAS

• COLOR TRANSFORMATIONS

• STATISTICAL UNDETECTABILITY

CAPÍTULO 1

1. INTRODUCCIÓN

1.1 Antecedentes

A lo largo de la historia, la ocultación de información ha estado presente en la humanidad con el fin de enviar mensajes secretos que únicamente pueden ser leídos por el receptor al que están dirigidos. Ejemplos de esto se dieron en la Antigua Grecia, como el caso de un personaje que tatúa un mensaje en la cabeza rasurada de su esclavo y espera que crezca su cabello para enviarlo a su respectivo destinatario. Un ejemplo más familiar es la escritura con tinta invisible, fabricada con zumo de limón u otros compuestos similares, incluso reacciones químicas se usaron para escribir mensajes visibles solamente tras la exposición al calor.

"De manera similar, durante la Segunda Guerra Mundial se hacían pequeñas perforaciones sobre las letras de interés de un periódico de tal forma que al sostenerlo a la luz se pueden observar todas aquellas letras seleccionadas e interpretarlas en forma de mensaje." (Instituto Nacional de Tecnologías de la Comunicación, 2013)

Tras esta guerra, se dieron considerables progresos en muchos aspectos de la ciencia. En la rama de la esteganografía, uno de los grandes avances es el uso de canales digitales como audio, video e imágenes, para encubrir mensajes. Esto se dio debido a que en la actualidad, gran parte de la información presente alrededor del mundo es digital, y transmitida por medios electrónicos. Dichos medios poseen protocolos de seguridad como métodos de encriptación, llaves de acceso, etc., pero a pesar de ello, la información es susceptible a ser interceptada por agentes distintos a su destino final.

Una imagen es un tipo de información muy usada en aplicaciones esteganográficas por su amplia difusión en Internet. Estas imágenes suelen contener información privada o confidencial que requiere estar oculta durante la transmisión de las mismas. Se han propuesto muchos métodos para asegurar la transmisión de imágenes. Los enfoques más comunes son la encriptación de imágenes y la ocultación de datos (esteganografía) (Ya-Lin & Wen-Hsiang, 2014).

La encriptación utiliza las propiedades características de la imagen para convertirla en una imagen ruidosa de tal manera que no se pueda visualizar su contenido, a menos que se disponga de la clave secreta para descifrarla (Satwinder & Varinder, 2015) (Patidar, Pareek, Purohit, & Sud, 2011). No obstante, el hecho de transmitir una imagen como ruido puede llamar la atención de cualquier atacante y advertir de la presencia de un mensaje oculto. Por lo tanto, una alternativa se encuentra en los métodos esteganográficos en los que la información secreta viaja en una portadora, en este caso, el medio más común para ocultar información son las imágenes por su alta difusión en la red.

El principal problema de ocultar información en una imagen es la capacidad de incrustación, por ejemplo, si se desea ocultar una imagen dentro de una imagen portadora del mismo tamaño, se debe comprimir la imagen secreta antes del proceso de incrustación. Esta compresión ocasiona grandes distorsiones y pérdidas en la imagen original (Hu, et al., 2013), lo que es muy perjudicial para imágenes con contenidos relevantes. En (Ya-Lin & Wen-Hsiang, 2014) se propone una nueva técnica como solución a este problema que transforma la imagen a transmitir o también llamada imagen secreta en una imagen mosaico de igual tamaño, la cual luce como una imagen portadora seleccionada arbitrariamente. Se aplica técnicas de transformación de color que permiten recuperar la imagen secreta con un bajo porcentaje de pérdidas. La información requerida suficiente para recuperar la imagen secreta se oculta mediante técnicas esteganográficas en el mosaico creado.

Existen diversas técnicas esteganográficas, pero las más usadas son las de domino espacial por su baja complejidad y su gran capacidad de incrustación, sin embargo, no son robustas ante las pérdidas por la compresión de imágenes. Se han propuesto varias técnicas que ocultan datos para este tipo en (Ya-Lin & Wen-Hsiang, 2014) (Nusrati, A., & Karimi, 2015) (Fridrich, Du, & Meng, 2000). En estas técnicas la cadena de bits del mensaje secreto se introduce en el flujo de bits de los píxeles de la imagen portadora.

La sustitución del bit menos significativo *(LSB o Least Significant Bit)* es una de las técnicas más sencillas, en donde se transforma la información en un flujo de bits y se oculta cada uno de ellos en el bit menos significativo de cada pixel. La variación

del color no es reveladora por lo que no se detecta visualmente. Sin embargo esta operación crea irregularidades estadísticas en la imagen, por lo que es susceptible a ataques de estego-analizadores (Westfeld & Pfitzmann, 1999), haciéndolo un método no seguro.

El avance del método de sustitución *LSB* se denomina *LSB matching*. Esta técnica es muy similar a la anterior, pero en lugar de sustituir el valor del *LSB* directamente por el bit a incrustar, se lo modifica cuando el bit a incrustar no coincide con el *LSB* del pixel correspondiente (Coltuc & Chassery, 2007). El resultado es una imagen con información oculta, difícilmente detectada visualmente y con irregularidades estadísticas no tan evidentes. Por lo que los estego-analizadores han recurrido a técnicas de *machine learning*, es decir, entrenar clasificadores para que sean capaces de diferenciar una imagen portadora de una imagen con mensaje oculto. Estos clasificadores detectan características de la imagen que se modifican habitualmente al ocultar información.

1.2 Justificación e Importancia

La esteganografía es una técnica que se encuentra en constante desarrollo, que además posee la capacidad para adaptarse a nuevas tecnologías. A medida que herramientas esteganográficas más avanzadas son creadas, las técnicas de estegoanálisis empleadas también incrementan su complejidad.

Esta ciencia tiene múltiples aplicaciones con fines constructivos en campos como la medicina, mecanismos de autenticación, entre otros. Por ejemplo, en lugar de tener una gran base de datos de los pacientes en un hospital, se puede insertar el historial de cada uno en su fotografía, en radiografías y similares.

En cualquier caso, la esteganografía debe ser combinada con otras técnicas para optimizar su utilidad, por sí sola puede ocultar el hecho de que una imagen portadora contenga información que se desea transmitir a un receptor en particular, pero si se descubre este intercambio, es posible que el adversario conozca el mensaje secreto o pueda modificar la estego-imagen para afectar el mensaje oculto (ataque activo) (Lerch Hostalot & Megías, 2014). Por lo que es importante buscar nuevas técnicas que dificulten la tarea de los estego-analizadores.

Muchos de los métodos esteganográficos actuales intentan mejorar el algoritmo de incrustación para lograr mayor robustez, sin embargo muy pocos se enfocan en las características del objeto contenedor donde se ocultará la información. Con el objetivo de reforzar la técnica de transmisión de imágenes propuesta en (Ya-Lin & Wen-Hsiang, 2014), en este trabajo se plantea un método esteganográfico combinado con un procedimiento de detección de las zonas más aptas para la incrustación de información. Estas zonas son seleccionadas por su naturaleza ruidosa que dificulta la extracción de características para el estego-análisis. La esteganografía se aplica únicamente a estas zonas, creando un mapa de la información oculta en la imagen que el atacante no conoce, y además se minimiza la creación de irregularidades estadísticas; de esta manera se obtiene una nueva técnica robusta estadísticamente ante ataques de estego-analizadores.

1.3 Alcance del Proyecto

Se han hecho estudios para encontrar lugares adecuados de la imagen portadora para ocultar información (Nusrati, A., & Karimi, 2015), donde se hace el menor número de cambios posible en los bits para tener la mínima modificación en el histograma de la imagen. Al usar estegoanálisis, se notó que existen zonas en las cuales es más difícil extraer características: texturas y bordes. Estas son las zonas ruidosas de la imagen, mismas que hacen bastante complicado extraer características adecuadas para entrenar a clasificadores encargados de realizar estegoanálisis y detectar imágenes con mensajes ocultos. En (Lerch Hostalot & Megías, 2014) se agrupan los píxeles vecinos en parejas y se establece un umbral que indica las zonas difíciles de modelar, de manera que solo se toman en consideración para ocultar información los píxeles cuya diferencia entre sí sea mayor o igual al umbral. Este método es bastante sencillo pero con baja inmunidad al ruido, lo que se puede mejorar usando operadores diferenciales que involucren más píxeles vecinos.

En el presente proyecto se propone un algoritmo que combina un método de transformación de imágenes utilizando las características del color y un nuevo método esteganográfico que utiliza operadores diferenciales y filtros para detectar las zonas más adecuadas de la imagen para incrustar información. Se empleará la técnica propuesta en (Ya-Lin & Wen-Hsiang, 2014), que consiste en transformar la imagen

secreta en una imagen mosaico de igual tamaño, la cual luce como una imagen portadora seleccionada arbitrariamente. La transformación se realiza mediante la división de la imagen secreta en bloques o fragmentos y la modificación de las características de sus colores en aquellos correspondientes a los fragmentos de la imagen portadora. Se aplica técnicas de transformación de color que permiten recuperar la imagen secreta con un bajo porcentaje de pérdidas.

La información requerida suficiente para recuperar la imagen secreta se oculta mediante técnicas esteganográficas en el mosaico creado, con la diferencia de que, para superar la desventaja que se presenta al trabajar con el *LSB* y ser menos vulnerable frente a estego-analizadores, los bits de los parámetros que contienen las claves para restaurar la imagen secreta como medias y desviaciones estándar de grupos de pixeles, y que son almacenados clandestinamente en el mosaico, se ocultarán esta vez en zonas ruidosas, o en otras palabras, en líneas y texturas, con lo que se mejora la indetectabilidad estadística (Lerch Hostalot & Megías, 2014). La detección de estas zonas y de los bordes de la imagen, se realizarán mediante operadores diferenciales y filtros que localizan transiciones significativas de valores de pixeles, además, se trabajará usando figuras a color, en contraste con la técnica de (Lerch Hostalot & Megías, 2014) que trabaja únicamente en escala de grises.

Para la evaluación de la calidad de la imagen obtenida se aplica el cómputo del valor RMSE *(Root mean square error)* y el valor de SSIM *(Structural similarity index)* entre la imagen portadora y el mosaico creado el cual se espera que sea pequeño para verificar la similitud de las imágenes.

Con la integración de técnicas de transformación de color y de esteganografía en zonas ruidosas de la imagen se espera un nuevo método de transmisión de imágenes que sean difícilmente detectadas por estego-analizadores. La efectividad del uso de esteganografía en zonas con un mayor nivel de ruido en la imagen se comprobará mediante el gráfico de los histogramas y usando el sistema de estegoanálisis Steg-Expose (Benedikt, 2014) para comparar la indetectabilidad estadística respecto al uso del sistema *LSB matching* en toda la imagen para la misma cantidad de información a ocultar.

1.4 Objetivos

1.4.1 General

Desarrollar un algoritmo para la transmisión de imágenes a través de procesos esteganográficos en zonas ruidosas de la imagen basados en transformaciones reversibles del color.

1.4.2 Específicos

- Documentar y realizar investigación bibliográfica acerca de técnicas esteganográficas en zonas de la imagen con un alto nivel de ruido.
- Aplicar la esteganografía en zonas ruidosas como un método más seguro en la incrustación de la información para recuperar la imagen secreta.
- Verificar la eficiencia del método mediante pruebas con estego-analizadores.

1.5 Contenido

El presente proyecto de grado está estructurado en cinco capítulos. La descripción del contenido de cada capítulo se presenta a continuación:

El Capítulo 1 presenta el escrito previo a la realización del trabajo como la descripción de antecedentes, justificación, alcance del proyecto y los objetivos principales, así como el resumen de los contenidos de cada capítulo para que el lector tenga una idea clara del tema del trabajo.

El Capítulo 2 detalla el marco teórico en el que se basa la investigación, describiendo y estudiando de forma básica el concepto y técnicas principales referentes a la esteganografía. Además se abordará de manera resumida y puntual temas importantes para este trabajo como el estegoanálisis, técnicas y herramientas, los parámetros fundamentales para el análisis de características de una imagen como desviación estándar, media e histogramas, métodos de detección de bordes y texturas basados en procesamiento de imágenes y por último métricas para evaluar la calidad de una imagen. Este capítulo permitirá comprender de mejor manera las técnicas usadas para el desarrollo del proyecto.

El Capítulo 3 es la base del algoritmo propuesto, ya que en este se describe detalladamente el método en el que se basa el tema propuesto y el procedimiento que se utilizó para la detección de bordes y texturas en la imagen. En este capítulo se encuentra el desarrollo del programa.

El Capítulo 4 muestra los resultados obtenidos después de implementar el algoritmo, la selección del tamaño óptimo de bloques y el análisis de mediciones con el cálculo de desviaciones estándar en cada canal RGB para el proceso de ordenamiento de bloques. Además describe las pruebas realizadas con imágenes diferentes para evaluar la calidad de las imágenes obtenidas (estego-imagen y la imagen secreta recuperada), y se verifica la efectividad del método propuesto ante ataques de estegoanalizadores mediante el análisis de histogramas y pruebas con la herramienta StegExpose.

El Capítulo 5 analiza las conclusiones, recomendaciones y describe las propuestas de trabajos futuros en la línea de la esteganografía.

CAPÍTULO 2

2. MARCO TEÓRICO

2.1 Esteganografía

La seguridad es una herramienta fundamental en cualquier sistema, es por ello que existen técnicas para brindar mayor robustez ante ataques externos de cualquier clase. Existen tres conceptos que están estrechamente relacionados dentro del dominio de seguridad (ver figura 1): esteganografía, criptografía y watermarking; sin embargo utilizan métodos diferentes para transmitir información de forma segura.

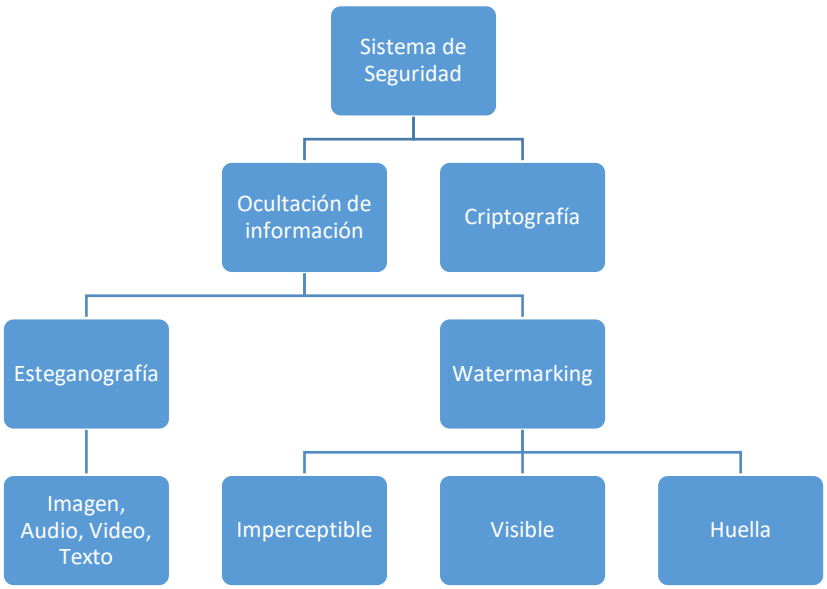


Figura 1 Clasificación de un sistema de seguridad

La palabra esteganografía viene de un vocablo griego que significa “escritura oculta”, lo que nos lleva a entenderla como la ciencia de ocultar mensajes en un objeto portador. El objeto portador puede ser cualquier contenido multimedia como imágenes, videos, archivos, etc, pero este caso se enfocará en su mayor parte a las imágenes, por su amplio uso en Internet para diversas aplicaciones, tales como álbumes personales en línea, archivos confidenciales de una empresa, sistemas de almacenamiento de documentos, sistemas de imágenes médicas, y bases de datos de imágenes militares (Ya-Lin & Wen-Hsiang, 2014). El objetivo principal de la esteganografía es una comunicación confidencial y segura, donde la información

oculta solo pueda ser visualizada mediante el uso de una clave secreta que solo conoce el receptor al que va dirigido el mensaje.

Los principales elementos de la esteganografía en imágenes se pueden observar en la figura 2:

- Imagen portadora: imagen donde se incrusta el mensaje oculto
- Mensaje oculto: información secreta que se desea ocultar
- Estego-imagen: combinación de la imagen portadora y el mensaje oculto, es la imagen que contiene información embebida
- Algoritmo de ocultación: método esteganográfico que se usa para incrustar información en el objeto portador

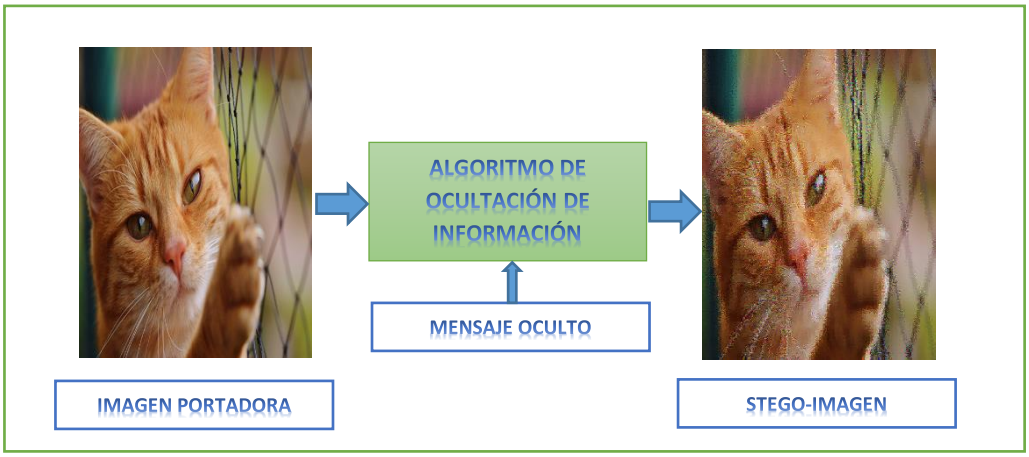


Figura 2 Principio de la Esteganografía

Por otro lado, la criptografía es una técnica que utiliza las características de la una imagen secreta para crear imágenes ruidosas que no pueden ser visualizada a menos que se aplique métodos para descifrar la imagen oculta mediante el uso de una clave (ver figura 3) como por ejemplo el método utilizado en (Prihandoko, 2013). En esta técnica se tiene la desventaja de no pasar inadvertido que existe un mensaje oculto, más bien la parte desconocida para cualquier intruso es el método de ocultación del mensaje, es decir el algoritmo de cifrado. Por lo tanto, en la criptografía se pretende mejorar los métodos de encriptación para hacerlos más indescifrables por atacantes.

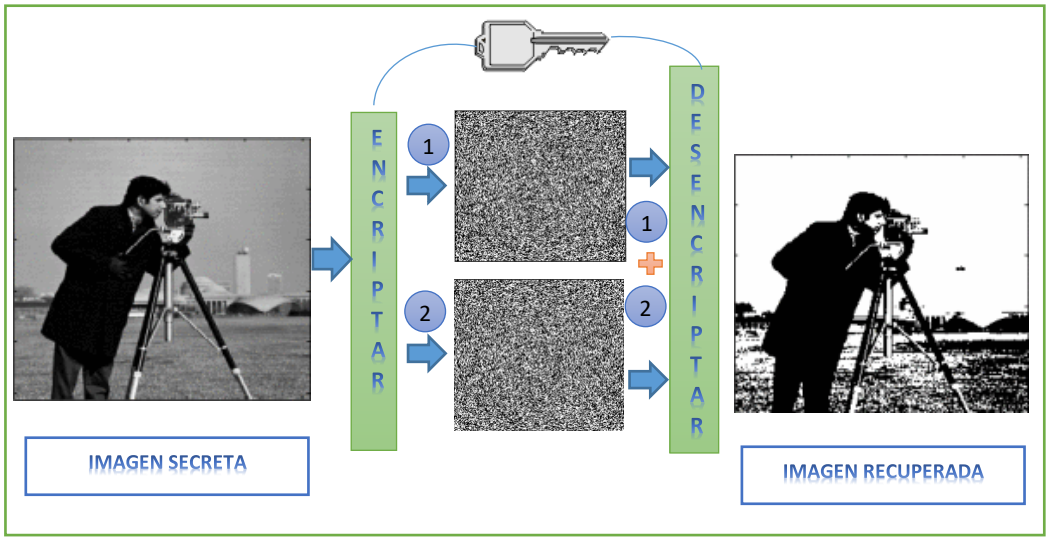


Figura 3 Principio de la Criptografía

Fuente: (Prihandoko, 2013)

A pesar de que las técnicas criptográficas son muy utilizadas y altamente confiables, el hecho de que se conozca que existe un mensaje oculto es un punto en contra en la seguridad de la información, por lo que las técnicas esteganográficas y watermarking son una alternativa a este problema. La diferencia entre estos dos métodos es que en el caso de la esteganografía la información secreta no debe ser visible al espectador, mientras que en el proceso de *watermarking* puede o no ser visible dependiendo de la aplicación (Satwinder & Varinder, 2015).

El *watermarking* es un método para ocultar información secreta e información adicional en la imagen portadora que más tarde puede ser extraída con varios propósitos como la autenticación, identificación del propietario, protección del contenido, etc. (Preeti & Rajeev, 2013)

2.1.1 Clasificación de la esteganografía

En la figura 4, se puede ver la esteganografía dividida en tres categorías según (Sumathi, Santaman, & Umamaheswari, 2013):

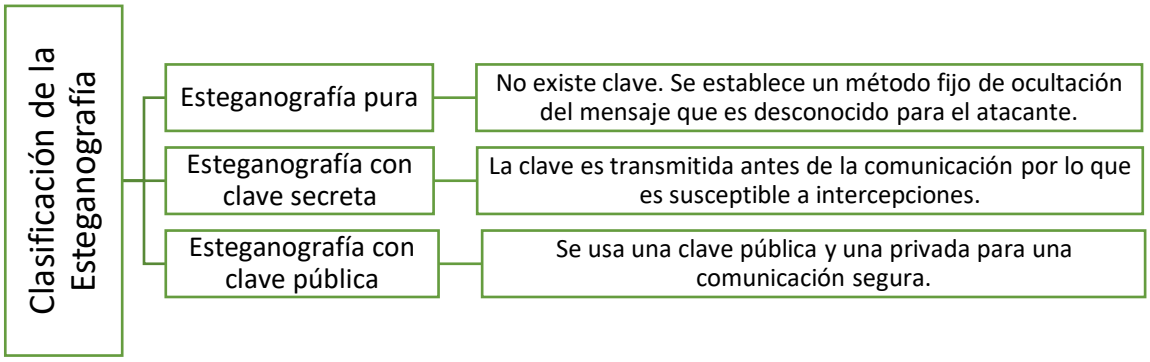


Figura 4 Clasificación de la Esteganografía

2.2 Tipos de técnicas esteganográficas en imágenes

Según (Satwinder & Varinder, 2015), se dividen en cuatro grupos: esteganografía en el dominio espacial, esteganografía en el dominio de la transformada, esteganografía spread spectrum, esteganografía adaptativa.

2.2.1 Esteganografía en el dominio espacial

Son técnicas muy usadas por su baja complejidad donde la información secreta es directamente embebida en los valores de los pixeles de la imagen, es decir, se modifica ligeramente los pixeles para ocultar un mensaje. Existen muchas técnicas de este tipo, a continuación se describen las más importantes y que aportaron al desarrollo de la técnica estenográfica propuesta:

A. Algoritmo de sustitución del Bit Menos Significativo (LSB)

Es el algoritmo más sencillo y consiste en sustituir el bit menos significativo de cada pixel por el bit del mensaje oculto. Para lo cual es necesario transformar el mensaje a ocultar en un flujo de bits y la imagen portadora a matrices de bits, 8 bits para cada pixel de imágenes en escala de grises y 24 bits para imágenes RGB. Esta técnica ofrece gran cantidad de incrustación sin embargo es susceptible a pérdidas de información por modificaciones en la imagen portadora por compresión o recorte. La alteración en la imagen portadora es mínima por lo que no es perceptible visualmente, sin embargo se crea una irregularidad estadística en la imagen por lo que puede ser susceptible a detecciones con estego-analizadores.

Para evitar la detección se ha mejorado el método *LSB* a un algoritmo llamado *LSB matching* que básicamente es una técnica muy similar a su predecesora, pero en lugar de sustituir el bit menos significativo se lo modifica cuando no coincide con el bit a ocultar. Con este método se obtiene una imagen con bits embebidos que visualmente son imperceptibles y menos detectables estadísticamente. Existen varios algoritmos basados en la técnica *LSB matching* entre los cuales se encuentran: *LSB Matching Revisited* (Mielikainen, 2006), *Lossless data hiding scheme based on lsb matching* (Quan & Zhang), *Very fast watermarking by reversible contrast mapping* (Coltuc & Chassery, 2007), entre otros.

B. Diferencia de valores de pixeles

Esta técnica es llamada *Pixel Value Differencing* (PVD) (El-Alfy & Al-Sadi, 2012) y consiste en dividir a la imagen portadora en bloques de dos pixeles consecutivos y evaluar la diferencia entre ellos que determina el número de bits que se pueden ocultar en el par de pixeles. Se basa en el hecho de que la vista de los seres humanos es más sensible a variaciones de la imagen en zonas lisas mientras que es más difícil que se detecte visualmente una alteración de la imagen en los bordes. Mientras mayor sea la diferencia entre pixeles, se tiene mayor capacidad de incrustación y esto se da normalmente en bloques ubicados en bordes, y cuando la diferencia se acerca a 0 quiere decir que se trata de un área lisa y por lo tanto tiene menor capacidad para ocultar bits. La stego-imagen obtenida es de mayor calidad y tiene mejores resultados visiblemente y estadísticamente es menos susceptible a ataques comparado con la técnica *LSB*, sin embargo su principal desventaja es su alto nivel de procesamiento. Una variación de esta técnica es la presentada en *Reversible Watermarking by Difference Expansion* (Tian, 2002) y en (Lerch Hostalot & Megías, 2014) donde se usa el principio de la diferencia de pixeles para clasificarlos como aptos para embeber información o no mediante la designación de un umbral y aquellos valores de diferencia mayores al umbral indican que los pixeles pertenecen a zonas ruidosas por lo que son útiles para ocultar bits.

C. Modificación de niveles de grises

Este método no esconde información, en su lugar realiza un mapa de ésta en la imagen portadora modificando el valor de niveles de grises en los pixeles. Para ello se

usan los números pares e impares y se selecciona mediante funciones matemáticas ciertos bits dentro de la imagen, los cuales se comparan posteriormente con el flujo de bits mapeado en la imagen. Esta técnica tiene como ventajas una gran capacidad de incrustación y menor complejidad de procesamiento.

D. Esteganografía basada en la textura

Se basa en la sustitución de bloques de pixeles, el primer paso consiste en dividir la imagen portadora y la imagen secreta en pequeños bloques del mismo tamaño. En cada bloque de la imagen secreta se determina un patrón de textura y se compara con los bloques de la imagen portadora con el objetivo de encontrar el más similar para reemplazarlo, creando una nueva imagen con la menor distorsión posible (Tiwari, Yadav, & Mittal, 2014). Esta técnica está enfocada en la ocultación de una imagen dentro de otra imagen.

E. Esteganografía basada en bordes

Este algoritmo se basa en la técnica *LSB*, sin embargo la información no se oculta en todos los pixeles de la imagen sino únicamente en aquellos que pertenecen a bordes. Para ello es necesario la detección previa de bordes, debido a que esta técnica tiene baja capacidad de incrustación actualmente se utiliza los 3 bits menos significativos de cada pixel.

F. Ventajas y desventajas de técnicas del dominio espacial

Se puede observar claramente las ventajas y desventajas de las técnicas en el dominio espacial (Tiwari, Yadav, & Mittal, 2014). Entre las principales ventajas se encuentran:

- Gran capacidad de incrustación
- Menor degradación de la calidad de la imagen original
- Bajo procesamiento matemático

Entre las desventajas se encuentran:

- Susceptibilidad a pérdidas de información por manipulación de la imagen
- Mas vulnerabilidad a ataques sencillos

2.2.2 Esteganografía en el dominio de la transformada

Una imagen puede ser transformada en el dominio de la frecuencia, dominio wavelet, etc. En este tipo de técnicas la información secreta no se oculta directamente en el valor de los pixeles como en la esteganografía en el dominio del espacio, se oculta en la transformada de la imagen portadora y posteriormente se puede volver a transformar al dominio del espacio. Un componente de la transformada puede abarcar varios pixeles de la imagen e incluso toda la imagen por lo que tiene la ventaja de mayor capacidad de incrustación y es más robusta ante pérdidas de información por recortes o compresiones de la imagen original. Pero es un tipo de técnica con una alta complejidad y mayor procesamiento matemático.

Las transformadas más usadas por ser las más sencillas en estos tipos de técnicas son las transformadas en frecuencia. Una imagen tiene componentes en frecuencias bajas y en frecuencias altas lo que se puede hacer la analogía con las zonas de la imagen; las frecuencias bajas se relacionan con las zonas de la imagen lisas donde cualquier variación de la imagen es más notable visualmente, y las frecuencias altas se asocian a los bordes o texturas donde es más difícil la percepción visual de alguna alteración en los valores de pixeles.

2.2.3 Esteganografía Spread Spectrum

Esta técnica consiste en modular una señal de banda estrecha con una señal de banda ancha mediante un generador de ruido pseudo-aleatorio (técnicas spread spectrum), de esta manera decrece la intensidad logrando una señal muy parecida al ruido y así no ser detectada. La señal modulada resultante se añade a la imagen portadora creando una estego-imagen robusta ante la detección y extracción (Tiwari, Yadav, & Mittal, 2014). Las ventajas de esta técnica es la calidad de la imagen, la robustez ante ataques y la gran capacidad de incrustación, sin embargo cualquier procesamiento digital de la imagen puede destruir el mensaje, por ejemplo filtros reductores de ruido. Para extraer el mensaje, es necesario que el generador de ruido esté sincronizado a la misma frecuencia que en el receptor.

2.2.4 Esteganografía Adaptativa

Esta técnica usa las características estadísticas de la imagen portadora, por lo cual también se la conoce como esteganografía basada en la estadística. Existen dos métodos: seleccionar los pixeles adaptativos al azar dependiendo de la imagen portadora y seleccionar los pixeles con mayor desviación estándar en la imagen (Tiwari, Yadav, & Mittal, 2014). La ventaja principal de este tipo de esteganografía es que no cambia las propiedades de la imagen con una distorsión visual insignificante para el sistema visual humano.

2.3 Estego-análisis

El estego-análsis es un conjunto de procedimientos para detectar información oculta dentro de un objeto multimedia (audio, imágenes, video). En este caso se abordará únicamente el estegoanálsis en imágenes, donde su principal objetivo es diferenciar entre una estego-imagen y una imagen portadora; conociendo poco o nada sobre el algoritmo de inserción (Manveer & Gagandeep, 2014). Luego de detectar información oculta dentro de la imagen, se busca extraer el mensaje oculto. En el cripto-análisis se conoce que existe un mensaje oculto encriptado; sin embargo, el objetivo de la esteganografía es que la existencia de un mensaje pase desapercibida ante los atacantes; por lo que el estego-análisis inicia con una serie de datos desconocidos que son reducidos mediante métodos estadísticos (Badr, Salam, Selim, & Khalil, 2014).

De la misma manera que las técnicas esteganográficas han ido evolucionando, el estegoanálisis ha ido a la par buscando nuevos algoritmos que recojan evidencia para demostrar que una imagen contiene información oculta. En (Meghanathan & Nayak, 2010) se presenta el estego-análisis dividido en dos grupos: específico y genérico. El estegoanálisis específico es diseñado para un tipo particular de algoritmo esteganográfico y tiene alta tasa de éxito si el mensaje fue embebido con el método para el cual fue creado. Por otro lado, el genérico o universal es independiente del método esteganográfico usado y es muy útil para detectar información oculta con métodos no convencionales o nuevos.

2.3.1. Estegoanálisis Específico

Adicionalmente los métodos estegoanalíticos específicos están divididos en: estegoanálisis basado en la firma y estegoanálisis estadístico. Esta clasificación se basa en que herramienta se usa para el análisis: la firma de la técnica esteganográfica o el uso de las estadísticas de la imagen (Badr, Salam, Selim, & Khalil, 2014).

A. Estegoanálisis basado en la firma

Las técnicas esteganográficas ocultan información y manipulan las imágenes de forma que se conserven imperceptibles al ojo humano. Sin embargo, al usar herramientas electrónicas para insertar información se modifican las características de la imagen portadora introduciendo formas de degradación o patrones inusuales. Estas modificaciones en forma de patrones pueden hacer el papel de firmas que advierten la presencia de un mensaje oculto (Chhikara & Singh, 2013). Por lo que, las técnicas de estegoanálisis basado en la firma, buscan estos patrones repetitivos en las imágenes sospechosas desenmascarando la herramienta esteganográfica que se utilizó. Esta técnica proporciona buenos resultados cuando el mensaje es incrustado de forma secuencial, pero es difícil de automatizar (Manveer & Gagandeep, 2014).

B. Estegoanálisis estadístico

La esteganografía introduce anomalías en las estadísticas de la imagen al incrustar información. Por lo tanto, el estegoanálisis estadístico se basa en analizar las estadísticas y descubrir estas alteraciones para detectar información oculta. Este tipo de estegoanálisis es más eficaz que el basado en la firma, ya que las técnicas matemáticas son más sensibles que la percepción visual (Chhikara & Singh, 2013). Se clasifica basado en las técnicas de incrustación en dos tipos: dominio espacial y dominio de la transformada. En este trabajo es de nuestro interés la clasificación en el dominio espacial debido a que la herramienta utilizada para el estego-análsis usa técnicas pertenecientes a esta clase.

a. Estegoanálisis estadístico en el dominio espacial

- Ataque *Chi-square*: (Westfeld & Pfitzmann, 1999) desarrollaron por primera vez la idea de este ataque enfocado a técnicas esteganográficas *LSB*, el cual se basa en la frecuencia de distribución de los pares de valores (POVs). Un POV

está compuesto por dos valores de pixeles que difieren únicamente en el bit menos significativo, es decir, es el par de valores cuyos *LSBs* son intercambiados durante el proceso de incrustación de la información. Antes de este proceso la distribución de estos pares es desigual, por el contrario después de embeber bits la ocurrencia de los dos pixeles del POV llegarán a ser iguales si el mensaje a ocultar posee una distribución uniforme de bits. El ataque consiste en comparar la distribución de frecuencias teórica esperada con la distribución de muestras obtenidas de la imagen portadora sospechosa mediante una prueba *chi-square*.

- *Sample Pairs*: La técnica está basada en una máquina de estados finita en la que sus estados son multisets seleccionados de pares de muestras, llamados *trace multisets*. Algunos de estos multisets son iguales en sus cardinalidades esperadas, si los pares de muestras son dibujados desde una señal continua digitalizada. Además, las estadísticas de los pares de muestras son altamente sensitivos a los *LSB* embebidos, incluso cuando el mensaje embebido sea muy corto. (Dumitrescu, Wu, & Wang, 2003)
- *RS analysis (Regular and Singular Group)*: Esta técnica utiliza estadísticas dualmente sensibles derivadas de correlaciones espaciales en imágenes. La imagen es dividida en grupos disjuntos de silueta fija. Dentro de cada grupo, el ruido es medido por el valor absoluto medio de las diferencias entre pixeles adyacentes. Cada grupo es clasificado como "regular" o "irregular" dependiendo de si el ruido del pixel dentro del grupo se incrementa o después de mover los *LSBs* de un set fijo de pixeles dentro de cada grupo usando una "máscara". La clasificación es repetida para un cambio de tipo dual. (Chhikara & Singh, 2013)
- *Primary sets*: Este método es utilizado para información embebida de forma aleatoria o secuencial y puede detectar con precisión la longitud del mensaje oculto, para lo cual la imagen es dividida en subconjuntos de pixeles cuyos tamaños pueden cambiar cuando se incrusta información. Se basa en la cuantificación de estas variaciones de tamaños mediante una identidad estadística de ciertos subconjuntos. (Dumitrescu, Xiaolin, & Nasir, On steganalysis of random LSB embedding in continuos-tone images, 2002)

b. Estegoanálisis estadístico en el dominio de la transformada

En este tipo de análisis existen varias técnicas que se basan en el estudio de los coeficientes DCT (*Discrete Cosine Transform*) de las imágenes. Se observa que los coeficientes DCT cuantificados se distribuyen simétricamente alrededor de cero para imágenes portadoras, mientras que cuando se incrusta un mensaje estas distribuciones se ven alteradas. Para detectar una estego-imagen en este caso se usa estadísticas *chi-square* (Zhang & Ping, 2003). También existen otros ataques basados en el análisis de las pérdidas de simetría del histograma de coeficientes después de la incrustación de información.

2.3.2. Estegoanálisis Universal

- Las técnicas de estego-análisis específico tienen buenos resultados cuando se usa el método de incrustación para el cual fueron diseñadas. Por el contrario, el estego-análisis universal o general supera las deficiencias de las técnicas específicas al requerir poco o nada de información a priori sobre los métodos esteganográficos usados. Este tipo de estego-análisis se considera una estrategia de aprendizaje basada en la clasificación de patrones para catalogar la imagen sospechosa como una imagen portadora o como una estego-imagen. En general, la clasificación se compone de dos partes: extracción de características y clasificación de patrones. Una buena característica para el análisis esteganográfico debe contener información clave que represente las propiedades de la imagen sobre los cambios efectuados por la información que se ocultó más que por el contenido de la imagen (Chhikara & Singh, 2013).

2.4 Herramienta de estegoanálisis StegExpose

Las herramientas de estegoanálisis juegan un papel muy importante en la detección de información oculta en una imagen, lo que quiere decir ahorro de tiempo y automatización de procesos. Para este trabajo en particular se usó la herramienta StegExpose[1], la cual es una herramienta de estegoanálisis especializada en la detección del uso de técnicas esteganográficas *LSB* en imágenes.

[1] La herramienta StegExpose se puede encontrar en https://github.com/b3dk7/StegExpose

2.4.1 Métodos usados

Según (Benedikt, 2014), para la elaboración de esta herramienta se realizó el estudio de varios métodos de estegoanálisis: *Primary sets* (Dumitrescu et al., 2002), *Chi square* (Westfeld & Pfitzmann, 1999), *Simple pairs* (Dumitrescu et al., 2003), *Rs analysis* (Fridrich, Miroslav, & Rui, 2001) y un método basado en el cálculo *chi-square* pero en el dominio de la transformada (Zhang & Ping, 2003). Todos estos métodos fueron explicados en la sección anterior.

La característica principal de este estegoanalizador es el uso de técnicas de fusión, es decir la combinación de varios métodos con el fin de lograr una herramienta más fuerte, tomando en cuenta todos los enfoques de la fusión como el proceso de clasificación y reglas de la fusión. La clasificación está dividida en dos etapas: pre-clasificación y post-clasificación. En la pre-clasificación la imagen es catalogada como limpia o estego por cada detector de forma independiente. En la segunda etapa las salidas de los detectores son combinadas antes de tomar una decisión. Y por último se elige una regla de la fusión que es una propiedad estadística tomada por un grupo de detectores para obtener un resultado final. (Benedikt, 2014)

Dentro de las técnicas de fusión de esta herramienta se encuentra una fusión estándar y una fusión rápida. La fusión rápida se creó con el objetivo de mejorar los tiempos de detección, especialmente cuando se analiza una gran cantidad de imágenes. En lugar de descartar por completo los detectores más lentos, este algoritmo intenta acelerar la clasificación de las imágenes limpias e invertir mayor tiempo en las imágenes sospechosas. Consta de cuatro etapas debido a los cuatro métodos usados, en cada etapa se añade un detector y se obtiene la media de los resultados, si la media es mayor que el umbral establecido pasa a la siguiente etapa de detección, de lo contrario el archivo es considerado como limpio. Si la imagen pasa por las cuatro etapas y el resultado sigue por encima del umbral, la imagen es clasificada como stego. (Benedikt, 2014)

En (Benedikt, 2014) se evaluó el rendimiento de la herramienta StegExpose mediante el cálculo de la precisión de los métodos usados, para lo cual se utilizó la gráfica de la curva ROC (*receiver operating characteristic*) con 23 coordenadas, que consiste en la tasa de los falsos positivos vs la tasa de falsos negativos y luego se

calculó el área bajo la curva llamada AUC *(Area under curve)* para todos los métodos y las técnicas de fusión (ver figura 5). Los resultados del AUC se observan en la figura 6, donde se puede notar que el método más robusto es la fusión estándar.

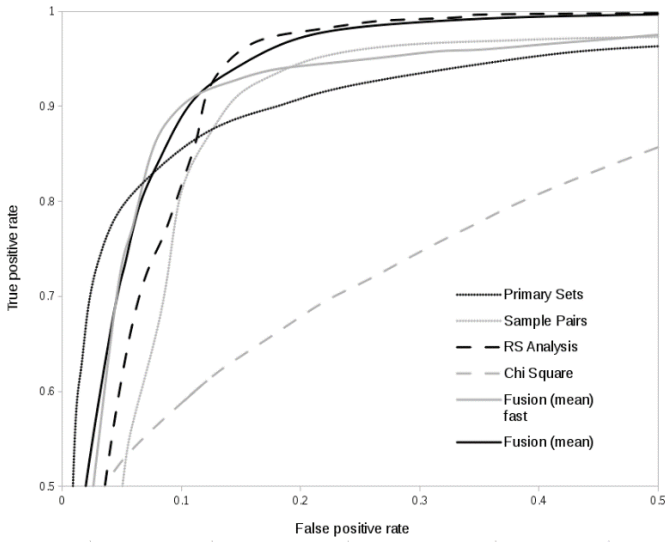


Figura 5 Curvas ROC para métodos de detección de StegExpose

Fuente: (Benedikt, 2014)

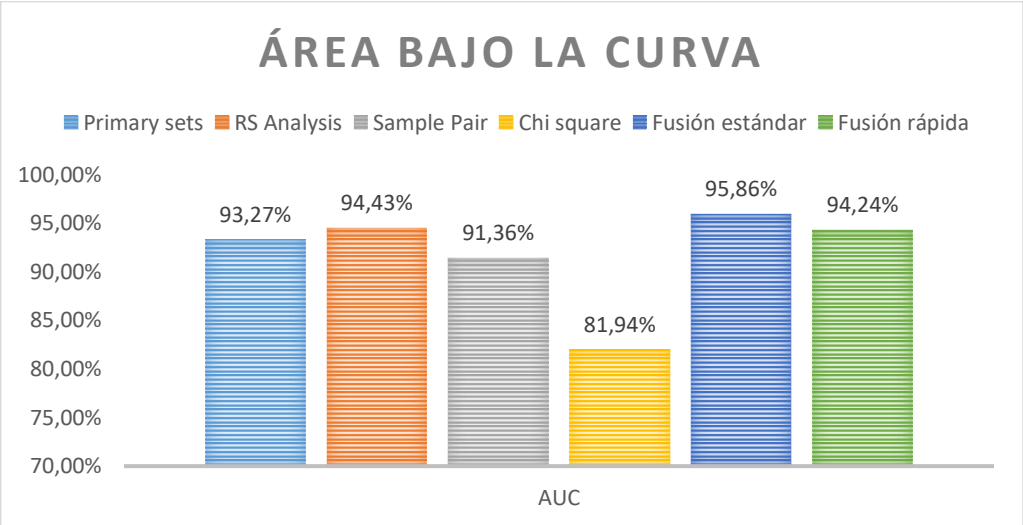


Figura 6 AUC para métodos de detección

Fuente: (Benedikt, 2014)

2.4.2 Funcionamiento

En (Benedikt, 2014) se indica el funcionamiento de StegExpose, herramienta diseñada en una interfaz de líneas de comando, por lo que es necesario una versión de

Java 1.6 o superior. Para su ejecución es necesario el siguiente comando, donde únicamente el primer argumento es obligatorio:

java -jar StegExpose.jar [directory] [speed] [threshold] [csv file]

- [directory]: directorio que contiene las imágenes que serán analizadas
- [speed]: (opcional) existen dos opciones: default o fast, en el modo por defecto el algoritmo ejecutará todos los detectores mientras que en el modo rápido evitará los detectores más complejos en el caso de que los detectores más simples logren determinar si la imagen está limpia.
- [threshold]: (opcional) el valor por defecto es 0.2 y determina el nivel en el cual los archivos son considerados como sospechosos de ocultar información.
- [csv file]: (opcional) nombre del archivo generado por el programa. Si se deja en blanco, el análisis será mostrado únicamente por consola.

2.5 Imágenes digitales

Una imagen puede definirse como una función f(x,y) con un conjunto finito de elementos (puntos) dados por coordenadas espaciales (x,y), cada elemento tiene una amplitud particular conocida como intensidad y una ubicación definida, donde cada punto es denominado pixel (Moreira, Valencia, & Chávez, 2009). Una imagen digital está compuesta por un conjunto de pixeles homogéneos resultantes de la discretización de una imagen natural (González Aguilera, 2011). Si está compuesta por una sola matriz de pixeles (8 bits por pixel), cuyas intensidades son llamadas niveles de grises es una imagen monocromática. Mientras una imagen a color está compuesta por tres matrices o canales *R, G, B (Red, Green, Blue)*, es decir, cada pixel está compuesto por 24 bits.

2.5.1 Parámetros de la imagen

En una imagen se pude obtener varios parámetros estadísticos útiles que describen la distribución de los valores de pixeles; para su análisis, modificación y manipulación, como desviación estándar, media, histograma. A continuación se explicará cada término:

- Media: es el nivel de gris medio en el caso de imágenes en escala de grises y nivel medio de color en imágenes RGB, indica la luminosidad o brillo de una imagen (González Aguilera, 2011).
- Desviación estándar: es una medida del contraste o variación de la información dentro de la imagen, si una imagen tiene poco contraste o información reducida el valor de la desviación será pequeño; mientras que si una imagen tiene mucha información o gran cantidad de contraste el valor será alto (González Aguilera, 2011).
- Histograma: es una representación gráfica de una imagen, en la cual se puede observar la distribución de valores de pixeles o niveles de color y la variación o contraste en la imagen. Según (González Aguilera, 2011), "se define el histograma de una imagen como la curva que a lo largo de uno de sus ejes representa cada uno de los posibles niveles de gris, y en el otro la frecuencia relativa de aparición del mismo en la imagen." (p. 7).

  Para obtener el histograma simplemente es necesario contar el número de cada valor de pixel presente en la imagen, es decir, obtener la frecuencia de cada valor de pixel. Cuando una imagen presenta tonos oscuros el histograma tiende a la izquierda, y para colores claros tiende a la derecha. En la figura 7 se puede ver un histogramas de imágenes a color compuesto de tres curvas, una para cada color, cada canal tiene un rango de niveles de intensidad de 0 a 255 (8 bits).

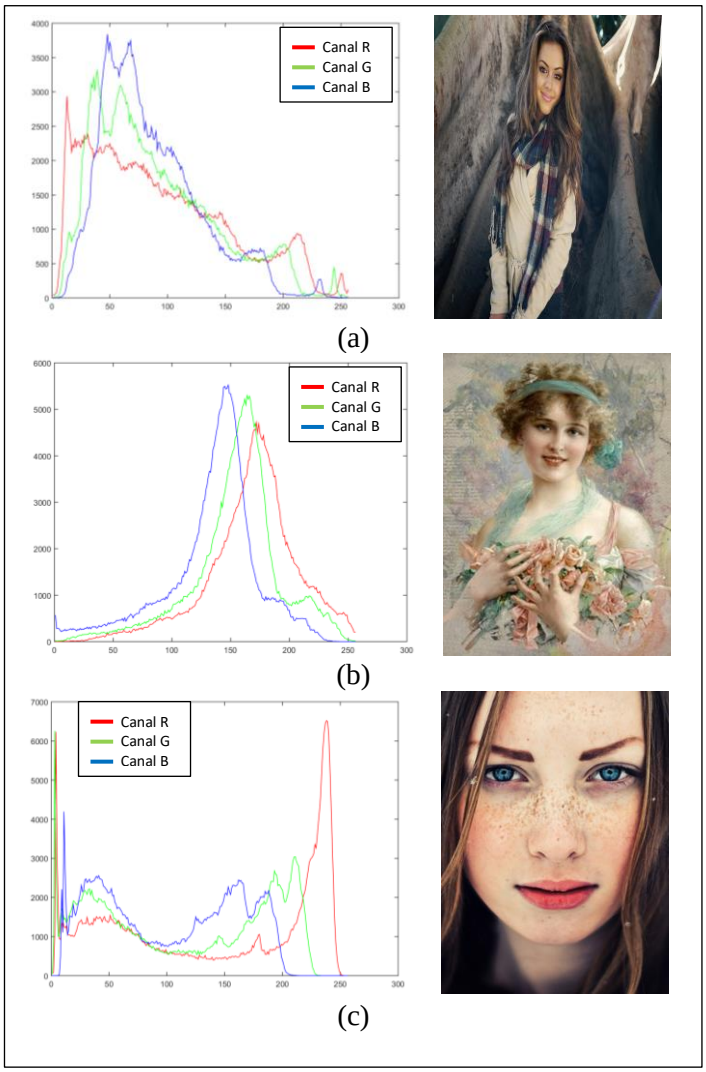


Figura 7 (a) Imagen e histograma de tonos oscuros (b) Imagen e histograma con tonos uniformes (c) Imagen e histograma de tonos claros

2.5.2 Detección de bordes y texturas en imágenes

En este trabajo se propone un método que incluye la búsqueda de zonas adecuadas para incrustar información, estas zonas se caracterizan porque son difíciles de modelar para entrenadores de clasificadores usados en estegoanálisis y están conformadas por bordes y texturas. Por lo cual es necesario la aplicación de filtros y operadores diferenciales usados en procesamiento de imágenes para localizar estas zonas en una imagen.

A. Detección de bordes

Un borde es considerado como una discontinuidad en la función de intensidad en una imagen, es decir una fuerte variación o cambio de la intensidad ocasionada por las fronteras entre objetos. Se utilizan las derivadas de la función imagen para detectar estas discontinuidades, por ejemplo la primera derivada es cero en todos los puntos de la imagen a excepción de los máximos de intensidad o transiciones de nivel, y el signo de la segunda derivada informa si los puntos están en el borde oscuro (signo positivo) o en el claro (signo negativo) (ver figura 8) (Departamento de Electrónica y Automática).

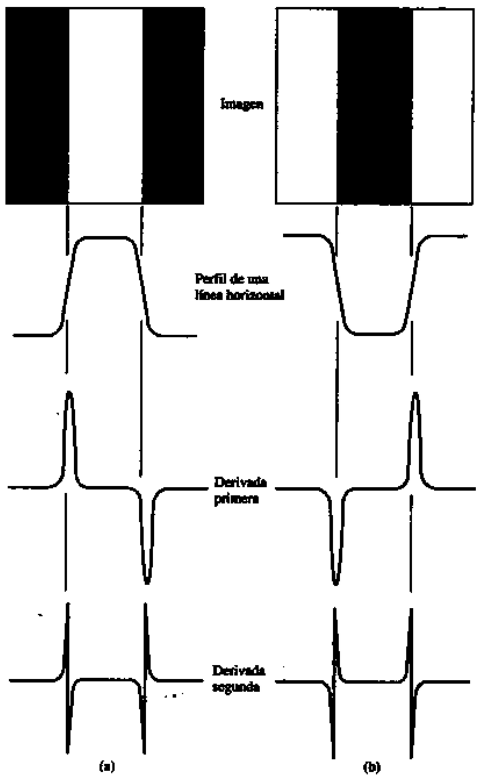


Figura 8 Primera derivada y segunda derivada de una imagen (a) Transición de borde oscuro a claro (b) Transición de borde claro a oscuro

Fuente: (Departamento de Electrónica y Automática)

Según lo mencionado anteriormente, los métodos para detectar bordes se basan en la primera y segunda derivada, la derivada de primer orden para una imagen es el operador gradiente por lo que existen algoritmos basados en este operador como: Sobel, Prewitt, Roberts. Estos métodos son muy usados en el ámbito de procesamiento

de imágenes y muestran los bordes como máximos en la derivada de la función intensidad de la imagen (González Aguilera, 2011).

En la tabla 1 se emite un resumen de los operadores gradiente para la detección de bordes, junto con las ventajas y desventajas del uso de cada uno.

Tabla 1

Resumen de operadores gradiente para detección de bordes

Operadores Gradiente	Ventajas	Desventajas
Roberts $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$ $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	- Buena respuesta en bordes horizontales y verticales. - Buena localización. - Simpleza y rapidez de cálculo.	- Mala respuesta en bordes diagonales. - Sensible al ruido. - Empleo de máscaras pequeñas. - No da información acerca de la orientación del borde. - Anchura del borde de varios píxeles.
Sobel $\frac{1}{4}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$ $\frac{1}{4}\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$	- Buena respuesta en bordes horizontales y verticales. - Diversidad de tamaños en las máscaras. - Poco sensible al ruido.	- Mala respuesta en bordes diagonales. - Lentitud de cálculo. - No da información acerca de la orientación del borde. - Anchura del borde de varios píxeles.
Prewitt $\frac{1}{3}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$ $\frac{1}{3}\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	- Buena respuesta en bordes horizontales y verticales. - Poco sensible al ruido. - Proporciona la magnitud y dirección del borde.	- Mala respuesta en bordes diagonales. - Lentitud de cálculo. - Anchura del borde de varios píxeles.

Fuente: (González Aguilera, 2011)

Se evidencia que existen varias desventajas en los operadores gradiente, principalmente porque el gradiente es dependiente de la dirección del barrido por lo que los bordes que se encuentren en dirección paralela no son detectados fácilmente y además tiene dificultades en detección de bordes en esquinas. Por estas razones, existe también un método de detección de bordes llamado Canny. Este detector es considerado el más efectivo debido a que no depende de las direcciones, lo que significa que no tiene problemas para detectar bordes verticales, horizontales y

diagonales. Para lograr un alto rendimiento, este detector trabaja en varias fases descritas en (Moreira, Valencia, & Chávez, 2009):

1. La imagen original es suavizada mediante un filtro gaussiano para reducir el ruido en la imagen como texturas o detalles sin importancia, la máscara del filtro es definida por la desviación estándar dada por el usuario.
2. Se aplica el operador gradiente en dirección horizontal y vertical con el objetivo de realzar los bordes y obtener una imagen con los cambios de intensidad como referencia para trabajar en la próxima fase.
3. En la imagen anterior se realiza una depuración de puntos no máximos, se compara un punto máximo con sus vecinos y en base a un umbral se decide si debe ser suprimido o no, obteniendo bordes más delgados.
4. Por último, se conecta los pixeles llamados "candidatos débiles" próximos a los pixeles seleccionados mediante un criterio de asignación de umbrales.

B. Detección de texturas

Una textura puede ser definida como un arreglo espacial de patrones visuales de una imagen con características constantes, invariables o periódicas como densidad, uniformidad, aspereza, linealidad (Hajek, Dezortova, Materka, & Lerski, 2006). Para su estudio es necesario el análisis de un pixel y sus pixeles vecinos, considerando las variaciones en la intensidad de la imagen (contraste) y variaciones de brillo.

Existen algunos filtros o características de la imagen para la detección de texturas como la entropía, la desviación estándar, rango de valores:

- Entropía: es una medida de la cantidad de información en una imagen, es decir es la métrica de la distribución de probabilidad de los valores de pixel. La distribución de probabilidad se puede obtener mediante el conteo de las veces de aparición de cada valor de pixel y su posterior división para el número total de pixeles. Por lo que, si en un grupo de pixeles tiene una alta aleatoriedad se va a tener una alta entropía, mientras que si existe una distribución más uniforme de los valores de gris se tendrá una baja entropía (Riera & Tacón, 2011).
- Desviación estándar: como ya se mencionó anteriormente la desviación estándar, al igual que la entropía mide la variación de la información dentro de

la imagen, es decir la diversidad existente dentro de entorno compuesto de un pixel y sus vecindades.

- Rangos de valores: es una medida del rango de valores en un grupo de pixeles, es decir la diferencia entre el máximo valor y el mínimo valor obtenidos mediante funciones morfológicas que dependen del pixel analizado y su relación con las vecindades.

2.5.3 Mediciones de la calidad de la imagen

Una imagen digital está expuesta a modificaciones e inserción de distorsiones debido a la manipulación, procesamiento y transmisión. En este caso el hecho de que una imagen secreta sea cambiada para crear un mosaico similar a la imagen portadora y además se oculte información en ella, ocasionará importantes modificaciones en la calidad de la imagen haciéndola más propensa a ataques estegoanalíticos visuales. Por esta razón es necesario evaluar la calidad visual de la imagen mediante algunas métricas usadas en este ámbito. Las medidas de calidad de la imagen pueden ser clasificadas según la disponibilidad de la imagen original (libre de distorsiones), la cual es comparada con la imagen distorsionada. En la mayoría de las aplicaciones prácticas no existe una imagen para comparar por lo cual se obtiene una calidad sin referencia, también existen referencias parcialmente disponibles, es decir solo se dispone de ciertas características extraídas de la imagen que ayudan en la evaluación de la calidad. Por último se encuentran las medidas de calidad con imágenes de referencia totalmente disponibles, las cuales se describirán a continuación:

- RMSE *(Root Mean Square Error)*: es una medida objetiva de la calidad de la imagen mediante el cálculo de la raíz cuadrada de la diferencia cuadrática media entre los valores de los píxeles de la imagen distorsionada y la imagen referencia, este valor representa un error por lo que un menor valor demuestra mayor similitud de las imágenes y por ende mayor calidad (Wang, Bovik, Sheikh, & Simoncelli, 2004).
- PSNR *(Peak Signal-to-Noise Ratio):* es una métrica de la cantidad de ruido en una imagen mediante el cálculo del radio entre la potencia máxima posible de la señal y la potencia del ruido en la imagen. Debido al amplio rango de valores máximos y mínimos de valores de pixeles en una imagen, el PSNR se calcula

tomando en cuenta el máximo valor de pixel y es expresado en la escala logarítmica de decibeles (National Instruments, 2013). El propósito es tener un valor alto de PSNR que quiere decir que la potencia de la señal es mayor a la potencia del ruido indicando una mejor calidad.

- SSIM *(Structural Similarity Index Measurement):* Los anteriores parámetros estiman errores percibidos para cuantificar degradaciones en la imagen, pero no toman en cuenta el tipo de error y la calidad visual que produce. La medida de calidad SSIM mas bien considera las degradaciones de la imagen como cambios en la información estructural que es lo que extrae la visión humana HSV *(Human Visual System).* Según (Wang et al., 2004), la información estructural de una imagen es definida como las propiedades que representan la estructura de los objetos en la escena, independiente del promedio de luminancia y contraste.

  El funcionamiento del sistema de medición SSIM se basa en tres comparaciones por separado: luminancia, contraste y estructura (ver figura 9). El sistema funciona con la entrada de dos señales de imágenes x e y, una de ellas es una imagen libre de distorsiones, es decir, se supone como la medida de referencia de calidad perfecta con la que es comparada la otra señal ingresada. Finalmente, luego de la comparacion de los tres parametros se realiza una combinación para obtener un resultado final de similitud (Wang et al., 2004).

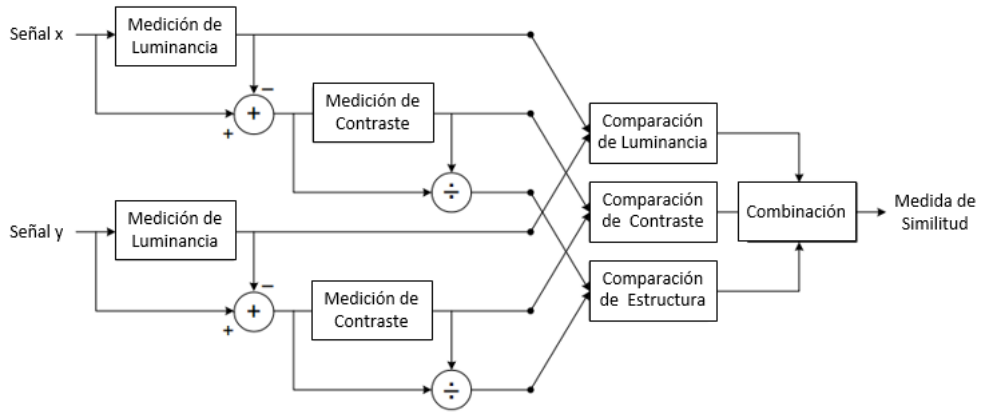


Figura 9 Diagrama del sistema de medida de la similitud estructural (SSIM)

Fuente: (Wang et al., 2004)

2.5.4 Capacidad de incrustación

La capacidad de incrustación es la métrica para determinar cuánta información puede ser embebida dentro de la imagen; según el método de incrustación en (Coltuc & Chassery, 2007) utilizado en este trabajo, la cantidad de información por pixel es determinada por la siguiente fórmula:

$$B = \frac{2D-F}{2F}\ [bpp] \tag{1}$$

Donde *F* es el número total de pares, *D* el número total de pares con información oculta y *B* la tasa de bits provista por el método en bpp (bits por pixel). El algoritmo brinda espacio de incrustación si al menos la mitad del número total de pares es utilizado, es decir $F > D/2$. Por lo tanto la capacidad máxima del sistema es 0.5 bpp cuando $F \approx D$. Para tener mayor capacidad de ocultación de datos, se puede ejecutar múltiples iteraciones del algoritmo en los pixeles de la imagen, alternando direcciones (horizontal, vertical). Al barrido de la imagen en filas o columnas en el proceso de incrustación se lo define como una iteración. Sin embargo, se debe tomar en cuenta que con cada iteración la distorsión en la imagen aumenta y el número de pares utilizables disminuye (Coltuc & Chassery, 2007).

CAPÍTULO 3

3. DISEÑO DEL PROGRAMA

3.1 Descripción general del programa

El algoritmo propuesto consiste en la combinación de un método de transformación de imágenes utilizando las características del color (Ya-Lin & Wen-Hsiang, 2014) y un nuevo método esteganográfico que utiliza filtros y operadores diferenciales para detectar las zonas más adecuadas de la imagen donde se ocultará la información clave en la recuperación de la imagen secreta. Dentro del programa se observa cuatro factores fundamentales:

1. Imagen secreta: imagen que se quiere ocultar
2. Imagen portadora: imagen base para la creación del mosaico
3. Imagen mosaico: imagen similar a la portadora y contiene a la imagen secreta
4. Algoritmo empleado

Dentro del algoritmo se puede notar dos secciones principales: la creación de un mosaico y la recuperación de la imagen secreta; estas secciones se pueden observar a continuación en el diagrama de la figura 10.

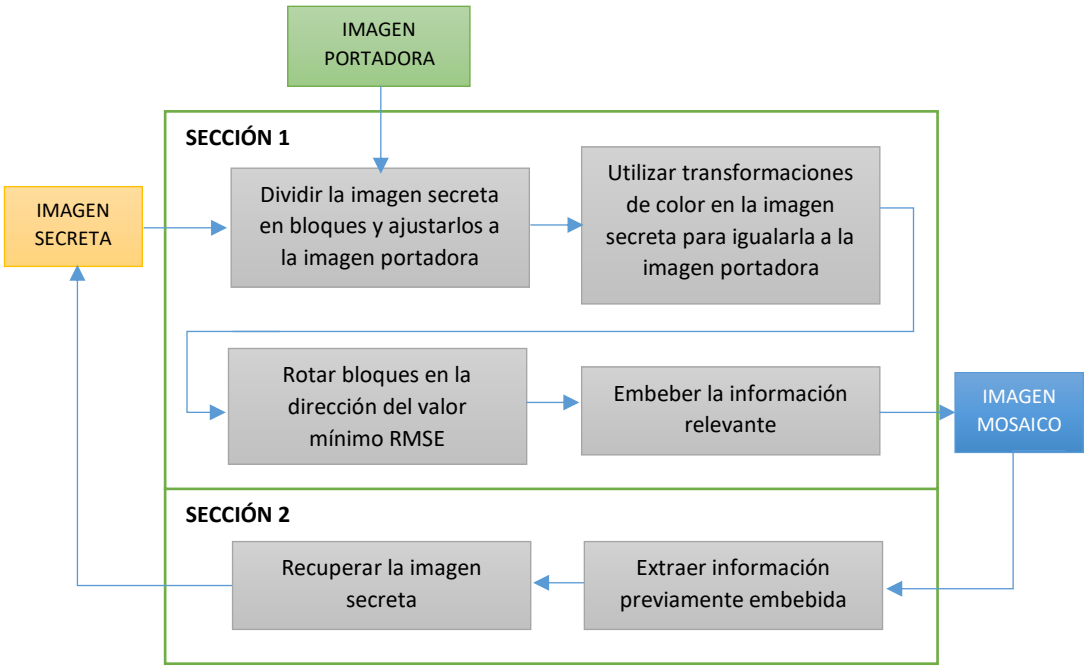


Figura 10 Diagrama de bloques de algoritmo Ya-Lin – Wen-Hsiang

El proceso de obtención del mosaico y recuperación de la imagen secreta se realizó mediante el método propuesto en (Ya-Lin & Wen-Hsiang, 2014), el cual se describe brevemente a continuación, por el contrario, en la sección 3.3 se aportará una amplia explicación de la nueva técnica utilizada para detectar zonas adecuadas y embeber la información relevante y su posterior extracción para recuperar la imagen secreta.

La primera parte del proceso descrito en la figura 10 tiene como objetivo obtener una imagen mosaico que luce similar a la imagen portadora mediante transformaciones reversibles del color de la imagen secreta. En la imagen mosaico se oculta información clave acerca de las características del color (medias y desviaciones estándar) de la imagen secreta mediante métodos esteganográficos. A diferencia de la técnica original que usa el método *LSB* en toda la imagen mosaico, se utilizó un nuevo método que mediante filtros y operadores diferenciales busca las zonas más ruidosas (texturas y bordes) de la imagen para el proceso de incrustación. Esta modificación se realizó con el propósito de obtener una imagen mosaico estadísticamente más imperceptible ante estego-analizadores.

La imagen mosaico es transmitida y en el lado del receptor se aplica el procedimiento de la segunda sección del algoritmo que consiste en la recuperación de la imagen secreta, para ello se realiza los procesos inversos de la primera sección utilizando las transformaciones reversibles de color y los parámetros como medias y desviaciones estándar ocultos en la imagen mosaico. Cabe mencionar que este método se realizó en imágenes a color por lo que todos los procedimientos aplicados se realizaron en cada una de las tres matrices de color *(Red, Green, Blue)*.

3.2 Procedimiento de transformación de color en imágenes

3.2.1 Sección 1: Creación de la Imagen Mosaico

Como primer paso en la creación del mosaico se debe verificar que las imágenes portadora *P* y secreta *S* sean del mismo tamaño, de no ser el caso se procede a establecer un tamaño común como se puede ver en la figura 11, en este caso se utilizó un tamaño de 640x480 debido a que son cantidades con capacidad de divisibilidad para 4, 5, 8 pixeles (tamaños de bloques usados en la creación de la imagen mosaico).


(a)

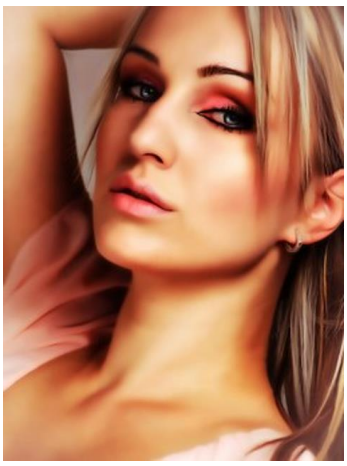
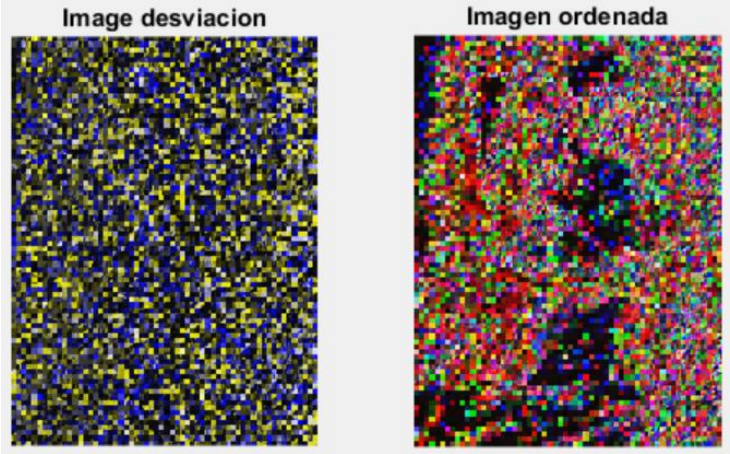

(b)

Figura 11 (a) Imagen secreta (b) Imagen portadora

Como segundo paso, se dividió en bloques del mismo tamaño a ambas imágenes. Para las pruebas se utilizaron bloques de 8x8 pixeles, tamaño que dio mejores resultados en relación al error RMSE, la cantidad de pixeles a ocultar y el tiempo de ejecución, como se demuestra en la sección de análisis de resultados.

El objetivo de dividir la imagen secreta en bloques es tener una correspondencia con los fragmentos de la imagen portadora para hacer que sus distribuciones del color luzcan similares. Evidentemente las características de color entre las dos imágenes son diferentes, por lo que es necesario realizar transformaciones de color previamente mediante parámetros estadísticos que describen la distribución del color en la imagen.

Los parámetros principales usados en el proceso de transformaciones de color son la desviación estándar y la media. Primero se calculó las desviaciones estándar de cada bloque de la imagen *P* y de la imagen *S* en los tres canales de color RGB y se ordenó los fragmentos de las dos imágenes de forma descendente de acuerdo a sus desviaciones estándar, luego se realizó la correspondencia entre los bloques de la imagen secreta y los de la imagen portadora (ver figura 12) y se guardó las posiciones iniciales para posteriormente ocultar esta información dentro del mosaico y obtener la imagen ordenada en la segunda sección. Se obtuvo mejores resultados realizando el proceso independientemente en cada matriz RGB y no obteniendo un promedio de las tres desviaciones estándar como se menciona en (Ya-Lin & Wen-Hsiang, 2014).



(a) (b)

Figura 12 Bloques de imagen secreta (a) Orden descendente (b) Orden de acuerdo a desviación estándar

Para conseguir que los nuevos colores luzcan similares a los de los bloques de la imagen portadora se aplicaron procesos de transformaciones de color reversible, que requieren del cálculo de la media y desviación estándar de cada bloque en cada canal RGB con las siguientes ecuaciones expresadas en (Ya-Lin & Wen-Hsiang, 2014):

Imagen secreta	Imagen portadora	
$\mu_c = \frac{1}{n}\sum_{i=1}^{n} c_i$	$\mu'_c = \frac{1}{n}\sum_{i=1}^{n} c'_i$	(2)
$\sigma_c = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - \mu_c)^2}$	$\sigma'_c = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(c'_i - u'_c)^2}$	(3)

En donde μ_c *y* μ'_c son las medias y σ_c *y* σ'_c son las desviaciones estándar para la imagen secreta y la imagen portadora respectivamente; *c* corresponde a los valores de pixel del bloque en cada canal *c* = *R*, *G*, o *B* y *n* es el número total de pixeles en cada bloque. Después se procede a obtener el nuevo valor de color de cada pixel c''_i con la Ecuación 4:

$$c''_i = q_c(c_i - \mu_c) + \mu'_c \quad (4)$$

Donde $q_c = \sigma'_c/\sigma_c$ corresponde al coeficiente de desviación estándar y *c* = *R*, *G* o *B* se refiere a los canales de la imagen a color. Se puede obtener fácilmente el color original del pixel mediante la ecuación inversa de la Ecuación 4:

$$c_i = \left({}^{1}/_{q_c}\right)(c_i'' - \mu'_c) + \mu'_c \quad (5)$$

Con la Ecuación 5 se puede recuperar los valores originales de la imagen secreta, para ello es necesario ocultar en la imagen mosaico la información necesaria. Según la Ecuación 5, se requiere ocultar los coeficientes de desviación estándar q_c y las medias μ_c de cada bloque, sin embargo estos son números reales y para ser incrustados deben ser limitados a un número fijo de bits. En este caso se permite que los valores de las medias μ_c estén representados por 8 bits, es decir se encuentren dentro del rango de 0 a 255 y los coeficientes de desviación q_c se limitan a 7 bits, es decir a un rango de 0.1 a 12.8. Si los valores son superiores o inferiores a los rangos dados se debe truncar el valor para que esté dentro de los rangos establecidos, debido a esto se crean pérdidas que son mínimas al recuperar la imagen original.

Después del proceso de transformación del color, algunos de los valores transformados podrían tener desbordamientos o subdesbordamientos, es decir valores mayores o menores al rango de 0 a 255. Para solucionar este problema, estos valores son cambiados a 255 si son mayores o a 0 si son menores y se obtiene la diferencia o residuos en el espacio no transformado. Estos residuos son guardados junto con una tabla de codificación Huffman (Huffman, 1952) que indica su frecuencia, para utilizarlos posteriormente en la recuperación del color original.

Para ajustar de mejor manera los bloques de la imagen secreta después de la transformación de las características de color, se rota el bloque en las 4 direcciones 0º, 90º, 180º y 270º, obteniendo un nuevo bloque rotado al ángulo óptimo para tener el menor error RMSE con respecto al bloque de la portadora (ver figura 13).

(a) (b)

Figura 13 (a) Imagen después de las transformaciones de color (b) Imagen después de la rotación de los bloques

El siguiente paso es la incrustación de información relevante en el mosaico que es requerida en la sección 2 para recuperar la imagen secreta. Este proceso consta de dos etapas: detección de zonas adecuadas e incrustación de la información, el resultado de la primera etapa es un mapa de bordes y texturas M_{bt} que indica los pixeles donde se aplicará la incrustación de la información. Este paso será descrito ampliamente en la sección 3.3.

La información que se debe ocultar en el mosaico es:

1. Los índices de las posiciones originales de la imagen portadora
2. El ángulo de rotación de los bloques de la imagen mosaico
3. Las medias y el coeficiente de desviaciones estándar de cada bloque dentro del rango establecido
4. Los residuos de desbordamiento y subdesbordamiento

Como se dijo anteriormente se debe establecer un número fijo de bits para cada ítem a ocultar: el número de bits para los índices se obtiene mediante el logaritmo en base dos del número de bloques totales de la imagen, en nuestro caso se tiene en total 4800 bloques por lo que se requiere 13 bits para cada índice; para el ángulo se requieren 2 bits ya que son 4 opciones de ángulos a rotar; para las medias se requiere 16 bits y para el coeficiente 14 bits en total para un bloque de la imagen secreta y uno de la portadora, y 8 bits son requeridos para los residuos en el espacio no transformado. Todos estos ítems de todos los bloques se deben concatenar formando un solo flujo de

bits llamado *bits_M* para cada canal, es decir se forman 3 flujos que se ocultarán independientemente en su matriz del canal correspondiente. Como se puede notar la cantidad de bits a ocultar es alta, por lo que pueden ser necesarias varias iteraciones en el proceso de incrustación, mientras el flujo de bits a ocultar sea mayor que los bits disponibles en el mapa de texturas y bordes.

Después de ocultar los tres flujos en los tres canales, se requiere otra información importante para el proceso de recuperación de la imagen; se debe crear otro flujo de bits llamado *bits_I* con los siguientes datos: el número de iteraciones requeridas en cada canal, el número de pixeles usados en la última iteración en cada canal, y la tabla Huffman utilizada para la codificación de los residuos. Además se requiere ocultar el mapa M_{bt} que se obtiene en la sección 3.3.1, para conocer las zonas donde se ocultó información al momento de extraer la información relevante. Estos dos flujos se ocultarán en el canal *R* y en el canal *G* respectivamente, con el método usado en (Ya-Lin & Wen-Hsiang, 2014) obteniendo finalmente una estego-imagen a transmitir (ver figura 14). El número de iteraciones y pixeles usados al ocultar el flujo *bits_I* y el mapa M_{bt} se utilizarán como una clave *K* para una mejor protección contra ataques. Esta clave debe ser conocida por el receptor para recuperar la imagen secreta.

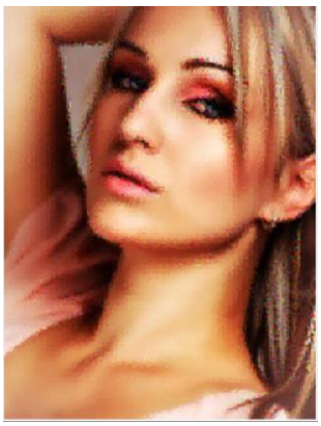


Figura 14 Imagen mosaico con información embebida

3.2.2 Sección 2: Recuperación de la imagen secreta

Para recuperar la imagen secreta, primero se debe extraer la información relevante, para ello es necesario tener la clave *K* formada en la anterior sección. Con esta clave se puede obtener el flujo *bits_I* y el mapa de bordes y texturas M_{bt}, el proceso de extracción de los bits está descrito en (Coltuc & Chassery, 2007), donde se usa las

Ecuaciones 7 descritas en la sección 3.3.2 para obtener los pixeles originales y continuar con la recuperación de toda la información oculta en varias iteraciones.

Para obtener el flujo *bits_M* de cada canal es necesario el flujo *bits_I* donde se encuentran el número de iteraciones y pixeles utilizados y el mapa M_{bt} para conocer las zonas donde se ocultaron los bits. Una vez identificado los pixeles alterados en el proceso de incrustación, se puede aplicar el algoritmo inverso de la técnica en (Coltuc & Chassery, 2007) solo en las zonas especificadas y así obtener el mensaje oculto y la imagen mosaico original antes del proceso de incrustación como se puede ver en la figura 15.

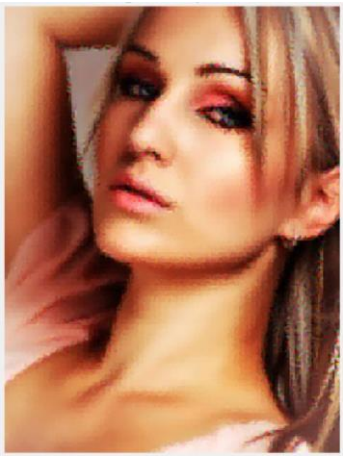


Figura 15 Imagen mosaico recuperada

Luego se procede a descomponer el flujo *bits_M* de cada canal según el número de bits utilizados en cada ítem que compone el flujo de bits: los índices, el ángulo de rotación, las medias y coeficientes de desviación y los residuos que son decodificados con la tabla de Huffman que se extrajo previamente. Con toda la información anterior se inicia el proceso inverso de la sección 3.2.1:

1. Rotar los bloques en la dirección inversa de los ángulos óptimos para obtener el bloque original
2. Aplicando la Ecuación 5, usando las medias y coeficientes y sumando los residuos donde sea necesario se obtiene el valor de pixel original.
3. Por último se ordena los bloques de acuerdo a los índices que indican las posiciones iniciales y se obtiene la imagen secreta recuperada llamada *Srec* (ver figura 16).

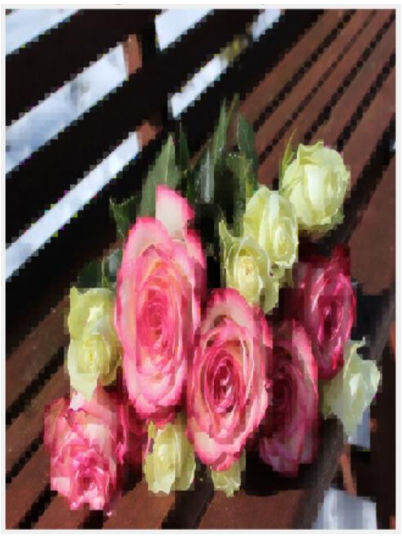


Figura 16 Imagen secreta recuperada

3.3 Algoritmo esteganográfico propuesto

Se puede notar claramente que para la recuperación de la imagen secreta es necesario tener cierta información relevante oculta dentro del mosaico creado. Esta información es la clave para la recuperación de la imagen original y es embebida mediante la técnica descrita en (Coltuc & Chassery, 2007). Esta técnica utiliza transformaciones integrales aplicadas a pares de pixeles para incrustar un flujo de bits, a diferencia de los métodos comunes *LSB* donde se reemplaza directamente los bits a ocultar.

En el método propuesto por Ya-Lin y Wen-Hsiang anteriormente descrito, la información relevante es ocultada en toda la imagen sin tomar en cuenta el tipo de áreas o superficies. El presente trabajo propone la búsqueda de zonas adecuadas para ocultar información, con el objetivo de tener imágenes esteganográficas difícilmente detectadas mediante estego-análisis estadístico. Este nuevo método surgió por la forma en que trabajan los estegoanalizadores: para detectar el uso de las técnicas más complejas conocidas como *LSB matching*, los estegoanalistas utilizan una base de datos de imágenes y entrenan clasificadores que detectan cuáles imágenes contienen información oculta y cuáles no. Para ello es necesario buscar que características de la imagen son susceptibles a ser alteradas al incrustar información y modelar estas características para entrenar al clasificador.

"Una de las lecciones aprendidas en los últimos años de investigación en estegoanálisis usando clasificadores es que existen zonas que son mucho más difíciles de modelar que otras: los bordes y las texturas. Estas zonas contienen mucho ruido y, en ellas, es muy difícil extraer características adecuadas para entrenar al clasificador". (Lerch Hostalot & Megías, 2014) (p. 173)

Los sistemas modernos de esteganografía centran sus investigaciones, mayoritariamente, en métodos que permiten ocultar información en estas zonas difíciles de modelar para los clasificadores usados por estegoanalistas. Por ejemplo los métodos de estegoanálisis descritos en (Lerch-Hostalot & Megías, 2013) y (Pevný, Bas, & Fridrich, 2009) modelan las características de la imagen como diferencias de pixeles vecinos. En el caso de la técnica PPD *(patterns of pixel diferences)* en (Lerch-Hostalot & Megías, 2013) se utilizan cinco pixeles vecinos, tomando como base uno de ellos para restar de los demás, es decir, se obtiene un vector con cuatro diferencias posibles que es tomado como un patrón.

En el caso más sencillo de una imagen en escala de grises de 8 bits se tendría valores de pixeles de 0 a 255. Por lo que la diferencia más grande entre dos pixeles es 255, y se podrían obtener hasta 255^4 características, cantidad no práctica para los clasificadores. Tomando en cuenta que no es frecuente que un pixel de valor 0 sea vecino de un pixel con valor 255, debido a que las imágenes cambian de valores gradualmente, se puede excluir las diferencias grandes sin perjudicar los sistemas de estegoanálisis. Es por esto que los métodos de extracción de características basados en la diferencia de pixeles utilizan un umbral para disminuir el número de características a una cantidad manejable que permita modelar la imagen.

Otro motivo para reducir la dimensión de las características es la insuficiencia de número de muestras. En el proceso de extracción de características, se ubican las muestras de la imagen en el patrón correspondiente. Al tener un número de muestras fijo, cuando aumenta el número de patrones disminuye su frecuencia, es decir, los patrones menos frecuentes (con baja cantidad de número de muestras) no serán significativos para el clasificador afectando su entrenamiento. Por estas razones existe un límite en el umbral a partir del cual los métodos de estegoanálisis dejan de ser efectivos.

Basado en el procedimiento de los clasificadores descrito, lo que se pretende en el algoritmo propuesto, es explotar las debilidades dimensionales en la extracción de características para entrenar clasificadores. Las zonas uniformes de las imágenes presentan cambios de color paulatinos, mientras que cambios bruscos de color representan bordes y texturas, es decir pixeles vecinos cuya diferencia sea un valor alto, por lo que pueden no ser tomados en cuenta en el proceso de modelado de imágenes formando zonas aptas para ocultar información.

A continuación se describe el algoritmo propuesto, el cual consta de dos etapas principales: (1) detección de zonas adecuadas e (2) incrustación de la información relevante.

3.3.1 Detección de zonas adecuadas

En (Lerch Hostalot & Megías, 2014) se plantea una técnica para encontrar las zonas donde existen bordes y texturas en la imagen, para ello se establece un umbral *U* y se trabaja con pares de pixeles cuya diferencia debe ser mayor o igual al umbral para ser tomados en cuenta en la incrustación de información. En este caso únicamente se usa dos pixeles, sin embargo en el algoritmo propuesto en este trabajo (ver figura 17) se utilizó operadores diferenciales y filtros que consideran más pixeles vecinos para detectar variaciones significativas en los valores de pixeles.

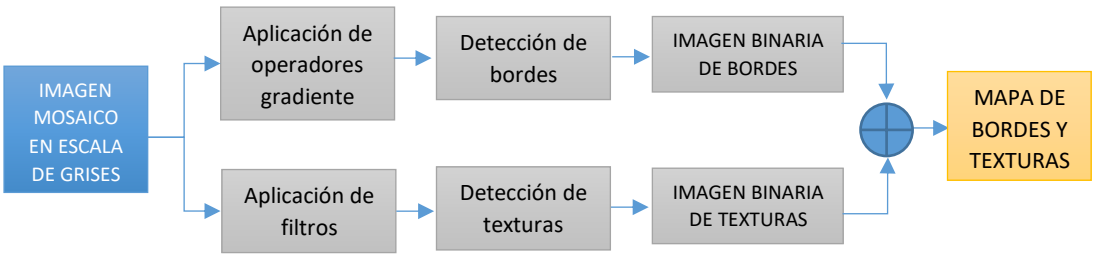


Figura 17 Diagrama de bloques de algoritmo de búsqueda de zonas adecuadas para la incrustación de información

Al trabajar con imágenes RGB se tiene tres matrices de pixeles con diferentes valores, lo que implica dificultades para detectar cambios en la imagen. Es necesario primero convertir la imagen a escala de grises (ver figura 18), para detectar las transiciones significativas y luego se procede a aplicar los algoritmos de detección de bordes y texturas.

Figura 18 Imagen mosaico en escala de grises

Para el caso de detección de bordes se usó operadores diferenciales. Como se mencionó anteriormente existen varios algoritmos basados en operadores gradiente como: Sobel, Prewitt y Roberts para la detección de bordes (González Aguilera, 2011) y además el detector Canny (Moreira, Valencia, & Chávez, 2009). En la figura 19 se puede ver los diferentes métodos que pueden ser usados para la detección, donde los bordes de la imagen se pueden observar como una imagen en binario con valor de 1 para representar las zonas donde el algoritmo encontró bordes.

(a) (b) (c) (d)

Figura 19 Métodos de detección de bordes (a) Canny (b) Sobel (c) Prewitt (d) Roberts

Sin embargo después de las pruebas realizadas, para el algoritmo propuesto se utilizó el operador Canny, ya que combina un operador diferencial con un filtro gaussiano, es decir consta de dos etapas principales: la primera utiliza el filtro gaussiano para reducir el ruido en la imagen (texturas y detalles no significativos), y

la segunda usa operadores diferenciales para detectar los bordes en todas las direcciones (horizontal, vertical y diagonal), a diferencia de los otros métodos.

Para el caso de detección de texturas se aplicó filtros basados en las características de un grupo de pixeles. El estudio de texturas se refiere a la capacidad de discernir diversas regiones de la imagen por sus características de superficie física. Este análisis se realiza en un pixel y su grupo de pixeles vecinos en donde se puede determinar algunas características para detectar texturas en la imagen como la entropía de la distribución de los niveles de grises, la desviación estándar (distribución de niveles de gris en una región), rango de valores (valores máximos y mínimos) (The MathWorks, Inc., s.f.). El resultado de los filtros mencionados se puede observar en la figura 20.

(a) (b) (c)

Figura 20 Filtros para detección de texturas (a) Entropía (b) Desviación estándar (c) Rango de valores

En el algoritmo propuesto en este trabajo se requiere una gran capacidad de incrustación es por ello que se eligió el filtro entropía que depende de la cantidad de niveles de grises y la frecuencia de cada nivel, es decir que si en un grupo de pixeles existe una gran cantidad de ellos con la misma intensidad se va a tener una alta entropía, mientras que si existe una distribución más uniforme de los valores de gris se tendrá una baja entropía (Riera & Tacón, 2011). Por lo que las zonas con mayor entropía representan las texturas en la imagen.

Para la obtención de las zonas con texturas primero se debe aplicar el filtro entropía (Gonzalez, 2003) cuya salida contiene el valor de entropía de una matriz 9x9 de pixeles vecinos alrededor del pixel seleccionado en la imagen de entrada. Esta salida debe ser

convertida a binario para tener un mapa de los pixeles que pueden ser utilizados en el proceso de incrustación de información. Por ello se realizan conversiones de la imagen previas a la obtención del mapa observado en la figura 20: primero se convierte los valores en una imagen en escalas de grises y luego se convierte la imagen a binario. Esta imagen binaria contiene unos (1s) que representan las texturas de la imagen y en donde se puede ocultar información y ceros (0s) para zonas uniformes.

Después de obtener dos imágenes binarias con los bordes y texturas de la imagen mosaico se realizó una simple operación OR o suma de las dos imágenes para obtener el mapa total requerido como se observa en la figura 21 para el algoritmo de incrustación de la información relevante.

Figura 21 Mapa de bordes y texturas de imagen mosaico

3.3.2 Incrustación de información relevante

En esta etapa se procede a aplicar la técnica presentada en (Coltuc & Chassery, 2007) en las parejas de pixeles indicadas con 1s en el mapa de bordes y texturas M_{bt} obtenido anteriormente. Para ello se requiere el uso de dos ecuaciones fundamentales donde (e,f) representan un par de pixeles del mosaico y (e',f') el par de pixeles transformados después de la aplicación de la técnica:

$$e' = 2e - f, \qquad f' = 2f - e \tag{6}$$

$$e = \left[\frac{2}{3}e' + \frac{1}{3}f'\right], \qquad f = \left[\frac{1}{3}e' + \frac{2}{3}f'\right] \tag{7}$$

Para ocultar información el procedimiento consiste en recorrer la imagen en dirección horizontal o vertical tomando cada vez una pareja diferente de pixeles

vecinos. Las parejas se forman sin solapamiento, de manera que, dados cuatro pixeles vecinos, (e,f,g,h), se formarían las parejas (e,f) y (g,h), siempre y cuando los pares de pixeles sean parte del mapa M_{bt}. Se puede obtener el valor original de los pixeles antes de la incrustación de información mediante las Ecuaciones 7.

Este método proporciona altas capacidades de incrustación debido a que permite varias iteraciones y esto es posible porque además de extraer el mensaje oculto, se puede recuperar el valor de los pixeles originales, con la complejidad operacional más baja.

3.4 Funciones principales

Como se indica en el diagrama de la figura 10, existen cuatro etapas principales en el algoritmo para la creación del mosaico:

1. Dividir la imagen secreta en bloques y ajustarlos a la imagen portadora de acuerdo a sus desviaciones estándar
2. Utilizar transformaciones de color en la imagen secreta para crear una imagen similar a la imagen portadora
3. Rotar los bloques de la imagen mosaico en la dirección del valor mínimo RMSE
4. Embeber la información relevante

El funcionamiento de cada uno de ellas dentro del programa será descrito a continuación mediante el diagrama de flujo de las funciones principales. La programación de todo el algoritmo se encuentra en el anexo 1.

3.4.1 Etapa 1: Funciones Bloques y Ordenar

Para el primer paso se requieren dos funciones principales: Bloques y Ordenar; la primera busca dividir la imagen portadora y la imagen secreta en bloques cuadrangulares del mismo número de píxeles *d*, y la segunda busca ordenar los bloques de la imagen secreta de acuerdo a las desviaciones estándar de la imagen portadora.

El diagrama de flujo de la función Bloques se puede observar en la figura 22. Como primer paso se lee la imagen a dividir y el tamaño del bloque dado por el número de pixeles *d*, se obtiene el alto c_1 y el ancho c_2 de la imagen y se trabaja en dos bucles

anidados para hacer un barrido vertical en la imagen y tomar *d* pixeles a lo ancho y *d* pixeles a lo largo formando un bloque de *dxd* pixeles que es guardado en la posición *m* de la celda *bloques.* La variable *m* incrementa su valor en uno hasta obtener todos los bloques de la imagen.

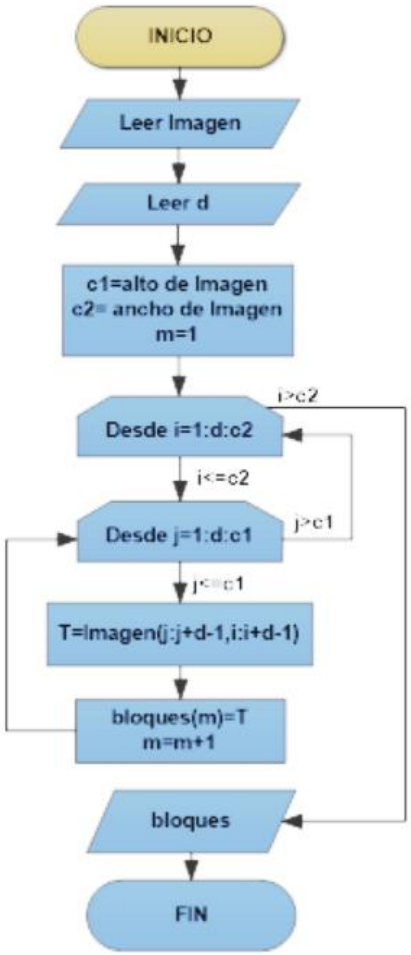


Figura 22 Diagrama de flujo de función Bloques

El diagrama de flujo de la función Ordenar se puede ver en la figura 23, el cual consiste en leer los bloques de la imagen portadora T_b, los bloques de la imagen secreta S_b y el número total de bloques *z* de la imagen; luego se utiliza un ciclo de repetición para obtener los vectores de desviaciones estándar *desT* y *desS* respectivamente, compuestos por el valor de desviación de cada bloque guardado en las celdas T_b y S_b. Se procede a ordenar de forma descendente los vectores *desT* y *desS* y se obtienen las posiciones o índices de los vectores ordenados *indT, indS*. A continuación se realiza

otro bucle de repetición porque se trabaja con vectores y se ordena los bloques S_b de acuerdo a las posiciones *indT*, es decir, según el orden de las desviaciones de la imagen portadora obteniendo una matriz ordenada *Bor* y las posiciones de los bloques S_b.

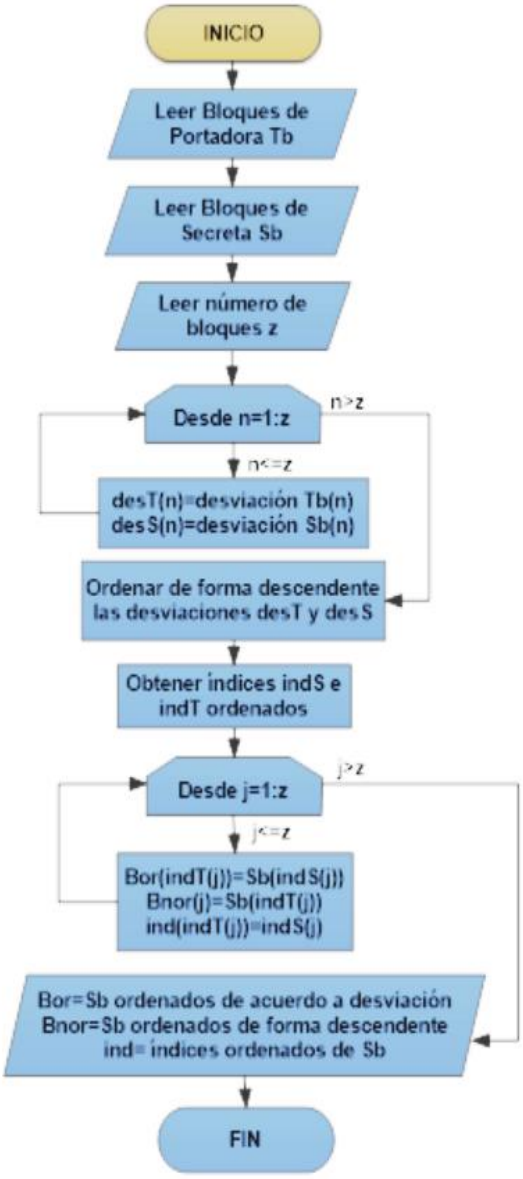


Figura 23 Diagrama de flujo de función Ordenar

3.4.2 Etapa 2: Función Nuevo pixel

El segundo paso es descrito en la función "Nuevo pixel". Esta función modifica el valor de los pixeles de los bloques de la imagen secreta ordenados con el fin de la imagen secreta luzca similar a la imagen portadora creando la imagen mosaico. El diagrama de flujo del funcionamiento de esta función se observa en la figura 24:

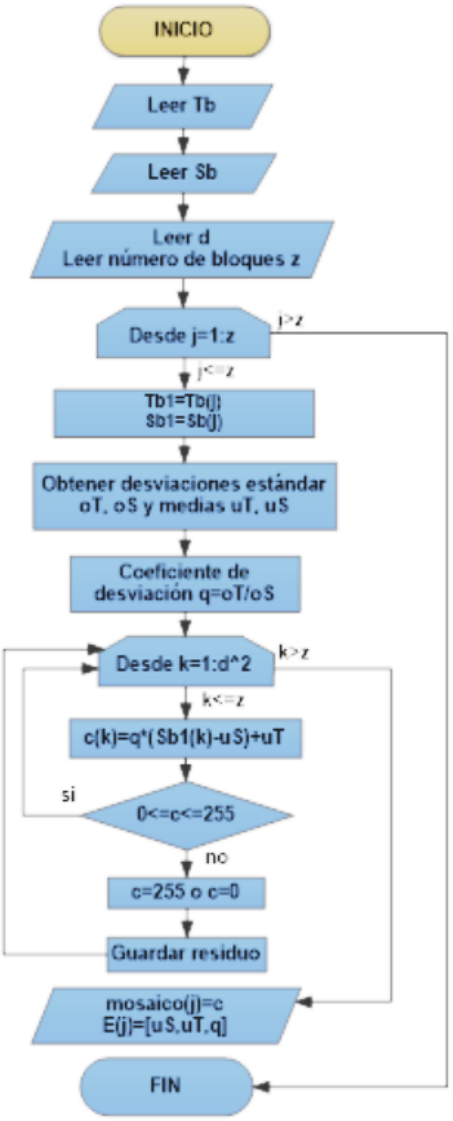


Figura 24 Diagrama de flujo de función Nuevo pixel

El proceso de la función Nuevo pixel consiste en leer los datos iniciales T_b, S_b *y z* usados en la figura 23, luego se crea un bucle de repetición para trabajar con vectores y obtener las desviaciones estándar *oT* y *oS*, y las medias *uT* y *uS* de cada bloque de T_b y S_b, de esta manera se puede utilizar la Ecuación 4 dentro de otro ciclo de repetición que permite moverse dentro del bloque de S_b para obtener el nuevo valor de cada pixel. Si el valor del pixel obtenido es mayor a 255 o menor a 0, se asume el valor más cercano 255 o 0 para ese pixel y se guarda la diferencia o residuo. Finalmente se obtiene un mosaico compuesto de pixeles con valores transformados, parecido a la imagen portadora, y un flujo llamado E compuesto por la medias *uS*, *uT* y el coeficiente de desviación *q*.

3.4.3 Etapa 3: Función Rotar

La función "Rotar" necesita como datos de entrada los bloques de la imagen mosaico *M*, los bloques de la imagen portadora T_b y el número de bloques *z*, después se utiliza un ciclo para leer los bloques de la imagen portadora y mosaico T_1 y M_1. Dentro del bucle de repetición se realiza el proceso de girar en las cuatro direcciones: 0º, 90º, 180º y 270º cada bloque M_1 de la imagen mosaico y realizar el cálculo del error RMSE con el fragmento de la imagen portadora T_1 correspondiente. El bloque es rotado el ángulo con el que se obtenga el menor error y guardado en la matriz *R*, este corresponde al ángulo óptimo de rotación que debe ser guardado en la matriz *a* para su posterior uso en la recuperación de la imagen secreta. Se puede ver en la figura 25 el funcionamiento de la función:

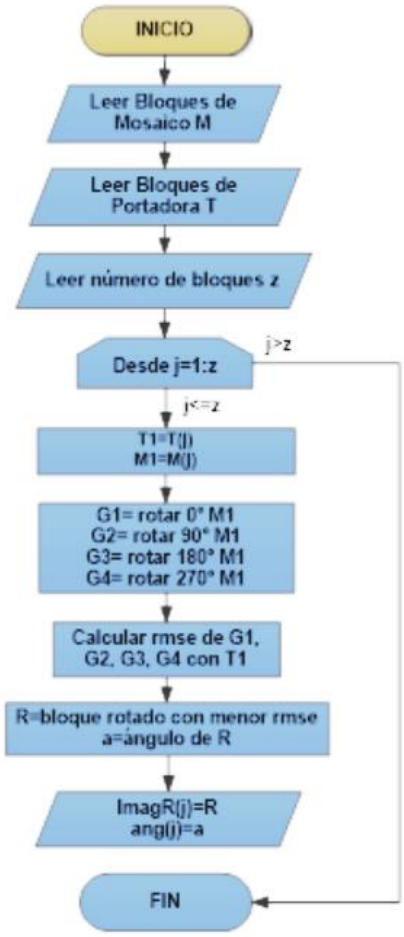


Figura 25 Diagrama de flujo de función Rotar

3.4.4 Etapa 4: Función Ocultar

Esta función utiliza el mapa de texturas y bordes generado para seleccionar los píxeles que serán usados en el proceso de incrustación. El algoritmo esteganográfico para embeber información relevante es descrito detalladamente en (Coltuc & Chassery, 2007). A continuación se puede observar en la figura 26 un diagrama de bloques del funcionamiento de la función Ocultar:

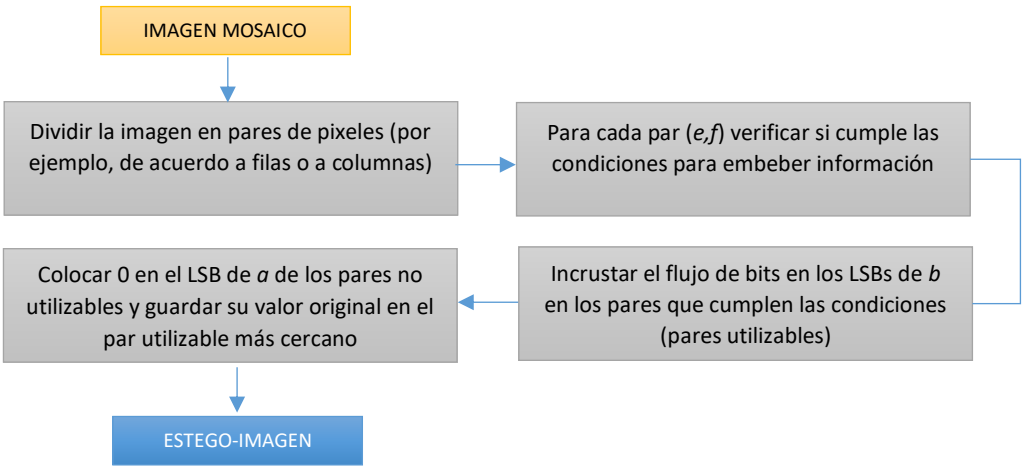


Figura 26 Diagrama de bloques de función Ocultar

CAPÍTULO 4

4. IMPLEMENTACIÓN, PRUEBAS Y ANÁLISIS

4.1 Pruebas realizadas

Se efectuó una serie de experimentos para probar el algoritmo propuesto. Las pruebas se realizaron con 70 imágenes extraídas de una base de datos de uso libre (Pixabay, s.f.) seleccionadas arbitrariamente, con las cuales se realizó un total de 100 pruebas con imágenes de un tamaño preestablecido de 640x480. Entre las 100 pruebas se combinó varias formas de seleccionar imágenes portadoras e imágenes secretas: una imagen portadora con varias imágenes secretas, una imagen secreta con varias imágenes portadoras e imágenes portadoras en correspondencia uno a uno con imágenes secretas.

El programa está automatizado para que se seleccione imágenes de una carpeta y se ejecute las funciones de creación de la estego-imagen y recuperación de la imagen secreta del algoritmo propuesto y posteriormente del algoritmo de Ya-Lin - Wen-Hsiang, en los dos casos se recopilan los valores de tiempos de ejecución, y número de bits incrustados. Además se almacena en carpetas las imágenes portadoras, imágenes secretas utilizadas, estego-imágenes obtenidas e imágenes recuperadas con una numeración de acuerdo al experimento realizado. Finalmente, se obtienen mediciones de similitud SSIM, nivel de ruido en las imágenes PSNR, error RMSE y se guardan los histogramas de la imagen portadora y de las estego-imágenes obtenidas con los dos algoritmos. Todos estos datos son almacenados en un archivo .csv creado previamente.

4.2 Análisis de resultados

En esta sección se realizó un análisis de los resultados obtenidos, primero se realizó una comparación de resultados con diferentes tamaños de bloques para seleccionar el mejor tamaño, luego se realizó un análisis de los resultados obtenidos en el proceso de ordenamiento de bloques según la forma de calcular las desviaciones estándar. Adicionalmente, se evaluó el rendimiento del algoritmo propuesto mediante la medición de diferentes parámetros como: tiempo de ejecución, capacidad de incrustación, ruido en la imagen, error RMSE y la similitud entre imágenes (valor

SSIM) en cada uno de los 100 experimentos realizados; además se comparó los resultados e histogramas obtenidos con el método utilizado en (Ya-Lin & Wen-Hsiang, 2014). A continuación se presenta el análisis realizado con los valores obtenidos al aplicar los dos métodos.

4.2.1 Selección del tamaño de bloque

El tamaño de bloque escogido se determinó en base a los parámetros descritos en la tabla 2 para un tamaño de 5x5, 8x8, 10x10, 16x16 y 32x32 pixeles.

Tabla 2

Comparación de tamaño de bloques

	Tamaño de bloques				
Parámetros de medición	**5x5**	**8x8**	**10x10**	**16x16**	**32x32**
RMSE de estego-imagen obtenida	27,709	13,487	14,014	18,676	28,981
Cantidad de bits a incrustar	1455506	575596	376311	186640	144033
Tiempo de ejecución (min)	6,819	2,644	1,824	1,061	0,875

Los resultados indican que el error RMSE con los tamaños de bloques 16x16 pixeles y 32x32 pixeles tienen altos valores, lo que muestra una calidad de la imagen baja con grandes distorsiones introducidas en el proceso de transformaciones de color, pero con un tiempo de ejecución bajo. En el caso del tamaño de 5x5 pixeles se observa un valor RMSE también alto debido a las distorsiones ocasionadas por la gran cantidad de información embebida y además un tiempo de ejecución mucho más alto en comparación a las demás pruebas. Los mejores resultados se obtuvieron con 8x8 y 10x10 pixeles, sin embargo se seleccionó 8x8 pixeles porque es el tamaño con el que se obtuvo mejor valor RMSE, que es un parámetro de calidad que se debe tomar muy en cuenta. A continuación se observa las estego-imágenes obtenidas con cada uno de los tamaños a prueba (ver figura 27).

(a) (b) (c) (d) (e)

Figura 27 Estego-imágenes (a) 5x5 (b) 8x8 (c) 10x10 (d) 16x16 (e) 32x32 pixeles

4.2.2 Proceso de ordenamiento de bloques según desviación estándar

Como se mencionó anteriormente, para el proceso de ordenamiento de bloques se utilizó el cálculo de las desviaciones estándar independientemente en cada canal RGB, en lugar del promedio de las desviaciones de los tres canales. Las mediciones de calidad de las estego-imágenes obtenidas con las dos opciones (ver tabla 3), pueden demostrar que el cálculo de la desviación en los canales RGB independientes proporciona mejores resultados en la calidad de la estego-imagen.

Tabla 3

Ordenamiento de bloques según desviación estándar

	Cálculo de desviación estándar	
Parámetros de medición de estego-imagen	**Canales RGB independientes**	**Promedio en canales RGB**
RMSE	15,670	16,001
PSNR	24,223	24,044
SSIM	0,916	0,915

4.2.3 Tiempo de ejecución

Las pruebas fueron realizadas en una laptop Lenovo ideapad 500, con sistema operativo Windows 10, procesador Intel Core i5-6200U CPU de 2.30 GHz y memoria RAM de 6GB. En la tabla 4 se observa el promedio de los valores medidos del tiempo transcurrido t_1 para la sección 1 (creación del mosaico) y del tiempo t_2 para la sección 2 (recuperación de la imagen secreta), aplicando los dos procedimientos: algoritmo propuesto y algoritmo Ya-Lin - Wen-Hsiang.

Tabla 4

Tiempo de ejecución promedio

ALGORITMO	PROPUESTO			YA-LIN – WEN-HSIANG		
Tiempo de Ejecución	t_1 (min)	t_2 (min)	TOTAL (min)	t_1 (min)	t_2 (min)	TOTAL (min)
Promedio	3,644	3,788	7,433	2,714	2,655	5,369

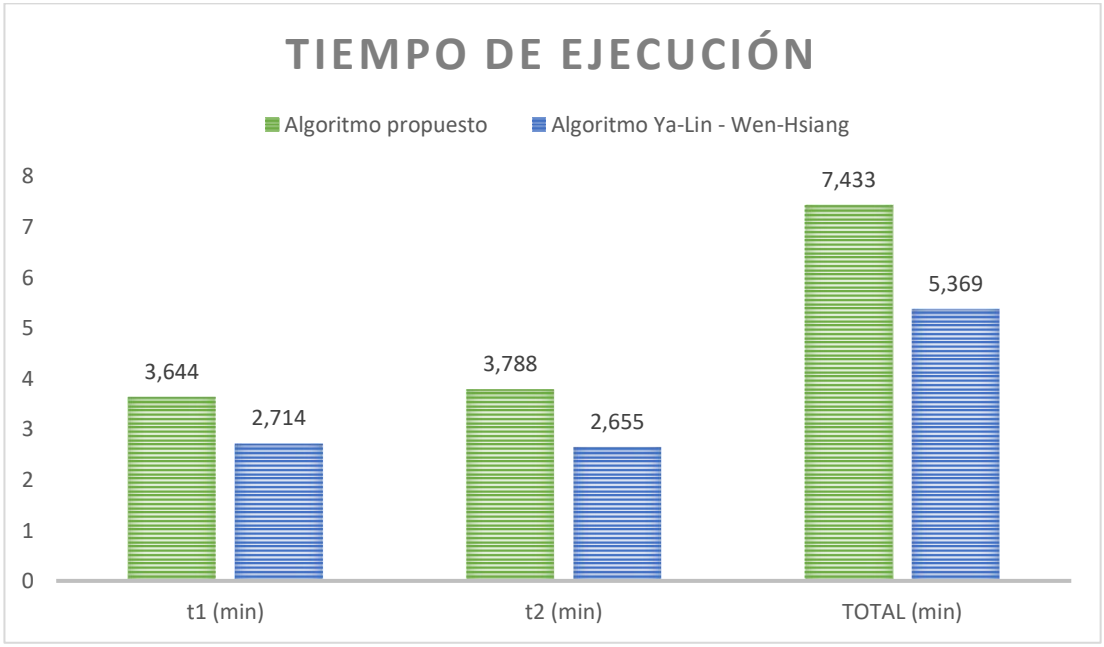


Figura 28 Tiempo de ejecución promedio

Después de obtener el tiempo promedio de ejecución de los 100 experimentos para los dos métodos (ver figura 28): para el algoritmo propuesto se obtuvo un tiempo de ejecución total promedio de 7,433; mientras que para el algoritmo de Ya-Lin - Wen-Hsiang se obtuvo un tiempo promedio de 5,369. Como se puede notar el tiempo de ejecución total promedio del algoritmo propuesto es mayor con 2 minutos aproximadamente; esto se debe principalmente al mayor número de iteraciones requeridas por el método propuesto como se observa en la tabla 5.

Tabla 5

Número de iteraciones totales

ALGORITMO	PROPUESTO	YA-LIN – WEN-HSIANG
Promedio de número de iteraciones	9	6

4.2.4 Capacidad de incrustación

La capacidad de incrustación permite determinar cuánta información puede ser oculta en una imagen. En la tabla 6 se indica la capacidad de incrustación promedio para cada método calculada mediante la Ecuación 1, para lo cual se usó el canal *R* de cada una de las imágenes mosaico y para F la cantidad de pares del mapa de bordes y texturas en el algoritmo propuesto, para el algoritmo de Ya-Lin - Wen-Hsiang F es el número total de pares de toda la imagen que en este caso es 1536000 para cada imagen como se muestra en el siguiente ejemplo con los datos obtenidos en el experimento 1 y usando la Ecuación 1: $B = (2D - F)/2F\ [bpp]$

Algoritmo propuesto

$$B = \frac{2(108899) - 129488}{2(129488)} = 0{,}3409\ [bpp]$$

Algoritmo Ya-Lin – Wen-Hsiang

$$B = \frac{2(130951) - 153600}{2(153600)} = 0{,}3525\ [bpp]$$

Tabla 6

Capacidad de incrustación

ALGORITMO	PROPUESTO			YA-LIN - WEN-HSIANG		
Capacidad de Incrustación	D (nº de pares)	F (nº de pares)	B(bpp)	D (nº de pares)	F (nº de pares)	B(bpp)
Promedio	83847,600	90351,160	0,428	145066,360	153600	0,444

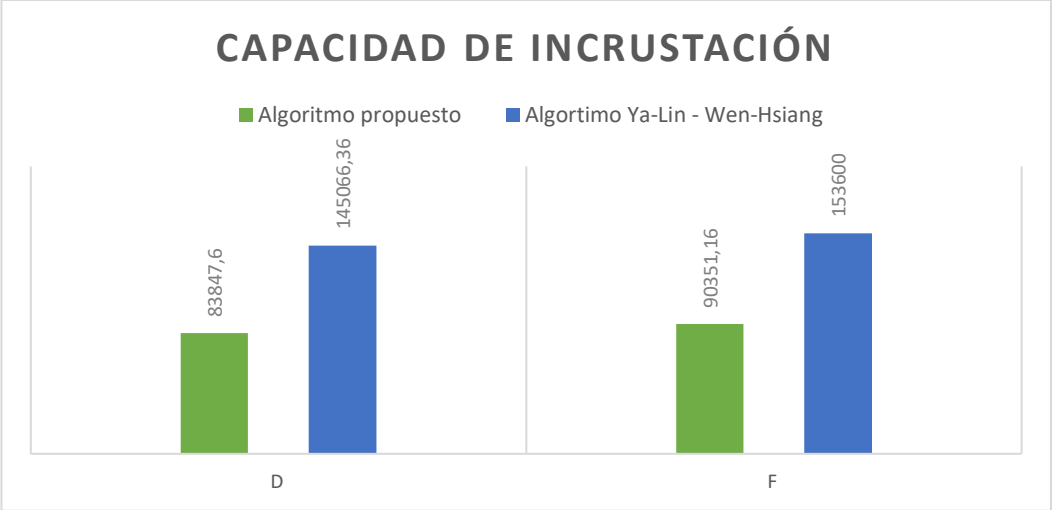


Figura 29 Parámetros para el cálculo de la tasa de bits B promedio en bpp

Si se observa la tabla 6 y se compara el promedio de la tasa de bits *B* entre los dos algoritmos: con 0,428 bpp para el algoritmo propuesto y 0,444 para el segundo algoritmo, se puede notar que tienen una mínima diferencia de 0,016 con respecto a la capacidad de incrustación; sin embargo como se puede ver en la figura 29 para el cálculo de *B* en el algoritmo propuesto se tomó únicamente el número de pares del mapa (en promedio 90351,16 pares), a diferencia del segundo algoritmo donde se tomó en cuenta el número de pares de toda la imagen (153600 pares). Por esta razón para el método propuesto en este trabajo, en una iteración se incrusta menor cantidad de bits en la imagen y es necesario mayor cantidad de iteraciones (ver tabla 5) que en el método de Ya-Lin y Wen-Hsiang.

4.2.5 PSNR *(Peak Signal to Noise Ratio)*

Una forma de medir la cantidad de ruido en una imagen con respecto a otra imagen de referencia es el cálculo del valor PSNR. En este trabajo se pretende medir el ruido de la estego-imagen con respecto a la imagen portadora y de la imagen secreta con respecto a la imagen recuperada. En la tabla 7 se indican los valores promedio obtenidos en las mediciones realizadas para las 100 imágenes. Un valor más alto indica una mejor calidad de la imagen:

Tabla 7

PSNR *(Peak Signal Noise to Ratio)*

	PSNR IMAGEN PORTADORA (dB)		PSNR IMAGEN SECRETA (dB)
ALGORITMO	**PROPUESTO**	**YA-LIN – WEN-HSIANG**	**AMBOS ALGORITMOS**
Promedio PSNR	18,839	21,734	24,285

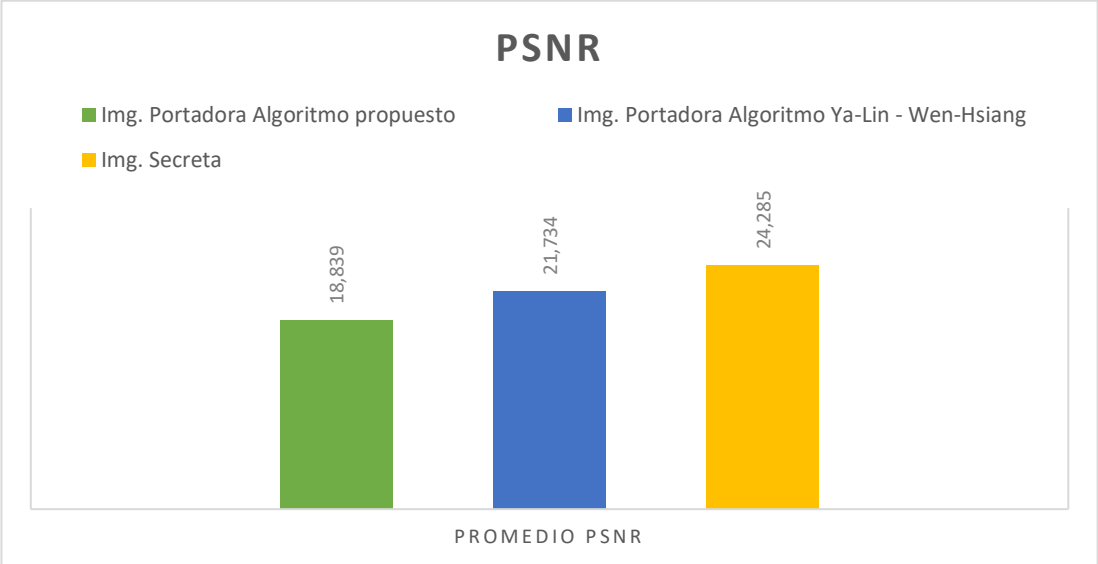


Figura 30 PSNR promedio en dB

Al comparar las tres mediciones promedio de PSNR realizadas en la figura 30, se evidencia claramente que la imagen secreta recuperada tiene el mayor valor promedio, lo que significa una mejor calidad en la imagen ya que existe menos ruido. Con respecto a las estego-imágenes, se obtuvo un valor mejor con el algoritmo de Ya-Lin y Wen-Hsiang, esto se debe a que el ruido es directamente proporcional al número de iteraciones ocasionando una mayor distorsión visual en la estego-imagen con el algoritmo propuesto.

4.2.6 RMSE *(Root Mean Square Error)*

En la tabla 8 se observa los valores RMSE promedio usados como medida de calidad de las imágenes obtenidas, se realizaron tres mediciones: el error RMSE entre la imagen portadora y la estego-imagen obtenida con el algoritmo propuesto y con el algoritmo Ya-Lin - Wen-Hsiang, y el RMSE entre la imagen secreta y la imagen secreta recuperada; este error es igual para los dos algoritmos.

Tabla 8

RMSE *(Root Mean Square Error)*

	RMSE IMAGEN PORTADORA		RMSE IMAGEN SECRETA
ALGORITMO	**PROPUESTO**	**YA-LIN – WEN-HSIANG**	**AMBOS ALGORITMOS**
Promedio RMSE	29,342	21,647	16,625

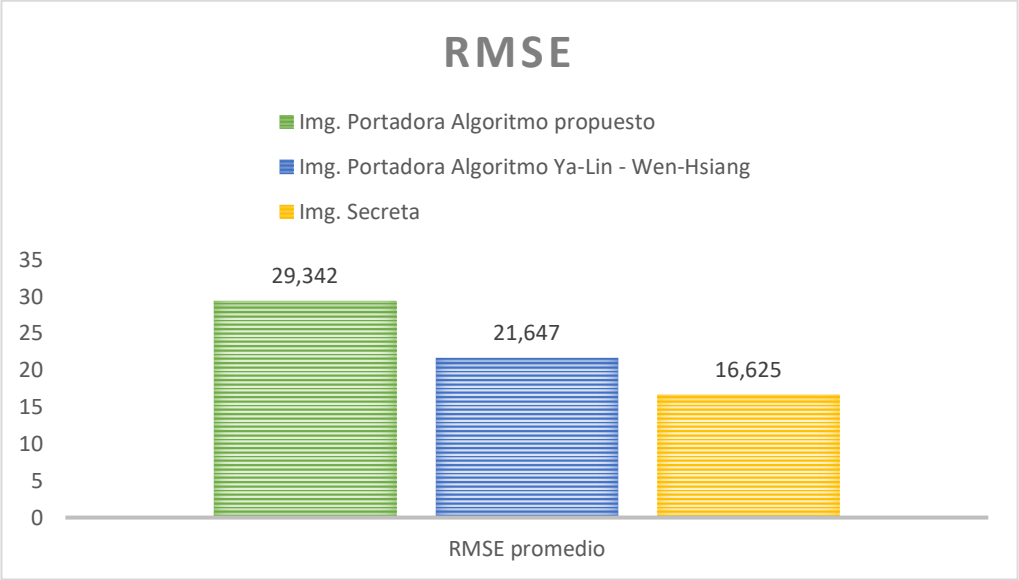


Figura 31 RMSE promedio

Según las mediciones RMSE obtenidas en la tabla anterior y en la figura 31, se observa en la estego-imagen obtenida con respecto a la imagen portadora que el valor RMSE promedio es de 29,342 para el algoritmo propuesto y 21,647 para el algoritmo de Ya-Lin y Wen-Hsiang; es decir, el error RMSE es menor para el segundo método debido a la distribución de bits embebidos en toda la imagen, a diferencia del método propuesto donde influye la mayor cantidad de bits incrustados únicamente en ciertas zonas; no obstante los valores en las dos técnicas están dentro de un rango aceptable en comparación con los resultados de RMSE obtenidos en (Ya-Lin & Wen-Hsiang, 2014). Con respecto a la imagen secreta los resultados serán los mismos para los dos métodos con un promedio de 16,625 que es un error menor a los valores obtenidos para las imágenes portadoras.

4.2.7 SSIM *(Structural Similarity Index)*

Las mediciones anteriores evalúan errores en la imagen como una medida objetiva, sin embargo con el valor SSIM se mide la similitud entre imágenes según el sistema visual humano, el rango de esta métrica es de 0 a 1, donde 0 corresponde a una pérdida total de la similitud estructural y 1 corresponde a una copia exacta de la imagen original. En la tabla 9 se observa los valores SSIM promedio obtenidos entre la estego-imagen y la imagen portadora para los dos métodos y entre la imagen secreta y la imagen recuperada.

Tabla 9

SSIM *(Structural Similarity Index)*

	SSIM IMAGEN PORTADORA		SSIM IMAGEN SECRETA
ALGORITMO	**PROPUESTO**	**YA-LIN – WEN-HSIANG**	**AMBOS ALGORITMOS**
Promedio SSIM	0,647	0,726	0,905

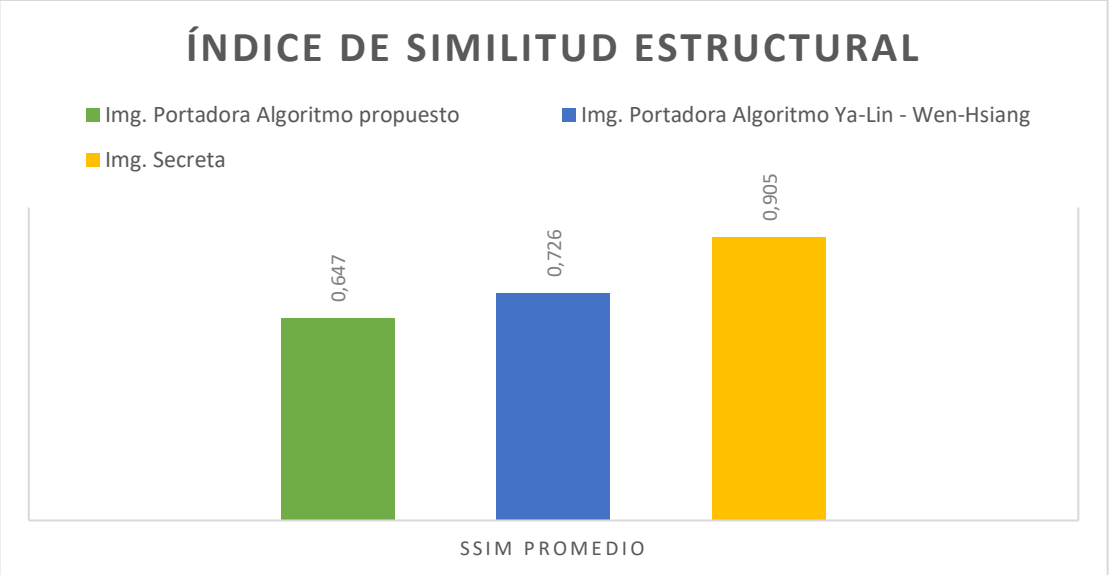


Figura 32 SSIM

Se puede evidenciar en la figura 32, que el promedio de los valores SSIM obtenidos en las 100 pruebas realizadas con respecto a la imagen secreta recuperada es de 0,905; cercano a 1. En el caso de las estego-imágenes se obtuvo un valor promedio de 0,647 para el método propuesto y 0,726 para el método planteado en (Ya-Lin & Wen-Hsiang, 2014), que indica que estructuralmente la similitud de la estego-imagen y la imagen portadora no es cercana a la imagen original. Sin embargo, esto no representa un gran problema en este trabajo, ya que no se busca similitud de imágenes en detalle, sino crear un mosaico visualmente similar de forma global cuyo principal objetivo es no ser detectado por estego-analizadores que utilizan métodos estadísticos.

4.3 Indetectabilidad estadística

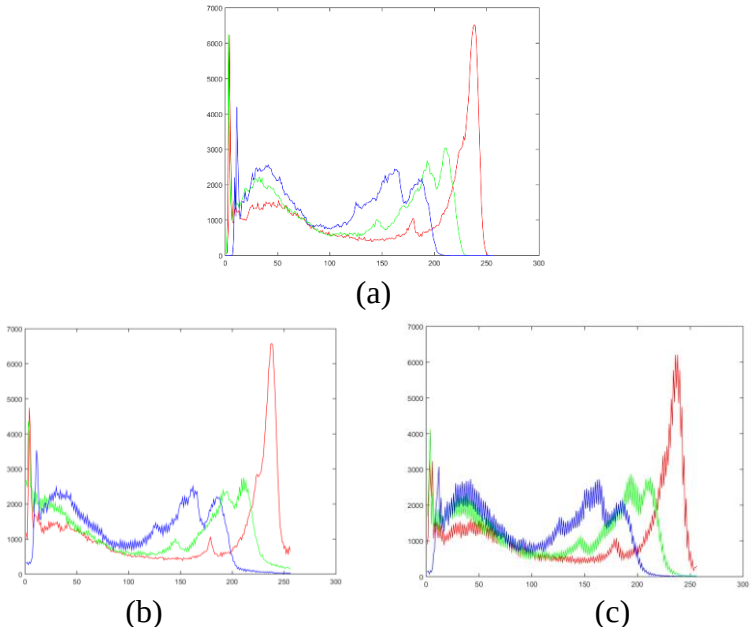
En la sección anterior se detallaron los valores promedio obtenidos del ruido en la imagen, el error RMSE y la similitud SSIM con el objetivo de evaluar la calidad de la imagen, es decir se obtuvo medidas de las degradaciones en la estego-imagen con

respecto a la imagen portadora original calculando errores percibidos y cambios en la información estructural que se refiere a la percepción visual de la imagen, lo que se puede interpretar como una evaluación frente al estego análisis visual. Pero el objetivo principal de este trabajo es evaluar el rendimiento con respecto al estego análisis estadístico que es el más común cuando se analiza una gran cantidad de imágenes, para lo cual se realizó un análisis de los histogramas obtenidos y pruebas con un programa de estegoanálisis estadístico.

4.3.1 Histogramas

A continuación se presentan los histogramas obtenidos en los 5 mejores experimentos de la imagen portadora y de la estego-imagen con el algoritmo propuesto y el algoritmo Ya-Lin – Wen-Hsiang (ver figuras 33-37). Los histogramas presentan tres curvas correspondientes a los canales RGB, la curva de color rojo corresponde al canal *R*, la de color verde al canal *G* y la de color azul al canal *B*. En el eje *x* se tiene el número de valores de pixeles que en este caso están en el rango de 0 a 255, mientras que el eje *y* corresponde a las frecuencias de cada valor de pixel.

Experimento 1

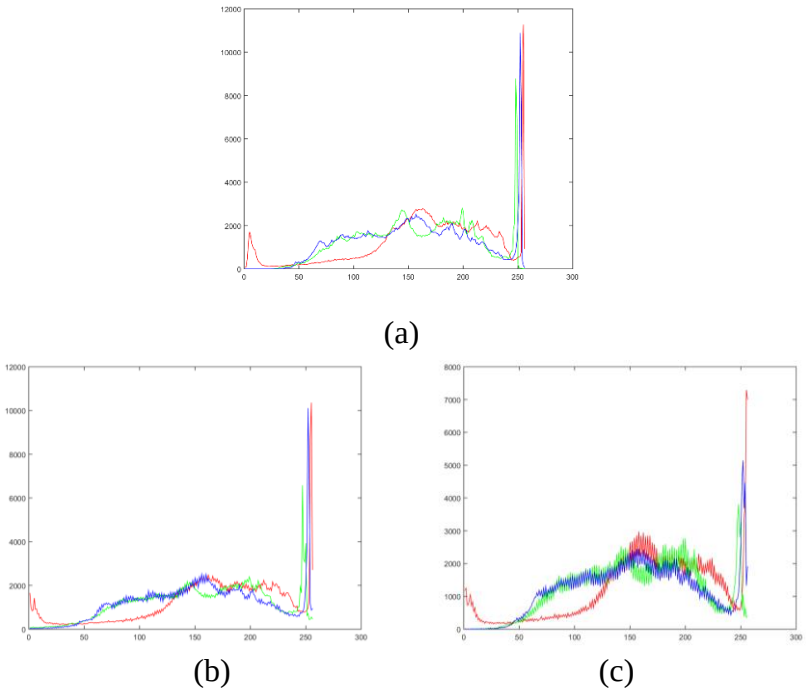


(a)

(b) (c)

Figura 33 Histogramas 1: (a) Imagen portadora (b) Estego-imagen algoritmo propuesto (c) Estego-imagen algoritmo Ya-Lin –Wen-Hsiang

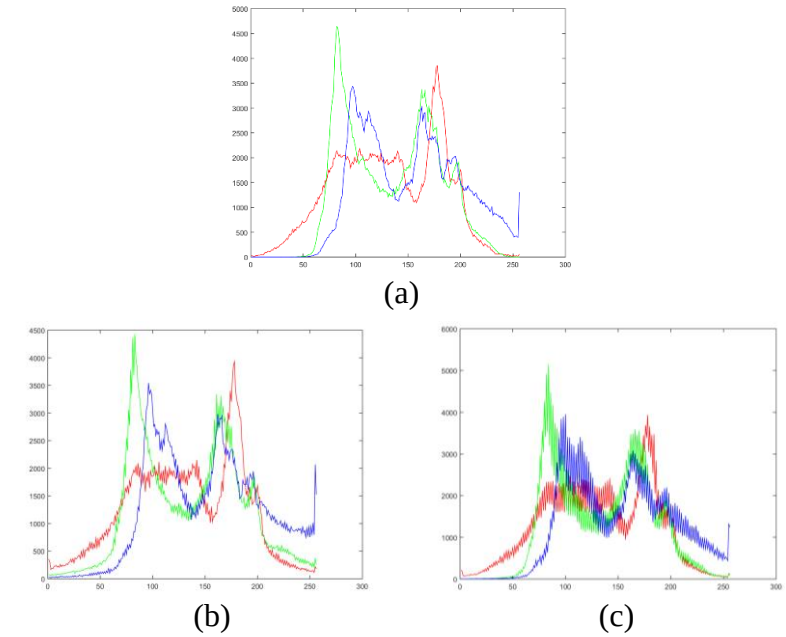
Experimento 2



(a) (b) (c)

Figura 34 Histogramas 2: (a) Imagen portadora (b) Estego-imagen algoritmo propuesto (c) Estego-imagen algoritmo Ya-Lin –Wen-Hsiang

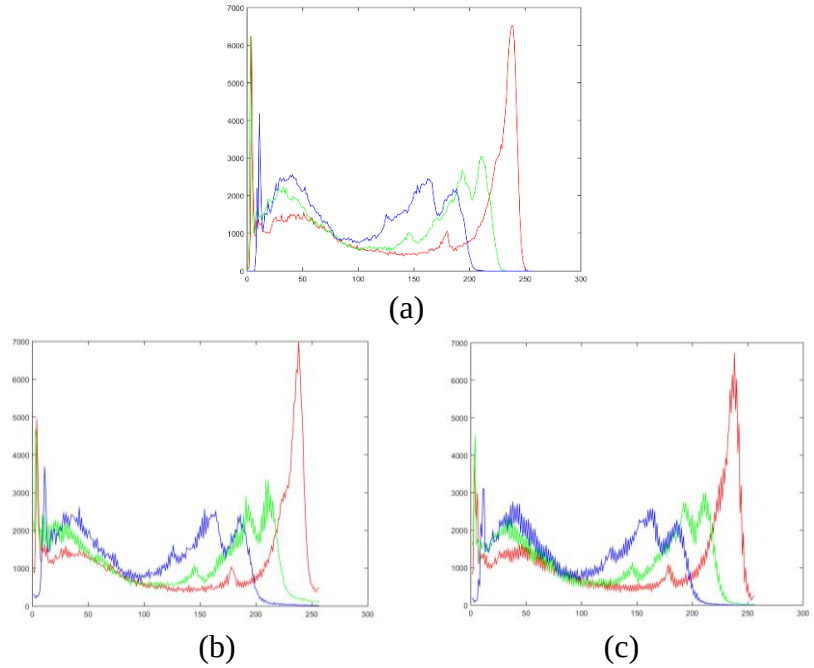
Experimento 3



(a) (b) (c)

Figura 35 Histogramas 3: (a) Imagen portadora (b) Estego-imagen algoritmo propuesto (c) Estego-imagen algoritmo Ya-Lin –Wen-Hsiang

Experimento 4

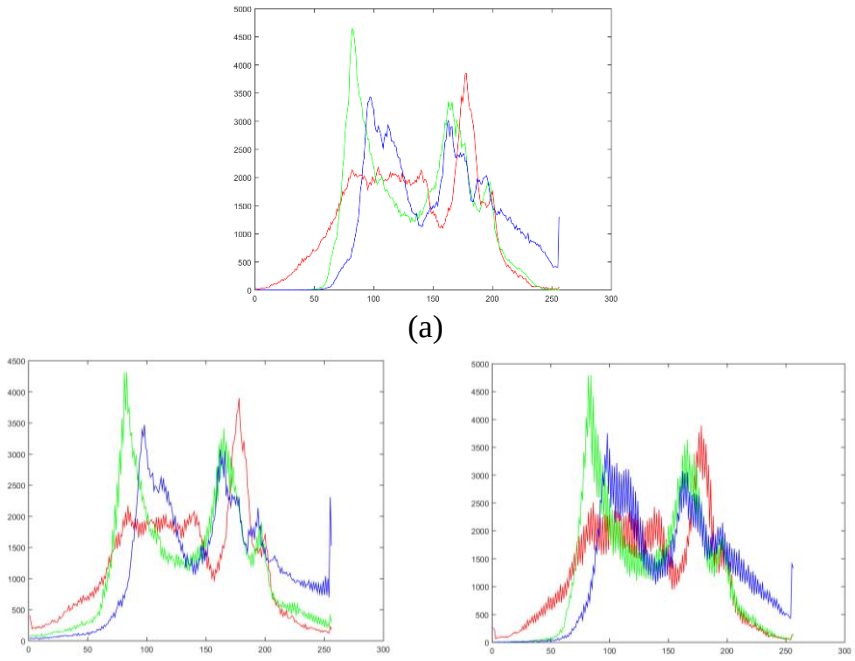


(a)

(b) (c)

Figura 36 Histogramas 4: (a) Imagen portadora (b) Estego-imagen algoritmo propuesto (c) Estego-imagen algoritmo Ya-Lin –Wen-Hsiang

Experimento 5



(a)

(b) (c)

Figura 37 Histogramas 5: (a) Imagen portadora (b) Estego-imagen algoritmo propuesto (c) Estego-imagen algoritmo Ya-Lin –Wen-Hsiang

Se evidencia claramente en los histogramas anteriormente mostrados, que el histograma obtenido en el algoritmo propuesto (ver figuras literal b) tiene mayor similitud al histograma de la imagen portadora original, ya que tiene una distribución más uniforme de valores de pixeles debido a que el algoritmo de incrustación no es aplicado a toda la imagen sino únicamente a las zonas más adecuadas, con lo cual se logró una mayor robustez ante ataques de estegoanálisis estadístico.

En el caso de los histogramas obtenidos con el algoritmo Ya-Lin –Wen-Hsiang se puede notar grandes variaciones en la distribución de pixeles ocasionando irregularidades estadísticas que tienen mayor probabilidad de ser detectadas por estegoanalizadores.

4.3.2 Pruebas con la herramienta Steg-Expose

Para este trabajo se analizó 200 estego-imágenes obtenidas: 100 imágenes con el algoritmo propuesto y 100 con el algoritmo de Ya-Lin – Wen-Hsiang. En la figura 38 se puede observar el comando insertado en el cmd (símbolo del sistema) para ejecutar la herramienta.

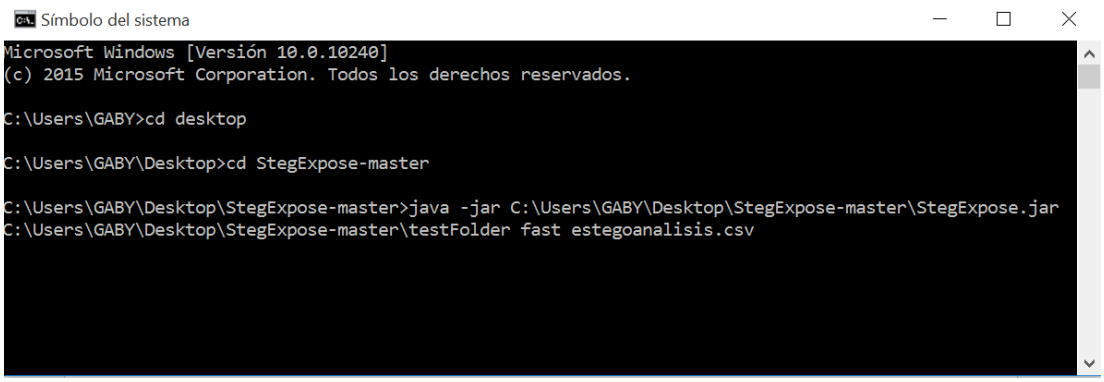


Figura 38 Comando de ejecución de StegExpose

Los resultados del estego-análisis se observan en la figura 39 donde las estego-imágenes con el algoritmo propuesto son llamadas Estego.png y las estego-imágenes con el algoritmo Ya-Lin – Wen-Hsiang son llamadas Imtransmitida.png. Las pruebas se realizaron con las mismas 100 imágenes portadoras pero utilizando los dos algoritmos diferentes.

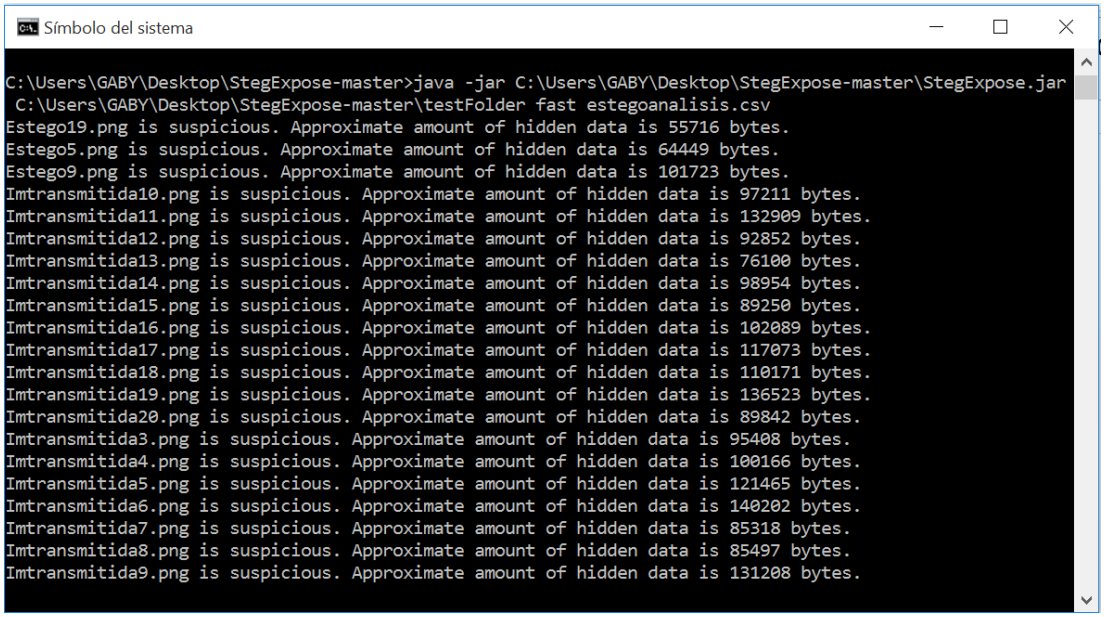


Figura 39 Resultados del estegoanálisis

Además de los resultados presentados en consola, también se genera un archivo que proporciona más información acerca del estegoanálisis, donde se muestra los valores obtenidos para cada método usado y un resultado final que fusiona todos los métodos para una mejor detección, si este resultado final está por encima del umbral por defecto de la herramienta 0,2 la imagen es considerada como estego-imagen (imagen con información oculta). Adicionalmente se cuenta con un estegoanálisis cuantitativo (determina la longitud del mensaje oculto) para las estego-imagenes. En la tabla 10 se evidencia un resumen de los resultados obtenidos después del estego-análisis con la herramienta Steg-Expose, además en el anexo 2 se observa la información que se encuentra en el archivo csv generado.

Tabla 10

Resumen de estego-análisis

	Algoritmo propuesto	Algoritmo Ya-Lin – Wen-Hsiang
Imágenes detectadas	16	97
Imágenes no detectadas	84	3

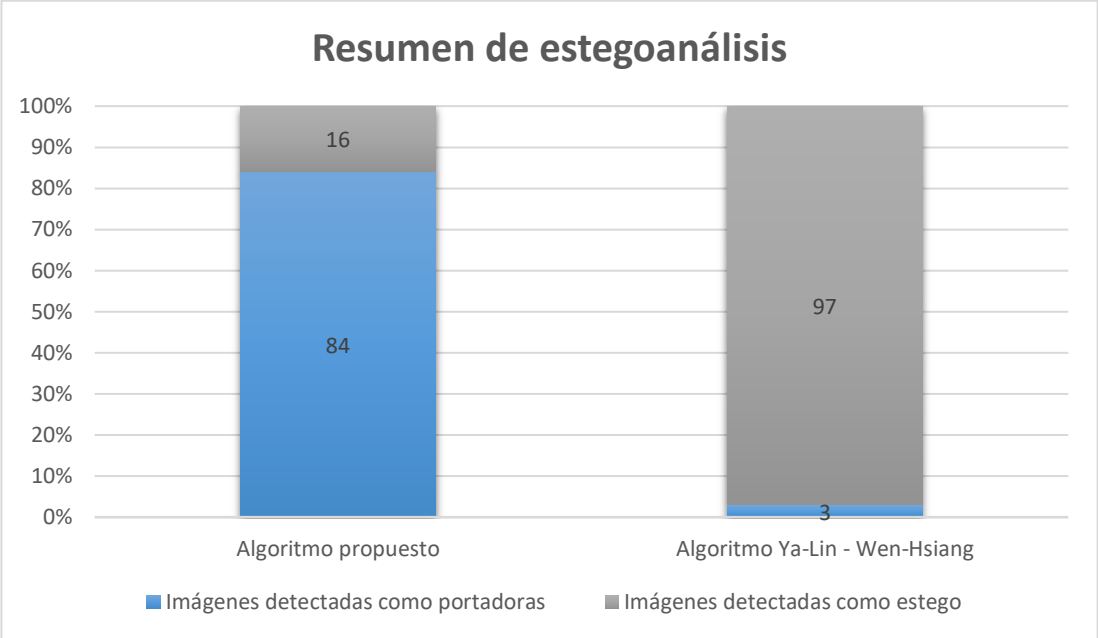


Figura 40 Resumen de estegoanálisis

La figura 40 nos demuestra que existe un 84% de las 100 estego-imágenes obtenidas con el algoritmo propuesto que fueron detectadas como imágenes portadoras (sin información oculta) mediante la herramienta StegExpose y un 16% fueron detectadas como imágenes estego (con información oculta) en el estegoanálisis. Mientras que para las mismas imágenes obtenidas con el método de Ya-Lin – Wen-Hsiang se obtuvo un 3% de imágenes detectadas como portadoras y un 97% de imágenes detectadas como estego por la herramienta. Estos resultados muestran que el método propuesto es más robusto estadísticamente ante ataques de estegoanalizadores.

A continuación se presenta en la tabla 11 los resultados de las 16 imágenes obtenidas con el algoritmo propuesto detectadas como estego-imágenes por el estego-analizador Steg-Expose:

Tabla 11

Resultados de estego-análisis para imágenes detectadas

File name	Above stego threshold ?	Secret message size in bytes (ignore for clean files)	Primary Sets	Chi Square	Sample Pairs	RS analysis	Fusion (mean & fast)
Estego81.png	true	61324	NaN	0.08137	0.31893	0.32517	0.24182
Estego83.png	true	58513	NaN	0.10759	0.31840	0.31368	0.24656
Estego58.png	true	62572	0.24000	0.07759	0.33010	0.33892	0.24665
Estego25.png	true	65020	NaN	0.07286	0.33850	0.35850	0.25662
Estego35.png	true	65790	NaN	0.08313	0.34312	0.35836	0.26154
Estego33.png	true	64056	0.23002	0.02806	0.35789	0.44981	0.26645
Estego61.png	true	71076	0.25318	0.13369	0.36055	0.35778	0.27630
Estego82.png	true	71987	NaN	0.12856	0.35198	0.35451	0.27835
Estego41.png	true	67509	NaN	0.04926	0.40258	0.41093	0.28759
Estego38.png	true	69608	NaN	0.05655	0.35371	0.45987	0.29004
Estego31.png	true	75964	NaN	0.13868	0.35069	0.42209	0.30382
Estego34.png	true	88388	NaN	0.10164	0.44919	0.45367	0.33483
Estego24.png	true	82601	NaN	0.15370	0.39786	0.46465	0.33874
Estego23.png	true	98703	NaN	0.14016	0.48693	0.49988	0.37566
Estego39.png	true	100696	NaN	0.10247	0.52905	0.51743	0.38298
Estego27.png	true	113666	NaN	0.14050	0.28785	0.79334	0.40723

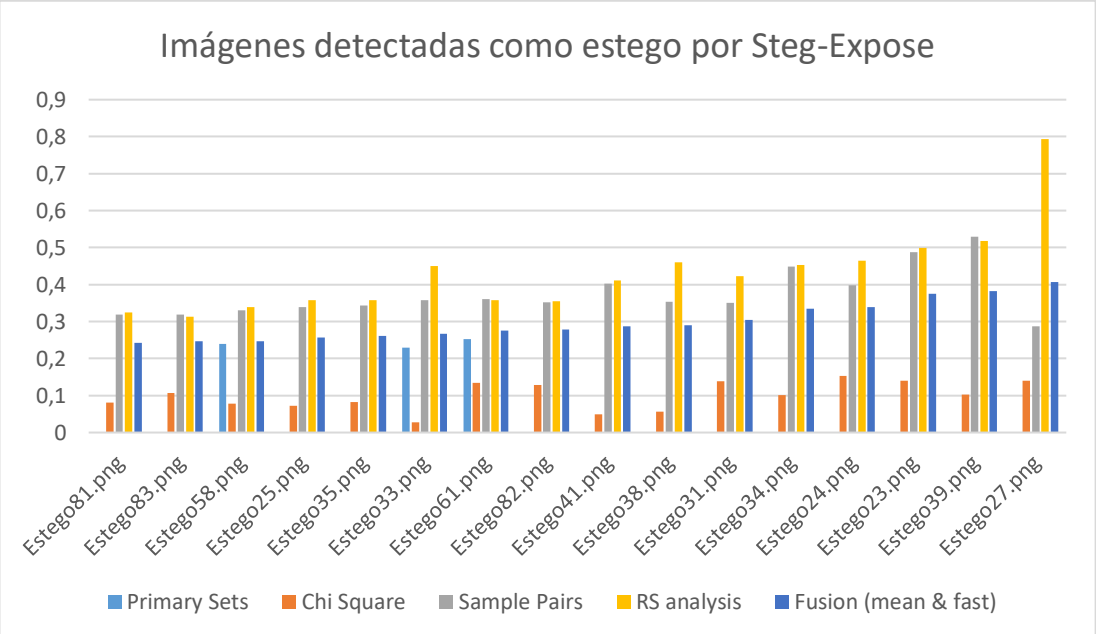
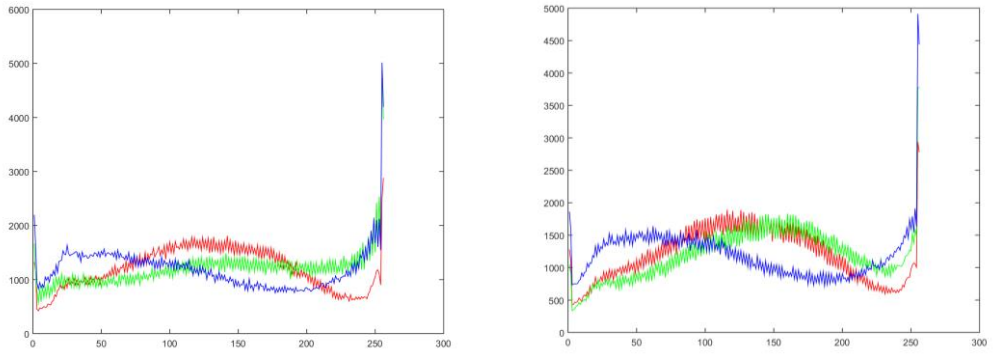


Figura 41 Detección de imágenes obtenidas por el algoritmo propuesto

En la figura 41 se puede ver los resultados obtenidos para las estego-imágenes detectadas con información oculta con cada método estegoanalítico usado por la herramienta Steg-Expose, los métodos que tuvieron más alto valor de detección fueron *Sample Pairs* y *RS analysis* que son métodos que miden las estadísticas de un grupo

de pixeles. La estego-imagen detectada con el mayor valor fue Estego27.png con 0,40723; aun así las estego-imágenes obtenidas con el algoritmo propuesto tienen valores más bajos de detección que las imágenes obtenidas con el algoritmo Ya-Lin – Wen-Hsiang que están en el rango de 0,42-0,76.

En la figura 42 se observa los histogramas de la imagen Estego27.png, donde se puede notar que no existe mucha diferencia entre la estego-imagen obtenida con el método propuesto y con el método Ya-Lin – Wen-Hsiang, esto se debe principalmente a que el mapa de bordes y texturas en este caso tiene una alta cantidad de pixeles (138126 pares de pixeles), muy cercano a los pixeles totales de la imagen, es decir los resultados obtenidos con los dos métodos son similares porque las zonas seleccionadas para incrustar información abarcan casi toda la imagen (ver figura 43).



(a) (b)

Figura 42 Histogramas: (a) Estego-imagen algoritmo propuesto (b) Estego-imagen algoritmo Ya-Lin –Wen-Hsiang

(a) (b)

Figura 43 (a) Imagen portadora (b) Mapa de bordes y texturas

El caso anterior puede aplicarse en las otras estego-imágenes detectadas cuyos mapas de bordes y texturas poseen una alta cantidad de pixeles. A continuación se presenta la tabla 12 con el número de pares de pixeles para cada mapa de bordes y texturas de las estego-imágenes detectadas, con lo que se puede corroborar que la robustez ante ataques está estrechamente relacionada con las zonas de incrustación de información.

Tabla 12

Pares de pixeles del mapa de bordes y texturas

Estego-imagen	Número de pares del mapa de bordes y texturas
Estego81.png	101462
Estego83.png	111090
Estego58.png	111419
Estego25.png	83349
Estego35.png	120653
Estego33.png	91225
Estego61.png	108792
Estego82.png	110698
Estego41.png	92094
Estego38.png	93127
Estego31.png	81673
Estego34.png	107974
Estego24.png	90205
Estego23.png	137720
Estego39.png	128556
Estego27.png	138126

4.4 Imágenes obtenidas

Después de realizar todo el proceso esteganográfico utilizando el algoritmo propuesto, a continuación se presenta las estego-imágenes obtenidas y las imágenes secretas recuperadas de los 5 mejores experimentos realizados con respecto a la indetectabilidad estadística (ver figuras 44-48):

Experimento 1

(a) (b) (c) (d)

Figura 44 Experimento 1: (a) Imagen secreta (b) Imagen portadora (c) Estego-imagen (d) Imagen secreta recuperada

Experimento 2

(a) (b) (c) (d)

Figura 45 Experimento 2: (a) Imagen secreta (b) Imagen portadora (c) Estego-imagen (d) Imagen secreta recuperada

Experimento 3

(a) (b) (c) (d)

Figura 46 Experimento 3: (a) Imagen secreta (b) Imagen portadora (c) Estego-imagen (d) Imagen secreta recuperada

Experimento 4

(a) (b) (c) (d)

Figura 47 Experimento 4: (a) Imagen secreta (b) Imagen portadora (c) Estego-imagen (d) Imagen secreta recuperada

Experimento 5

(a) (b) (c) (d)

Figura 48 Experimento 5: (a) Imagen secreta (b) Imagen portadora (c) Estego-imagen (d) Imagen secreta recuperada

CAPÍTULO 5

5.1 Conclusiones

- Se desarrolló e implementó un algoritmo esteganográfico de imágenes a color, basado en el método de Ya-Lin & Wen-Hsiang, que utiliza transformaciones reversibles del color, adicionalmente se utilizaron operadores diferenciales y filtros para seleccionar las zonas más adecuadas para ocultar información, creando un nuevo método más robusto ante ataques de estego-análisis estadístico.
- Se hicieron pruebas con el estegoanalizador StegExpose, basado en un conjunto de métodos de estegoanálisis estadístico y de técnicas de fusión que lo convierte en una herramienta más fuerte. Se logró una detección del 16% de estego-imágenes (imágenes con información oculta) con el método propuesto en este trabajo y del 97% con el método Ya-Lin & Wen-Hsiang, es decir, se comprueba la robustez del nuevo algoritmo ante ataques estadísticos.
- Las zonas más adecuadas para incrustar mensajes son los bordes y texturas de la imagen, consideradas como zonas con alto nivel de ruido, por lo que en ellas, es más difícil extraer características usadas para entrenar clasificadores dentro del estego-análisis.
- En la detección de bordes se utiliza el operador diferencial Canny, mismo que resultó ser el más eficiente, ya que es una combinación de dos etapas: un filtro para suavizar la imagen y reducir el ruido y, la unión de operadores diferenciales de detección de bordes en las tres direcciones de barrido horizontal, vertical y diagonal.
- El parámetro de entropía se usó para la detección de texturas, al tratarse de una medida de la distribución de niveles de color dentro de la imagen, en donde se toma en cuenta un pixel y sus pixeles vecinos, a diferencia del método de Lerch Hostalot & Megías, donde únicamente se toman pares de pixeles.
- La incrustación de información se llevó a cabo mediante la técnica de *LSB matching*, por su baja complejidad y su indetectabilidad visual.
- Se realizaron varias pruebas para evaluar el método propuesto en este trabajo y el método Ya-Lin – Wen-Hsiang, los resultados obtenidos con respecto al

tiempo de ejecución indican que mi método requiere más tiempo por cada experimento debido a la cantidad de iteraciones por imagen, pero esto se relaciona también con la capacidad de incrustación donde se tiene menor cantidad de pixeles utilizables dependiendo de la imagen portadora por lo tanto existe menor capacidad de incrustación y mayor número de iteraciones.

- En el caso de las mediciones de calidad de la estego-imagen obtenida y la imagen portadora en los dos métodos, se observa que el método Ya-Lin – Wen-Hsiang tiene un promedio de PSNR en la imagen mayor y un error RMSE menor, aseverando que la imagen obtenida con el método propuesto en este trabajo tiene una calidad menor. Para el caso de la imagen secreta recuperada, se obtuvieron mejores resultados de PSNR y RMSE que en las mediciones de calidad entre la estego-imagen y la imagen portadora para ambos métodos.
- La medición de calidad visual realizada, toma en cuenta la información estructural de la imagen, en la que se basa el sistema visual humano mediante el parámetro SSIM. Las estego-imágenes obtenidas con los dos métodos tienen un SSIM en el rango de 0,6 - 0,7 aproximadamente. Por el contrario, con respecto a la imagen secreta, se obtuvo un promedio de 0,905 en la medición de este parámetro.
- El desempeño de mi método propuesto en relación a la calidad objetiva y subjetiva de la imagen, es menor que en el método Ya-Lin & Wen-Hsiang. Sin embargo, estas métricas no son fundamentales para este trabajo, principalmente porque el objetivo buscado fue la robustez ante el estegoanálisis estadístico, donde se obtuvo excelentes resultados con el método propuesto. Adicionalmente, las medidas de calidad mencionadas requieren una imagen de referencia para su obtención, la que usualmente no está disponible en aplicaciones prácticas.
- Tras realizar una evaluación de la indetectabilidad estadística, se obtuvieron los histogramas de la imagen portadora y de las estego-imágenes para ambos métodos, se observó una notable diferencia entre el histograma de la imagen portadora y la estego-imagen obtenida con el método Ya-Lin – Wen-Hsiang donde se hacen presentes grandes variaciones en la distribución de los niveles de color. Por el contrario, el histograma obtenido con el método propuesto en

este trabajo presenta una distribución más uniforme, muy similar a la del histograma de la imagen portadora, resultado reflejado en las pruebas con estegoanalizadores.

- En el caso de detección de imágenes usando mi método propuesto, se aprecia que los valores de detección más altos del algoritmo detector corresponden a *RS analysis*, parámetro que mide el ruido como la diferencia de pixeles. Esta detección está relacionada estrechamente con la imagen portadora y el mapa de bordes y texturas obtenido, en estos casos se observa que el mapa tiene gran cantidad de pares de pixeles, bastante cercana a la cantidad total de la imagen, es decir, las zonas seleccionadas para incrustar información abarcan casi toda la imagen, ocasionando histogramas similares al otro método con grandes diferencias en la distribuciones de valores de pixeles, medidas como ruido por la herramienta StegExpose.

5.2 Recomendaciones

- Para un mejor desempeño del método propuesto, es necesario que las imágenes portadoras no tengan gran cantidad de texturas, con la finalidad de seleccionar únicamente las zonas adecuadas y no gran parte de la imagen, pues esto influye en la robustez ante ataques.
- Es recomendable el uso de la técnica de fusión del tipo *fast* en la herramienta StegExpose en caso de análisis de gran cantidad de imágenes, ya que es solamente 0,19% menos preciso que el mejor método detector (*RS analysis*) y es 3,16 veces más rápido.
- Es importante tomar en cuenta el umbral de binarización de bordes y texturas de la imagen, en este caso se usó un umbral de 0,7 debido a la gran cantidad de información a ocultar. Se puede incrementar este valor para definir mejor las zonas de texturas y bordes o disminuirlo en caso de requerir mayores zonas ruidosas.
- El resultado obtenido, además de ser una técnica con alto porcentaje de indetectabilidad estadística, requiere seguridad adicional para descifrar el mensaje oculto, es por ello que se recomienda la encriptación del mensaje con una clave secreta que solo el destinatario conozca.

5.3 Trabajos futuros

Como trabajo futuro se propone enfocarse en la calidad visual de la estego-imagen, ya que al implementarse el algoritmo propuesto se evidencia la presencia de distorsiones a causa de la información embebida si se tiene una imagen de referencia. Se podría mejorar esta característica mediante técnicas de esteganografía de otro tipo, que tengan mayor capacidad de incrustación y menor degradación de la imagen, como los métodos de diferencia de valores de pixeles, que poseen mayor espacio de incrustación pero también mayor complejidad y procesamiento matemático. Para esta aplicación es necesario incrustar grandes cantidades de información, sin embargo, se puede proponer la técnica expuesta en este trabajo para otras aplicaciones que requieren menos información embebida, es decir, para las que busquen transmitir información fundamental como desviaciones y medias. Por otro lado, en la parte de recepción, se propone utilizar una paleta de colores para reconstruir la imagen, con el fin de requerir menor cantidad de iteraciones y disminuir las distorsiones.

Referencias

Badr, S., Salam, G., Selim, G., & Khalil, A. (2014, September). A Review on Steganalysis Techniques: From Image. *International Journal of Computer Applications, 102*(4), pp. 11-19.

Benedikt, B. (2014, Octubre 24). *StegExpose - A Tool for Detecting LSB Steganography.* Retrieved from https://arxiv.org/abs/1410.6656

Chhikara, R., & Singh, L. (2013, Octubre). A review on digital image steganalysis techniques categorised by features extracted. *International Journal of Engineering and Innovative Technology (IJEIT)*, pp. 203-2013.

Coltuc, D., & Chassery, J.-M. (2007, April). Very fast watermarking by reversible contrast mapping. *IEEE Signal Proccess, 14*(4), 255-258.

Departamento de Electrónica y Automática. (n.d.). *Departamento de Electrónica y Automática (DEA).* Retrieved from Universidad Nacional de San Juan: http://dea.unsj.edu.ar/imagenes/recursos/capitulo5.pdf

Dumitrescu, S., Xiaolin, W., & Nasir, M. (2002). On steganalysis of random LSB embedding in continuos-tone images. *IEEE international Conference on Image Processing. 3*, pp. 641-644. IEEE.

Dumitrescu, S., Xiaolin, W., & Zhe, W. (2003). Detection of LSB steganography via sample pair analysis. *Signal Processing IEEE Transactions on 51.7*, 1995-2007.

El-Alfy, E.-S., & Al-Sadi, A. (2012). *Academia.* Retrieved from https://www.academia.edu/

Fridrich, J., Du, R., & Meng, L. (2000). Steganalysis of LSB Encoding in Color Images. *Proceedings IEEE International Conference on Multimedia and Expo.* New York City.

Fridrich, J., Miroslav, G., & Rui, D. (2001). Reliable detection of LSB steganography in color and grayscale images. *Proceedings of the 2001 workshop on Multimedia and security*, 27-30.

González Aguilera, D. (2011, Abril 01). *OpenCourseWare de la Universidad de Salamanca.* Retrieved from Procesamiento avanzado de imagenes digitales: http://ocw.usal.es

Gonzalez, R. (2003). *Digital Image Processing Using MATLAB.* New Jersey: Prentice Hall.

Hajek, M., Dezortova, M., Materka, A., & Lerski, R. (2006). *Texture Analysis for Magnetic Resonance Imaging.* Prague, Czech Republic: Med4publishing S.R.O.

Hu, X., Zhang, W., Hu, X., Yu, N., Zhao, X., & Li, F. (2013, May). Fast estimation of optimal marked-signal distribution for reversible data hiding. *IEEE Transactions on Information Forensics and Security, 8*(5), 779-788.

Huffman, D. (1952). A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the Institute of Radio Engineers*, 1098-1101.

Instituto Nacional de Tecnologías de la Comunicación. (2013). *egov.* Retrieved from http://www.egov.ufsc.br/portal/sites/default/files/esteganografia1.pdf

Lerch Hostalot, D., & Megías, D. (2014). Esteganografía en zonas ruidosas de la imagen. *RECSI*.

Lerch-Hostalot, D., & Megías, D. (2013, February). LSB matching steganalysis based on patterns of pixel differences and random embedding. *Computers & Security, 32*, 192-206.

Manveer, K., & Gagandeep, K. (2014). Review of Various Steganalysis Techniques. *International Journal of Computer Science and Information Technologies*, pp. 1744-1747.

Meghanathan, N., & Nayak, L. (2010, January). Steganalysis algorithms for detecting the hidden information in image, audio and video cover media. *International Journal of Network Security & Its Application (IJNSA), 2*(1), pp. 43-55.

Mielikainen, J. (2006, May). LSB Matching Revisited. *IEEE Signal Processing Letters, 13*(5), 285-287.

Moreira, J., Valencia, V., & Chávez, P. (2009). *DSpace en ESPOL.* Retrieved from https://www.dspace.espol.edu.ec/

National Instruments. (2013, September 11). *National Instruments*. Retrieved from http://www.ni.com/white-paper/13306/en/

Nusrati, M., A., H., & Karimi, R. (2015). Steganography in Image Segments Using Genetic Algorithm. *5th IEEE International Conference on Advanced Computing & Communication Technologies (ACCT).* Haryana.

Patidar, V., Pareek, N. K., Purohit, G., & Sud, K. K. (2011, September). A robust and secure chaotic standard map based pseudorandom permutation substitution scheme for image encryption. *Optics Communications, 284*(19), pp. 4331-4339.

Pevný, T., Bas, P., & Fridrich, J. (2009). Steganalysis by Subtractive Pixel Adjacency Matrix. *ACM Multimedia and Security Workshop*, 75-84.

Pixabay. (n.d.). (B. &. GbR, Producer) Retrieved from Commons Creative CC0: https://pixabay.com/es/

Preeti, P., & Rajeev, K. S. (2013). A Survey: Digital Image Watermarking Techniques. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 7*(6), pp. 111-124.

Prihandoko, A. C. (2013, February 27). *Anton's MathCrypt*. Retrieved from Visual Cryptography (Visual secret sharing scheme): https://antoniuscpilkom.wordpress.com

Quan, X., & Zhang, H. (n.d.). *Lossless data hiding scheme based on lsb matching.* Retrieved from http://sdiwc.net/

Riera, M., & Tacón, J. (2011). *Entropía*. Retrieved from Tlmag: http://iie.fing.edu.uy

Satwinder, S., & Varinder, K. A. (2015). State of the art Review on Steganographic Techniques. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 8*(7), pp. 161-170.

Sumathi, C. P., Santaman, P., & Umamaheswari, G. (2013, December). A Study of Various Steganographic Techniques Used for Information Hiding. *International Journal of Computer Science & Engineering Survey (IJCSES), 4*(6), pp. 9-25.

The MathWorks, Inc. (n.d.). *MathWorks*. Retrieved from Texture Analysis: https://www.mathworks.com/help/images/texture-analysis.html

Tian, J. (2002). *Digimarc.* Retrieved from Watermarking by Difference Expansion: https://www.digimarc.com/

Tiwari, A., Yadav, S. R., & Mittal, N. K. (2014, January). A Review on Different Image Steganography. *International Journal of Engineering and Innovative Technology (IJEIT), 3*(7), pp. 121-124.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. (2004, April). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing, 3*(4), 600-612.

Westfeld, A., & Pfitzmann, A. (1999). Attacks on steganographic systems. In A. Westfeld, & A. Pfitzmann, *Information Hiding* (pp. 61-76). Dresden: Springer-Verlag Berlin Heidelberg.

Ya-Lin, L., & Wen-Hsiang, T. (2014, April 4). A New Secure Image Transmission Technique via Secret-Fragment-Visible Mosaic Images by Nearly Reversible Color Transformations. *IEEE Transactions on circuits and systems for video technology, 24*(4).

Zhang, T., & Ping, X. (2003). A fast and effective steganalytic technique against JSteg-like algorithms. *ACM Symposium on Applied Computing.* Florida.