Visión Electrónica Vol. 12 No. 2 (2018) • Julio-Diciembre • p.p. 198-207 • ISSN 1909-9746 • ISSN-E 2248-4728 • Bogotá (Colombia)

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Visión Electrónica
Más que un estado sólido
https://revistas.udistrital.edu.co/index.php/visele

Visión Electrónica

VISIÓN INVESTIGADORA

Marcado de audio mejorado en el dominio wavelet

Improved audio watermarking in the wavelet domain

Erika Pinto[1], Dora Maria Ballesteros[2], Diego Renza[3]

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Enviado: 17/07/2018
Recibido: 09/08/2018
Aceptado: 25/09/2018

Palabras clave:
Marcado de agua
Procesamiento de imágenes
Transformada wavelet discreta

Keywords:
Watermarking
Image processing
Discrete wavelet transform

RESUMEN

El marcado de agua en imágenes ha sido ampliamente usado con fines de protección de derechos de autor. Uno de los métodos utilizados para insertar la marca corresponde a QIM, y típicamente se realiza directamente sobre la imagen. La principal desventaja consiste en que el nivel de seguridad es bajo, y la marca se puede eliminar fácilmente si se conocen el valor del paso de cuantización utilizado. Con el propósito de aumentar la seguridad, proponemos una mejora al método QIM tradicional, aplicando la inserción en el dominio wavelet de la imagen, y solamente en algunos coeficientes de alta amplitud seleccionados de forma aleatoria por el sistema. De acuerdo a las pruebas realizadas, se pudo verificar que, si la imagen marcada es atacada de forma intencional, la marca recuperada tiene contenido legible. Adicionalmente, aumentamos la seguridad, dado que la clave de recuperación no solo depende del paso de cuantización, sino de la selección de los coeficientes wavelet.

ABSTRACT

Image watermarking has been widely used for copyright protection purposes. One of the methods used to insert the marker corresponds to QIM, and it is typically done directly on the image. The main disadvantage is that the level of security is low, and the mark can be easily removed if the value of the quantization step used is known. In order to increase security, we propose an improvement to the traditional QIM method, applying the insertion in the wavelet domain of the image, and only in some high amplitude coefficients randomly selected by the system. According to the tests carried out, it was verified that, if the marked image is attacked intentionally, the recovered mark has a readable content. Additionally, we have increased security, considering the recovery key depends not only on the quantization step, but also on the selection of the wavelet coefficients.

[1] Ing. en Telecomunicaciones de la Universidad Militar Nueva Granada, Bogotá D.C, Colombia. Egresada. Correo electrónico: u1400954@unimilitar.edu.co, ORCID: https://orcid.org/0000-0003-3621-7554

[2] Ing. Electrónica de la Universidad Industrial de Santander, MSc en Ingeniería Electrónica de la Universidad de los Andes, PhD. en Ingeniería Electrónica de la Universitat Politecnica de Catalunya. Profesor Titular, Universidad Militar Nueva Granada, Bogotá, Colombia. Correo electrónico: dora.ballesteros@unimilitar.edu.co, ORCID: https://orcid.org/0000-0003-3864-818X

[3] Ing. Electrónico de la Universidad Sur Colombiana, MSc en Ingeniería de Telecomunicaciones de la Universidad Nacional de Colombia, PhD. en en Computación Avanzada para Ciencias e Ingeniería de la Universidad Politécnica de Madrid, Profesor Asociado, Universidad Militar Nueva Granada, Bogotá, Colombia. Correo electrónico: diego.renza@unimilitar.edu.co, ORCID: https://orcid.org/0000-0001-8073-3594

Citar este artículo como: E. Pinto, D. M. Ballesteros y D. Renza, "Marcado de audio mejorado en el dominio wavelet". Visión Electrónica, algo más que un estado sólido, Vol. 12, No. 2, 198-207, julio-diciembre 2018.. https://doi.org/10.14483/22484728.14071

1. Introducción

Actualmente el contenido multimedia es almacenado en forma digital, debido a que es más sencillo compartir información de este tipo en internet para poder llegar a muchos usuarios a nivel mundial. Uno de los principales problemas de compartir información multimedia es la exposición del contenido a ser duplicado, modificado o distribuido de forma ilegal, quebrantando las políticas de los derechos de autor. Dentro de los sectores más afectados por la copia ilegal de contenido se encuentra la industria musical, por ello se hace necesario brindar protección a los derechos de autor y el uso autorizado de la información, buscando mecanismos en donde se pueda identificar el propietario del contenido. Uno de los mecanismos más utilizados en la protección de contenido musical es el marcado de agua, ya que permite "firmar" el audio sin que se deteriore su calidad [1,2].

En el marcado de agua, se deben satisfacer principalmente tres condiciones: imperceptibilidad, robustez y capacidad de ocultamiento (HC: *hiding capacity*). La imperceptibilidad es un factor importante, ya que la marca incrustada en el audio debe pasar inadvertida y no debe generar deterioro en la calidad del audio que se pretende proteger. En el caso de la robustez, se busca que la marca permanezca dentro del audio aún si se realiza manipulación sobre el mismo, dentro de un ataque pasivo [3]. En la capacidad de ocultamiento, se cuantifica la cantidad de bits que se pueden ocultar por cada segundo de duración del audio. Un método con baja capacidad de ocultamiento soportará marcas de tamaño muy pequeño. Adicional a los parámetros anteriores, se requiere que la marca no sea fácilmente identificable y removible por un tercero no autorizado, dentro de un ataque activo [4].

Se han realizado diferentes trabajos en cuanto al marcado de agua digital, con diferentes técnicas y métodos de marcado, cada una de ellas con sus ventajas y desventajas. En [3] se pueden encontrar diferentes métodos de ocultamiento comunes en el marcado de agua, así como la clasificación de los mismos dependiendo de si son realizados en el dominio del tiempo o en el dominio de la frecuencia. Los métodos basados en el dominio del tiempo permiten insertar la marca directamente en el archivo de audio, mientras que los métodos en el dominio de la frecuencia, insertan la marca modificando los coeficientes de frecuencia. Dentro de las ventajas de los métodos en frecuencia se tiene mayor robustez frente a ataques de manipulación del audio marcado, aunque su implementación puede llegar a ser más complicada [5]. Algunos de los trabajos reportados en la literatura se enfocan en una de las condiciones de diseño, por ejemplo, en [6] se utiliza la DWT y un cifrado adicional a la marca para aumentar la seguridad del proceso de marcado. En [7], se prevalece la capacidad de ocultamiento, a expensas de no cumplir plenamente con la condición de imperceptibilidad. En [8], la robustez frente a ataques pasivos, como compresión MP3 es el principal criterio de diseño. Adicionalmente, en algunos esquemas de ocultamiento de datos se propone ensanchar la tasa de datos de la señal, con el fin de alcanzar mayor robustez, por ejemplo, con el uso de códigos OVSF (Orthogonal Variable Spreading Factor) [9], o mediante un proceso de binarización basado en la conjetura de Collatz [10]. En general, los métodos propuestos cumplen con las condiciones de imperceptibilidad y robustez, pero tienen una capacidad de ocultamiento muy baja. Por ejemplo, de 45.9 bps (DWT) [4], 43 bps (dominio del tiempo) [11], 25.6 bps (DWT) [12], 27.1 bps (FFT) [13], 4.2 bps (dominio del tiempo) [14] y 2 bps (dominio del tiempo) [15].

Un método que satisface alta imperceptibilidad, alta capacidad de ocultamiento y moderada resistencia frente a algunos tipos de ataque es el denominado QIM (*Quantization Index Modulation*), sin embargo, tiene baja seguridad frente a la detección de la marca, ya que si un tercero conoce el método y el valor de cuantización (delta), puede detectar y eliminar la marca, desprotegiendo al archivo de los derechos de autor. De tal forma, en este trabajo de investigación se presenta una mejora al método QIM enfocada en aumentar en dos niveles el nivel de seguridad del método, manteniendo las fortalezas de su versión original.

El artículo está organizado de la siguiente forma: en la sección 2, se explican brevemente conceptos para la realización del método propuesto, en la sección 3 se presenta el método propuesto de marcado de agua con imágenes binarias, en la sección 4 se presentan los resultados y su análisis; y finalmente el método se concluye en la sección 5.

2. Conceptos básicos

2.1. Transformada Wavelet Discreta (DWT)

En el procesamiento de audio, una de las transformadas más utilizadas corresponden a la

Transformada Wavelet Discreta, la cual consiste en la aplicación de filtros de cuadratura, permitiendo obtener los coeficientes de aproximación de salida del filtro pasa-bajos y los coeficientes de detalle de salida del filtro pasa-altos. Este proceso se conoce como descomposición. De los dos grupos de coeficientes, el que contiene la mayor parte de la energía de la señal original, corresponde a los coeficientes de aproximación. A partir de los coeficientes de aproximación y de detalle, es posible recuperar la imagen original, aplicando un proceso conocido como la re-construcción. Sin embargo, es necesario que se conserven los mismos parámetros utilizados en la fase de descomposición [16,17].

2.2.Modulación del índice de cuantización

Dentro de los métodos utilizados en la literatura para el ocultamiento de información dentro de una señal huésped, se encuentra QIM (Quatization Index Modulation), el cual es útil tanto para fines de esteganografía, como de marcado de agua. Este método consiste en modificar la amplitud de la muestra de la señal de audio, por medio de un valor específico de paso de cuantización (Δ) y dependiendo si lo que se desea ocultar es un bit de valor 1 o 0 [18].

La ecuación (1) se utiliza para la inserción binaria de información, mientras que la ecuación (2) para la identificación o extracción del bit que se ocultó.

$$S = \begin{cases} \Delta\left\lfloor \frac{H}{\Delta} \right\rfloor & w = 0 \\ \Delta\left\lfloor \frac{H}{\Delta} \right\rfloor + \frac{\Delta}{2} & w = 1 \end{cases} \quad (1)$$

$$w_r = \begin{cases} 1 & \frac{\Delta}{4} < \left| S - \Delta\left\lfloor \frac{S}{\Delta} \right\rfloor \right| < \frac{3\Delta}{4} \\ 0 & en\ otro\ caso \end{cases} \quad (2)$$

De las ecuaciones anteriores se tiene que el parámetro *S* corresponde al audio marcado, *H* es el audio original (o huésped), *w* es la marca binaria y *wr* es la marca recuperada. Si el audio marcado no sufre ningún ataque pasivo (ej. filtrado, compresión, adición de ruido), se espera que la marca recuperada sea igual a la marca incrustada.

2.3.Imperceptibilidad

Una condición necesaria en el marcado de audio consiste en que la marca debe ser invisible, es decir, que dicha marca no debe generar ruidos o artefactos que degraden la calidad del audio. Si el audio es una canción a la que se quiere proteger su copyright, es indispensable que la marca sea "imperceptible".

Uno de los parámetros matemáticos que permiten medir la imperceptibilidad corresponde a SPCC (*Squared Pearson Correlation Coefficient*), el cual es un indicador de similitud de dos secuencias de datos, en este caso el audio original y el audio marcados. Idealmente, el valor de SPCC deberá ser 1 o muy cercano a este valor, se queremos que la marca sea imperceptible.

Matemáticamente, el valor de SPCC entre el audio original (*H*) y la señal marcada (*S*), se calcula por medio de (3).

$$\rho(H,S) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{H_i - \mu_H}{\sigma_H}\right)\left(\frac{S_i - \mu_S}{\sigma_S}\right) \quad (3)$$

En donde ρ es el valor de SPCC entre los dos audios; μ_H y σ_H son la media y la desviación estándar de H, respectivamente; μ_S y σ_S son la media y la desviación estándar de S; y N corresponde al total de muestras de cada señal de audio.

3. Método de marcado de agua propuesto

Nuestra propuesta de marcado de agua en señales de audio con fines de protección de derechos de autor incluye dos módulos: el primero, correspondiente a la inserción de la marca, la cual es una imagen binaria; el segundo, correspondiente a la extracción de la marca a partir del audio marcado.

3.1.Módulo en el transmisor: inserción de la marca

En la Figura 1 se presenta el diagrama de bloques correspondiente al módulo de inserción de la marca.

El primer paso consiste en la lectura tanto del audio a marcar, como de la marca a utilizar. Es necesario calcular la cantidad de bits que se deben insertar en el audio, para esto, se multiplica el total de filas de la marca por el total de columnas. El resultado se asigna a la variable P.

Posteriormente, al audio se le aplica la descomposición wavelet, para un solo nivel de descomposición. Como resultado, se obtienen los coeficientes de aproximación y los coeficientes de detalle. En esta propuesta, solamente la inserción se aplica en los coeficientes de aproximación de la señal, por lo que, los coeficientes de detalle permanecen intactos.

A partir de los coeficientes de aproximación, se aplica un proceso de ordenamiento por amplitud de forma descendente. Como resultado, se obtiene un vector con los coeficientes de aproximación ordenados, y otro vector, con las posiciones originales de los coeficientes.

Del total de coeficientes de aproximación, se selecciona un subconjunto de coeficientes, correspondiente a diez veces el valor de P. Con este subconjunto de coeficientes, de forma pseudo-aleatoria, se seleccionan P coeficientes de aproximación.

De forma paralela al procesamiento que sufre la señal de audio, a la marca también se le aplica un procesamiento. Esta imagen se convierte en un vector de valores binarios, de longitud P. De tal forma que, a los coeficientes de aproximación seleccionados, se les inserta los bits provenientes de la marca, aplicando la ecuación (1).

Como último paso, los coeficientes de aproximación cuantizados retornan a su posición inicial, y junto con los coeficientes de detalle, se reconstruye obteniendo el audio marcado.

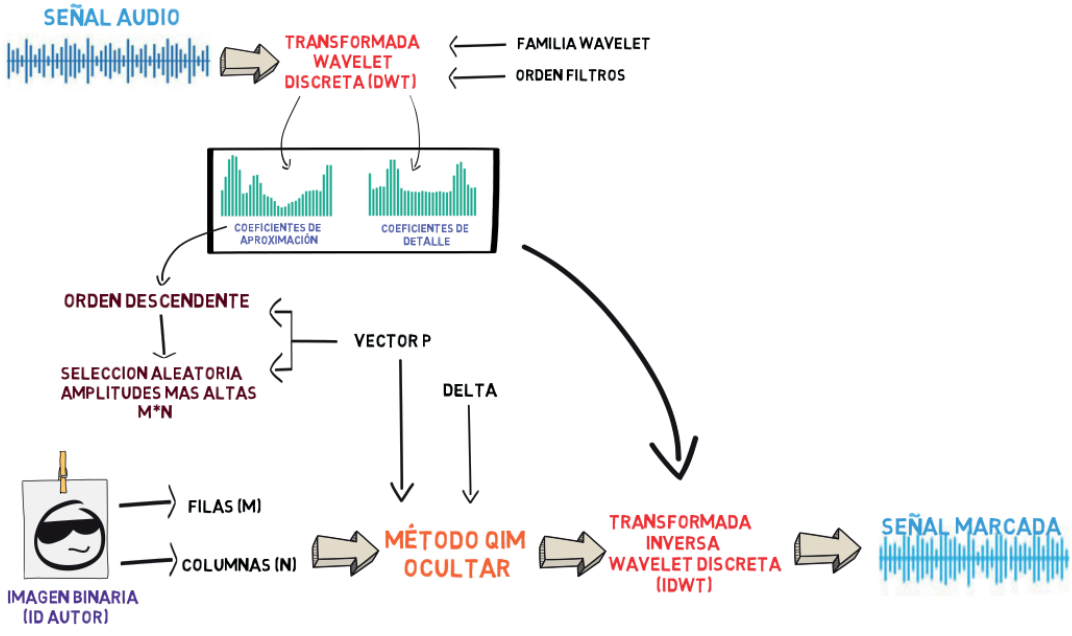
De forma ilustrativa, se presenta un ejemplo. Suponga que el audio que se quiere proteger posee una longitud de 60 segundos, con un total de muestras de 480 K. Por otro lado, se tiene una marca correspondiente al logo de la empresa, con un. tamaño de 100 * 100 pixeles. Entonces, dado que la marca es binaria, el total de bits que se necesitan incrustar en el audio, corresponde a 10 K (es decir, P=10 K). Si el audio se descompone utilizando la base haar, se tendrán 240 K coeficientes de aproximación y 240 K coeficientes de detalle. De los coeficientes de aproximación, se pre-seleccionan los 10 * P coeficientes de mayor amplitud, es decir, 100 K coeficientes. De este subconjunto, de forma aleatoria se escogen 10 K coeficientes, en los cuales a cada uno de ellos se les insertará un bit, proveniente de la marca binaria. Los coeficientes modificados (cuantizados), retornan a su posición original, y en conjunto con los coeficientes de detalle se reconstruye el audio marcado.

3.2.Módulo de recepción: extracción de la marca

El transmisor envía al receptor la siguiente información: audio marcado y clave. La clave está conformada por paso de cuantización (delta, Δ), la base wavelet, la semilla para seleccionar P valores de los coeficientes de aproximación, el total de filas (m) y de columnas (n) de la marca.

En la Figura 2, se presenta el diagrama para la extracción de la marca. Inicialmente, en el receptor se leer el archivo marcado, y se descompone el audio utilizando la misma base wavelet del módulo transmisor. Una vez obtenidos los coeficientes de

Figura 1. Proceso de inserción de la marca de agua



Fuente: elaboración propia.

aproximación del audio marcado, se seleccionan P valores de acuerdo a la información contenida en la semilla. Para cada uno de estos P coeficientes de aproximación, se aplica la extracción de bits del método QIM, correspondiente a la ecuación (2). Por cada coeficiente, se obtiene un bit, es decir, un pixel binario de la marca. Una vez recuperados todos los bits, se procede a ordenarlos en m filas y n columnas, de acuerdo a la información contenida en la clave.

4. Pruebas y resultados

En esta sección, se presentan los resultados de marcado de agua de audio (música) con una imagen binaria. Se evalúa la imperceptibilidad del audio marcado y la robustez de la marca frente a tres tipos de ataques: ruido aditivo, compresión MP4, filtrado.

Las características del protocolo de pruebas son (Tabla 1).

La imagen binaria utilizada como marca tiene un tamaño de (150 x 150) pixeles, para un total de 22550 píxeles (Figura 3).

Figura 3. Marca de agua binaria

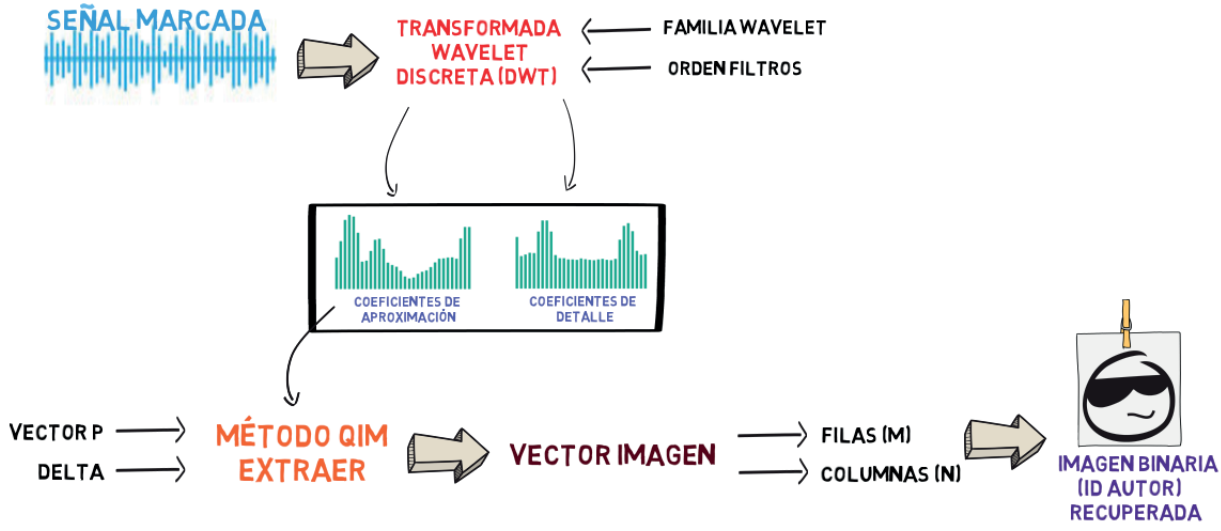
Fuente: elaboración propia.

4.1. Imperceptibilidad

Utilizando los 45 audios de prueba junto con los 3 valores de cuantización, se obtienen 135 audios marcados. Cada audio marcado se compara con su audio original, por medio del parámetro SPCC, descrito en la sección 2.3. Como resultado se tiene que, en todos los casos, el valor de SPCC es mayor a 0.99

De forma ilustrativa, se presenta la Figura 4, en la cual la señal superior corresponde al audio original, y la señal inferior al audio marcado. Se evidencia la alta similitud entre las dos señales.

Figura 2. Proceso de extracción de la marca de agua



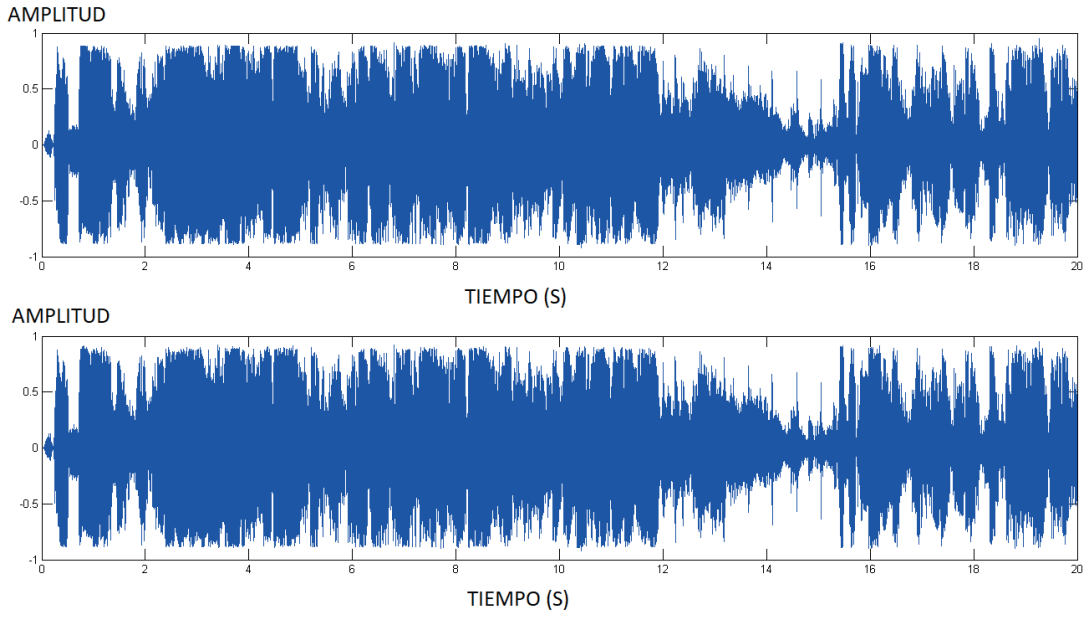
Fuente: elaboración propia.

Tabla 1. Protocolo de pruebas

Tiempo del audio	fs del audio	Bit rate del audio	Cantidad de audios	Valores de cuantización
20 segundos	44100 Hz	705600 bits	45	3

Fuente: elaboración propia.

Figura 4. Ejemplo de señal de audio original y marcada



Fuente: elaboración propia.

Por lo anterior, nuestro método cumple exitosamente el criterio de imperceptibilidad.

4.2.Robustez

Cuando un tercero quiere vulnerar el copyright de una imagen, puede atacarla de forma intencionada buscando eliminar la posible marca que contenga. El objetivo de cualquier método de marcado con fines de protección de derechos de autor consiste en ser robusto ante estos ataques intencionados, de tal forma que, se puede recuperar la imagen y demostrar su copyright.

Con el propósito de medir que tan robusto es nuestro método, hemos seleccionado tres tipos de ataque, que son fácilmente implementados por un usuario malintencionado.

Las características de los ataques son:

- Ruido aditivo: este ataque consiste en la adición de ruido en toda la señal de voz, con el propósito de "cubrir" la posible marca que contenga. En esta prueba, se utilizó un ruido blanco gaussiano de amplitud correspondiente a la cuarta parte del valor del paso de cuantización.
- Filtrado: en este ataque, las componentes altas del audio se eliminan, utilizando para ello un filtro Butterworth pasa-bajo, con fc= 18 KHz y de 2do orden.
- Compresión con pérdida de información: este ataque consiste en la transformación del formato del audio, pasando de un formato WAV a uno comprimido con pérdida de información, específicamente MP4. Es importante mencionar que el audio original tiene un *bit rate* de 705.6 kbps, dado que la fs=44.1 K Hz y la resolución es de 16 bits. El audio marcado comprimido tiene un *bit rate* de 192 kbps.

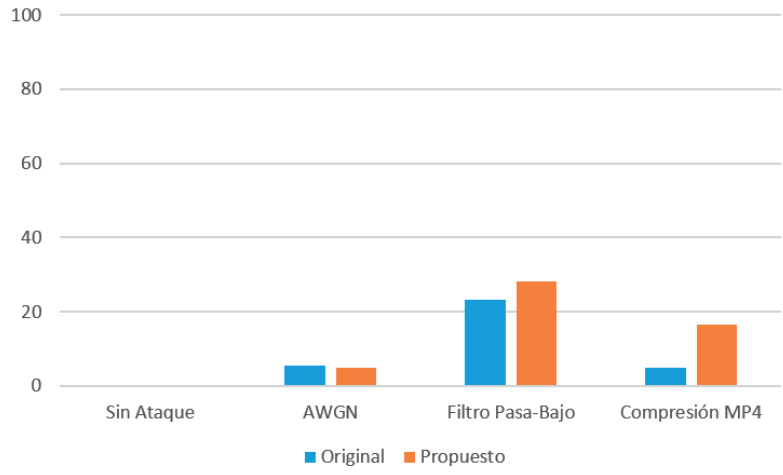
Para evaluar la similitud entre la marca original y la marca recuperada después del ataque se tuvo en cuenta el parámetro HD (Hamming Distance), el cual consiste en la comparación bit a bit de los valores de la imagen binaria original (marca) con los valores de la marca recuperada. Cada vez que difieren, aumenta el valor de conteo. Una vez se han comparado todos los bits, se divide el valor de conteo entre el total de pixeles de la marca, y se expresa el resultado en porcentaje. Por ejemplo, suponga que la marca es de tamaño 150 * 150, para un total de 22,5 K pixeles (o valores binarios); si 1000 pixeles difieren entre la marca original y la marca recuperada, entonces el valor de HD es de 4.4%. Nótese que un alto HD indica que la marca recuperada no es similar a la marca original; mientras que, un bajo HD indica que existe una alta similitud y que el audio marcado fue robusto ante el ataque.

En la Figura 5 se presentan los resultados en términos de HD para un audio específico, el cual fue

atacado con ruido aditivo, filtrado y compresión MP4. De acuerdo a los resultados presentados, cuando no se ataca el audio marcado, en ambos casos, con el método QIM original y con el propuesto, el valor de HD es de 0. En el caso de adición de ruido, con el método propuesto se obtiene una pequeña mejora en términos de HD. Por otro lado, con los ataques de filtrado y de compresión, el valor de HD es mejor en el método original. Sin embargo, en todos los casos, el valor de HD se mantuvo por debajo del 30%, es decir, que al menos el 70% de los pixeles de la marca son exitosamente recuperados.

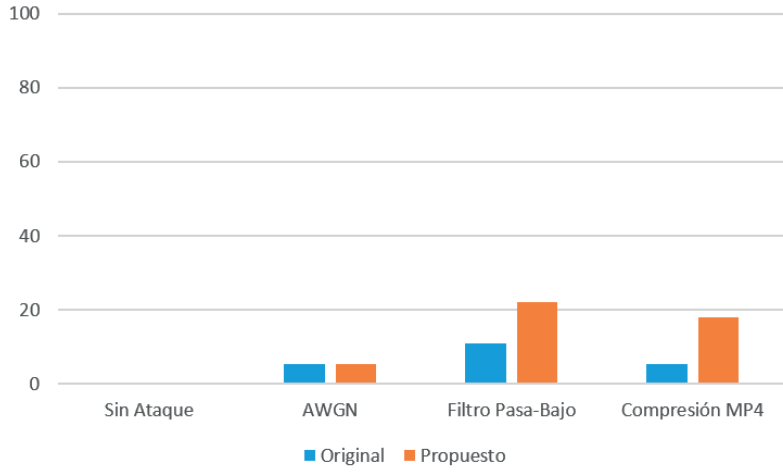
A cada uno de los audios seleccionados para esta prueba, se les aplicaron los tres tipos de ataque, obteniendo 45 audios por ataque (135 audios en total). Para cada caso, se detectó la marca y se comparó con la original, por medio del parámetro HD. El consolidado de la prueba se presenta en la Figura 6, con un comportamiento similar al de la Figura 5. Es decir, cuando no existe ataque la marca recuperada es exactamente igual a la marca original; el ataque de ruido afecta de "igual" manera a los dos métodos; en los ataques de filtro pasa-bajo y compresión MP4, afecta de mayor manera al método propuesto.

Figura 5. Resultados HD de un caso específico: método original vs propuesto



Fuente: elaboración propia.

Figura 6. Resultados HD de 135 audios: método original vs propuesto



Fuente: elaboración propia.

4.3. Capacidad de ocultamiento (HC: Hiding Capacity)

La capacidad de ocultamiento mide la cantidad de bits que se ocultan por cada segundo del audio. A mayor capacidad de ocultamiento, se podrá ocultar una marca de mayor tamaño o varias marcas pequeñas, para generar redundancia de información y proteger en mayor medida al audio.

Debido a que el método propuesto trabaja con los coeficientes de aproximación y al efecto de sub-muestreo por la DWT, la máxima capacidad de ocultamiento de este método es igual a la mitad de la frecuencia de muestreo del audio, si se trabaja con la familia Haar, o en forma general, a la cantidad de coeficientes de aproximación por segundo. Por ejemplo, si el audio tiene una fs de 44.1 kHz y se utiliza la familia Haar, el valor de HCmax es de 22.05 kbps. En el caso de QIM tradicional, como se aplica directamente sobre el audio en el dominio del tiempo, el valor máximo de HC es igual a la frecuencia de muestreo de la señal.

En la Tabla 2 se compara el valor de HC del método propuesto con otros trabajos reportados en la literatura. La unidad de medida es bps (bits por segundo). Se puede inferir que el método propuesto tiene un valor de HC sensiblemente mayor al de otros métodos y sólo la mitad del método QIM original. Este alto valor de HC le permite proteger en mejor medida el audio.

5. Análisis en términos de la seguridad

Este apartado se constituye como el de valor agregado en relación al método QIM tradicional. Se analizan los siguientes aspectos: tamaño del conjunto de claves, sensibilidad a la clave.

5.1. Tamaño del conjunto de claves.

A medida que aumenta la cantidad de claves para realizar el proceso de inserción de la marca, aumenta la dificultad para un atacante de eliminar la marca. En el método QIM tradicional, e conjunto de claves posibles se enfocaba únicamente a la cantidad de valores diferentes que podía tomar el paso se cuantización. Supongamos por ejemplo que el audio tiene una resolución de 16 bits, entonces la cantidad de valores (amplitudes) diferentes que puede tener una muestra es de 2^{16}. A partir de este valor, determinamos la máxima cantidad de valores del paso de cuantización, como la cuarta parte, obteniendo hasta 2^{14} valores. Nótese que es necesario dividir en cuatro, para que el valor de la muestra cuantizada no supere el máximo valor posible que puede tener la muestra original, al aplicar la ecuación (1) presentada en este artículo.

Por otro lado, en nuestra propuesta, los parámetros que pertenecen a la clave, son: paso de cuantización, semilla para selección de los coeficientes de aproximación en los cuales se oculta la marca. Teniendo en cuenta que, se seleccionan solamente P coeficientes de aproximación para ocultar los P bits que hacen parte de la marca y que la selección se realiza de forma aleatoria, el total de opciones diferentes de selección se calcula a partir de una fórmula de variaciones sin repetición:

$$V_N^P = \frac{N!}{(N-P)!} \tag{4}$$

Con N como el total de coeficientes de aproximación del audio a marcar y P el total de pixeles de la marca binaria. Retomando el ejemplo presentado en la

Tabla 2. Cuadro comparativo en términos de HC

Referencia	Método	Capacidad de Ocultamiento
Propuesto	QIM Mejorado	Hasta 0.5 bits/ muestra Si *fs*=44.1 kHz, entonces HC_{max}~22.05 kbps Valor recomendado: 1.1 kbps
QIM original		Hasta 1 bit/muestra Si *fs*=44.1 kHz, entonces HC_{max}=44.1 kbps
Bhat K, Sengupta and Das [4]	Adaptative DWT SVD	45,9 bps
Cvejic and Seppanen [10]	Spread Spectrum	27,1 bps
Li et al. [11]	Content-based	4,2 bps
Xiang et al. [12]	Histogram-based	2 bps

Fuente: elaboración propia.

sección 3.1., se tendría que P= 100 K, N= 240 K, entonces, la cantidad de variaciones, son:

$$V_N^P = \frac{240000!}{(240000-100000)!} = \frac{240000!}{140000!} \quad (5)$$

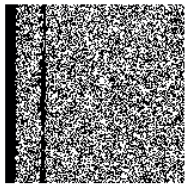
Cuyo resultado supera con creces el valor obtenido de posibles claves del método QIM tradicional. Por lo tanto, la seguridad del método propuesto es sustancialmente mayor.

5.2. Sensibilidad a la clave

En adición al análisis anterior, es importante establecer si es posible recuperar la marca con una clave que no sea correcta. Es decir, definir qué tan sensible es el sistema a la clave. A medida que la sensibilidad sea mayor, se puede tener mayor seguridad en la protección exitosa del copyright, dado que, si la clave no es fácilmente identificable por un tercero no autorizado, tampoco será la adulteración de la misma.

A manera de ejemplo, se presenta la marca recuperada por un usuario no autorizado, quien conoce el método utilizado para la inserción de la marca, el paso de cuantización y el tamaño de la marca, pero desconoce cuáles coeficientes de aproximación fueron seleccionados para insertar la información binaria. Si este usuario, selecciona los P coeficientes de aproximación de mayor amplitud, la información que recupera, es la que se presenta en la Figura 6. Al comparar esta imagen con la presentada en la Figura 3, se evidencia el alto nivel de sensibilidad a la clave, factor deseable en un sistema de marcado con fines de protección de derechos de autor.

Figura 7. Marca extraída con clave errónea



Fuente: elaboración propia.

6. Conclusiones

En este artículo se presentó un método de ocultamiento de marcas binarias en audio, enfocado a la protección de derechos de autor. Como principal característica del método propuesto consiste en la mejora de la seguridad en términos de: tamaño del conjunto de claves y sensibilidad a la marca.

En relación al tamaño del conjunto de claves, se puedo establecer teóricamente que, con el método propuesto, se aumenta de forma significativa la seguridad de la marca, de tal forma que, un tercero no autorizado tendría que gastar mucho más tiempo intentado adivinar la clave correcta, utilizando, por ejemplo, un ataque de fuerza bruta. Esta mejora obedece a la inclusión de un proceso de aleatoriedad en la selección de los coeficientes de aproximación en los cuales se inserta la marca binaria. Por otro lado, en términos de sensibilidad a la clave, se pudo validar experimentalmente que, si un usuario no autorizado utiliza una clave errónea o incompleta, no podrá detectar de forma exitosa la marca contenida en el audio.

Como toda solución, existe un costo asociado al mejoramiento de la calidad, que en nuestro caso corresponde a la disminución en la robustez de la marca, al comparar los resultados con el método QIM original. Sin embargo, aun cuando el valor de similitud entre la marca recuperada y la marca original disminuye, se garantiza que al menos el 70% de los pixeles se recuperan de forma correcta.

7. Reconocimientos

Esta investigación fue financiada por la Vicerrectoría de Investigaciones de la Universidad Militar Nueva Granada bajo el proyecto INV-ING-1910 de 2015.

Referencias

[1] R. Wang and Y. Xiong, "A novel audio aggregation watermarking for copyright protection," in 2010 *IEEE International Conference on Intelligent Systems and Knowledge Engineering*. IEEE, nov 2010, pp. 166-172. https://doi.org/10.1109/ISKE.2010.5680816

[2] S. P. S. Chauhan and S. A. M. Rizvi, "A survey: Digital audio watermarking techniques and applications," in *2013 4th International Conference on Computer and Communication Technology (ICCCT)*. IEEE, sep 2013, pp. 185–192. https://doi.org/10.1109/ICCCT.2013.6749625

[3] D. Renza, D. M. Ballesteros L., and C. Lemus, "Authenticity verification of audio signals based on fragile watermarking for audio forensics," *Expert Systems with Applications*, vol. 91, pp. 211–222, jan 2018. https://doi.org/10.1016/j.eswa.2017.09.003

[4] R. Martínez Noriega, "Mejoras de los métodos qim mediante códigos turbohadamard para el marcado de agua en audio digital," Ph.D. dissertation, Instituto Politécnico Nacional, 2007.

[5] V. B. K, I. Sengupta, and A. Das, "An adaptive audio watermarking based on the singular value decomposition in the wavelet domain," *Digital Signal Processing*, vol. 20, no. 6, pp. 1547–1558, dec 2010, https://doi.org/10.1016/j.dsp.2010.02.006

[6] A. R. Elshazly, M. M. Fouad, and M. E. Nasr, "Secure and robust high quality DWT domain audio watermarking algorithm with binary image," in *2012 Seventh International Conference on Computer Engineering & Systems (ICCES)*. IEEE, nov 2012, pp. 207–212. https://doi.org/10.1109/ICCES.2012.6408514

[7] K.-C. Choi and C.-M. Pun, "High capacity digital audio reversible watermarking," in *2013 IEEE International Conference on Computational Intelligence and Cybernetics (CYBERNETICSCOM)*. IEEE, dec 2013, pp. 72–75. https://doi.org/10.1109/CyberneticsCom.2013.6865784

[8] M. Salma, C. Maha, and B. A. Chokri, "A robust audio watermarking technique based on the perceptual evaluation of audio quality algorithm in the multiresolution domain," in *The 10th IEEE International Symposium on Signal Processing and Information Technology*. IEEE, dec 2010, pp. 326–331. https://doi.org/10.1109/ISSPIT.2010.5711803

[9] D. Renza, D. M. Ballesteros, and H. D. Ortiz, "Text hiding in images based on QIM and OVSF," *IEEE Latin America Transactions*, vol. 14, no. 3, pp. 1206–1212, mar 2016. https://doi.org/10.1109/TLA.2016.7459600

[10] D. Renza, C. Lemus, and D. M. Ballesteros, "Audio authenticity and tampering detection based on information hiding and collatz p-bit code," Journal of Information Hiding and Multimedia Signal Processing, vol. 8, no. 6, pp. 1294–1304, 2017.

[11] W.-N. Lie and L.-C. Chang, "Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification," *IEEE Transactions on Multimedia*, vol. 8, no. 1, pp. 46–59, feb 2006, https://doi.org/10.1109/TMM.2005.861292

[12] A. Jadhav and M. Kolhekar, "Digital watermarking in video for copyright protection," in *2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies*. IEEE, jan 2014, pp. 140–144. https://doi.org/10.1109/ICESC.2014.29

[13] N. Cvejic and T. Seppnen, "Spread spectrum audio watermarking using frequency hopping and attack characterization," *Signal Processing*, vol. 84, no. 1, pp. 207–213, jan 2004, https://doi.org/10.1016/j.sigpro.2003.10.016

[14] W. Li, X. Xue, and P. Lu, "Localized audio watermarking technique robust against time-scale modification," *IEEE Transactions on Multimedia*, vol. 8, no. 1, pp. 60–69, feb 2006, https://doi.org/10.1109/TMM.2005.861291

[15] S. Xiang, H. J. Kim, and J. Huang, "Audio watermarking robust against time-scale modification and MP3 compression," *Signal Processing*, vol. 88, no. 10, pp. 2372–2387, oct 2008, https://doi.org/10.1016/j.sigpro.2008.03.019

[16] A. Jensen y A. la Cour-Harbo, "Ripples in mathematics: the discrete wavelet transform", Springer Berlin Heidelberg, 2001, https://doi.org/10.1007/978-3-642-56702-5

[17] D. M. Ballesteros, D. Renza, and L. F. Pedraza, "Hardware design of the discrete wavelet transform: an analysis of complexity, accuracy and operating frequency," *Ingeniería y Ciencia*, vol. 12, no. 24, pp. 129–148, 2016. https://doi.org/10.17230/ingciencia.12.24.6

[18] B. Chen and G. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423–1443, may 2001, https://doi.org/10.1109/18.923725