UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Visión Electrónica
Más que un estado sólido
https://doi.org/10.14483/issn.2248-4728

VISIÓN ELECTRONICA
Visión Investigadora

Esteganografía de voz basada en cuantización mejorada

Speech steganography based on enhancement QIM

Carlos Steven Vargas Hernández [1] , *Dora María Ballesteros Larrotta* [2] , *Diego Renza Torres* [3]

INFORMACIÓN DEL ARTICULO

Historia del articulo
Enviado: 09/01/2018
Recibido: 13/03/2018
Aceptado: 28/06/2018

Palabras clave:
Capacidad de ocultamiento
Cuantización
Esteganografía
Señal huésped
Transparencia

RESUMEN

Una de los mecanismos para transmitir información segura consiste en la utilización de técnicas esteganográficas, entre las cuales, QIM sobresale por proporcionar alta transparencia. Sin embargo, la capacidad de ocultamiento es relativamente baja. En este artículo, se presenta una mejora al método QIM con el objetivo de mantener una alta transparencia en la señal transmitida, pero cuadriplicando su capacidad de ocultamiento. Se realizaron numerosas pruebas para medir el efecto de cuantización tanto en la señal esteganográfica, como en el mensaje secreto recuperado.

Keywords:
Hiding capacity,
Quantization,
Audio steganography,
Host signal,
Imperceptibility.

ABSTRACT:

One of the mechanisms for transmitting secure information consists of the use of steganographic techniques, among which, QIM provides high transparency. However, its hiding capacity is relatively low. In this paper, an improvement to the QIM method is presented in order to maintain a high transparency in the transmitted signal (i.e. stego signal), but increasing the hiding capacity (HC), by a factor of four times. Several tests were performed in order to measure the quantization effect on both the stego signal and the recovered secret message.

1 Ingeniero en Telecomunicaciones, Universidad Militar Nueva Granada, Bogotá, Colombia, Correo electrónico: u1400965@unimilitar.edu.co. ORCID: https://orcid.org/0000-0002-4671-6448.
2 Ingeniera Electrónica, Universidad Industrial de Santander, Colombia. MSc. En Ingeniería Electrónica, Universidad de los Andes, Bogotá, Colombia. Ph.D. En Ingeniería Electrónica, Universitat Politècnica de Catalunya, España. Docente Universidad Militar Nueva Granada, Bogotá, Colombia. Correo electrónico: dora.ballesteros@unimilitar.edu.co. ORCID: https://orcid.org/0000-0003-3864-818X.
3 Ingeniero Electrónico, Universidad Sur Colombiana, Colombia. MSc. En Ingeniería de Telecomunicaciones, Universidad Nacional de Colombia. Ph.D. En Computación Avanzada para Ciencias e Ingeniería de la Universidad Politécnica de Madrid, España. Docente 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: C. S. Vargas-Hernández, D. M. Ballesteros-Larrotta y D. Renza-Torres, "Esteganografía de voz basada en cuantización mejorada", *Visión electrónica, algo más que un estado sólido*, vol. 1, no. 1, Edición especial, enero-junio 2018. DOI revista: https://doi.org/10.14483/issn.2248-4728

1. Introducción

Actualmente, la cantidad de información que se transmite por canales públicos, como internet, va en aumento. Esta información puede ser de carácter confidencial, caso en el cual se espera que solamente sea conocida por los usuarios autorizados en la comunicación. Sin embargo, al transmitir datos por un canal público no se puede garantizar que se resguarde la privacidad de la información. Un mecanismo para aumentar el nivel de privacidad consiste en ocultar la información confidencial en un medio no confidencial (o no sensible), técnica conocida como esteganografía. De esta forma, si un intruso intercepta la información transmitida, no tendrá acceso al contenido secreto y solamente conocería la información de carácter no sensible [1].

En el anterior sentido, cuando la información secreta corresponde a mensajes de voz, un medio para ocultarla puede ser un archivo de audio el cual funciona como el huésped del contenido secreto. No obstante, para conservar la privacidad del contenido secreto, es necesario que el audio resultante -conocido como audio stego- no presente indicios sobre la existencia del contenido secreto (imperceptibilidad); y, adicionalmente, que el usuario autorizado pueda recuperar exitosamente el mensaje secreto a través de una clave privada.

Una de las dificultades típicas en el ocultamiento de señales de voz corresponde a la capacidad de ocultamiento del método (HC: Hiding Capacity), dado que si el método tiene un bajo HC requerirá de señales huésped de larga duración temporal para ocultar mensajes de voz de baja duración. Por el contrario, si el HC es alto la duración en tiempo entre las señales huésped y secreta puede ser similar.

Es así que un primer grupo de métodos de ocultamiento de voz se encuentra basado en sustitución de información binaria (LSB: Least Significant Bit) de la señal huésped, con un valor máximo de HC de 8 bits/muestra para señales huésped con resolución de 16 bits [2-4]. Los esquemas basados en LSB (en el dominio temporal, frecuencial o tiempo-frecuencia) tienen un bajo costo computacional, lo que facilita su implementación en dispositivos hardware [5]. Como desventaja se tiene que el método no brinda seguridad, ya que un intruso podría acceder fácilmente al contenido secreto con solo extraer los bits menos significativos de la señal stego, [6].

Un segundo grupo de métodos que permite ocultar voz en audio se basa en la manipulación del espectro de la señal huésped por medio de la sustitución de parte del espectro de esta señal con el espectro de la señal de voz, conocido como espectro desplazado (Shift Spectrum) [7], o la dispersión del espectro del mensaje secreto dentro del espectro del mensaje huésped (Spread Spectrum), [8]. En ambos casos, la capacidad de ocultamiento se puede ajustar de acuerdo a parámetros de diseño, pero la principal desventaja es que el mensaje secreto se puede remover con una pequeña manipulación de la señal stego (ej. filtrado, re-muestreo).

Un tercer grupo de métodos se basa en el enmascaramiento de la señal secreta por parte de la señal huésped, utilizando una amplitud en la señal huésped mucho mayor a la amplitud del mensaje secreto para que exista un efecto de cubrimiento (enmascaramiento) [9, 10]. Aunque este método puede tener un alto HC, el mensaje recuperado puede presentar un ligero deterioro en relación al mensaje secreto original. Una mejora a ese método es el enmascaramiento eficiente wavelet (EWM: Efficient Wavelet Masking) [11], el cual presenta alto HC y alta imperceptibilidad, pero alto costo computacional.

Por otro lado, un método que tradicionalmente se ha utilizado para watermarking es el método QIM [12, 13], el cual se basa en modificar la amplitud de la señal huésped (host) -de acuerdo a la información (bits) del mensaje secreto- por medio de un proceso de cuantización. Este método también se puede utilizar en aplicaciones de esteganografía que requieran baja capacidad de ocultamiento y alta imperceptibilidad. Sin embargo, cuando el mensaje secreto corresponde a una señal de voz, el método QIM requerirá de señales huésped de larga duración, ya que solamente oculta un bit por muestra [14]. Algunas modificaciones al método QIM incluyen cuantización no lineal [15, 16], la cual mejora la robustez frente a ataques pasivos; en cualquier caso, el costo computacional aumenta.

Por lo expuesto, este documento propone y valida, de manera experimental, una mejora al método QIM enfocada en aumentar la capacidad de ocultamiento de 1 a 4 bits por muestra, utilizando cuantización lineal. Se pretende conservar las ventajas del método QIM en términos de imperceptibilidad, bajo costo computacional y alta calidad del mensaje recuperado; a su vez que mejorara la baja capacidad de ocultamiento del método QIM original.

2. Quantization Index Modulation (QIM)

Entre los métodos ampliamente utilizados para insertar información binaria dentro de señales huésped, como audio, imagen, voz, o video, se encuentra el método QIM, el cual consiste en modificar la amplitud del dato por medio de una fórmula de cuantización. Dependiendo del valor binario a ocultar, el dato cuantizado será par o impar. Adicionalmente, utilizando también un proceso de cunatización, pero reversado, se puede extraer el contenido binario insertado. Tanto el transmisor como el receptor deben utilizar el mismo paso de cuantización para poder recuperar de forma exitosa la información insertad [12].

La fórmula de cuantización utilizada en el transmisor, corresponde a la ecuación (1), con Δ como paso de cuantización, H es la señal huésped, ST es la señal esteganografica y w es el valor vinario a ocultar.

$$ST = \begin{cases} \Delta \left| \frac{H}{\Delta} \right| & W = 0 \\ \Delta \left| \frac{H}{\Delta} \right| + \frac{\Delta}{2} & W = 1 \end{cases} \quad (1)$$

Como contraparte, en el receptor se aplica la ecuación (2) con el propósito de extraer el contenido secreto bit a bit. En este caso, w_r es el valor vinario recuperado

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

3. Método propuesto

En el método propuesto, se realiza una modificación a la técnica conocida como QIM para ocultar datos provenientes de una señal de voz, dentro de una señal huésped, que para nuestro caso también es una señal de voz. Una vez la señal secreta se ha ocultado, la señal resultante, conocida como señal esteganográfica, es transmitida por un canal no seguro a su destinatario final. Por otro lado, la clave, que permite recuperar el mensaje secreto, se transmite utilizando otro canal de comunicación. El primer proceso, se conoce como ocultamiento, y el segundo, como recuperación. En la Figura 1 se presenta el diagrama general

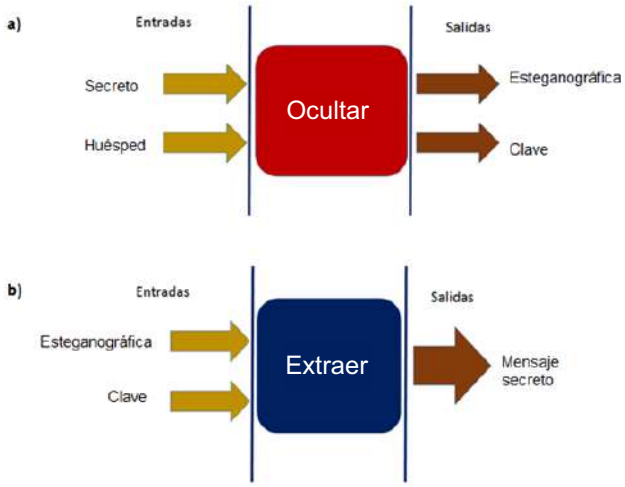


Figura 1. Diagrama de bloques esquema propuesto: a) Módulo transmisor, b) módulo receptor. Fuente: elaboración propia.

A continuación, se explican cada uno de los módulos.

3.1. En el módulo de transmisión

Paso 1: Cada una de las muestras de la señal de voz secreta se representa utilizando 16 bits. La amplitud máxima se codifica con "111...11" y la amplitud mínima con el código "000...00". Posteriormente, los códigos de cada una de las muestras de la señal de voz se concatenan formando un único string de P valores binarios.

Paso 2: El código obtenido en el paso anterior se divide en bloques de 4 bits cada uno; es decir, se obtiene un total de G bloques que corresponde a P dividido en 4.

Paso 3: se realiza la inserción de los G grupos de valores binarios en cada una de las muestras de la señal huésped. A diferencia del QIM clásico que solo permite insertar un bit a la vez, en nuestra propuesta se pueden insertar cuatro bits a la vez.

La fórmula de inserción corresponde a la ecuación (3), donde H corresponde a la señal huésped, St es la señal esteganográfica, y el índice d corresponde al equivalente decimal del grupo de 4 bits..

$$St = \Delta \left| \frac{H}{\Delta} \right| + \left(\frac{d * \Delta}{16} \right) \quad \text{para} \quad d = 0,1,\ldots 15 \quad (3)$$

La cantidad de muestras de la señal huésped que son cuantizadas es igual al total de grupos de 4 bits, es decir, G.

Por ejemplo, si se quiere ocultar el código "1111", el valor d corresponde a 15, el cual al reemplazarlo en la ecuación (3), queda de la forma:

$$St = \Delta \left| \frac{H}{\Delta} \right| + \left(\frac{15 * \Delta}{16} \right)$$

Ahora bien, si se quiere ocultar el código "1010", el valor decimal es 10, y la fórmula de cuantización se establece como:

$$St = \Delta \ \frac{H}{\Delta} \ + \left(\frac{10 * \Delta}{16} \right)$$

Una de las ventajas del método QIM y que se conservan en el método mejorado, consisten en que el valor máximo y mínimo de la señal no se supera, dado que la longitud en bits se conserva.

Paso 4: en este paso se escribe en la clave el valor de delta utilizado en la cuantización de las muestras de la señal de voz huésped y el valor de G. Sin la información contenida en la clave, en el receptor no se podría recuperar el contenido secreto a partir de la señal esteganográfica.

Paso 5: una vez se ha obtenido la señal esteganográfica y la clave, se realiza la transmisión utilizando dos canales diferentes, uno para la señal esteganográfica y otro para la clave.

3.2. En el módulo de recepción

En este módulo, el receptor conoce dos tipos de datos: uno correspondiente a la señal esteganográfica, y el otro a la clave. Cada uno obtenido de un canal de comunicación diferente.

Paso 1: el primer paso consiste en la extracción de cada grupo de 4 bits a partir de G muestras de la señal huésped. Para ello, se lee el valor de delta y de G guardados en la clave, y utilizando la señal esteganográfica se aplica la siguiente fórmula de extracción:

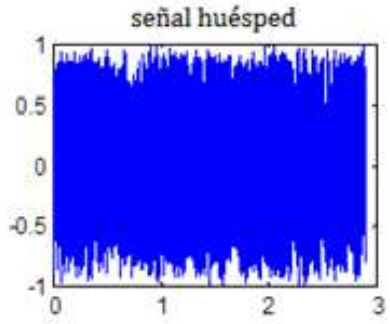
$$w_r = \begin{cases} d & \frac{2d-1 \ * \Delta}{32} < \left| St - \Delta \left| \frac{St}{\Delta} \right| \right| < \frac{2d+1 \ * \Delta}{32} \\ 0 & \text{en otro caso} \end{cases} \quad (4)$$

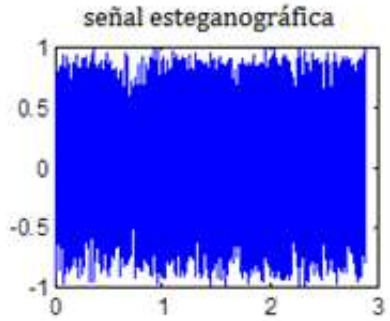
Donde *d* corresponde al valor decimal del código de 4 bits que se quiere detectar y wr es el código extraído de la señal esteganográfica..

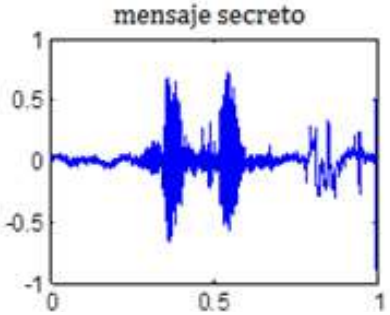
Paso 2: Con los G grupos binarios de 4 bits cada uno, el siguiente paso consiste en re-agrupar la información en longitudes de 16 bits. Posteriormente, se realiza el mismo proceso de mapeo entre valores decimales y binarios del módulo de ocultamiento, para en este caso, convertir el código binario de 16 bits en un número decimal en el rango de -1 a 1.

4. Influencia del valor delta (Δ)

Una de las variables importantes para el correcto funcionamiento del método propuesto consiste en el valor delta, dado que influye directamente en la calidad de la señal esteganográfica. Un valor alto implica que el valor cuantizado se aleje en gran medida del valor original y se generen distorsiones perceptuales, efecto no deseado en una señal esteganográfica. Por otro lado, un valor muy bajo de delta puede conllevar a solapamientos entre valores cuantizados y dificultaría el proceso de recuperación de la información. Por lo anterior, se hace necesario estudiar la influencia del valor delta tanto en la calidad de la señal esteganográfica, como del mensaje secreto recuperado. Los resultados se presentan en las Figuras 2, 3, y 4.







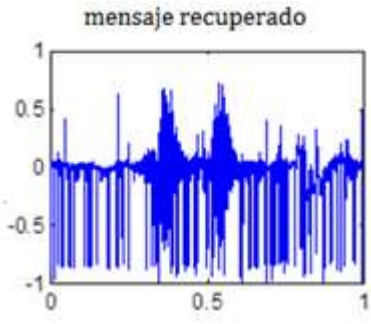


Figura 2. Ejemplo ocultamiento y extracción, delta igual a $9/(2^{16}-1)$.
Fuente: elaboración propia.

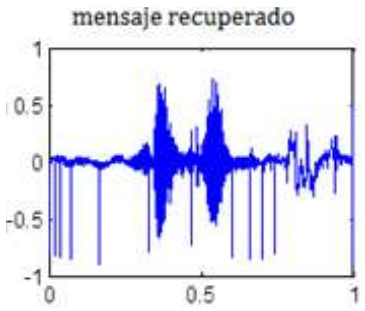
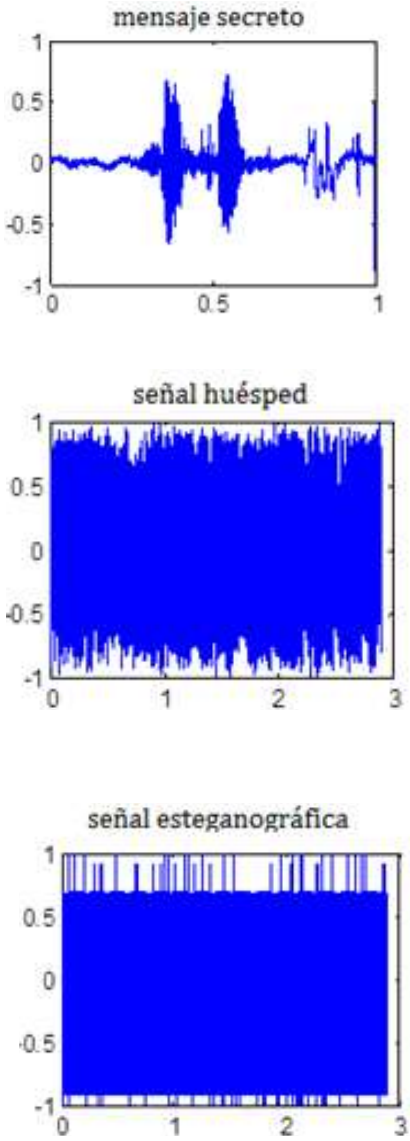
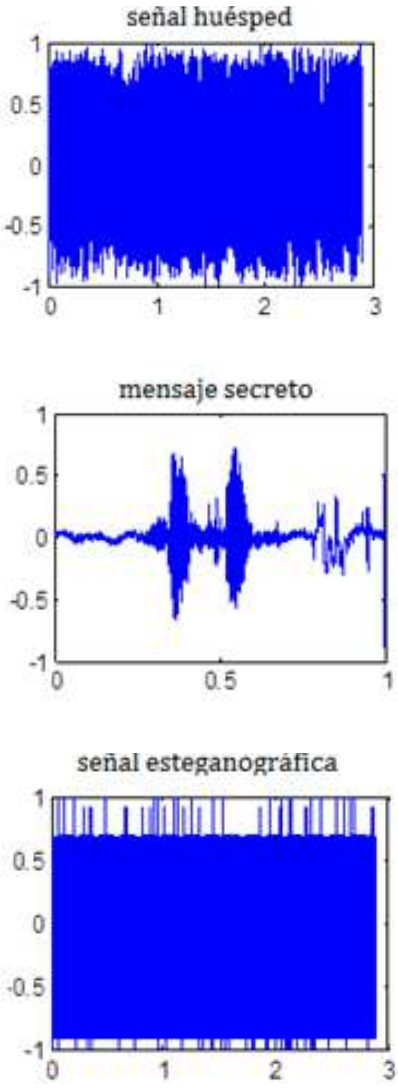


Figura 3. Ejemplo ocultamiento y extracción, delta igual a $40/(2^{16}-1)$.
Fuente: elaboración propia.





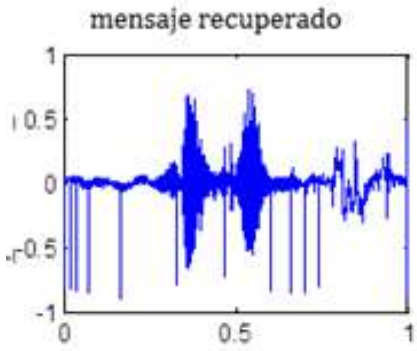


Figura 4. Ejemplo ocultamiento y extracción, delta igual a $3000/(2^{16}-1)$.
Fuente: elaboración propia.

De acuerdo a los resultados presentados en las Figuras 2-4, el valor de delta influye apreciablemente en la calidad de la señal esteganográfica y/o señal secreta recuperada. Si el valor de delta es muy pequeño, la señal esteganográfica es de muy buena calidad (altamente similar a la señal huésped original), pero, el mensaje recuperado es de muy mala calidad (con distorsiones fácilmente identificables). Cuando el valor delta es muy alto, la señal esteganográfica se ve apreciablemente afectada. Para un valor de delta adecuado, como corresponde a la Figura 3, tanto la señal esteganográfica, como el mensaje recuperado son de muy buena calidad.

5. Selección del valor delta

Para la selección del valor delta adecuado, se aplicó el siguiente protocolo de pruebas:

- Mensaje secreto: se utilizaron cinco mensajes secretos de contenido diferente. La resolución del mensaje secreto es de 16 bits/muestra y frecuencia de muestreo de 8 K Hz.
- Mensaje huésped: se utilizaron cinco mensajes huésped de contenido diferente. Estos mensajes tiene una duración de al menos cuatro veces la duración de los mensajes secretos. La resolución de la señal huésped es de 16 bits/muestra y frecuencia de muestreo de 8 K Hz.
- Delta: se utilizaron diez valores diferentes de delta.
- Total de pruebas: cada uno de los mensajes secretos se ocultó en cada una de las señales huésped, para cada valor de delta. El total de pruebas se obtiene de multiplicar el número de mensajes secretos diferentes por el número de señales huésped diferentes, por el número de valores delta, obteniendo 250 resultados.

5.1. Transparencia

La transparencia consiste en medir el nivel de similitud entre la señal huésped original y la señal esteganográfica. Un buen sistema de esteganografía debe proporcionar señales esteganográficas altamente similares a las señales huésped de que provienen.
El primer parámetro utilizado para calcular la similitud es el coeficiente de correlación cuadrático de Pearson (SPCC: Squared Pearson Correlation Coefficient) entre la señal huésped y la señal esteganográfica. La selección de este parámetro obedece a que su resultado no depende del rango dinámico de las señales que se comparan, a diferencia de otros parámetros como SNR. Si el SPCC es cercano a 1, indica que las señales son altamente correlacionadas (similares) y si está cercano a 0 indica una baja (o nula) correlación. La Figura 5 presenta el valor de SPCC para 10 valores de delta, con rango de confianza del 95%.

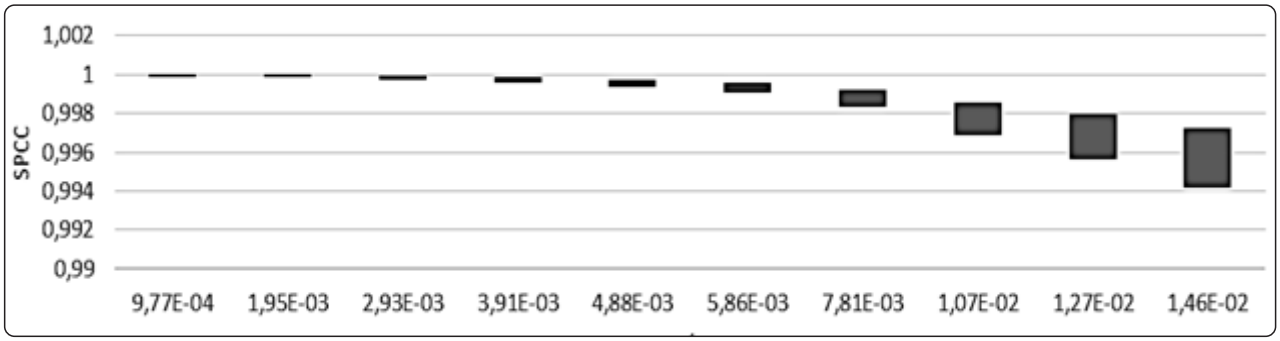


Figura 5. Intervalos de confianza de la correlación entre la señal host y la señal esteganográfica.
Fuente: elaboración propia.

De acuerdo a la Figura 5, si el valor de Δ aumenta la similitud entre la señal huésped y la señal esteganográfica disminuye. Es decir, en términos de transparencia, es mejor trabajar con valores de delta bajos.

Adicionalmente, se compararon los valores estadísticos entre la señal huésped y la señal esteganográfica, específicamente en términos de la desviación estándar, la curtosis . La selección de estas estadísticas obedeció a que tienen alta influencia en la distribución de probabilidad de los datos, de tal forma que si estos valores son bajos (~1%) las variaciones son imperceptibles [11]. La Figura 6 presenta el porcentaje de error existente entre la desviación estándar de la señal host y la desviación estándar de la señal esteganográfica.

La Figura 7 presenta lo correspondiente para la curtosis.

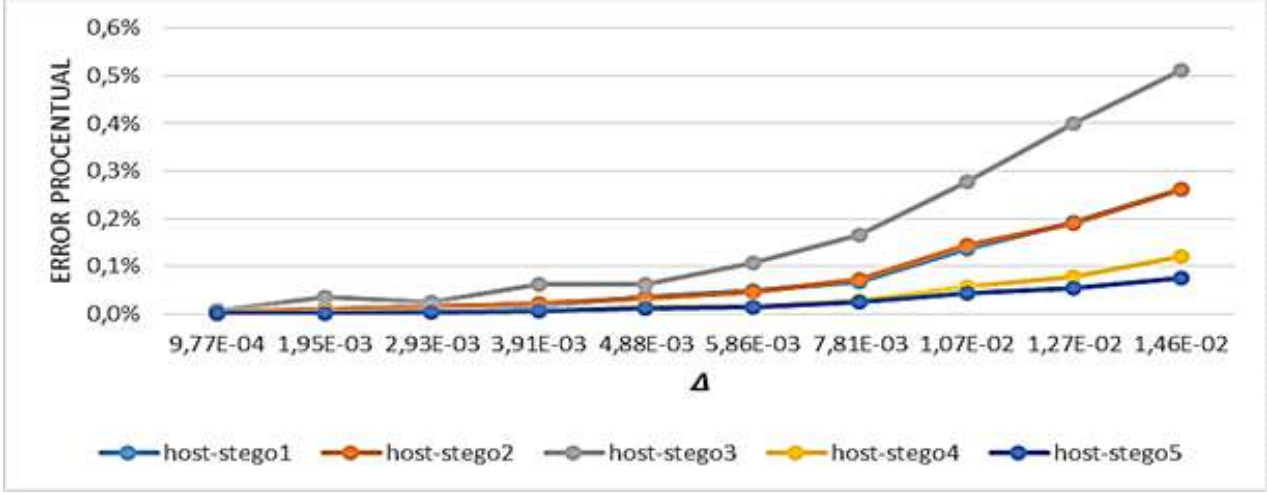


Figura 6. Error porcentual calculado entre la desviación estándar de la señal host y señal esteganográfica. Fuente: elaboración propia.

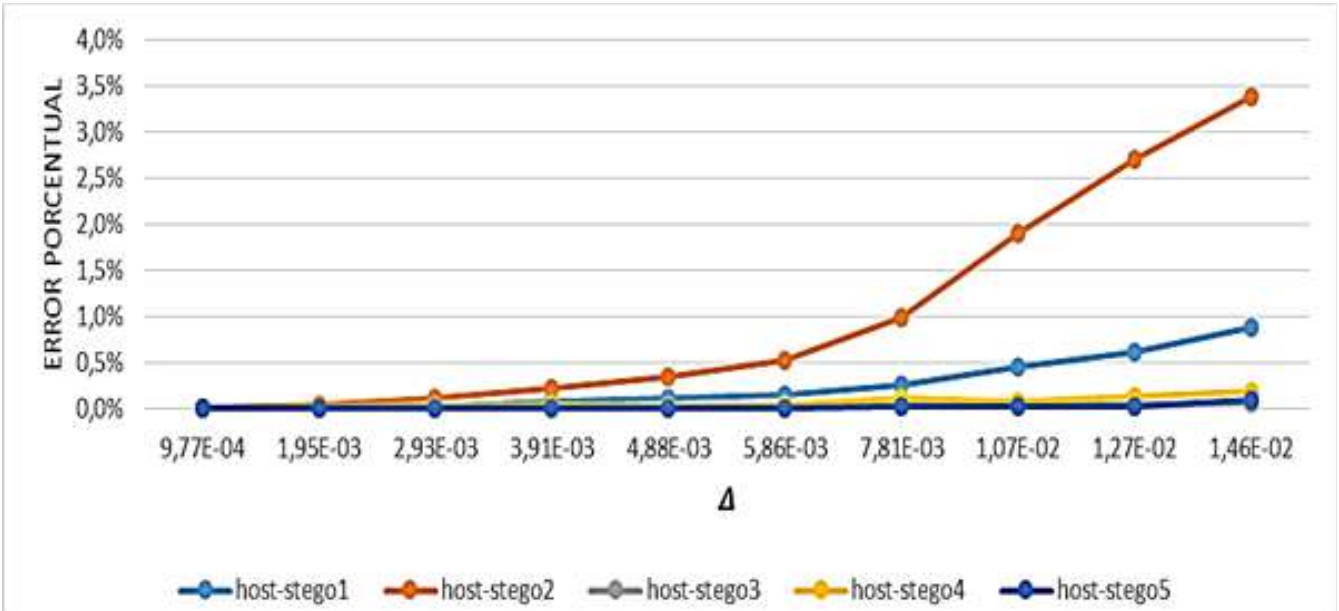


Figura 7. Error porcentual calculado entre la curtosis de la señal host y señal esteganográfica. Fuente: elaboración propia.

4 La curtosis es el cuarto momento estadístico estándar, e indica el grado de concentración de la distribución de probabilidad de una variable aleatoria real alrededor del origen.

En términos de la desviación estándar (Figura 6), todas las señales esteganográfica generadas mantuvieron un error estadístico por debajo del 0.6% para los 10 valores de Δ seleccionados. En el caso de la curtosis (Figura 7), el máximo porcentaje de error fue menor al 4%, pero la mayoría de los resultados estuvieron por debajo al 1%. De tal forma, se puede concluir que a nivel estadístico se garantiza la imperceptibilidad del mensaje secreto en la señal esteganográfica.

5.2. Mensaje recuperado

En la segunda parte de la validación, se analiza la calidad del mensaje de voz recuperado en términos de su similitud con el mensaje de voz original. Idealmente este valor debe ser 1, pero se debe considerar tanto en el módulo del transmisor como en el módulo del receptor, se realiza conversión A/D y D/A, respectivamente, generando leves errores. Los resultados se presentan en la Figura 8.

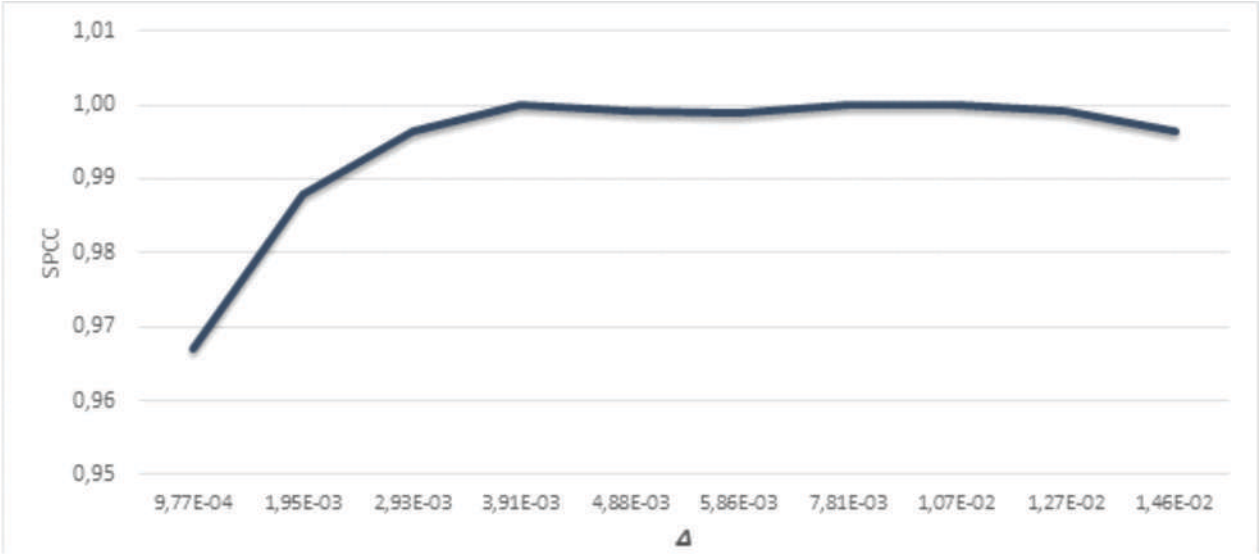


Figura 8. Promedio de la correlación entre el mensaje secreto original y el recuperado.
Fuente: elaboración propia.

5.3. Capacidad de ocultamiento de voz en audio

Lo primero a tener en cuenta para implementar el método QIM de cuatro bits es que la señal huésped debe tener una resolución de por lo menos 16 bits/muestra, que permita ocultar los datos secretos sin generar mayor distorsión.

De forma general, la ecuación (4) permite obtener el tiempo de duración de la señal huésped necesario para ocultar el mensaje secreto. La única restricción, es que la señal huésped tenga de resolución 16 bits/muestra. De la ecuacion Th es el tiempo de la senal huesped, Ts es la duracion del tiempo del mensaje secreto y Fh es la frecuencia de muestreo de la señal huesped.

$$Th = 4 \quad \frac{T_S * F_S}{F_h} \tag{4}$$

Por ejemplo, si se quiere ocultar un audio que tiene una duración de 4 segundos, una resolución de 16 bits/muestra y una frecuencia de muestreo de 16KHz, dentro de una señal huésped con frecuencia de muestreo de 44,1KHz, es necesario que el audio huésped tenga una duración de por lo menos 5,805 segundos. Si se quisiera ocultar la misma señal dentro de la misma señal huésped aplicando el método QIM tradicional, sería necesario que la duración fuese de por lo menos 23,22 segundos.

6. Conclusiones

En este artículo se propuso un esquema de esteganografía de voz en audio utilizando QIM mejorado. La mejora consistió en ampliar la capacidad de ocultamiento del método QIM con cuantización lineal de cuatro bits.

De acuerdo a la validación experimental realizada, se encontró que el valor de delta influye levemente en la calidad de la señal esteganográfica, particularmente para valores de delta menores a $1.4*10^{-2}$. Para este rango

de valores, el SPCC entre la señal host y la señal esteganográfica disminuye a medida que el delta aumenta, pero siempre es mayor a 0,994. A nivel estadístico, el error porcentual en la desviación estándar y la curtosis entre las dos señales es en la mayoría de los casos menor a 1%, valor que es lo suficientemente bajo para garantizar imperceptibilidad en la señal esteganográfica.

En la valoración del mensaje de voz recuperado, la influencia de delta es mayor que en la señal esteganográfica y el valor de SPCC llega a ser de 0,93. En este caso, el SPCC mejora en valores de delta medios, es decir, no se recomienda que el delta sea muy bajo (menor a $0{,}4*10^{-2}$) o muy alto (mayor a $1{,}1*10^{-2}$). Cuando el delta no es el adecuado, se tiene distorsión en el mensaje recuperado en las zonas de silencio (efecto click).

En conclusión, para satisfacer simultáneamente la imperceptibilidad de la señal esteganográfica y la calidad del mensaje de voz recuperado se recomienda utilizar un delta entre $0{,}4*10^{-2}$ y $1{,}1*10^{-2}$, para una señal de voz de 16 bits de resolución. Con estos valores se conservan las ventajas del método QIM (alta imperceptibilidad, bajo costo computacional), con una capacidad de ocultamiento cuadruplicada.

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] F. Djeebar, B. Ayad, K. Abed Meraim y H. Hamam, "Comparative study of digital audio steganography techniques", *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2012, no. 25, 2012, pp. 1-16. DOI: https://doi.org/10.1186/1687-4722-2012-25.

[2] N. Cvejic y T. Seppänen, "A wavelet domain LSB insertion algorithm for high capacity audio steganography", IEEE Proceedings of Digital Signal Processing Workshop, and the 2nd Signal Processing Education Workshop, 2002, pp. 53-55. DOI: https://doi.org/10.1109/DSPWS.2002.1231075.

[3] N. Cvejic y T. Seppänen, "Increasing the capacity of LSB-based audio steganography", IEEE Workshop on Multimedia Signal Processing, 2002, pp. 336-338. DOI: https://doi.org/10.1109/MMSP.2002.1203314.

[4] R. Sridevi y A Damodaram, "Efficient method of audio steganography by modified LSB algorithm and strong encryption key with enhanced security", *Journal of Theoretical and Applied Information Technology*, vol. 5, no 6, 2009.

[5] D. M. Ballesteros y J. M. Moreno, "Real-time, speech-in-speech hiding scheme based on least significant bit substitution and adaptive key", *Computers and Electrical Engineering*, vol. 39, no. 4, 2013, pp. 1192-1203. DOI: https://doi.org/10.1016/j.compeleceng.2013.02.006.

[6] S. Dumitrescu, X. Wu y Z. Wang, "Detection of LSB steganography via sample pair analysis", *IEEE Transactions on Signal Processing*, vol. 51, no. 7, 2003, pp. 1995-2007. DOI: https://doi.org/10.1109/TSP.2003.812753.

[7] D. E. Skopin, I. M. El-Emary, R. J. Rasras y R. S. Diab, "Advanced algorithms in audio steganography for hiding human speech signal", IEEE 2nd International Conference on Advanced Computer Control, vol. 3, 2010, pp. 29-32. DOI: https://doi.org/10.1109/ICACC.2010.5486735.

[8] H. Matsuka, "Spread spectrum audio steganography using sub-band phase shifting", IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2006, pp. 3–6. DOI: https://doi.org/10.1109/IIH-MSP.2006.265106.

[9] F. Djebbar, H. Hamam, K. Abed-Meraim y D. Guerchi, "Controlled distortion for high capacity data-in-speech spectrum steganography", IEEE Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 212-215. DOI: https://doi.org/10.1109/IIHMSP.2010.60.

[10] O. Yilmaz y S. Rickard, "Blind separation of speech mixtures via time-frequency

masking" IEEE transactions on Signal Processing, vol. 52, no. 7, 2004, pp. 1830-1847. DOI: https://doi.org/10.1109/TSP.2004.828896.

[11] D. M. Ballesteros y A. Moreno, "Highly transparent steganography model of speech signals using Efficient Wavelet Masking", *Expert Systems with Applications*, vol. 39, no. 10, 2012, pp. 9141-9149. DOI: https://doi.org/10.1016/j.eswa.2012.02.066.

[12] B. Chen y G. W: 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, 2001, pp. 1423-1443. DOI: https://doi.org/10.1109/18.923725.

[13] D. Renza, D. M. Ballesteros y C. Lemus. "Authenticity verification of audio signals based on fragile watermarking for audio forensics", *Expert systems with applications*, vol. 91, 2018, pp. 211-222. DOI: https://doi.org/10.1016/j.eswa.2017.09.003.

[14] D: M. Ballesteros, D. Renza y R. Rincón, "Gray-scale images within color images using similarity histogram-based selection and replacement algorithm", *Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, no. 6, 2015, pp. 1156-1166.

[15] N. K. Kalantari y S. M. Ahadi, "A logarithmic quantization index modulation for perceptually better data hiding", *IEEE Transactions on Image Processing*, vol. 19, no. 6, 2010, pp. 1504-1517. DOI: https://doi.org/10.1109/TIP.2010.2042646.

[16] N. K. Kalantari y S. M. Ahadi, "Logarithmic quantization index modulation: A perceptually better way to embed data within a cover signal", IEEE International Conference on Acoustics, Speech and Signal Processing, 2009, pp. 1433-1436. DOI: https://doi.org/10.1109/ICASSP.2009.4959863.