Evaluación de técnicas de reducción de ruido basadas en wavelets orientadas al procesamiento de señales bioacústicas

Alejandro Gómez Echavarría, alejandrogomezechavarria@gmail.com

Trabajo de Grado presentado para optar al título de Ingeniero de Sonido

Asesor: Juan Pablo Ugarte Macías, Doctor (PhD) en Ciencias Médicas.

Universidad de San Buenaventura Colombia
Facultad de Ingenierías
Ingeniería de Sonido
Medellín, Colombia
2018

Bibliotecas Universidad de San Buenaventura

- Biblioteca Fray Alberto Montealegre OFM - Bogotá.
- Biblioteca Fray Arturo Calle Restrepo OFM - Medellín, Bello, Armenia, Ibague.
- Departamento de Biblioteca - Cali.
- Biblioteca Central Fray Antonio de Marchena – Cartagena.

Universidad de San Buenaventura Colombia

Universidad de San Buenaventura Colombia - http://www.usb.edu.co/
Bogotá - http://www.usbbog.edu.co
Medellín - http://www.usbmed.edu.co
Cali - http://www.usbcali.edu.co
Cartagena - http://www.usbctg.edu.co
Editorial Bonaventuriana - http://www.editorialbonaventuriana.usb.edu.co/
Revistas - http://revistas.usb.edu.co/
Dedicatoria

Dedico este trabajo a mi padre y a mi madre quienes apoyaron incondicionalmente mi formación personal y académica. Crecer y madurar en un hogar complaciente, lleno de amor y cariño ha sido un privilegio que me ha permitido emprender en el camino del conocimiento como mi mayor pasión. También dedico este proyecto a nuestro planeta Tierra el cual nos ha albergado cuidadosamente a pesar del inclemente egoísmo de la humanidad. Este trabajo proyecta la capacidad intelectual y creativa de nuestra raza para plantear estrategias que contribuyan al cuidado de otras especies. Es un deber de la ciencia aportar ideas que contrarresten el desequilibrio generado por nuestros sistemas y actividades insostenibles.

Agradecimientos conjuntos

Los más sinceros agradecimientos para mi maestro y asesor Juan Pablo Ugarte Marcías quien acompaña y oriente de forma excepcional el desarrollo de este proyecto. Agradezco también a mi maestro Diego Murillo quien contribuyó significativamente en la gestación de esta investigación junto, con Claudia Isaza del grupo SISTEMIC quien proveyó la base de datos para este proyecto.
<table>
<thead>
<tr>
<th>Sección</th>
<th>Página</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLA DE CONTENIDO</td>
<td>1</td>
</tr>
<tr>
<td>RESUMEN</td>
<td>10</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>11</td>
</tr>
<tr>
<td>I. INTRODUCCIÓN</td>
<td>12</td>
</tr>
<tr>
<td>II. PLANTEAMIENTO DEL PROBLEMA</td>
<td>13</td>
</tr>
<tr>
<td>III. JUSTIFICACIÓN</td>
<td>14</td>
</tr>
<tr>
<td>IV. ESTADO DEL ARTE</td>
<td>15</td>
</tr>
<tr>
<td>V. OBJETIVOS</td>
<td>19</td>
</tr>
<tr>
<td>A. Objetivo general</td>
<td>19</td>
</tr>
<tr>
<td>B. Objetivos específicos</td>
<td>19</td>
</tr>
<tr>
<td>VI. MARCO TEÓRICO</td>
<td>20</td>
</tr>
<tr>
<td>A. Wavelet denoising</td>
<td>20</td>
</tr>
<tr>
<td>B. La transformada discreta wavelet</td>
<td>21</td>
</tr>
<tr>
<td>C. La transformada discreta wavelet no diezmada</td>
<td>23</td>
</tr>
<tr>
<td>D. Thresholding</td>
<td>24</td>
</tr>
<tr>
<td>E. Consideraciones para el ruido</td>
<td>25</td>
</tr>
<tr>
<td>A. Clases de ruido</td>
<td>26</td>
</tr>
<tr>
<td>VII. METODOLOGÍA</td>
<td>27</td>
</tr>
<tr>
<td>A. Selección del wavelet</td>
<td>27</td>
</tr>
<tr>
<td>B. Estimación del número de niveles de descomposición</td>
<td>28</td>
</tr>
<tr>
<td>C. Estrategias de denoising</td>
<td>29</td>
</tr>
<tr>
<td>1). Last approximation standard deviation (LAstd)</td>
<td>30</td>
</tr>
<tr>
<td>2). Level dependent standard deviation (LDstd)</td>
<td>31</td>
</tr>
<tr>
<td>3). Level dependent median absolute deviation (LDmad)</td>
<td>31</td>
</tr>
<tr>
<td>4). Método de Matlab (SURE)</td>
<td>32</td>
</tr>
</tbody>
</table>
D. Señales de prueba .................................................................................................................................................. 33
F. Medidas de desempeño ........................................................................................................................................ 37
   1). Relación señal/ruido (SNR) ......................................................................................................................... 37
   2). Entropía espectral (SEN) ............................................................................................................................... 37
VIII. RESULTADOS .................................................................................................................................................... 38
   A. Selección del wavelet ....................................................................................................................................... 38
   B. Denoising en cantos de búhos ..................................................................................................................... 38
   C. Denoising en los registros de ecosistemas colombianos ........................................................................... 43
IX. DISCUSIÓN .......................................................................................................................................................... 45
X. CONCLUSIONES ..................................................................................................................................................... 52
   A. Limitaciones .................................................................................................................................................... 52
   B. Trabajos futuros .............................................................................................................................................. 53
   C. Productos ......................................................................................................................................................... 53
REFERENCIAS ........................................................................................................................................................... 54
LISTA DE TABLAS

TABLA I. GRUPOS DE DATOS CON DIFERENTES PERFILES DE RUIDO ADITIVO GENERADOS AL CONTAMINAR LOS 10 DIFERENTES CANTOS DE BÚHOS SELECCIONADOS..........................................................34
LISTA DE FIGURAS

Fig. 1. Descomposición y reconstrucción usando la DWT.................................................................22
Fig. 2. Descomposición y reconstrucción usando la DWT.................................................................23
Fig. 3. Descomposición y reconstrucción usando la UDWT...............................................................23
Fig. 4. *hard thresholding* en el recuadro izquierdo y *soft thresholding* en el recuadro derecho...25
Fig. 5. Metodología basada en wavelets para la eliminación de ruido.............................................27
Fig. 6. Wavelets................................................................................................................................28
Fig. 7. Aproximaciones en una DWT de seis niveles de una señal ruidosa.........................................29
Fig. 8. Estrategia *LAstd* con dos niveles de descomposición...........................................................30
Fig. 9. Estrategia *LDstd* con dos niveles de descomposición............................................................31
Fig. 10. Desempeño de la función *wden* con los diferentes criterios de estimación del nivel de *threshold*.................................................................................................................................33
Fig. 11. Espectrogramas de los cantos sin ruido de las 10 especies de búhos en estudio.................35
Fig. 12. Espectrogramas de los eventos seleccionados de Colustethus y Norandinus......................36
Fig. 13. Efectos del aliasing que provocan algunos wavelets sobre la señal después realizar la DWT, aplicar un leve *thresholding* y reconstruir la señal nuevamente.........................................................38
Fig. 14. Mediana de la SNR de las señales contaminadas con ruido Gaussiano blanco de diferentes amplitudes, para diferentes valores de *q*..................................................................................................................39
Fig. 15. Mediana de la SNR de las señales contaminadas con ruido rosa de diferentes amplitudes, para diferentes valores de *q*..................................................................................................................40
Fig. 16. Mediana de la SNR de las señales contaminadas con ruido browniano de diferentes amplitudes, para diferentes valores de *q*..................................................................................................................40
Fig. 17. Mediana de la SEN de las señales contaminadas con ruido Gaussiano blanco de diferentes amplitudes, para diferentes valores de *q*..................................................................................................................41
Fig. 18. Mediana de la SEN de las señales contaminadas con ruido rosa de diferentes amplitudes, para diferentes valores de *q*..................................................................................................................41
Fig. 19. Mediana de la SEN de las señales contaminadas con ruido browniano de diferentes amplitudes, para diferentes valores de *q*..................................................................................................................42
Fig. 20. Mediana de la SEN de los eventos bioacústicos.....................................................................43
Fig. 21. Mediana de la SEN de Colustethhus y Norandinus. .........................................................44
Fig. 22. Resultado del denoising en el espectrograma con $q = 1$. .................................................45
Fig. 23. Resultado del denoising en el espectrograma con $q = 1$. ...................................................46
Fig. 24. Resultado del denoising en el espectrograma con $q = 0,25$. ..............................................47
Fig. 25. Resultado del denoising en el espectrograma con $q = 1$ y amplitud del ruido de 30%. ...48
Fig. 26. Resultado del denoising en el espectrograma con $q = 1$. ...................................................49
Fig. 27. Resultado del denoising en el espectrograma con $q = 1$. ...................................................50
Fig. 28. Resultado del denoising en el espectrograma con $q = 0,3$. ...............................................51
Lista de abreviaciones

**WT** Wavelet Transform

**DWT** DiscreteWavelet Transform

**UDWT** Undecimated DiscreteWavelet Transform

**SNR** Signal to Noise Ratio

**SEN** Spectral ENtropy
RESUMEN

Este proyecto de investigación evalúa y propone diferentes estrategias de reducción de ruido basadas en la transformada discreta wavelet no diezmada para procesar señales bioacústicas ruidosas. Las estrategias propuestas se basan en diferentes criterios de estimación del nivel de ruido presente en la señal para realizar un soft thresholding sobre los coeficientes de detalle obtenidos de la descomposición wavelet. Los algoritmos propuestos, denominados $L_{Astd}$, $L_{Dstd}$ y $L_{Dmad}$, usan la desviación estándar de la última aproximación, la desviación estándar de cada nivel de detalle y la desviación media absoluta para estimar el nivel de ruido respectivamente. Las estrategias son comparadas con el método de estimación de threshold de riesgo imparcial de Stein (SURE) implementado con la función de Matlab $wden$ del Wavelet Toolbox, la cual usa la transformada discreta wavelet.

Para probar los algoritmos se usaron tres conjuntos de datos. En un primer acercamiento se evaluaron las estrategias con distintos cantos de búhos con diferentes perfiles de ruido aditivo y amplitudes. Para probar la tolerancia de los algoritmos al ruido ambiente, se usó otro conjunto de datos formado por fragmentos de señales grabadas en campo con eventos acústicos de alta actividad biológica. Finalmente, se probaron los algoritmos con las grabaciones completas de los ecosistemas colombianos los cuales contenían los eventos acústicos. Los resultados obtenidos fueron cuantificados usando la relación señal a ruido y la entropía espectral. Para complementar los resultados se hizo un análisis visual de los espectrogramas de las señales procesadas. La metodología $L_{Astd}$ obtuvo un excelente desempeño al procesar los cantos de búhos contaminados con ruido Gaussiano blanco, sin embargo, no es tolerante a ruido de color y de banda angosta. Por otro lado, las metodologías $L_{Dstd}$ y $L_{Dmad}$ demuestran un mejor desempeño en señales con ruido de color, demostrando su capacidad para procesar señales grabadas en campo con ruido ambiente. La implementación de un parámetro “$q$” de suavizado del thresholding en las metodologías propuestas permitió ajustar el procesamiento para evitar la pérdida de información importante, otorgando a los algoritmos versatilidad para desempeñarse bien en distintos escenarios.

**Palabras clave:** Transformada wavelet, wavelet denoising, bioacústica.
ABSTRACT

This research project assesses and proposes different strategies for noise reduction based on the undecimated discrete wavelet transform to process noisy bioacoustic signals. The proposed strategies are based on different criteria for estimating the level of noise present in the signal to perform a soft thresholding on the detail coefficients obtained from the wavelet decomposition. The proposed algorithms, called \textit{LASTd}, \textit{LDstd} and \textit{LDmad}, use the standard deviation of the last approximation, the standard deviation of each detail at each level and the median absolute deviation to estimate the noise level respectively. The strategies are compared with Stein's unbiased risk threshold estimation method (SURE) implemented with the Matlab function \texttt{wden} from the wavelet toolbox which uses the discrete wavelet transform.

Three sets of data were used to assess the algorithms. In a first approach, the strategies were evaluated using different owl calls with different additive noise profiles and amplitudes. To test the tolerance of the algorithms to ambient noise, a data set of fragments with acoustic events of high biological activity were extracted from field recordings. Finally, the algorithms were tested with the full recordings of the Colombian ecosystems which contained the acoustic events. The results were quantified using the signal-to-noise ratio and the spectral entropy. To complement the results, a visual analysis of the spectrograms of the processed signals was made. The methodology \textit{LASTd} obtained an excellent performance when processing the owl calls contaminated with white Gaussian noise, however it is not tolerant to colored noise and narrow band noise. On the other hand, the methodologies \textit{LDstd} and \textit{LDmad} show a better performance in signals with colored noise, demonstrating their ability to process field recordings with ambient noise. The implementation of a smoothing parameter “q” for the thresholding in the proposed methodologies allowed to adjust the processing to avoid the loss of important information, giving to the algorithms versatility to perform well in different scenarios.

\textbf{Keywords:} Wavelet transform, wavelet denoising, bioacoustics.
I. INTRODUCCIÓN

El campo de la bioacústica propone estrategias para el estudio y supervisión de especies mediante parámetros acústicos que se extraen de diferentes ecosistemas con el propósito de identificar, establecer relaciones y tendencias entre especies animales y el entorno [1], [2]. Sin embargo, las señales acústicas son raramente observadas en forma aislada y transportan consigo información indeseada (ruido) que disminuye la capacidad de análisis y dificulta el procesamiento y extracción de información relevante [1]. La reducción de ruido en señales bioacústicas ha sido un problema que se ha abordado de diferentes formas con el fin de extraer la información deseada para facilitar el análisis y caracterización de señales [1], [2], [3]. El desarrollo de los wavelets abre camino al procesamiento de señales en este dominio, donde la localización en tiempo y frecuencia son posibles gracias a las características y versatilidad de estas formas de onda de duración limitada.

El acercamiento fundamental para la reducción de ruido mediante wavelets consiste en efectuar una descomposición de la señal para obtener los coeficientes wavelet, los cuales representan una relación de similitud ente la señal y el wavelet seleccionado. Los coeficientes de menor amplitud están asociados al ruido. Para eliminar estos componentes se realiza un proceso de thresholding donde se conservan los coeficientes más representativos de la señal deseada. Posteriormente se reconstruye una estimación de la señal limpia a partir de los coeficientes modificados.

Este proyecto de investigación evalúa y propone diferentes estrategias de reducción de ruido basadas en la transformada discreta wavelet y la transformada discreta wavelet no diezmada para procesar señales bioacústicas ruidosas. Para obtener una estimación del nivel de ruido en el dominio wavelet, se plantean dos estrategias basadas en la desviación estándar y una estrategia basada en la desviación media absoluta. Para evaluar las estrategias se usó un grupo de datos con diferentes perfiles de ruido aditivo y dos grupos de señales ruidosas grabadas en campo de dos ecosistemas diferentes.
II. PLANTEAMIENTO DEL PROBLEMA

El uso de parámetros acústicos para la supervisión y el estudio de poblaciones biológicas juega un papel importante cuando el contacto directo con las especies en estudio no es deseado. Existen distintas especies que emiten sonidos con un propósito cooperativo y/o competitivo [4]. Estas señales de origen biológico son consideradas indicadores importantes del estado de los ecosistemas y es por esto que su análisis es de gran utilidad en áreas de la biología y la conservación, y sirven para evaluar el impacto que causan algunas actividades humanas en el medio ambiente [1].

La adquisición de información es un proceso importante cuando se plantean estrategias automáticas o semiautomáticas para extraer la información deseada. Estos sistemas son susceptibles al fallo cuando deben reconocer segmentos en la grabación donde las especies en estudio se encuentran contaminadas por ruido [5], [6], [7], [8]. El uso de micrófonos direccionales es una práctica común cuando se busca una menor influencia de algunos sonidos que hacen parte del paisaje sonoro [9], [2]. Sin embargo, el uso de este tipo de micrófonos es insuficiente para atenuar algunas fuentes consideradas como ruido, como la geofonía (sonidos naturales no biológicos) y antrofonía (sonidos originados por el hombre) las cuales pueden enmascarar especies en estudio, dificultando la extracción de la información [7], [10], [11]. Algunos autores hacen uso de filtros pasabanda con anchos de banda optimizados para especies específicas [10], [12], [13]. Sin embargo este método no es capaz de discriminar fuentes indeseadas en la misma banda de estudio [14].

Para alcanzar una mejor reducción de ruido es necesario plantear etapas de pre-procesamiento robustas y selectivas, tanto en tiempo como en frecuencia, con el fin de realizar las regiones del espectro con información relevante y así mejorar el desempeño de etapas posteriores que involucran el análisis y caracterización de las señales que son llevadas a procesos de selección y clasificación de especies. Esta investigación se enfoca en el uso de algoritmos de reducción de ruido basados en wavelets como etapa de pre-procesamiento debido a la capacidad que tiene esta herramienta matemática para analizar señales en tiempo-frecuencia.
III. JUSTIFICACIÓN

Existen diversas técnicas en el procesamiento de señales que permiten mejorar la relación señal/ruido (SNR). Algunas de las más usadas son, el filtrado mediante filtros pasabanda, la sustracción espectral, el filtrado Wiener y el filtrado de Ephraim–Malah. Estas técnicas hacen uso de la transformada de tiempo corto de Fourier (STFT) como herramienta matemática, asumiendo que las señales en estudio son señales estacionarias [15]. Sin embargo, la reducción de ruido mediante wavelets, también llamada Wavelet denoising, tiene la capacidad de procesar señales no estacionarias gracias a la versatilidad de los wavelets como herramienta matemática [15].

El uso de los wavelets como herramienta de procesamiento ha sido enfocado en campos como la compresión de imágenes, procesamiento de señales médicas y reconocimiento de patrones [16], [17], [18]. Es por esto que, la presente investigación amplía el espectro de aplicaciones de esta herramienta contribuyendo en la mejora de procesos de análisis y caracterización, importantes para la segmentación y clasificación de especies animales por medio de parámetros acústicos.

En situaciones donde el ruido afecta la calidad del procesamiento de las señales en estudio, es necesario implementar una etapa de pre-procesamiento donde la información deseada se realza con el fin de reducir la interferencia que genera el ruido sobre un proceso de extracción de la información. Esta etapa de pre-procesamiento se conoce generalmente como reducción de ruido o "denoising", y su objetivo es aumentar la relación señal/ruido de la señal contaminada.
IV. ESTADO DEL ARTE

Uno de los acercamientos fundamentales en la reducción de ruido mediante wavelets fue propuesto por Donoho y Johnstone en 1994, quienes desarrollaron un método para reconstruir una función desconocida \( f \) a partir de datos ruidosos en el dominio wavelet trasladando todos los coeficientes wavelet de la señal ruidosa hacia cero aplicando un proceso denominado \textit{thresholding} [19]. En 1995 Donoho en su trabajo \textit{De-noising by Soft-Thresholding} propone una interpretación formal del término \textit{denoising} y muestra cómo las transformadas wavelet deben ser usadas de forma óptima para realizar la eliminación de ruido. El procedimiento para reconstruir la función \( f \) se resume en los siguientes pasos descritos en [20]:

1. Aplicar el algoritmo de filtrado piramidal adaptado a intervalos desarrollado por Cohen, Daubechies, Jawert y Vial [21] a los datos medidos, para obtener los coeficientes empíricos wavelet.
2. Aplicar el \textit{thresholding} de forma coordinada con los coeficientes wavelet con un valor de \textit{threshold} especialmente elegido mediante \( t_n = \gamma \sigma \sqrt{2 \log(n)/n} \). Siendo \( \gamma \) una constante definida en [20], \( \sigma \) el nivel de ruido y \( n \) la longitud de la señal wavelet.
3. Invertir el filtrado piramidal para reconstruir la función.

Los trabajos de Donoho y Johnstone se convirtieron en las bases fundamentales para desarrollar aplicaciones de eliminación de ruido mediante la descomposición wavelet en otros campos de la ciencia.

Várady, en 2001 [17], propone un método de eliminación de ruido basado en wavelets para fonocardiografías. Este método usa señales de dos canales y una técnica adaptativa para el \textit{thresholding} de los coeficientes wavelet. Este método consiste en cuatro pasos importantes, donde primero se realiza la descomposición wavelet de ambos canales, luego se realiza la cancelación de ruido externo mediante el \textit{cross-channel thresholding} adaptativo de los coeficientes de la señal, para después hacer un proceso adaptativo de \textit{thresholding} a la señal resultante del proceso anterior. Finalmente, se reconstruye la señal deseada a partir de los coeficientes wavelet resultantes.
La eliminación de ruido mediante wavelets en combinación con diferentes técnicas, resulta ser útil para mejorar la extracción de la información, haciendo del denoising un proceso más robusto, así como lo presentan Aminghafari et al. [22] en el 2006, y posteriormente Yang et al. [23] en el 2011 quienes proponen un procedimiento para la eliminación de ruido en señales multivariadas donde combinan wavelets y el análisis de componentes principales (PCA). Este método exhibe un comportamiento promisorio en las señales estándar diseñadas por Donoho y Johnstone, además de desempeñarse bien en grabaciones multicanal. Por otro lado, Patil en [24] combina la transformada wavelet y la descomposición de valores singulares (SVD) para la eliminación de ruido y concluye que la unión de estos métodos reduce el error entre la señal original y la señal procesada.

Algunos autores han desarrollado métodos para establecer el número de niveles de descomposición wavelet adecuado para una señal, además de nuevas formas de calcular los valores de threshold para la eliminación de ruido. Cetin y Mohammad [25], en su nota de lectura “Projection-based Wavelet Denoising”, desarrollan un método de eliminación de ruido en el dominio wavelet que consiste en realizar proyecciones ortogonales de las señales wavelet de la señal ruidosa sobre una región en un espacio multidimensional con forma de pirámide invertida formada por planos horizontales llamados $L_1$-ball. Este método lleva a estimación automática de los valores de threshold para las señales wavelet. Aquí se presenta una mejora en la relación señal a ruido de una señal corrupta por ruido Gausiano blanco respecto a otros métodos implementados en Matlab como lo son el Minimaxi, Rigrsure y el Wavelet Multivariado descrito en [22].

Srivastava et al. [26] proponen un nuevo método de selección de niveles de descomposición y threshold para la eliminación de ruido en señales unidimensionales. Este método provee una forma de seleccionar el número de niveles de descomposición y usa una nueva fórmula para calcular el valor de threshold que no requiere la estimación del ruido, usa diferentes valores de threshold para coeficientes positivos y negativos, aplica denoising sobre los componentes de aproximación y permite ajustar los umbrales del ruido.

Jing-yi et al. [27] proponen una nueva función de threshold para optimizar la eliminación de ruido. En su trabajo exponen los problemas existentes con los métodos tradicionales de thresholding e introduce un factor de ajuste para construir la nueva función de threshold a partir del soft
thresholding. Este nuevo método de thresholding introduce una exponencial compleja en la expresión, la cual otorga mayor adaptabilidad y se ajusta a diferentes señales.

También existen aplicaciones en el dominio wavelet para la reducción de ruido en señales de audio como lo demuestra Ghanbari [28], quien propone un sistema de mejoramiento de una señal de voz basado en un thresholding adaptativo del paquete wavelet (WP), ya que la asunción del ruido Gausiano blanco propuesta en los algoritmos tradicionales de reducción de ruido basados en wavelets muestra algunos defectos en segmentos sin información, mala calidad para la audición, entre otros. El método evaluado demuestra mejoras en el realce de la voz. Por otro lado, Messaoud et al. [29] consideran la reducción de ruido sobre una señal de audio de un solo canal aplicando una técnica basada en la reconstrucción de los coeficientes de la transformada wavelet en combinación con el análisis de componentes principales (PCA). Con el fin de detectar zonas ruidosas en tiempo y frecuencia, se realiza una técnica de postprocesamiento basada en la descomposición subespacial. Este método se desempeña mejor sobre los métodos comparados en el documento.

El uso de técnicas wavelet para la reducción de ruido en señales bioacústicas se considera un avance emergente y poco común. Sin embargo, en 2007, Gur y Niezrecki [30] presentan un método de reducción de ruido para vocalizaciones de manatíes usando la transformada discreta wavelet redundante (UDWT) sobre señales corruptas por ruido de botes de amplio espectro, demostrando la gran capacidad de los métodos no lineales para mejorar la relación señal a ruido.

Posteriormente Ren et al. [15] proponen una nueva forma de eliminación de ruido en señales bioacústicas que puede ser usada en un amplio rango de especies. Esta metodología se basa en el paquete de descomposición wavelet (WPD) usando una función de escala Greenwood de especies específicas. Este nuevo acercamiento es comparado con otras técnicas que incluyen el filtrado básico pasa banda, la sustracción espectral, el filtrado Wiener, y el filtrado de Ephraim-Malat. La evaluación de estos métodos se realizó sobre vocalizaciones de diferentes especies con ruido Gausiano blanco y ruido ambiental, demostrando que el método propuesto tenía mejor desempeño sobre los otros métodos en comparación.
Uno de los trabajos más recientes de reducción de ruido en señales bioacústicas es presentado por Priyadarshani et al. [14], quienes describen un método de \textit{denoising} combinando el filtrado pasa banda y el paquete de descomposición wavelet. En este trabajo se hace uso de la entropía de Shannon para escoger el número de niveles de descomposición del (WPT), así como para la selección del \textit{threshold}. 
V. OBJETIVOS

A. Objetivo general

Evaluar técnicas de reducción de ruido basadas en wavelets para el procesamiento de señales bioacústicas.

B. Objetivos específicos

- Identificar diferentes técnicas wavelet para la reducción de ruido basándose en el desempeño que estas tengan sobre señales de un solo canal.
- Desarrollar un algoritmo de reducción de ruido basado en wavelets aplicado al tratamiento de señales bioacústicas.
- Evaluar el algoritmo desarrollado mediante su aplicación a un caso de estudio, usando una base de datos de diferentes especies de aves.
VI. MARCO TEÓRICO

A. Wavelet denoising

La Transformada Wavelet (WT) es una herramienta potente para el procesamiento de señales que ha sido ampliamente usada en diferentes campos de la ciencia gracias a su capacidad para analizar las características de una señal tanto en tiempo como en frecuencia. La WT es ideal para aplicaciones donde se analizan señales no estacionarias que pueden variar instantáneamente en el tiempo [31], [32], [33]. En comparación con la Transformada de Fourier de Tiempo Corto (STFT), que provee información en tiempo-frecuencia, la Transformada Continua wavelet (CWT) proporciona una representación en tiempo-escala, obteniendo un análisis más detallado de las señales en estudio [34], [35].

La CWT usa una función prototipo, también llamada wavelet madre, especialmente seleccionada considerando la señal en estudio para formar una expansión wavelet donde la señal es comparada con un diccionario wavelet o funciones base, construidas al dilatar y contraer la función prototipo. Esta transformación descompone la señal en diferentes escalas con diferentes niveles de resolución, donde el parámetro de escala está indirectamente relacionado con la frecuencia cuando se considera la frecuencia central del wavelet madre [36].

El wavelet madre debe tener un valor medio igual a cero para permitir que el kernel de transformación soporte de manera compacta la localización en tiempo, y así poder capturar los cambios instantáneos abruptos de la señal en periodos de tiempo corto [33], [37], [38].

Un diccionario wavelet $D$ se construye a partir de un wavelet madre $\psi$ el cual es dilatado mediante un parámetro de escala $s$ y desplazado mediante $u$ [39].

$$D = \left\{ \psi_{u,s} = \frac{1}{\sqrt{s}} \psi \left( \frac{t - u}{s} \right) \right\}_{u \in \mathbb{R}, s > 0}$$ (1)

De este modo, la CWT de una función $f$ se expresa como:
La expresión de Wavelet es una expansión bidimensional del tiempo y frecuencia. Esto implica que la mayor cantidad de energía presente en la señal será representada por unos pocos coeficientes [36]. Cuando $s$ se realiza en potencias de dos y $\psi$ tiene propiedades ortogonales, se le llama a este tipo de transformación como transformada Wavelet ortogonal diódrica o transformada discreta Wavelet (DWT) [40], [41]. Debido a estas propiedades de ortogonalidad, la DWT no presenta redundancias y provee el algoritmo para descomponer una señal en escalas con diferente resolución en tiempo-frecuencia [31], [37].

Una de las aplicaciones principales de la DWT es el proceso de denoising, el cual busca remover partes en la señal que se asumen como ruido [42], [43]. Toda señal que se obtiene de un medio físico, lleva consigo ruido que afecta la calidad de la información. Con los métodos de hard thresholding y soft thresholding propuestos por Donoho y Johnstone los proceso de wavelet denoising se presentan como gran herramienta para el procesamiento de señales de una y dos dimensiones.

### B. La transformada discreta wavelet

La DWT se implementa con un banco de filtros de reconstrucción perfecta usando una familia de wavelets ortogonales con el fin de descomponer la señal en sub-señales con diferente contenido en frecuencia [36]. Los wavelets más comunes que proporcionan las propiedades de ortogonalidad son los Daubechies, Symlets, Coiflets y Meyer discreto y producen una reconstrucción no redundante en algoritmos rápidos [44]. En su forma diédrica el diccionario wavelet $D$ se construye para potencias de dos para el escalamiento $s$, manteniendo el parámetro de desplazamiento $u$ [39].

$$ D = \left\{ \psi_{u,s} = \frac{1}{\sqrt{s}} \psi \left( \frac{t - u}{2^j} \right) \right\}_{u \in \mathbb{R}, s > 0} , $$

con $j = 1, 2, ..., L$.

El resultado de la DWT es una descomposición multinivel, en la cual la señal se descompone en coeficientes de aproximación y detalle para cada nivel en un proceso equivalente a filtrar mediante
filtros pasa bajos y pasa altos respectivamente [37]. La figura 1 ilustra la estructura de una DWT de dos niveles de descomposición.

En la figura 1 se puede observar el proceso de downsampling y upsampling para cada nivel característico de la DWT. Los filtros pasa bajos $L_j(\omega)$ y pasa altos $H_j(\omega)$ producen los coeficientes de aproximación $cA_j$ y detalle $cD_j$ de cada nivel $j = 1, 2, \ldots, L$ respectivamente. La franja gris en medio de la descomposición y reconstrucción representa cualquier proceso de modificación para los coeficientes.

Este proceso también es llamado codificación por sub-bandas y su distribución en frecuencia se puede observar en la figura 2 para un ejemplo de tres niveles de descomposición.
C. La transformada discreta wavelet no diezmada

La UDWT posee características similares a la DWT, sin embargo en su estructura la señal no pasa por un proceso de downampling en cada nivel de descomposición, esto introduce redundancia al no eliminar datos [45]. La figura 3 ilustra la UDWT para dos niveles de descomposición.

Al no realizar el proceso de downampling, la UDWT realiza un upsampling diádico de cada filtro en cada nivel de descomposición eliminando los efectos de aliasing [45].
**D. Thresholding**

Para eliminar el ruido usando los coeficientes resultantes de la descomposición de la señal, es necesario eliminar los coeficientes más pequeños, los cuales se asocian al ruido. Después de eliminar estos coeficientes, una estimación de la señal limpia $\hat{S}$ se obtiene mediante la reconstrucción. Usualmente se asume que el ruido tiene principalmente componentes de alta frecuencia, por lo que la eliminación de los coeficientes pequeños aplica generalmente en los coeficientes de detalle [36]. La idea principal en la eliminación de ruido mediante wavelets, consiste en estimar adecuadamente el nivel de ruido para realizar un proceso de *thresholding* usando este nivel.

Donoho y Johnstone proponen una estrategia no lineal de *thresholding* la cual se puede implementar mediante *hard* o *soft thresholding*, también llamada *wavelet shrinkage* [19]. En el *hard thresholding*, los coeficientes wavelet por debajo del nivel de *threshold* son llevados hacia cero, mientras que en el *soft thresholding* todos los coeficientes reducen su amplitud en proporción al nivel de *threshold*. En la figura 4 se ilustran los dos tipos de *wavelet shrinkage* los cuales se pueden expresar de forma analítica como

$$\delta_H = \begin{cases} \hat{w}_i = w_i & si \ |w_i| > \lambda \\ \hat{w}_i = 0 & si \ |w_i| \leq \lambda \end{cases}$$

(4)

$$\delta_s = \begin{cases} \hat{w}_i = w_i & si \ |w_i| \leq \lambda \\ \hat{w}_i = w_i - \lambda & si \ w_i \ > \lambda \\ \hat{w}_i = w_i + \lambda & si \ w_i \ < -\lambda \end{cases}$$

(5)

donde $\delta_H$ y $\delta_s$ son *hard* y *soft thresholding* respectivamente, $w_i$ es la señal wavelet de entrada, $\hat{w}_i$ es la señal wavelet después de aplicado el *thresholding* y $\lambda$ es el nivel de *threshold*. 
E. Consideraciones para el ruido

Krause [4] describe los componentes del paisaje sonoro representado por todos los sonidos en un ambiente particular captados por un receptor. Cada sonido puede clasificarse entre tres grandes grupos dependientes de la fuente acústica que los genera, estos son: Biofonía, Geofonía y Antrofonía.

- **La Biofonía** corresponde al grupo de sonidos generados por los organismos vivos en un territorio, este es el más complejo y amplio de los grupos. En biomas ricos en densidad y diversidad de voces de especies, los organismos estructuran acústicamente sus vocalizaciones en una relación cooperativa o competitiva donde cada voz es distinguible para reducir los efectos de enmascaramiento, esta relación recibe el nombre de la hipótesis de nicho.

- **La Geofonía** corresponde al grupo de los sonidos naturales emitidos por fuentes no biológicas en determinado hábitat. Este grupo se divide generalmente en cuatro subgrupos que corresponden al viento, agua, clima y fuerzas geofísicas.

- **La Antrofonía** es el grupo de sonidos producto de las actividades humanas, y pueden ser de cuatro tipos que son, electromecánicos, fisiológicos, sonidos controlados e incidentales.

Ya que este proyecto se enfoca en estudiar la actividad acústica biológica, los otros grupos que comprenden el paisaje sonoro son categorizados como ruido. Sin embargo, las grabaciones de
campo siempre contienen estos grupos de manera conjunta contaminando la información acústica deseada. En ocasiones el ruido puede enmascarar o distorsionar la señal deseada, haciendo más difícil su identificación.

A. Clases de ruido

Existe toda una teoría para el ruido en el área del procesamiento digital de señales, aquí se define el ruido como la señal no deseada que interfiere con la comunicación o medición de otra señal. Dependiendo de su espectro en frecuencia y sus características temporales, el ruido puede clasificarse como se explica detalladamente en [46] de la siguiente forma.

- **El ruido blanco** es ruido completamente aleatorio con función de autocorrelación caracterizada por un impulso y un espectro de potencia plano. En teoría el ruido blanco tiene igual potencia en todas las frecuencias.
- **El ruido blanco de banda limitada** es ruido con espectro en frecuencia plano que usualmente cubre la banda de donde se encuentra la señal de interés. La autocorrelación de este ruido tiene forma de función sinc.
- **El ruido de banda angosta** corresponde al ruido con un ancho de banda aproximado de 50 o 60 Hz. Generalmente el ruido inherente a los micrófonos tiene estas características.
- **El ruido de color** no tiene espectro de potencia uniforme y tiene energía generalmente decreciente en proporción con la frecuencia. Ejemplos de este tipo de ruidos son, el ruido rosa y el ruido café.
- **El ruido impulsivo** este ruido consiste en pulsos de corta duración con amplitud y tiempo de ocurrencia aleatoria.
VII. METODOLOGÍA

La metodología para la eliminación de ruido mediante la DWT o la UDWT se representa en la figura 5 donde una señal ruidosa $S$ se descompone para obtener los coeficientes wavelet, los cuales son modificados mediante una estrategia de thresholding para posteriormente construir una estimación de la señal limpia $\hat{S}$ mediante la transformada inversa correspondiente (IDWT o IUDWT).

![Fig. 5. Metodología basada en wavelets para la eliminación de ruido.](image)

Con base en la metodología descrita en la figura 5 se construyeron tres algoritmos cuyo desempeño se comparó junto con la función wden del Wavelet Toolbox de Matlab. Los algoritmos propuestos hacen uso de la UDWT para realizar la descomposición wavelet de las señales sin causar efectos de aliasing y presentan diferentes estrategias de thresholding, las cuales se explican en la sección “C”.

Para realizar un proceso de denoising exitoso mediante la DWT o la UDWT es necesario seleccionar un wavelet y determinar el número de niveles de descomposición apropiado que se ajusten al tipo de señales en estudio. De este modo, en la sección “A” se describe un procedimiento para seleccionar el wavelet y en la sección “B” se describe la estrategia implementada para estimar el número de niveles de descomposición de forma automática para las diferentes señales de prueba descritas en la sección “D”.

A. Selección del wavelet

Debido al downsampling en la descomposición de la DWT, es necesario usar un filtro wavelet apropiado que se ajuste a las características de la señal para producir una reconstrucción sin aliasing
Para seleccionar un wavelet apropiado, se realizó la descomposición de una señal de referencia (sin ruido) usando la DWT. Después de obtener los coeficientes wavelet, se eliminó el último nivel de aproximación y se reconstruyó la señal a partir de los coeficientes modificados. De este modo mediante inspección visual del espectrograma se determinó el tipo de wavelet que produce una mejor reconstrucción de la señal [14]. En el análisis se usaron los wavelets Haar, Daubechies 8, Daubechies 10, Symlet 8, Coiflet 5 y Meyer discreto. Algunos de ellos se muestran en la figura 6.

![Wavelets](image)

**B. Estimación del número de niveles de descomposición**

Para estimar el número adecuado de niveles de descomposición, se implementó un criterio de entropía de Shannon [14]. La entropía de Shannon proporciona una medición estándar de incertidumbre o desorden presente en un conjunto de datos y se define como [47]:
\[ ShEn = \sum_{i} s_i^2 \ln(s_i^2), \]  
(6)

donde \( s_i \) es la \( i \)-ésima muestra de la señal y se asume la convención \( 0 \log 0 = 0 \) [48].

Mediante este criterio se calcula el nivel de descomposición \( L \) analizando los coeficientes de aproximación ya que estos representan aproximadamente los rasgos característicos de la señal deseada como se muestra en la figura 7. Comenzando por \( L = 1 \), se realiza la descomposición mediante la DWT hasta que la \( ShEn \) en el nivel \( L + 1 \) sea mayor que la \( ShEn \) en el nivel \( L \), siendo \( L + 1 \) el nivel de descomposición estimado.

---

**C. Estrategias de denoising**

![Figura 7. Aproximaciones en una DWT de seis niveles de una señal ruidosa.](image-url)
Las siguientes estrategias de denoising fueron implementadas en Matlab en forma de funciones. De los métodos disponibles en Matlab para wavelet denoising se escogió el método de estimación de riesgo imparcial de Stein, el cual proporcionó un mejor desempeño. A continuación, se describen las estrategias implementadas.

1) Last approximation standard deviation (LAstd)

La estrategia LAstd usa la UDWT para realizar la descomposición multinivel de la señal para obtener los coeficientes de detalle y aproximación. La figura 8 ilustra la estrategia LAstd usando dos niveles de descomposición. La señal de entrada $S$ se descompone en $L$ niveles, donde los coeficientes de detalle $cD_j$ con $j = 1, 2, \ldots, L$ son soft thresholded con un nivel de threshold $\lambda$ el cual se calcula mediante la ecuación 7.

$$
\lambda = q\lambda_u \quad \text{con} \quad \lambda_u = \sigma \sqrt{\frac{2 \ln(N)}{N}}
$$

(7)

donde $\lambda_u$ es el threshold universal propuesto por Donoho [19], $\sigma$ es la desviación estándar de los coeficientes de aproximación $cA_L$, $N$ es el número de muestras en $S$ y $q$ es un parámetro de control que sirve para modificar el nivel de threshold, cuando $q > 0$. Para obtener una estimación de la señal deseada $\hat{S}$ los coeficientes de aproximación $cA_L$ se hacen cero y se realiza la reconstrucción de la señal mediante la IUDWT [17].

![Fig. 8. Estrategia LAstd con dos niveles de descomposición.](image)
2). Level dependent standard deviation (LDstd)

La estrategia LDstd, al igual que LAstd, usa la UDWT para realizar la composición multinivel de la señal. La figura 9 describe la estrategia LDstd con dos niveles de descomposición. De forma general, la señal $S$ se descompone en $L$ niveles donde los coeficientes de detalle $cD_j$ son *soft thresholded* aplicando un *threshold* nivel dependiente $\lambda_j$ con $j = 1, 2, \ldots, L$ el cual se calcula mediante la ecuación 8

$$\lambda_j = q \lambda_{ju} \quad \text{con} \quad \lambda_{ju} = \frac{\sigma_j \sqrt{2 \ln(N)}}{\ln(j + 1)},$$

donde $\lambda_{ju}$ es el *threshold* dependiente de cada escala $j$ [49], $\sigma_j$ es la desviación estándar de los coeficientes de detalle $cD_j$, $N$ es el número de muestras en $S$ y $q$ es el parámetro de control ya descrito. Al igual que en LAstd, los coeficientes de aproximación $cA_L$ se hacen cero para realizar la reconstrucción wavelet y obtener una estimación de la señal deseada $\hat{S}$ [17].

![Fig. 9. Estrategia LDstd con dos niveles de descomposición.](image)

3). Level dependent median absolute deviation (LDmad)

La estrategia LDmad comparte la misma estructura de LDstd a diferencia del *threshold* nivel dependiente $\lambda_j$, el cual se calcula mediante la ecuación 9.
\[ \lambda_{ju} = q \frac{MAD(|cD_j|)\sqrt{2\ln(N)}}{\ln(j + 1)}, \]  
(9)

donde \(MAD(| \cdot |)\) es el operador de la desviación media absoluta.

4. Método de Matlab (SURE)

Para hacer una comparación de las estrategias propuestas con un método disponible de denoising se usó la función \texttt{wden} del Wavelet Toolbox de Matlab, la cual usa la DWT para descomponer la señal en coeficientes de detalle y aproximación. La función \texttt{wden} permite realizar un proceso de denoising usando diferentes criterios de estimación del threshold. Los parámetros de la función se describen a continuación:

\[
XD = \texttt{wden}(X, \text{TPTR}, \text{SORH}, \text{SCAL}, N, 'wname')
\]

- \(X\) es la señal ruidosa.
- \(\text{TPTR}\) es la regla de selección del threshold.
  - '\texttt{rigrsure}' para el criterio de riesgo imparcial de Stein (SURE).
  - '\texttt{heursure}' para una versión heurística de SURE.
  - '\texttt{sqtwoLog}' para estimar el threshold con el método de Donoho propuesto en [20].
  - '\texttt{minimaxi}' para minimax thresholding [19].
- \(\text{SORH}\), 'h' para hard thresholding y 's' para soft thresholding.
- \(\text{SCAL}\) define la forma de reescalar el nivel de threshold.
  - 'one' para no reescalar
  - 'sln' para reescalar haciendo una única estimación del nivel de ruido basándose en los coeficientes del pimer nivel.
  - 'mln' para un thresholding nivel dependiente.
- \(N\) es el número de niveles de descomposición.
- '\texttt{wname}' es el filtro wavelet seleccionado.

Para seleccionar un criterio de estimación de threshold se realizó un proceso de denoising nivel dependiente sobre una señal de prueba con el fin de establecer el criterio con mejor desempeño.
mediante inspección visual del espectrograma. La figura 10 resume los resultados obtenidos al procesar la señal cuyo espectrograma se muestra en la figura 12.a.

\[\text{Fig. 10. Desempeño de la función wden con los diferentes criterios de estimación del nivel de threshold.}\]

De los criterios disponibles se seleccionó la estimación del \textit{threshold} de riesgo imparcial de Stein (SURE) nivel dependiente ya que los otros métodos eliminan información relevante en la señal. A diferencia de las estrategias propuestas, la función \texttt{wden} no cuenta con un parámetro de control \( q \) y para suavizar el proceso de \textit{denoising}.

\textbf{D. Señales de prueba}

Los algoritmos de reducción de ruido fueron probados con tres conjuntos de datos diferentes, todos con una frecuencia de muestreo de 44.1 kHz. El primer conjunto está compuesto por 10 cantos de diferentes especies de búhos con un rango en frecuencia que varía entre los 350 Hz y los 15 kHz como se puede observar en los espectrogramas de la figura 11. Estos cantos fueron tomados de \texttt{www.owling.com} y se les adicionó ruido con diferentes perfiles de manera controlada considerando el 10%, 20% y 30% de la amplitud de la señal limpia, formando así 10 grupos de datos que se resumen en la tabla I. Este conjunto de datos se construyó con el fin de poder evaluar la relación señal/ruido de las señales procesadas, teniendo conocimiento a priori de las señales limpias.
El segundo conjunto de datos está formado por 23 fragmentos (cada uno de un minuto de duración) de grabaciones bioacústicas registradas en dos ecosistemas colombianos brindadas por el grupo SISTEMIC de la Universidad de Antioquia. Este conjunto contiene señales de diferentes especies de aves, anuros, murciélagos e insectos, así como ruido ambiente consecuencia de la geofonía. Las 23 señales están divididas en dos grupos: siete señales denominadas Colustethus y 16 señales denominadas Norandinus, debido a las especies de anuros que se encuentran en ellas.

El tercer conjunto de datos está formado por 10 fragmentos de 5.9443 segundos de duración, extraídos del segundo conjunto de datos correspondientes a eventos con importante actividad bioacústica en diferentes partes del espectro. Los espectrogramas de estos eventos se muestran en la figura 12.

<table>
<thead>
<tr>
<th>Grupo</th>
<th>Tipo de ruido</th>
<th>Amplitud</th>
<th>Nombre del grupo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gaussiano blanco</td>
<td>10%</td>
<td>WGN 0.1</td>
</tr>
<tr>
<td>2</td>
<td>Rosa</td>
<td>10%</td>
<td>PNK 0.1</td>
</tr>
<tr>
<td>3</td>
<td>Browniano</td>
<td>10%</td>
<td>BRW 0.1</td>
</tr>
<tr>
<td>4</td>
<td>Gaussiano blanco</td>
<td>20%</td>
<td>WGN 0.2</td>
</tr>
<tr>
<td>5</td>
<td>Rosa</td>
<td>20%</td>
<td>PNK 0.2</td>
</tr>
<tr>
<td>6</td>
<td>Browniano</td>
<td>20%</td>
<td>BRW 0.2</td>
</tr>
<tr>
<td>7</td>
<td>Gaussiano blanco</td>
<td>30%</td>
<td>WGN 0.3</td>
</tr>
<tr>
<td>8</td>
<td>Rosa</td>
<td>30%</td>
<td>PNK 0.3</td>
</tr>
<tr>
<td>9</td>
<td>Browniano</td>
<td>30%</td>
<td>BRW 0.3</td>
</tr>
<tr>
<td>10</td>
<td>Banda angosta</td>
<td>10%</td>
<td>NBWGN</td>
</tr>
</tbody>
</table>
Fig. 11. Espectrogramas de los cantos sin ruido de las 10 especies de búhos en estudio.
Fig. 12. Espectrogramas de los eventos seleccionados de Colustethus y Norandinus.
F. Medidas de desempeño

Para evaluar las metodologías propuestas se usaron la relación señal/ruido y la entropía espectral.

1). Relación señal/ruido (SNR)

La relación señal/ruido en decibles se calcula con la expresión:

\[ SNR = 20 \log \left( \frac{\|S_{ref}\|}{\|S_{noise}\|} \right), \]  

donde \(S_{ref}\) y \(S_{noise}\) son la señal de referencia (señal sin ruido) y el ruido presente en la señal respectivamente, y \(||\cdot||\) es el operador de la norma Euclidiana.

2). Entropía espectral (SEN)

La entropía espectral es la entropía de Shannon (\(ShEn\)) de la densidad espectral de potencia normalizada y se calcula de la siguiente forma [50]:

\[ SEN = -\frac{\sum P_k \ln(P_k)}{\ln(N)}, \]

donde \(P_k\) es la potencia espectral normalizada, y \(N\) el número de bins de frecuencia.
A. Selección del wavelet

En la figura 13 se observan los efectos de aliasing que producen algunos de los wavelets usados en la metodología propuesta en la sección “A” del capítulo “VII” al procesar la señal representada en la figura 11.d. En el espectrograma de la figura 13.d se observa que el wavelet Meyer discreto obtuvo un mejor desempeño para este tipo de señales respecto a los otros wavelets.

Fig. 13. Efectos del aliasing que provocan algunos wavelets sobre la señal después realizar la DWT, aplicar un leve thresholding y reconstruir la señal nuevamente.

B. Denoising en cantos de búhos

Los resultados obtenidos para la SNR del primer conjunto de datos (cantos de los búhos) se condensan en los siguientes diagramas de barras, donde cada barra representa la mediana de la SNR de los 10 cantos de búhos para cada estrategia de denoising en contraste con la señal ruidosa de entrada. Los resultados obtenidos reportan la mediana de las cantidades ya que los datos no presentan una distribución normal (Gaussiana). Cada diagrama contiene los tres escenarios
propuestos para este conjunto de datos, donde la amplitud del ruido varía considerando el 10%, 20% y 30% de la amplitud de la señal original.

En la figura 14 se reportan los datos obtenidos para las muestras contaminadas con ruido Gaussiano blanco. Para las estrategias $LDstd$ y $LDmad$ se observa un aumento en la SNR al disminuir $q$. Los métodos en estudio mejoran la SNR respecto a las señales ruidosas, a diferencia del $LDstd$, el cual comienza a mejorar su desempeño cuando $q = 0.25$.

Los diagramas de barras de la figura 15 reportan los datos obtenidos al procesar las señales contaminadas con ruido rosa de diferentes amplitudes. Este procedimiento se realizó con el fin de probar la tolerancia de los algoritmos a diferentes perfiles de ruido. Para este escenario, los algoritmos propuestos mejoran la SNR para amplitudes de ruido mayores al 10% y al disminuir $q$ aumenta la SNR.
La figura 16 reporta los resultados obtenidos para ruido Browniano. Para este caso, el algoritmo \textit{LDmad} mejora la SNR de las señales ruidosas superando el desempeño de los demás.
Los siguientes diagramas de barras representan la mediana de la SEN del conjunto de los cantos de búhos para cada estrategia de *denoising* para cada tipo y amplitud del ruido. Este análisis evalúa el comportamiento relativo de cada estimación de la SEN de las señales procesadas en comparación con la SEN de la señal ruidosa. De este modo, un decremento en la SEN representa un aumento en la regularidad de los datos.

La figura 17 reporta la SEN para las señales con ruido Gaussiano blanco. Para \( q = 1 \) y \( q = 0.75 \) la SEN demuestra mayor organización de los datos procesados con las estrategias LAstd, LDstd y LDmad, indicando una menor presencia de componentes ruidosos. Sin embargo, con \( q = 0.25 \), la SEN no proporciona información concluyente para LDstd, LDmad y SURE.

Los diagramas de la figura 18 representan la SEN para las señales contaminadas con ruido rosa. Con \( q = 1 \), y amplitud del ruido del 30%, la estrategia LAstd presenta una SEN igual a cero, lo que indica que el método elimina toda la información en las señales de este grupo. Los métodos LDstd y LDmad producen datos más organizados para \( q \) mayores a 0.25.
Para el caso que se muestra en la figura 19 el algoritmo $LAstd$ elimina toda la información en las señales con ruido del 20% y 30% de amplitud con $q = 1$ y 30% con $q = 0.75$. 

Fig. 19. Mediana de la SEN de las señales contaminadas con ruido browniano de diferentes amplitudes, para diferentes valores de $q$. 

Fig. 18. Mediana de la SEN de las señales contaminadas con ruido rosa de diferentes amplitudes, para diferentes valores de $q$. 

(a) SEN con $q = 1$

(b) SEN con $q = 0.75$

(c) SEN con $q = 0.5$

(d) SEN con $q = 0.25$
Para los casos de las figuras 18.a, 19.a y 19.b, la SEN indica que el algoritmo $LA_{std}$ tiene poca tolerancia para señales con ruido de color.

**C. Denoising en los registros de ecosistemas colombianos**

En los conjuntos Colustethus, Norandinus, y por consiguiente en los eventos extraídos, no se cuenta con una estimación de la potencia del ruido presente en las grabaciones, por lo que solo se registra la mediana de la SEN en los diagramas de barras de las figuras 20 y 21.

![Fig. 20. Mediana de la SEN de los eventos bioacústicos.](image)
En las figuras 20 y 21 se logra observar que todos los métodos propuestos producen una SEN menor. Por lo tanto, la organización de los datos es mayor, indicando menor presencia de componentes ruidosos.
IX. DISCUSIÓN

De forma visual se pueden evaluar los resultados de las estrategias de denoising usando el espectrograma normalizado a 0 dB. En la figura 22 se visualizan los resultados al procesar con \( q = 1 \), una señal de la especie denominada como Northern Saw-whet del grupo WGN 0.1 cuya señal limpia se muestra en la figura 11.g. La figura 22.a demuestra el excelente desempeño de la metodología \( \text{LAstd} \) para esta señal. Por otro lado en las figuras 22.b y 22.c se evidencia pérdida de información debido a una inapropiada estimación del nivel de ruido. En la figura 22.d se observa un desempeño regular de la función \( \text{wden} \) (SURE) en comparación con los otros métodos, además de los efectos de aliasing causados por el uso de la DWT, generando distorsión indeseada.

El diagrama de barras de la figura 14.a reporta un desempeño en la SNR de la estrategia SURE que supera los demás algoritmos propuestos. Sin embargo, mediante inspección visual de los espectrogramas se puede confirmar el desempeño de los algoritmos para casos específicos. Otro criterio indicador del desempeño de los algoritmos es la SEN, con la cual se puede estimar la organización de los datos en el espectro de potencia. En los diagramas de barras de las figuras 17 a
19, los valores menores para SEN representan una mayor organización de los datos, indicando menor contenido de componentes ruidosos en la señal.

En señales con mayor contenido en frecuencia, los algoritmos LAstd y LDstd ocasionan perdida de información para un $q = 1$ como se observa en la figura 23, la señal limpia se muestra en la figura 11.d. De este modo componentes armónicos con baja amplitud son considerablemente afectados por el proceso de denoising. Por otro lado, las estrategias LDmad y SURE (figuras 23.c y 23.d) muestran un mejor desempeño cuando se cuenta con señales de este tipo.

Para contrarrestar la perdida de información en frecuencia, los algoritmos LAstd, LDstd y LDmad pueden ajustarse mediante su parámetro $q$ para suavizar el proceso de thresholding a costo de disminuir la SNR. La figura 24 demuestra el efecto de usar $q = 0.25$ para una señal del grupo WGN 0.1 (su señal limpia se representa en la figura 11.i).
En la figura 25 se demuestra la tolerancia de los algoritmos propuestos a una amplitud del 30% considerando el valor máximo en la señal. A pesar de que el proceso de denoising de los algoritmos propuestos produce perdida de información, el algoritmo $L_{std}$ proporciona un procesamiento sobresaliente respecto a los otros, para el caso de esta señal correspondiente a la especie de búho *Western Screech* (figura 11.b).

En la figura 25 se demuestra la tolerancia de los algoritmos propuestos a una amplitud del 30% considerando el valor máximo en la señal. A pesar de que el proceso de denoising de los algoritmos propuestos produce perdida de información, el algoritmo $L_{std}$ proporciona un procesamiento sobresaliente respecto a los otros, para el caso de esta señal correspondiente a la especie de búho *Western Screech* (figura 11.b).
Al probar los algoritmos con las señales del grupo NBWGN (señales contaminadas con ruido Gaussiano blanco de banda angosta), la estrategia LAstd presentan dificultades para estimar el nivel de ruido ya que, en su estructura, el nivel de ruido se estima de la última aproximación correspondiente a la banda más baja $cA_L$. Al no presentar contenido ruidoso en esta banda, no se estima apropiadamente el nivel del ruido y el proceso de denoising presenta un desempeño bajo respecto a las otras estrategias, las cuales presentan tolerancia a este tipo de ruido. En la figura 26 se representa este caso un canto correspondiente a la especie de búho Northern Pygmy (figura 11.f).
Las señales proporcionadas por el grupo SISTEMIC presentan una mayor densidad poblacional de especies. Estas muestras contienen señales bioacústicas generadas por diferentes especies de aves, ranas, murciélagos e insectos que interactúan en un mismo ecosistema. De este modo, se pone a prueba la capacidad de los algoritmos para eliminar ruido ambiente en un caso de estudio real.

Para establecer un mejor análisis del desempeño de los algoritmos planteados en el caso de los eventos bioacústicos representados en la figura 12, es necesario el análisis visual de espectrograma ya que la SEN no proporciona información sobre la degradación que las señales pueden sufrir en el proceso de denoising.

El desempeño de los algoritmos para este tipo de señales varía de acuerdo con las características de las mismas. De manera similar al caso de la figura 26, el desempeño de $LA_{std}$ para el caso representado en la figura 17 presenta dificultades para hacer un proceso de denoising adecuado ya que el ruido presente en esta señal tiene mayor energía en la banda que se encuentra entre los 5 kHz y 10 kHz. Por otro lado las estrategias $LD_{std}$ y $LD_{mad}$ muestran su tolerancia a este tipo de
El ruido, eliminando casi por completo los componentes ruidoso (la señal original se representa en la figura 12.h).

En casos donde el ruido no presenta una distribución normal (Gaussiana) a lo largo de todo el espectro (p.ej., la geofonía en un ecosistema), es necesario ajustar el parámetro $q$ para encontrar un balance entre una SNR y la degradación de la señal, cuando se busca continuar con una etapa posterior de procesamiento. En la figura 28 se usa $q = 0.3$ para aumentar la SNR sin eliminar componentes importantes de la señal representada en la figura 12.a.
En la figura 28 también se puede observar que los algoritmos LDstd y LDmad logran conservar componentes bioacústicos de poca amplitud en altas frecuencias, a diferencia de la estrategia LAstd, la cual elimina estos componentes. Por otro lado, la estrategia SURE conserva algunos de estos componentes, pero agrega distorsión y aliasing a la señal.

Es importante destacar que los algoritmos LDstd y LDmad se adaptan mejor a señales contaminadas con ruido ambiente, ya que ajustan el nivel de threshold para cada nivel de detalle (cada nivel es tratado de forma independiente). Por otro lado, el parámetro $q$ proporciona adaptabilidad a los algoritmos para procesar señales en distintos escenarios donde se desconoce el perfil de ruido.
X. CONCLUSIONES

- Cuando se procesan señales contaminadas con ruido aditivo Gaussiano Blanco (para el caso de los búhos), la estrategia LAstd permite una estimación muy aproximada de la señal original ya que el proceso de thresholding se realiza con base en la última aproximación donde todos los coeficientes representan componentes ruidosos. Esto permite una adecuada estimación del nivel de ruido para todos los niveles de detalle.
- En casos donde se cuenta con señales con ruido ambiente, el parámetro $q$ proporciona versatilidad a los algoritmos propuestos de denoising para adaptarse a diferentes escenarios donde se desconoce el perfil del ruido. En esta investigación se establecieron valores de $0 < q < 1$ ya que para valores mayores a 1 ocurriría perdida de información relevante. Este grado de libertad permite, además, enfocar los algoritmos para procesar señales de especies específicas. Un ejemplo de ello se representa en las figuras 27.b y 28.c donde un $q = 1$ revela los cantos predominantes en esta señal, mientras que un $q = 0.3$ conserva cantos con menor influencia.
- En cuanto a la valoración del desempeño de los algoritmos, el análisis visual del spectrograma permite identificar los casos donde la señal ha sido distorsionada o se ha perdido información debido al proceso de denoising, esto complementa el análisis cuantificado mediante la SNR y la SEN.
- Una de las deficiencias más evidentes de la estrategia LAstd resulta al procesar señales con ruido de banda angosta. Como se explica en el capítulo IX. La estrategia no logra estimar un nivel de ruido apropiado debido a la falta de información en el último nivel de aproximación. A pesar de lo anterior, esta estrategia se desempeña bien cuando se cuenta con información ruidosa en el último nivel de aproximación (banda más baja).

A. Limitaciones

- La entropía espectral (SEN) permite cuantificar el orden de los datos en la densidad espectral de potencia de la señal, sin embargo, esta medida no proporciona información relevante cuando en el proceso de denoising se pierde información o se distorsiona la señal. Para proporcionar una cuantificación más confiable es necesario implementar otros criterios que acompañen la medida que logren diferenciar la calidad del procesamiento.
- Estimar la relación señal/ruido (SNR) es un proceso simple cuando se cuenta con la señal en su forma aislada. Sin embargo, cuando se procesan señales grabadas en campo, cuantificar la SNR es un procedimiento complejo ya que la señal está embebida en un perfil de ruido desconocido. Debido a esto, es necesario buscar alternativas que logren cuantificar numéricamente la reducción en la amplitud del ruido respecto a la señal deseada.

- Para obtener medidas significativas del desempeño de los algoritmos propuestos, es necesario contar una base de datos más extensa que permita evaluar señales con características diferentes a la procesada en este proyecto. Sin embargo, la base de datos utilizada permitió evidenciar un desempeño promisorio que debe ser validado en estudio futuros.

B. Trabajos futuros

- Plantear la integración de los métodos de denoising con otras etapas de grabación y procesamiento. Un acercamiento a considerar, sería el uso de técnicas de grabación multicanal de señales en campo que permitan una mejor discriminación del paisaje sonoro y etapas de extracción de características de las señales.

- Plantear la implementación de mejores estrategias de evaluación de los algoritmos propuestos cuando se cuenta con señales con ruido ambiente.

- Implementar diferentes técnicas de thresholding y estimación del nivel de ruido para procesos de denoising robustos y automáticos.

C. Productos

- Participación en el Workshop on Engineering Applications (WEA 2018) cuyas actas de congreso serán publicadas en la serie Communications in Computer and Information Science con memorias indexadas en Scopus.

- Registro del software BWD (Bioacoustic wavelet denoising) basado en las estrategias LAstd, LDstd y LDmad.
REFERENCIAS


