

DESCOMPOSICIÓN DE SEÑALES BIOACÚSTICAS USANDO FAMILIAS DE POLINOMIOS ORTOGONALES

DECOMPOSITION OF BIOACUSTIC SIGNALS USING FAMILIES OF ORTHOGONAL POLYNOMIALS

Luis Eduardo Tobón¹. Carlos Andrés Posada²

¹ Departamento de Ingeniería Electrónica y Ciencias de la Computación Pontificia
Universidad Javeriana Cali. E-mail: letobon@javerianacali.edu.co

² Departamento de Ingeniería Electrónica y Ciencias de la Computación Pontificia
Universidad Javeriana Cali. E-mail: carlosaposada2002@gmail.com

Recibido: 3 Diciembre

Acceptado: 15 Diciembre

*Correspondencia del autor. Departamento de Ingeniería Electrónica y Ciencias de la Computación Pontificia Universidad Javeriana Cali.
E-mail: carlosaposada2002@gmail.com

RESUMEN

Este trabajo presenta una técnica de descomposición de señales bioacústicas empleando como funciones base dos familias de polinomios ortogonales. Esta técnica está orientada a implementarse en dispositivos de bajas capacidades de cómputo, como redes de sensores inalámbricos, por lo tanto debe ser rápida y eficiente, además de permitir una alta compresión de la información y una buena calidad en la reconstrucción. En este artículo se explica el fundamento de la descomposición usando familias de polinomios ortogonales, se muestran el procedimiento empleado para alcanzar la resolución adecuada, que permite una mayor compresión de la información y una mejor calidad en la reconstrucción de la señal. Por último, se muestran los resultados obtenidos al aplicar el método a funciones generadas artificialmente y a registros sonoros del páramo de Chingaza (Cundinamarca) suministrados por el instituto Alexander von Humboldt, cuyos resultados se comparan con la transformada de Fourier tradicional.

Palabras claves: Polinomios Ortogonales, Descomposición de Señal, Señales bioacústicas, Calidad de la reconstrucción.

ABSTRACT

This study presents a decomposition technique of Bio-acoustic signals using two families of orthogonal polynomials as basis functions. This technique is oriented to be implemented in low computing capability devices such as wireless sensor networks. Therefore, this technique should be quick and efficient and should allow high compression of information and a good quality in its reconstruction. In this article the foundation of decomposition using orthogonal polynomials families is explained. The procedure used is to achieve adequate resolution, which allows greater information compression and better quality in the reconstruction of the signal. Finally, the results obtained by applying the method to functions generated artificially and to sound recordings at Chingaza moor will be (Cundinamarca) exposed. Those sound recordings were provided by the Alexander von Humboldt Institute. The results are compared with the traditional Fourier Transform.

Palabras claves: Orthogonal polynomials , decomposition of Signals, Bioacoustic signals , quality reconstruction.

INTRODUCCIÓN

El Instituto Alexander von Humboldt monitorea y estudia la conservación ambiental en el territorio nacional mediante estaciones de observación. Estas estaciones se encuentran distribuidas en los diferentes ecosistemas de Colombia como: El páramo de Chingaza, la sierra nevada de Santa Marta, los bosques húmedos de Gorgona, Magdalena Medio y Chocó. Los cambios climáticos y la pérdida de cobertura vegetal han generado la necesidad de evaluar los cambios en la biodiversidad de estos ecosistemas (1). Una de las estrategias más útiles para detectar cambios en la biodiversidad es realizar estudios sobre el comportamiento de las aves, debido que estas se encuentran distribuidas en gran parte de los ecosistemas (2, 3, 4). Particularmente el territorio colombiano ofrece unas condiciones de megadiversidad únicas en el mundo, con más 162 especies de aves entre endémicas y casi endémicas, más de 266 especies acuáticas entre migratorias y residentes (5), convirtiendo a Colombia en el país con mayor diversidad de avifauna con más de 1989 especies de aves identificadas hasta enero del 2013(6). Estas condiciones han generado motivación para realizar estudios sobre el estado de conservación de las aves de Colombia. Tradicionalmente esos estudios se han realizado por medio de complejas técnicas de observación invasivas y no invasivas (7, 8); sin embargo en los últimos años han tomado gran importancia los estudios de conservación usando señales bioacústicas (9, 10, 11, 12, 2), para realizar detección, clasificación y estudios de conservación principalmente de aves y anfibios.

Muchos estudios sobre el comportamiento bioacústico de las aves utilizan técnicas de descomposición y extracción de características clásicas basadas en la transformada de Fourier (11, 9), sin embargo la transformada de Fourier no genera buenas representaciones de las vocalizaciones y llamados de algunas especies animales(2). Esto ha impulsado el uso de técnicas alternativas como descomposición en funciones wavelet (13) o análisis por series de tiempo (12).

Este artículo presenta una técnica de descomposición de señales utilizando diferentes familias de polinomios ortogonales, que está orientada a la caracterización de vocalizaciones de aves y caracterización bioacústica de los diferentes ecosistemas. Las secciones 2, 3 y 4 presentan la descripción general del método y su fundamentación matemática, la secciones 5 y 6

explican los análisis empíricos para afinar el funcionamiento de la descomposición y asegurar una mejor comprensión de la información respectivamente y finalmente las secciones 7 y 8 presentan el análisis de los resultados al compararlos con la descomposición en series de Fourier tradicional y las conclusiones del trabajo respectivamente.

2. Descomposición Ortogonal

Suponga una familia de funciones base $\mathbf{S}_n = \{\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_n\}$ ortogonales entre sí, donde cada función es representada por un vector de longitud L_p ; suponga también una señal representada por el vector X de longitud finita L_s .

Es posible encontrar una matriz $\mathbf{M}_{n \times m}$ donde $m = \frac{L_s}{L_p}$ y n es el número de funciones en la familia

\mathbf{S}_n . El m -ésimo coeficiente está dado por la siguiente ecuación.

$$M_{nm} = \int_{i=1}^{L_p} \{n(i)X[Im(i)]\} \quad (1)$$

donde I_m es un intervalo de la señal $X[n]$ y está definido como

$$I_k = kL_p, \dots, (k+1)L_p + 1 \quad k = 1, 2, \dots, \frac{L_s}{L_p} \quad (2)$$

Por otro lado, si se conoce la matriz $\mathbf{M}_{n \times m}$ y la familia de funciones base \mathbf{S}_n es posible reconstruir una aproximación $X^o[n]$ a la señal original por medio de la siguiente expresión.

$$X^o[Im] = \sum_{i=1}^n Z^i M_{im} \quad (3)$$

donde M_{im} corresponde a la i -ésima fila de la matriz $\mathbf{M}_{n \times m}$ y φ_i corresponde a la i -ésima función de la familia \mathbf{S}_n .

3. Error de Reconstrucción.

La descomposición ortogonal es una herramienta, que ayuda entre otras cosas, a extraer información importante de una señal, eliminar información poco relevante y reconstruir una versión de la señal aproximada a la original. Sin embargo las aproximaciones computacionales y la reducción en la cantidad de la información introduce cierto error en la señal reconstruida que debe ser cuantificado. Teniendo en cuenta lo anterior es importante determinar el error en la reconstrucción, para esto utilizamos la siguiente ecuación.

$$E = \frac{\sum_{i=1}^{L_y} (X[i] - X'[i])^2}{\sum_{i=1}^{L_y} X^2[i]} * 100 \quad (4)$$

que representa el error cuadrático medio de la señal normalizado respecto a la potencia de la señal original

4. Familias Ortogonales

Una familia ortogonal se define como un grupo de funciones que obedecen entre si la relación de ortogonalidad de la ecuación 5 sobre un dominio definido y bajo una función de peso $w(x)$.

$$\int_{\Omega} \phi_n(x) \phi_m(x) w(x) dx = \delta_{mn} C_n \quad (5)$$

La función de peso $w(x)$ tiene como función normalizar los polinomios bajo el dominio Ω , ϕ_n y ϕ_m corresponden a polinomios del orden m y n respectivamente y δ_{mn} es el delta de Kronecker y cuando C_n tiene el valor de 1, ϕ_n y ϕ_m se consideran ortonormales entre sí.

Para este trabajo se han seleccionado dos familias ortogonales como funciones base para la descomposición: polinomios de Legendre y funciones armónicas (14).

4.1. Polinomios de Legendre

Uno de los conjuntos de ortogonales más comunes son los polinomios de Legendre (15, 16, 14). Este conjunto es ortogonal en un dominio definido entre $D = [-1,1]$ con respecto a la función de peso $w(x) = 1$. Por definición los polinomios de Legendre requieren que $P_0 = 1$ y $P_1 = x$ (14).

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x) \quad (6)$$

La ecuación anterior se utiliza para generar los polinomios de Legendre de forma recursiva. Haciendo uso de esta expresión es posible obtener polinomios ortogonales de orden 1, 3, 5 y 7 como los que se observan en la Figura 1.

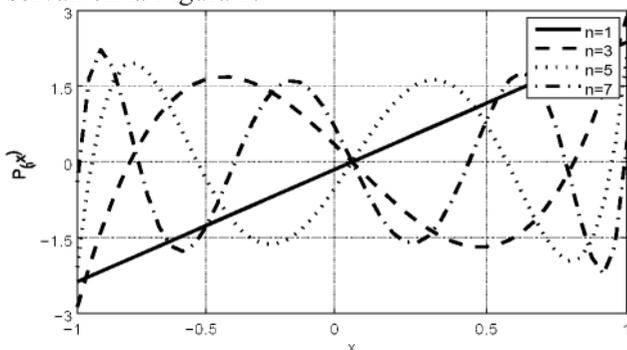


Figura 1: Polinomios de Legendre de orden 1, 3, 5 y 7.

4.2 Funciones Armónicas

El otro conjunto ortogonal seleccionado para este trabajo son las funciones armónicas, es decir un conjunto de funciones que pueden ser seno o coseno según el número de ceros deseados en un dominio definido $[-1,1]$. La siguiente ecuación muestra la forma de generar este conjunto ortogonal.

$$G_n(x) = \begin{cases} \cos\left(\frac{2}{n}nx\right) & \text{si } n \text{ es par} \\ \sin\left(\frac{2}{n}nx\right) & \text{si } n \text{ es impar} \end{cases} \quad (7)$$

Las funciones armónicas no son una familia de polinomios ortogonales como tal, sin embargo obedecen la relación de ortogonalidad (14, 15). En la figura 2 se observan algunas de las funciones armónicas para $n = 1, 2, 3$ y 4 que utilizamos para la descomposición ortogonal.

5. Resolución de la Descomposición

Con el fin asegurar el buen funcionamiento del método es necesario definir la resolución a la que este trabajará. Dicha resolución está dada tanto el número de vectores en la familia S_n como por la longitud L_p de esos vectores. Debido a esto se deben encontrar los valores de n y L_p adecuados para extraer la información necesaria sin incrementar el costo computacional y el error de reconstrucción. Para encontrar esta resolución se generó una señal artificial que se descompuso variando la longitud (L_p) de los vectores en S_n para cada familia ortogonal dando diferentes valores a n . Los resultados obtenidos al utilizar como S_n los polinomios Legendre y las funciones armónicas se observan en las Figuras 3a y 3b.

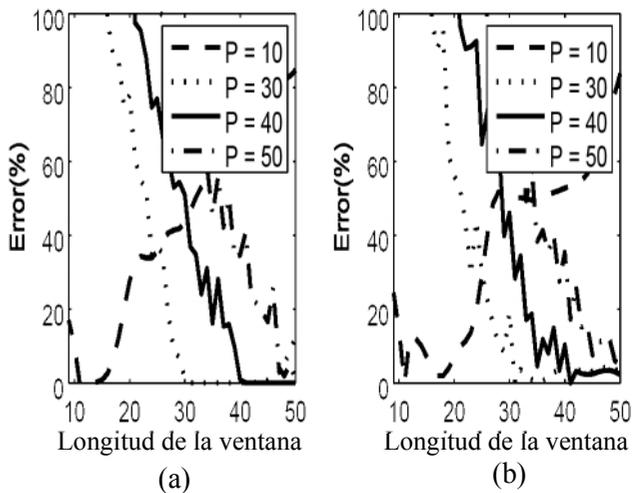


Figura 3: (a) Resolución para Polinomios de Legendre, (b) Resolución para Funciones Armónicas.

Como se observa en la figura 3, las dos familias ortogonales obtienen su mejor rendimiento cuando $L_p = n + 1$, es decir cuando la longitud de los polinomios es igual al número de polinomios en la familia S_n .

Para el buen funcionamiento del método se seleccionó una resolución de $L_p = 41$ y $n = 40$, estos valores aseguran una buena calidad en la reconstrucción, un bajo costo computacional y permiten extraer información de la señal (Figura 4a). Valores inferiores en el orden de polinomio obtienen un mejor error de reconstrucción y un menor costo computacional, mas no permiten extraer información de la señal, como se observa en la Figura 4b.

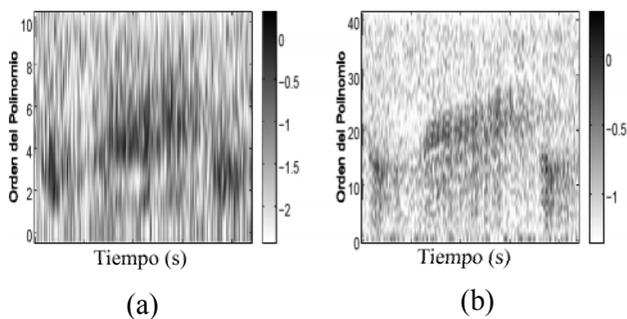


Figura 4: (a) Descomposición de señal con resolución 10x11. (b) Descomposición de señal con resolución 40x41.

Es importante mencionar que para este estudio asumimos que todas las señales fueron muestreadas a 44100kHz que es el estándar para señales de audio y esta es la frecuencia a la que se encuentran digitalizados los registros suministrados por el instituto Alexander von Humboldt.

6. Compresión de la Información

El objetivo de este procedimiento es reducir la cantidad de información disponible, aproximando a cero los coeficientes con valores muy pequeños en comparación con el resto de la matriz. Sin embargo para evitar que esto traiga un incremento significativo en el error de reconstrucción, la energía de los coeficientes eliminados se distribuye en los coeficientes que se conservan.

El criterio para decidir si un coeficiente es descartado está definido como $|M_{nm}| < K \times \max(|\mathbf{M}_{n \times m}|)$, es decir un coeficiente es descartado si es menor a K veces el valor del mayor coeficiente de toda la matriz y su energía será distribuida equitativamente entre los coeficientes de su columna que no fueron descartados. Análisis empíricos nos permitieron determinar que el

valor óptimo de K es de 0,05.

7. Resultados

Para determinar la calidad del método propuesto se analizaron tres factores: Densidad de la matriz de coeficientes $\mathbf{M}_{n \times m}$, error de reconstrucción y el tiempo que tarda en realizarse la descomposición y compresión. Es importante mencionar que el tiempo se midió usando la función *cputime de Matlab* en un computador con 8GB de memoria RAM y un procesador intel core i5 a 2,5GHz.

7.1. Señal Artificial.

La señal artificial $X[t]$ la definimos por las siguientes ecuaciones

$$X[t] = g^{t1}(t)S^{\omega1}(t) + g^{t2}(t)S^{\omega2}(t) \quad (8)$$

donde

$$g^{tn}[t] = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-t_n)^2}{2\sigma^2}} \quad (9)$$

$$S^{\omega n}(t) = \sin(\omega_n t) \quad (10)$$

Seleccionamos esta señal para observar la capacidad de cada familia ortogonal de representar por separado dos eventos de frecuencias diferentes ($\omega1$ y $\omega2$) que ocurren en un mismo instante de tiempo. La figura 5 muestra la señal en el dominio del tiempo y de la frecuencia.

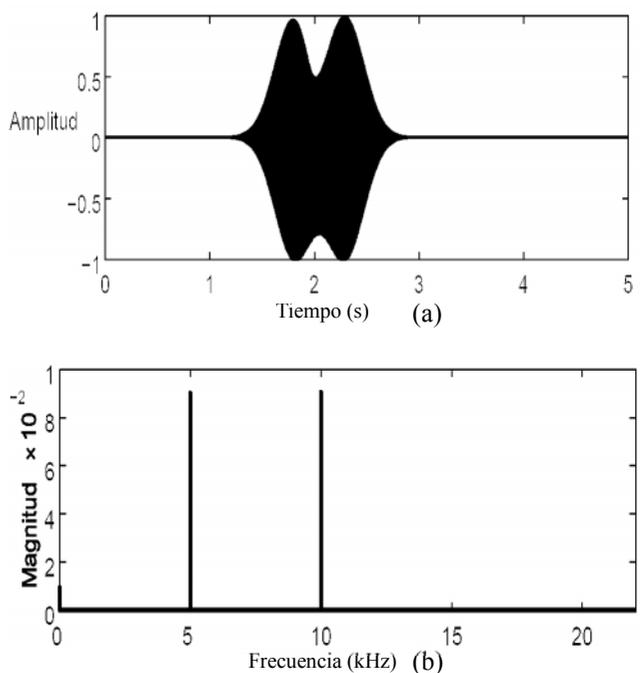


Figura 5: (a) Señal artificial en el dominio del tiempo. (b) Señal artificial en el dominio de la frecuencia.

Al realizar la descomposición por polinomios de Legendre, funciones armónicas y coeficientes de Fourier por el método tradicional se obtienen las representaciones que se observan en la Figura 6.

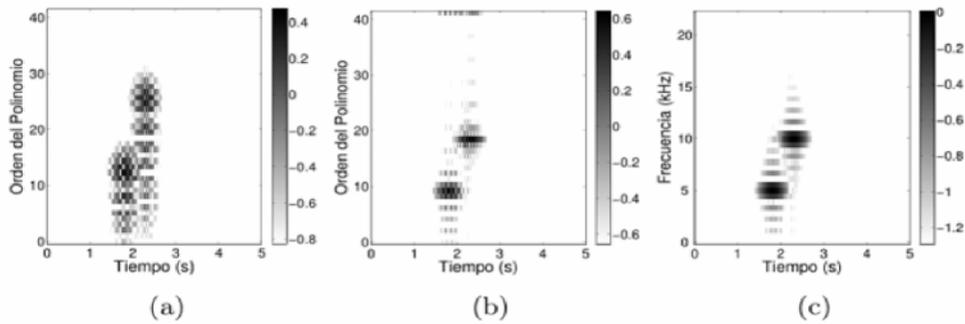


Figura 6: (a) Polinomios de Legendre, (b) Funciones Armónicas, (c) Coeficientes de Fourier.

A partir de estas representaciones se obtuvieron los resultados que se observan en la Tabla 1, en la cual se compara el rendimiento de las dos familias ortogonales con el método tradicional de Fourier.

Tabla 1: Resultados para descomposición ortogonal con compresión del 0.05 sobre la señal generada artificialmente.

Familia	Densidad	Error	Tiempo
Polinomios de Legendre	10.28%	0.75%	1.80 s
Funciones Armónicas	5.79%	2.60%	1.69 s
Coeficientes de Fourier	6.65%	2.69%	117.08 s

Se observa un bajo error para los polinomios de Legendre, así como un bajo tiempo de procesamiento. Las funciones armónicas, como era de esperar, tienen un desempeño similar a los coeficientes de Fourier en densidad y error, pero el procesamiento tarda menos del 90% cuando se utilizan funciones armónicas que cuando se utiliza la transformada de Fourier.

7.2. Registro del Páramo de Chingana

Esta señal corresponde a un registro grabado por el instituto Alexander von Humboldt en el páramo de Chingaza el 5 de febrero del 2013 a las 6:00 am. El registro tiene una duración total de 3:00 minutos. Sin embargo para la visualización en la Figura 8 se seleccionó un segmento de 0.44s de duración ya que contiene el canto de un ave. No obstante los resultados registrados en la Tabla 2 corresponden al análisis del registro completo. En la Figura 7 se observa la señal tanto en el dominio del tiempo como en el de la frecuencia.

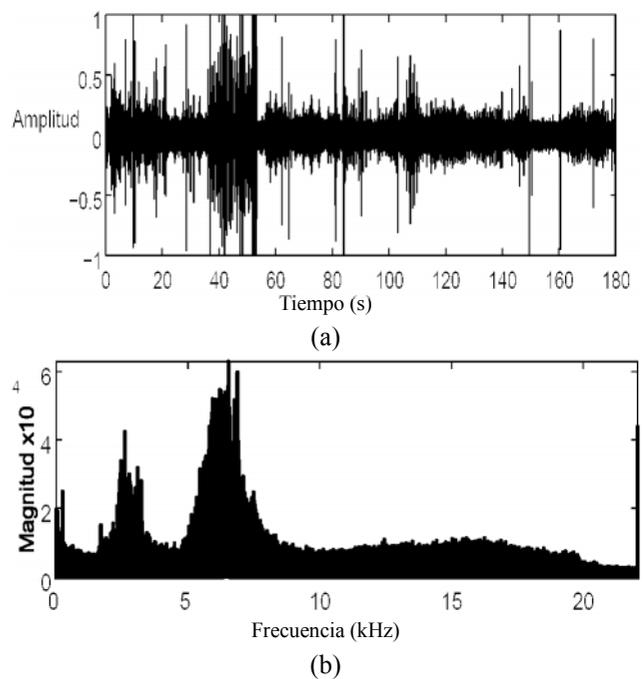


Figura 7: (a) Registro de Chingaza en el dominio del tiempo. (b) Registro de Chingaza en el dominio de la frecuencia.

Al igual que para la señal artificial obtuvimos representaciones de la señal al descomponerla usando el método tradicional de Fourier y las familias ortogonales de Legendre y armónicas. Dichas descomposiciones se observan en la Figura 8.

Partiendo de estas representaciones se obtienen los resultados mostrados en la tabla 2.

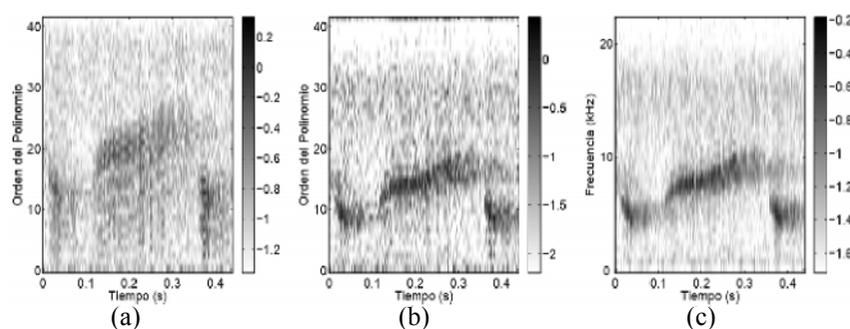


Figura 8: (a) Polinomios de Legendre, (b) Funciones Armónicas, (c) Coeficientes de Fourier.

Tabla 2: Resultados para descomposición ortogonal con compresión del 0.05 sobre un segmento del páramo Chingaza.

Familia	Densidad	Error	Tiempo
Polinomios de Legendre	1.53%	30.10%	51.80
Funciones Armónicas	0.64%	33.19%	51.05
Coeficientes Fourier	1.09%	31.04%	548.36

8. Conclusiones

Las familias de funciones ortogonales como los polinomios de Legendre o las funciones Armónicas se pueden considerar funciones base que generar una buena representación tanto de señales generadas artificialmente como de señales bioacústicas adquiridas de ambientes naturales. Las tablas 1 y 2 muestran que el método propuesto es cerca del 90% más rápido que la representación por series de Fourier tradicional sin incrementar considerablemente el error de reconstrucción y en algunas ocasiones presenta mejoras en este parámetro.

El método propuesto presenta una buena capacidad de compresión de la información sin comprometer de forma significativa el error de reconstrucción, esto abre paso a posibles aplicaciones en caracterización de ecosistemas, caracterización de especies, detectores y clasificadores de sonidos ambientales, medición de índices acústicos entre otros. La sección 3 muestra la posible existencia de una relación entre los diferentes órdenes de polinomios en la matriz S_n y los componentes de frecuencia presentes en la señal. En esta sección también se observa una posible relación entre la resolución de la descomposición y la frecuencia de muestro de la señal que aún debe ser demostrada.

Aunque los polinomios de Legendre presentaron un mejor error de reconstrucción en todos los casos las funciones armónicas presentaron una menor densidad en la matriz resultante y una mejor representación (figuras 6 y 8), cada familia de funciones ortogonales

presenta ventajas y desventajas que se deben tener en cuenta dependiendo de la aplicación en la se deseen utilizar.

9. Trabajo Futuro

Como parte del proyecto de investigación en este momento nos encontramos realizando pruebas finales a un detector de vocalizaciones de aves sobre registros del páramo Chingaza. Este detector se basa en la teoría de detección binaria y en reglas fijas de combinación para generar tantos detectores como número de polinomios existen en la matriz S_n y encontrar la combinación que permite realizar una mejor detección.

También hemos observado que partiendo de la descomposición ortogonal es posible caracterizar de forma acústica los diferentes ecosistemas de la geografía colombiana. Para esto proponemos medir el nivel de energía presente en cada orden polinomio en los diferentes momentos del día. En la actualidad se está desarrollando un trabajo relacionado que pretende utilizar la descomposición ortogonal para determinar índices bioacústicos de un ecosistema, área que es de gran interés para los biólogos.

La compresión de la información que permite la descomposición ortogonal abre la puerta al diseño de métodos para determinar cuál es la cantidad mínima de información que se necesita almacenar para obtener una buena reconstrucción de la señal. Esto también genera la necesidad de diseñar métricas para determinar la calidad de la señal que no dependan de la señal original y que permitan medir la calidad de la señal de forma consecuente con la percepción del usuario. El estudio presentado solo analiza dos de las cuatro familias ortogonales que se analizaron en la investigación. Sin embargo, existen un gran número de familias ortogonales, por lo tanto es posible que existan familias ortogonales que tengan una mejor funciona-

miento que las familias planteadas en este estudio.

Otros trabajos podrían centrarse en demostrar la relación entre los órdenes de polinomios y los componentes de frecuencia en la señal a descomponer, así como la relación entre la resolución de la descomposición y la frecuencia de muestreo de la señal

10. Agradecimientos

Queremos agradecer al instituto Alexander von Humboldt, en especial a Paula Caycedo y Oscar Laverde por suministrarnos una gran cantidad de registros sonoros del páramo de Chingaza, a Gustavo Kattan por suministrarnos la licencia de Raven Pro, a la fundación Alejandro Ángel Escobar por otorgar a este trabajo una de las becas Colombia Biodiversa y a la Pontificia Universidad Javeriana Cali.

BIBLIOGRAFÍA

1. Medina GF. Areas desprotegidas de Colombia. UNPeriodico. 2010;140.
2. Bardeli R, Wolff D, Kurth F, Koch M, Tauchert KH, Frommolt KH. Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognition Letters*. 2010;31(12):1524–1534.
3. Schroeder R, Medellín R, Ramírez O, Rojo A. La importancia de los objetivos de hábitat en los Planes de Manejo de las Unidades de Manejo para la Conservación de la Vida Silvestre (UMA). *Investigación Ambiental*. 2009;1(2):136–142.
4. Villareal H, Alvarez M, Córdoba S, Escobar F, Fagua G, Gast F, et al. Manual de métodos para el desarrollo de inventarios de biodiversidad. Programa de inventarios de biodiversidad Instituto de investigación de recursos biológicos Alexander von Humboldt Bogotá, Colombia. 2004;236:1981.
5. Cifuentes-Sarmiento Y, Cortés LFC. COLOMBIA-INFORME ANUAL 2013;.
6. Periodico EL Tiempo Colombia. Electronic References; 2013. Available from: <http://www.eltiempo.com> [cited 2013 abril 2013].
7. Gallardo RJ. El Contexto Regional para Valorar las Aves de Río Amarillo, Copán, Honduras. 2013;.
8. Mendoza LF, Carmona R. Distribución espacial y temporal de aves playeras (Orden: Charadriiformes) en Laguna San Ignacio, Baja California Sur, México. *Rev Biol Trop*. 2013;61(1):229–241.
9. Trifa V. A framework for bird songs detection, recognition and localization using acoustic sensor networks. Master's thesis, University of California Los Angeles and Ecole Polytechnique Fédérale de Lausanne. 2006;.
10. Juan Mayor LT Michael Martinez. Frogs species Classification using LPC and Classification Algorithms on Wireless Sensor Network Platform. 2009;.
11. Ulloa JS. Diseño y Evaluación de Algoritmos para Detectar Vocalizaciones de Aves en Registros Sonoros; 2012.
12. Kasten EP, McKinley PK, Gage SH. Ensemble extraction for classification and detection of bird species. *Ecological Informatics*. 2010;5(3):153–166.
13. Selin A, Turunen J, Tanttú JT. Wavelets in recognition of bird sounds. *EURASIP Journal on Applied Signal Processing*. 2007;2007(1):141–141.
14. Zarowski CJ. An introduction to numerical analysis for electrical and computer engineers. Wiley. com; 2004.
15. Trefethen LN, Bau III D. Numerical linear algebra. vol. 50. Siam; 1997.
16. Burden RL, D'iaz JID, Faires JD, Balderrama RI, Bocanegra FP. Análisis numérico; 1996.