



**UNIVERSIDAD
DE ANTIOQUIA**

**Identificación de Múltiples Especies en Paisajes
Acústicos Usando Técnicas de Aprendizaje No
Supervisado**

Autor

Maria José Guerrero Muriel

Universidad de Antioquia
Facultad de Ingeniería, Departamento de Ingeniería
Electrónica
Medellín, Colombia
2020



Identificación de Múltiples Especies en Paisajes Acústicos Usando Técnicas de Aprendizaje
No Supervisado

Maria José Guerrero Muriel

Informe de trabajo de grado como requisito para optar al título de:
Ingeniera Electrónica

Asesora

Claudia Victoria Isaza Narváez
Profesora Vinculada Universidad de Antioquia, PhD

Co-Asesor

Carol Bedoya Acevedo, PhD

Universidad de Antioquia
Facultad de Ingeniería, Departamento de Ingeniería Electrónica
Medellín, Colombia
2020.

Resumen

En este trabajo de grado, se propone un método no supervisado y multiclase para identificar llamadas de especies de diferentes animales en un paisaje sonoro. Esta metodología no requiere conocimiento previo de las fuentes bióticas para ser identificadas y es capaz de encontrar el grado de relación entre ellas. Para ello se usa un algoritmo de agrupamiento difuso (LAMDA-Learning Algorithm for multivariate analysis) con un operador Yager-Ribarov $\exists\pi$, cuya naturaleza difusa permite obtener el grado de pertenencia entre clases y así categorizar las fuentes de sonido en base a sus similitudes. Este método se presenta como una forma novedosa de monitoreo acústico el cual permite hacer seguimiento de manera pasiva de los cambios presentados en un ecosistema y su biodiversidad.

Palabras Claves: Clasificación, LAMDA, multiespecie, paisaje sonoro, monitoreo acústico



Contenido

| | |
|---|----|
| Contenido | 3 |
| Introducción | 5 |
| Objetivos | 6 |
| General..... | 6 |
| Específicos..... | 6 |
| Marco Teórico | 7 |
| Paisaje Sonoro | 7 |
| Monitoreo Acústico | 7 |
| Aprendizaje no Supervisado..... | 7 |
| Algoritmo de Clasificación - LAMDA..... | 7 |
| Materiales y Métodos | 8 |
| Base de datos | 8 |
| Metodología Propuesta | 8 |
| <input type="checkbox"/> Preprocesamiento y segmentación..... | 9 |
| <input type="checkbox"/> Extracción de Características..... | 10 |
| <input type="checkbox"/> Clasificación | 11 |
| Software AUREAS1.6..... | 12 |
| Resultados y análisis | 13 |
| <input type="checkbox"/> Entrenamiento..... | 13 |
| <input type="checkbox"/> Prueba | 16 |
| Conclusiones | 18 |
| Trabajo a Futuro | 18 |
| Referencias | 19 |

Lista de Figuras

| | |
|--|----|
| <i>Figura 1 Etapas implementadas para el monitoreo acústico e identificación multiespecie.</i> | 8 |
| <i>Figura 2 Espectrograma sin segmentación</i> | 9 |
| <i>Figura 3 Espectrograma segmentado con posibles sonotipos</i> | 10 |
| <i>Figura 4 Interfaz del software AUREAS con enfoque multiespecie</i> | 12 |
| <i>Figura 5 Tabla de Clasificación de segmentos de AUREAS</i> | 12 |
| <i>Figura 6. Detección automática de sonotipos por parte del clasificador</i> | 13 |
| <i>Figura 7. Significado biológico de sonotipos</i> | 13 |
| <i>Figura 8. Endógama de similitud entre clústeres</i> | 15 |
| <i>Figura 9 Grabación ruidosa y con error en segmentación</i> | 17 |
| <i>Figura 10. Grabación ruidosa que presenta error en la segmentación</i> | 17 |

Lista de Tablas

| | |
|---|----|
| <i>Tabla 1 Resultados en etapa de entrenamiento</i> | 14 |
| <i>Tabla 2 Resultados en etapa de prueba</i> | 16 |

Introducción

El Monitoreo de las poblaciones biológicas proporciona información sobre el estado general de la biodiversidad de un área, el número de especies, su estado de amenaza y la tendencia de la población [1]. Este monitoreo es conocido por ser una tarea compleja que involucra altos costos operativos, intervalos reproductivos desconocidos de las especies y difícil visualización de individuos [2]. Existen diversos métodos para monitorear las especies, entre ellos, el monitoreo acústico con ayuda de grabaciones de audio que recopilan la actividad de los paisajes sonoros. Este método surge como una ayuda para hacer seguimiento en periodos largos de tiempo a especies críticas. Para analizar los audios, la identificación manual de especies por parte de expertos se puede llevar a cabo analizando las grabaciones a partir de espectrogramas, pero para grandes cantidades de datos, es necesario facilitar el análisis mediante métodos de reconocimiento automático [3].

Se han propuesto varios métodos para la identificación de especies como aves, anuros y algunos mamíferos. Entre estos métodos se encuentra el uso de índices acústicos en combinación de un espectrograma de falso color de larga duración (LDFC) para identificar la actividad vocal de ranas y aves [4] [5]. También se han usado algoritmos de extracción de características en combinación de técnicas de machine learning para identificar a las especies mediante patrones de sonido analizando los parámetros de tiempo – frecuencia [2] [6]. Entre las técnicas de clasificación se encuentran los modelos ocultos de Márkov (HMM), máquinas de soporte vectorial (SVM), modelos de mezcla Gaussiana (GMM), *random forest* y *K- nearest neighbor* (KNN). Sin embargo, estos trabajos se han centrado en el reconocimiento de especies para un solo tipo de animal.

En este proyecto, se plantea una metodología que permite identificar automáticamente posibles llamadas de especies para diferentes animales por medio de los segmentos de una grabación. En cada segmento se asume que existe información relevante y se usa un procedimiento multiclase y sin supervisión para identificar sonotipos que se consideran pueden ser emitidos por diferentes animales. Para la identificación de los sonotipos se usará un algoritmo de agrupamiento difuso multiclase (*LAMDA-Learning Algorithm for multivariate analysis*), cuya naturaleza difusa permite obtener el grado de pertenencia entre clases y categorizar los datos. Este algoritmo no requiere un número de clases como parámetro de entrada a diferencia de otros algoritmos de agrupamiento, lo cual lo hace muy llamativo para identificar especies en una zona donde no se tiene conocimiento a priori de los animales que se encuentran. Para este problema en específico, se realizará una modificación del método LAMDA original, cambiando su conectivo difuso híbrido normal por un operador Yager-Ribalov 3π , el cual reduce el número de clases generadas [7] [8].

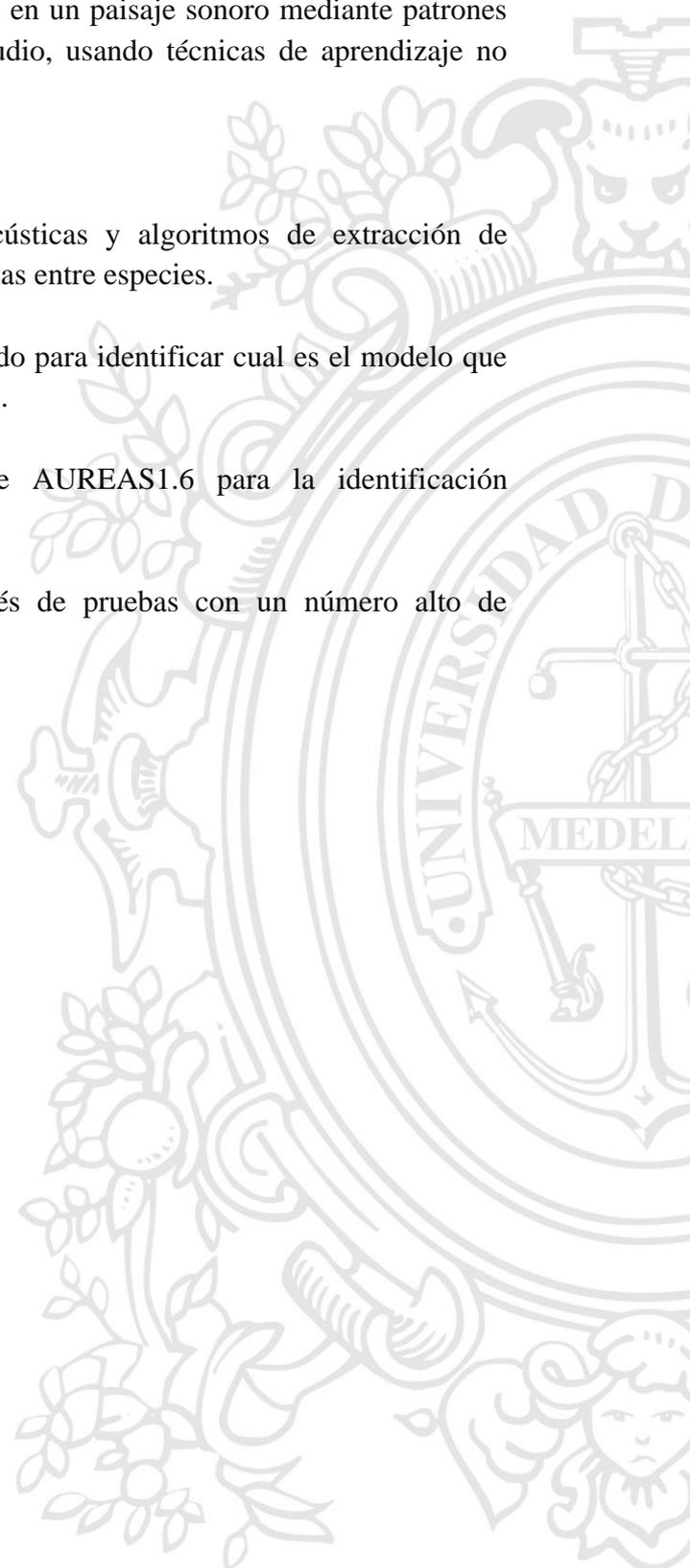
Objetivos

General

- Identificar múltiples especies que interactúan en un paisaje sonoro mediante patrones de sonido encontrados en grabaciones de audio, usando técnicas de aprendizaje no supervisado

Específicos

- Analizar patrones de sonido, señales bioacústicas y algoritmos de extracción de características que permitan detectar diferencias entre especies.
- Evaluar técnicas de aprendizaje no supervisado para identificar cual es el modelo que más se ajusta a la clasificación de las especies.
- Actualización de métodos en el software AUREAS1.6 para la identificación multiespecie en paisajes sonoros.
- Evaluar el desempeño del software a través de pruebas con un número alto de grabaciones.



Marco Teórico

En esta sección se presentan los conceptos relacionados con la temática del trabajo de grado.

Paisaje Sonoro

La confluencia de sonidos bióticos (biofonía), geofísicos (geofonía) y antropogénicos (antropofonía) genera lo que se conoce como paisaje sonoro. Las características acústicas de estos paisajes sonoros cambian según las interacciones de las comunidades ecológicas con el paisaje (físico/topográfico). El estudio de los paisajes sonoros permite evaluar y monitorear, de manera indirecta, la heterogeneidad del paisaje teniendo en cuenta a las comunidades de fauna vocalmente activa [9].

Monitoreo Acústico

Los sonidos de los animales proporcionan pistas a través de las cuales se puede evaluar la abundancia, ubicación y posibles riesgos en su entorno. A su vez, estos sonidos pueden ser usados para el reconocimiento de especies [4] [10]. Este reconocimiento se lleva a cabo mediante patrones de sonido (i.e., representaciones espectro-temporales de fuentes sonoras). Estos patrones provienen de los cantos o llamadas que realizan las especies para comunicarse entre ellas con fines reproductivos, de advertencia o territoriales. Dichos patrones son obtenidos mediante grabaciones de audio que son analizados con propósitos de monitoreo que puede facilitar el control de la biodiversidad. Este tipo de monitoreo es menos invasiva, más rentable y efectiva. [5]

Aprendizaje no Supervisado

En problemas como el monitoreo acústico, es posible no tener un conocimiento a priori de los datos debido a que el paisaje sonoro está lleno de fuentes bióticas de procedencia y número desconocido. Para encontrar una posible relación entre los datos en función de sus características, se usa aprendizaje no supervisado. Este ha adquirido popularidad debido a su capacidad de extraer información pertinente y encontrar relación entre los datos para asociarlos a un mismo grupo. Se usa en problemas de agrupación o *clustering* donde el objetivo es designar elementos similares a un mismo conjunto sin un conocimiento previo de la naturaleza de los datos. En casos como la clasificación de especies de ranas [2] y aves [11] se implementan algoritmos que utilizan este tipo de aprendizaje.

Algoritmo de Clasificación - LAMDA

LAMDA-*Learning Algorithm for Multivariate Data Analysis* (Algoritmo de aprendizaje para análisis multivariado de datos)- está basado en la teoría de lógica difusa, no requiere el número de clases como parámetro de entrada y encuentra el grado de adecuación global de cada dato estimando la contribución de las características de entrada. La modificación propuesta en el conectivo difuso (operador Yager-Ribalov 3π) [7] [8] permite fortalecer la relación entre el número de clústeres generados y el número de clases generadas asociadas con las fuentes sonoras bióticas. Esto evita generar clústeres con unidades acústicas que son variaciones

mínimas de una clase previamente establecida. Una de las grandes ventajas que proporciona este algoritmo es el reconocimiento de nuevas clases sin un entrenamiento previo.

Materiales y Métodos

Base de datos

Para el desarrollo del proyecto se contó con más de 150.000 grabaciones acústicas realizadas por el Grupo Herpetológico de Antioquia (que hace parte del Instituto de Biología de la Universidad de Antioquia). Estas grabaciones fueron realizadas en diferentes zonas geográficas del bosque húmedo colombiano

De las 150.000 se seleccionaron 150 para el reconocimiento multiespecie. Estas grabaciones fueron tomadas en la represa Jaguas en Alejandría – Antioquia. Cada grabación tiene una duración de 1 minuto y se tomaron cada 20 minutos durante 1 mes.

De las 150 grabaciones, 50 fueron usadas para entrenamiento y 100 para prueba.

Metodología Propuesta

Se describe a continuación, la metodología usada para extraer e identificar las unidades acústicas de las fuentes bióticas que interactúan en un paisaje sonoro. Consta de 4 etapas: preprocesamiento, segmentación, extracción de características y clasificación.

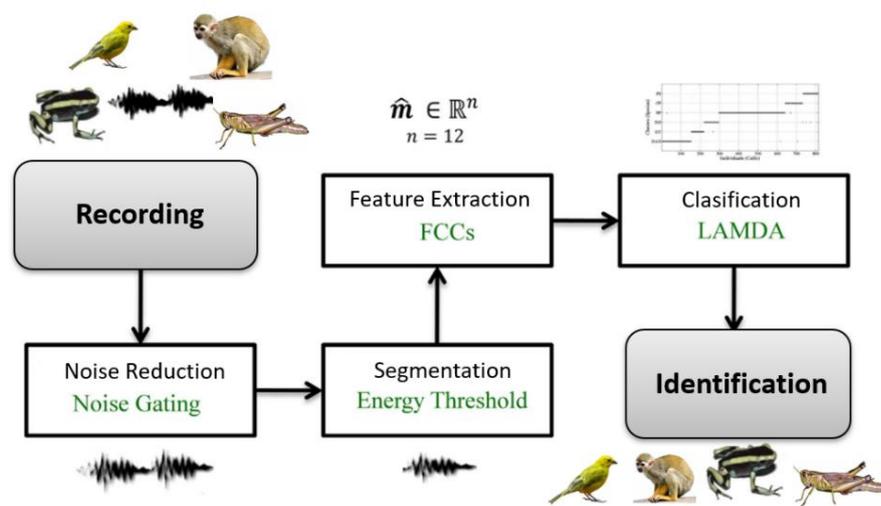


Figura 1 Etapas implementadas para el monitoreo acústico e identificación multiespecie.
Modificado de (Bedoya, 2014)

Se usó un selector de bandas de frecuencia para determinar las regiones del espectrograma con actividad acústica relevante (biofonías) y un segmentador basado en la amplitud para extraer las unidades acústicas detectadas. Luego, se utilizaron coeficientes cepstrales escalados linealmente para extraer la información de cada unidad y una metodología multiclase difusa no supervisada para su categorización.

- **Preprocesamiento y segmentación**

El preprocesamiento consiste en una etapa de eliminación de ruido seguido de una etapa de segmentación de la señal.

Inicialmente, se tiene el espectrograma sin segmentar como se muestra en la figura 2. En la segmentación, primero se encuentran las bandas de frecuencia cada cierto intervalo de tiempo, luego se seleccionan los puntos mínimos donde hay mayor probabilidad de no encontrar un canto. Después, para cada intervalo se hace una selección de bandas; para esto se divide nuevamente el espectrograma, esta vez en tres partes y se calcula el máximo valor de frecuencia para cada sección. Posterior a esto, se promedian los valores para obtener una proyección global del segmento y ver en cuales frecuencias hay valores máximos y mínimos para seleccionar las bandas. Finalmente, se suman todos los valores de intensidad a lo largo de las frecuencias para cada intervalo de tiempo, se encuentran los puntos mínimos para ubicar los puntos de corte y realizar el cuadro que rodea el segmento de vocalización como se observa en la figura 3.

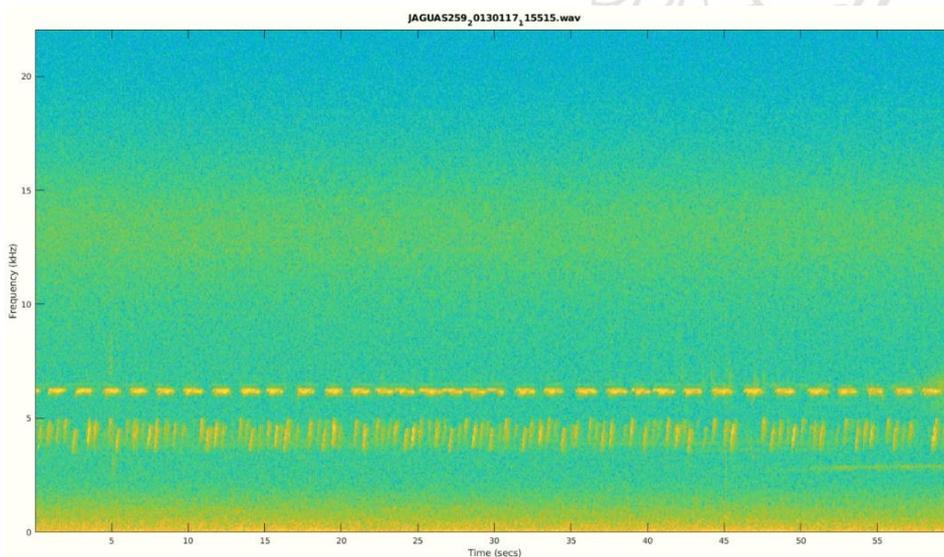


Figura 2 Espectrograma sin segmentación

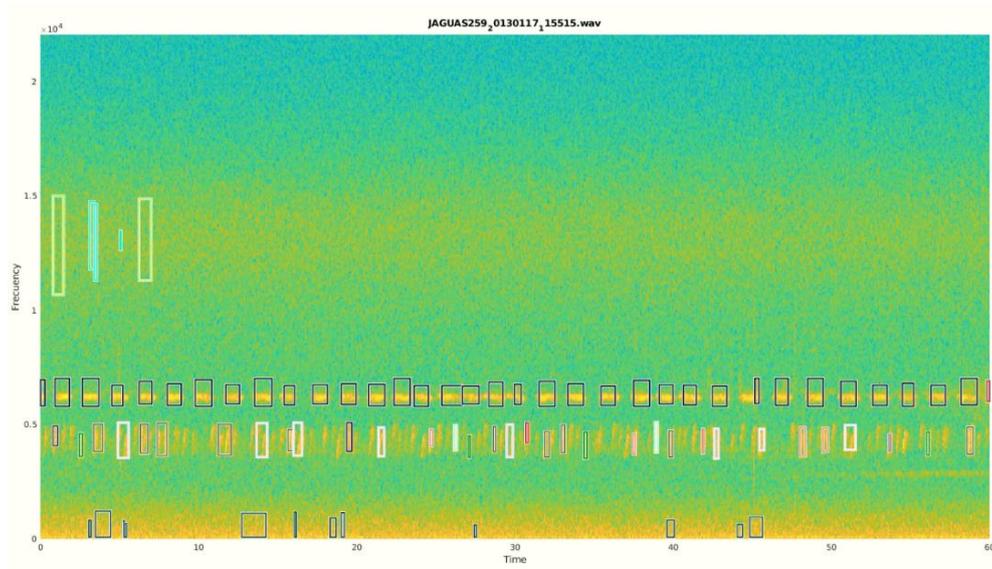


Figura 3 Espectrograma segmentado con posibles sonotipos

- **Extracción de Características**

Luego de la segmentación y de obtener los límites temporales de cada biofonía, se extrae del espectrograma cada llamada del animal y se establecen los límites espectro-temporales para cada biofonía.

El primer paso de la extracción de características consiste en estimar el logaritmo de la energía para el segmento que representa el sonido del animal. Esta operación extrae la información acústica importante del dominio temporal del segmento, luego lo redistribuye a través del dominio espectral de manera no lineal.

$$Q_{i,m} = \log \left(\sum_{j=m-w}^{m(w+1)} |H_{i,j}|^2 \right) \quad (\text{Ec. 1})$$

Donde H representa la matriz que contiene el sonido de un animal, w es el tamaño del inventariado (el paso) y m es el número de logaritmos calculados en cada ventana, Posterior a esto, se usa la transformada discreta de coseno (Ec. 2) del logaritmo calculado con el objetivo de reducir la dimensión y establecer una longitud común para el vector de características extraídas en todas las biofonías independientemente de su ancho de banda o duración.

$$Y_{k,m} = N_p \sum_{i=1}^{N_u} Q_{i,m} \cos \left(\frac{\pi}{2N_u} (2i-1)(k-1) \right) \quad (\text{Ec. 2})$$

Donde Q es la matriz con el logaritmo de las energías calculadas, m es el índice de la banda de frecuencia, N_k es el número de coeficientes, N_u es la longitud de la llamada segmentada en los dominios temporal y espectral y N_p es un factor de normalización utilizado para hacer ortogonal la matriz.

Luego, se normaliza el vector de características obtenido como paso requerido para la entrada al algoritmo de clasificación.

- **Clasificación**

Ya con la extracción de características, se procede a realizar la clasificación de cada biofonía haciendo uso del algoritmo LAMDA 3 π .

El primer paso consiste en calcular el grado de adecuación marginal MADs (Ec. 3) que son las contribuciones de las características calculadas en la extracción de características (coeficientes cepstrales) de cada unidad acústica a cada uno de los clústeres existentes

$$M_{c,f} = \rho_{c,f}^{\hat{y}_f} (1 - \rho_{c,f})^{1-\hat{y}_f} \quad (\text{Ec. 3})$$

Donde M es la matriz de valores MADs extraídos de cada elemento analizado, ρ es la matriz con los valores medios de las características de cada clúster, \hat{y} es un vector con los valores de las características de cada elemento analizado, N_f es el número de características y N_c es el número de clústeres existentes.

Inicialmente, solo se tiene definido el clúster de la clase sin información (NIC). Cada dato clasificado en él es considerado como no reconocido, por ello, el primer elemento se considera como no reconocido y es asignado a una nueva clase con parámetros NIC arbitrarios modificados por los valores del elemento (Ec. 4)

$$\rho_{1,f} = \frac{(\hat{y}_f + \rho_{0,f})}{2} \quad (\text{Ec. 4})$$

Cuando se analiza un elemento nuevo y se obtiene sus MADs, estos son combinados utilizando un conectivo de lógica difusa (Ec. 5) como operador de agregación para obtener el grado de adecuación global (GAD) de un elemento a un cluster.

$$G_c = \frac{\prod_{f=1}^{N_f} M_{c,f}}{\prod_{f=1}^{N_f} M_{c,f} + \prod_{f=1}^{N_f} (1 - M_{c,f})} \quad (\text{Ec. 5})$$

Una vez obtenidos los GADs de todos los clústeres, el elemento se clasifica en el clúster con máximo GAD y se pueden dar dos casos: el primero en el que el valor máximo corresponde al de la NIC, se crea un nuevo clúster (Ec. 4) y se inicializa con los parámetros de la NIC modificados por los valores de entrada. El segundo caso es cuando el dato es asignado a un clúster existente, allí el clúster al que fue asignado se actualiza con los valores de ese elemento con el propósito de agregar la contribución de la nueva entrada. (Ec. 6)

$$\rho_{c,f}^{(k)} = \rho_{c,f}^{(k-1)} + \frac{\hat{y}_f - \rho_{c,f}^{(k-1)}}{n_c^k} \quad (\text{Ec. 6})$$

Donde c es el cluster actual, n_c^k es el número actual de elementos clasificados en el cluster c y $\rho_{c,f}^{(k-1)}$ es el valor previo de $\rho_{c,f}$ antes de la actualización. El proceso continua hasta que se analicen todos los datos

Software AUREAS1.6

El software AUREAS se desarrolló como una herramienta para el análisis de cantos de anuros, cuyo objetivo era identificar, dentro de una grabación, los segmentos que correspondían a una vocalización y reconocer cuál especie emitió dicha vocalización. Con la metodología descrita previamente, se hace un cambio en el enfoque de este software para realizar un análisis multiespecie, identificando los sonotipos de diferentes especies de animales que se encuentran en el ecosistema evaluado. En la figura 4 se muestra la ventana principal del software. Allí es posible acceder al proceso de clasificación y reconocimiento de los sonotipos de uno o varios audios. Una de las modificaciones que se realizó fue agregar una barra de selección de banda de frecuencia, la cual le permite al experto, filtrar la búsqueda del sonotipo a través de su frecuencia.

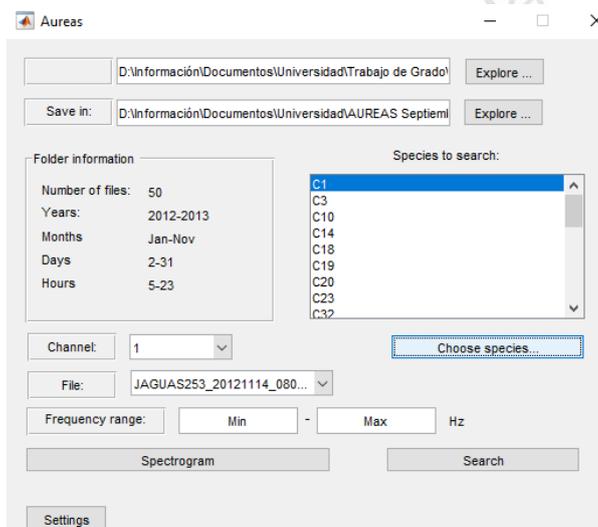


Figura 4 Interfaz del software AUREAS con enfoque multiespecie

En la figura 5 se tiene la tabla de clasificación, la cual muestra los diferentes clústeres (con diferentes colores) a los que fue asociado cada segmento del audio de manera que, uno o mas clústeres puede representar un sonotipo de interés.

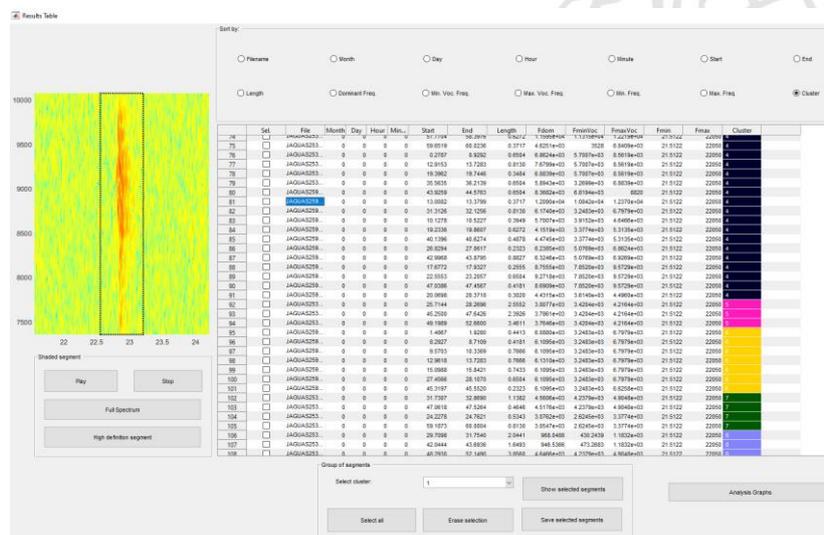


Figura 5 Tabla de Clasificación de segmentos de AUREAS

Resultados y análisis

• Entrenamiento

Se analizaron los 50 audios en el software AUREAS con el nuevo enfoque multiespecie. Se reconocieron 29 sonotipos usando la clasificación no supervisada. En la figura 6 se muestran algunos de los sonotipos detectados automáticamente por el clasificador del software para una de las 50 grabaciones. Los sonotipos son segmentos del espectrograma donde el algoritmo considera que posiblemente hay un canto. Los segmentos del mismo color representan segmentos similares o con un patrón parecido que puede corresponder al canto de una misma especie y los otros colores indican un patrón diferente que corresponde posiblemente al canto de otra especie del mismo animal o a otro animal.

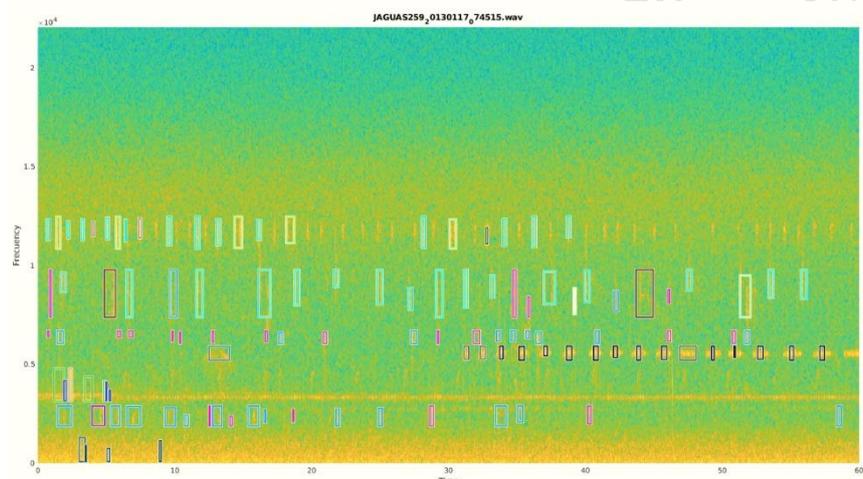


Figura 6. Detección automática de sonotipos por parte del clasificador

A cada sonotipo se le dio un significado biológico por parte de un biólogo experto que hace la asociación entre la clase y el significado del sonotipo. Un ejemplo de esto se muestra en la figura 7 donde se muestra en el recuadro rojo el sonotipo correspondiente a un grillo, en el rosado a una rana *Leucostethus Jota*, el azul oscuro a una rana *Diasporus Gularis* y los recuadros morado y azul claro a dos especies de aves diferentes.

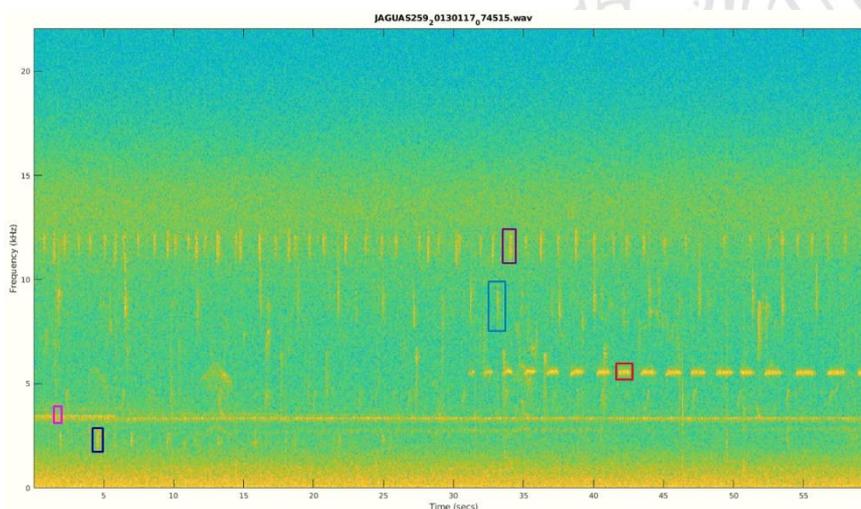


Figura 7. Significado biológico de sonotipos

Luego, se hizo el proceso de reconocer los sonotipos identificados sobre los mismos datos de entrenamiento y se hizo el análisis con ayuda de la Tabla 1.

Para cada una de las 50 grabaciones, un experto identificó las clases asociadas un sonotipo realizando una inspección visual del espectrograma, estos datos se observan en la segunda columna de la Tabla 1. En la tercera columna se tienen las clases detectadas automáticamente por el software. La columna detección de sonotipo hace referencia la identificación del sonotipo por el experto y por el software, siendo 1 como reconocido y 0 como no reconocido.

Tabla 1 Resultados en etapa de entrenamiento

| EXPERTO AUREAS | | | | | |
|-------------------------------|-------|-------|-----------|-----------------------|-------------------------|
| Archivo | Clase | Clase | Similitud | Detección de Sonotipo | Detección Segundo nivel |
| JAGUAS253_20121114_080107.wav | 3 | 3 | | 1 | 1 |
| | | 10 | | | |
| | | 14 | | | |
| | 18 | 18 | | 1 | 1 |
| | | 23 | | | |
| | | 41 | | | |
| | | 58 | | | |
| | 82 | 82 | | 1 | 1 |
| | | 125 | | | |
| | | 137 | | | |
| | | 179 | | | |
| | | 82 | NIC | | |
| JAGUAS253_20121114_081107.wav | 3 | 10 | 2-3 | 0 | 1 |
| | | 18 | | | |
| | | 32 | | | |
| | | 41 | | | |
| | 32 | 58 | | 1 | 1 |
| | | 72 | | | |
| | | 82 | | | |
| | | 136 | 2-3 | | |
| | 32 | 179 | | | |
| JAGUAS253_20121114_082107.wav | 3 | 10 | 2-3 | 0 | 1 |
| | | 19 | 1-32 | | |
| | | 42 | | | |
| | 32 | 58 | | 0 | 1 |
| | | 82 | | | |
| | 32 | 125 | | | |
| Porcentaje de aciertos | | 83,47 | | | |

Como se observa en la tercera columna, las clases detectadas por el software presentan mucha variación entre ellas. Esto se debe a que hay variación en el tipo de canto de una misma especie y que por tener grabaciones omnidireccionales se encuentra en ellas la interferencia de otros sonidos biofónicos y geofónicos. Gracias a que el algoritmo de clasificación es difuso, el cual permite obtener grados de pertenencia para cada una de las clases se puede usar esta información para establecer similitud entre clases. Al tener una representación de clústeres por su grado de pertenencia, cada clúster se puede representar como un conjunto difuso y con ayuda de la Ec 7 se puede obtener la similitud entre dichos conjuntos difusos. De esta manera, se puede saber que tan parecido son entre sí los clústeres y representarlos en un dendograma como se muestra en la figura 8.

$$G(A, B) = \frac{\sum_{n=1}^N (\mu_{An} \cap \mu_{Bn})}{\sum_{n=1}^N (\mu_{An} \cup \mu_{Bn})} \quad (\text{Ec. 7})$$

Donde μ_{An} corresponde al grado de pertenencia del dato n a la clase A y μ_{Bn} el grado de pertenencia del dato n a la clase B.

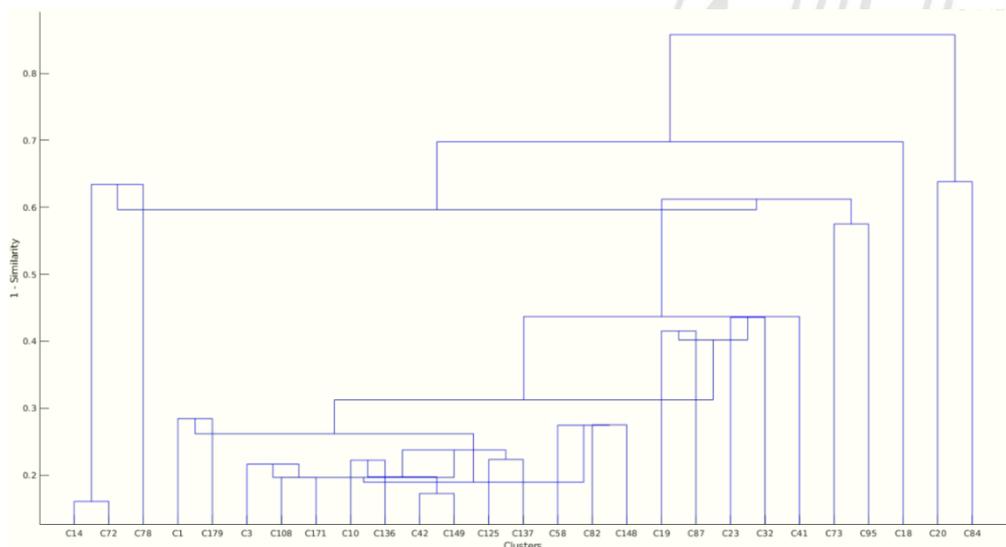


Figura 8. Endógama de similitud entre clústeres

Partiendo de esto, se realiza una detección de segundo nivel que representa si alguna de las clases identificadas por el software fue detectada o tiene similitud con otra clase identificada como sucede, por ejemplo, con la clase 10 y la clase 3. De esta forma, se obtuvo un porcentaje de aciertos igual a **83.47%**

- **Prueba**

Se hace el proceso de reconocer los 29 sonotipos en 100 grabaciones nuevas. Estas grabaciones fueron seleccionadas de manera aleatoria de la base de datos y varias de estas presentaron mucho ruido. El resultado se muestra en la tabla 2.

Tabla 2 Resultados en etapa de prueba

| Archivo | EXPERTO | | AUREAS | | Detección de Clase | Detección segundo nivel |
|-------------------------------|-------------------------------|-------|-----------|--|--------------------|-------------------------|
| | Clase | Clase | Similitud | | | |
| JAGUAS260_20121118_175500.wav | 3 | 3 | | | 1 | 1 |
| | | 10 | | | | |
| | | 14 | 1-72 | | | |
| | 18 | 18 | | | 1 | 1 |
| | | 32 | | | | |
| | | 41 | | | | |
| | | 58 | | | | |
| | 72 | 108 | | | 0 | 1 |
| | | 125 | | | | |
| | | 136 | | | | |
| | | 137 | | | | |
| | JAGUAS260_20121118_180500.wav | 10 | 10 | | | 1 |
| 18 | | | | | | |
| 18 | | 41 | | | 1 | 1 |
| | | 58 | | | | |
| 125 | | 108 | | | 1 | 1 |
| | | 125 | | | | |
| | | 137 | | | | |
| JAGUAS260_20121118_181500.wav | 18 | 3 | | | 1 | 1 |
| | | 10 | | | | |
| | | 18 | | | | |
| | | 41 | | | | |
| | 72 | 58 | | | 0 | 0 |
| | | 108 | | | | |
| | | 125 | | | | |
| | | 137 | | | | |
| | 171 | 149 | | | 1 | 1 |
| | | 171 | | | | |
| | | 179 | | | | |
| NIC | | | | | | |
| Porcentaje de aciertos | 68,98 | | | | | |

El porcentaje de aciertos en este caso fue de **68.98%** Esto se debe a las grabaciones con alto contenido de ruido que conllevan a una mala segmentación como se muestra en la figura 9 y posteriormente, a una mala clasificación bajando el rendimiento del software.

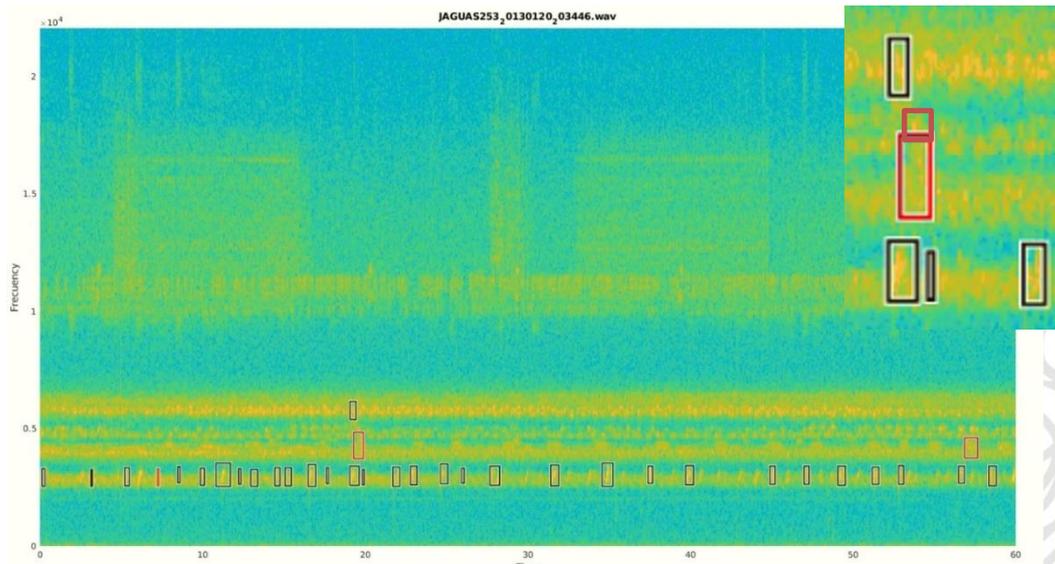


Figura 9 Grabación ruidosa y con error en segmentación

En la figura 10 se puede observar otro error cometido en la etapa de segmentación en donde se toman más frecuencias de las necesarias provocando una confusión en el clasificador.

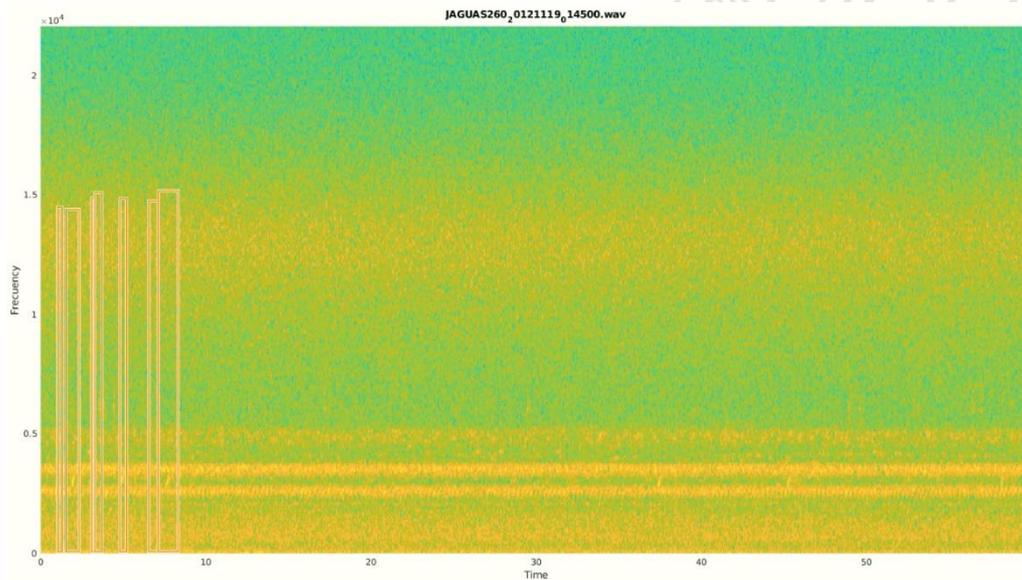


Figura 10. Grabación ruidosa que presenta error en la segmentación

Conclusiones

- Esta metodología muestra una nueva forma de generar evaluaciones de biodiversidad y monitoreo ecológico, utilizando patrones de las fuentes de sonido, realizando un análisis no invasivo y ayudando a un monitoreo más sencillo para los expertos.
- Este método es realmente útil para la evaluación de especies críticas en un ecosistema pues permite llevar un control de ellas en cuanto a su presencia o ausencia.
- Gracias al método de clasificación difuso se pudo estimar la similitud entre clases. Esto ofrece una ventaja en cuanto a la detección de sonotipos.
- El segmentador juega un papel fundamental en esta metodología pues de allí parte que el clasificador funcione de la manera esperada.
- Los resultados del clasificador se pueden ver afectados debido a la presencia de ruido que no se alcanza a eliminar en la etapa de preprocesamiento de la señal de audio.

Trabajo a Futuro

Como trabajo a futuro se proponen varias cosas:

- Realizar una mejora en la etapa de segmentación, esta etapa resulta de gran importancia pues de aquí se parte para una buena selección de características y clasificación
- Disminuir ruido de fondo de la señal
- Migrar el software a una plataforma con menos costo computacional y mas amigable para el usuario

Referencias

- [1] I. Potamits, S. Ntalampiras, O. Jahn, K. Riede, Automatic bird sound detection in long real-field recordings: Applications and tools, *Applied Acoustics*, vol. 80, pp. 1-9, 2014.
- [2] C. Bedoya, C. Isaza, J. Daza, J. López, Automatic recognition of anuran species based on syllable identification, *Ecological Informatics*, vol. 24, pp. 200-209, 2014.
- [3] J.S. Ulloa, T. Aubin, D. Llusia, C. Bouveyron, J. Sueur, Estimating animal acoustic diversity in tropical environments unsupervised multiresolution analysis, *Ecological Indicators*, vol. 90, pp. 346-355, 2018.
- [4] S. Brodie, S. Allen-Ankins, M. Towsey, P. Roe, L. Schwarzkopf, Automated species identification of frog choruses in environmental recordings using acoustic indices, *Ecological Indicators*, vol. 119, 2020.
- [5] H. Gan, J. Zhang, M. Towsey, A. Truskinger, D. Stark, B.J. van Rensburg, Y. Li, P. Roe, Data selection in frog chorusing recognition with acoustic indices, *Ecological Informatics*, 2020.
- [6] RR KVSN, J. Montgomery, S. Gatg, M. Charlesto, Bioacoustics Data Analysis - A Taxonomy, Survey and Open Challenges, *IEE Access*, vol. 8, pp. 57684-57708, 2020.
- [7] C. Bedoya, J. Waissman, C. Isaza, Yager-Ribalov Triple Pi Operator as a Means of Reducing the Number of Generated Clusters in Unsupervised Anuran Vocalization Recognition, *Nature-Inspired Computation and Machine Learning*, vol. 8857 of Lecture Notes in Computer Science, Springer International Publishing, pp. 382-391, 2014.
- [8] R. Yager, A. Rybalov, Full reinforcement operators in aggregation techniques, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, pp. 757-769, 1998.
- [9] P. Caycedo, C. González, Caracterización y Monitoreo de Aves y Paisajes Sonoros en tres microhábitats de la región de la Mojana, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt y Universidad de Córdoba, 23 Agosto 2019. [En línea]. Available: http://i2d.humboldt.org.co/ceiba/resource.do?r=puntoconteo_mojana_2018. [Último acceso: 16 Agosto 2020].
- [10] A. Fouquet, T. Tilly, A. Pašukonis, E.A. Courtois, P. Gaucher, J.S Ulloa, J. Sueur, Simulated chorus attracts conspecific and heterospecific Amazonian explosive breeding frogs, *Biotropica*, pp. 1-11, 2020.
- [11] O. Ovaskainen, U. Moliterno de Camargo, P. Somervu, Animal Sound Identifier (ASI): software for automated identification of vocal animals, *Ecology Letters*, vol. 21, pp. 1244-1254, 2018.