

# Reconocimiento automatizado de señales bioacústicas: Una revisión de métodos y aplicaciones

Paula Catalina Caycedo-Rosales<sup>1</sup>, José Francisco Ruiz-Muñoz<sup>2</sup> y  
Mauricio Orozco-Alzate<sup>3</sup>

Recepción: 14-05-2013, Aceptación: 30-07-2013

Disponible en línea: 05-11-2013

PACS: 43.60.-c, 43.60.Lq, 43.80.-n

---

## Resumen

Durante la última década se ha publicado una gran cantidad de estudios de investigación y aplicaciones sobre monitoreo bioacústico automatizado; no obstante, tales estudios están dispersos en la literatura de ingeniería y ciencias biológicas. En este artículo se presentan conceptos fundamentales sobre monitoreo bioacústico automatizado; se revisan, comparan y categorizan —en una taxonomía de técnicas de DSP/PR— las contribuciones de las investigaciones y las aplicaciones publicadas; se sugieren algunas direcciones para investigaciones futuras y se resaltan los retos y las oportunidades relacionados con la instalación de esta tecnología en Colombia.

---

<sup>1</sup> Magister en Ecología, [pcaicedo@humboldt.org.co](mailto:pcaicedo@humboldt.org.co), Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá-Colombia.

<sup>2</sup> Magister en Ingeniería, [jfruiзму@unal.edu.co](mailto:jfruiзму@unal.edu.co), Universidad Nacional de Colombia, Manizales-Colombia.

<sup>3</sup> Doctor en Ingeniería Automática, [morozcoa@unal.edu.co](mailto:morozcoa@unal.edu.co), Universidad Nacional de Colombia, Manizales-Colombia.

**Palabras clave:** monitoreo ambiental automatizado; bioacústica; procesamiento de señales acústicas; reconocimiento de patrones.

### Aspectos relevantes

- Se expone la importancia de las señales acústicas en los estudios de biología de la conservación y la necesidad de implementar este tipo de tecnología en Colombia.
- Se exponen métodos de representación y clasificación, aplicaciones y listado de software comercial y libre utilizados para el análisis de reconocimiento bioacústico.

---

## Automated Recognition of Bioacoustic Signals: a Review of Methods and Applications

---

### Abstract

During the past decade, numerous research studies and applications on automated bioacoustic monitoring have been published; however, such studies are scattered in the literature of engineering and life sciences. This paper presents a review on fundamental concepts of automated acoustic monitoring. Our aim is to compare and categorize—in a taxonomy of techniques DSP/PR—the contributions of published research studies and applications; in order to suggest some directions for future research and highlight challenges and opportunities related to the deployment of this technology in Colombia.

**Key words:** automated environmental monitoring; bioacoustics; acoustic signal processing; pattern recognition.

---

## 1 Introducción

La destrucción y degradación de la naturaleza conduce al empobrecimiento y, en muchos casos, a la pérdida de los beneficios que ésta nos suministra. Entender las consecuencias de los agentes tensionantes sobre la biodiversidad y los servicios ecosistémicos—de los cuales dependemos— requiere de una evaluación constante de las tendencias de cambio de los patrones y procesos ecológicos. Sin embargo, uno de los mayores desafíos en ecología y biología de la conservación, es la evaluación de la biodiversidad mediante técnicas de monitoreo eficaces, que permitan abarcar amplias escalas, tanto espaciales como temporales [1]. Este conocimiento debe ser la base para

construir criterios más robustos que conlleven a un mejor manejo del territorio, mediante un sistema de monitoreo centralizado que permita que los investigadores, tomadores de decisión, o personas interesadas, tengan una buena perspectiva general sobre los resultados arrojados por los programas de monitoreo a escala nacional [2].

Actualmente, el conocimiento de la gran mayoría de los sistemas ecológicos se basa en escalas espaciales y temporales muy limitadas [3]. Adicionalmente, en el ámbito nacional, las tareas de monitoreo están altamente descentralizadas —en el sentido de que no se cuenta con sistemas eficaces mediante los cuales se pueda compartir la información obtenida en programas de monitoreo aislados— lo que dificulta que los investigadores o planificadores de conservación puedan realizar análisis y tomar decisiones basándose en una perspectiva global [2]. Los programas de monitoreo —que generalmente incluyen detección, localización, clasificación y conteo de los animales en un sitio [4]— son fundamentales para conocer el estado y las tendencias de la biodiversidad, identificar especies vulnerables y en riesgo de extinción, evaluar el impacto de la actividad humana sobre ecosistemas naturales, hacer seguimiento de especies invasoras y parásitos, entre otros [2]. Los métodos tradicionales de levantamiento de información en inventarios, caracterizaciones y monitoreo de fauna dependen del grupo de estudio; a grandes rasgos se basan en observaciones, capturas-recapturas, trampas, huellas y registros auditivos [5],[6].

### **1.1 Importancia de las señales acústicas en los animales de vida silvestre**

Las señales acústicas emitidas por grupos taxonómicos como insectos, anuros, aves, peces y mamíferos refuerzan las diferencias entre las especies, ya que presentan características que permiten el reconocimiento entre individuos [7]. Las funciones más conocidas de estas señales son de carácter sexual, pues implican varios procesos comportamentales como la atracción y selección de pareja, y el mantenimiento y defensa de un territorio [8]; ambos aspectos son cruciales para el flujo genético, el mantenimiento específico de procesos implicados directamente en la evolución y patrones de distribución de las especies [9]. Estudios en biología de la conservación han utilizado las vocalizaciones de la fauna, ya que son muy eficientes para con-

firmar la presencia de una especie y mucho más en ambientes en donde la visibilidad es poca, como por ejemplo en los bosques tropicales, en los sistemas acuáticos y en el estudio de especies nocturnas [10],[11],[12],[13],[14]; en otros casos han sido empleadas para identificar el sexo, la edad e individuos [15],[16].

En las últimas décadas, el avance en tecnologías que buscan automatizar el monitoreo de la fauna —mediante el uso de unidades de grabación remota y paquetes analíticos de reconocimiento acústico automatizado de especies— está transformando la forma como se estudian las interacciones de los animales y sus hábitats [17],[18],[19],[13]. Comparada con las metodologías de muestreo acústico tradicionales, la toma de muestras automatizada permite recolectar los ambientes sonoros en diferentes puntos de la zona de estudio de manera simultánea y sistematizada en escalas temporales amplias, incrementando así el tamaño de la muestra y por consiguiente aumentando la representatividad del universo muestral [20].

## 1.2 Bioacústica y tecnología

Los avances tecnológicos y computacionales abren paso a nuevos enfoques para el censo y recolección de información bioacústica a escala mundial [10]. A través de hardware especializado (unidades de grabación remota) es posible grabar y almacenar enormes cantidades de información [21]. Los datos adquiridos se pueden transmitir, procesar y analizar usando herramientas computacionales. Esta prometedora herramienta bioacústica brinda la capacidad de ampliar la escala espacio-temporal de los muestreos biológicos y de realizar análisis que permiten el manejo de vastas cantidades de información para el reconocimiento de especies, lo que se traduce en una metodología ideal de monitoreo de la fauna colombiana. Sin embargo, a pesar de que la tecnología se presenta como una solución, existen desafíos que se deben enfrentar; entre ellos están la gran capacidad de almacenamiento y la velocidad de procesamiento que se requiere, teniendo en cuenta la magnitud de la información que se puede recolectar y, el desarrollo de algoritmos precisos para la clasificación y reconocimiento automatizado de las vocalizaciones específicas [17],[22].

Tradicionalmente, la tarea de la identificación de especies a través de sus vocalizaciones requiere de una persona con amplia experiencia. Usual-

mente, el uso de espectrogramas facilita la identificación, ya que se emplea no sólo la memoria auditiva, sino también la visual y de esta manera se refuerza el proceso de aprendizaje. Sin embargo, esta tarea es titánica si se requiere analizar el volumen de información bioacústica obtenida por las unidades de grabación remota. El procesamiento automatizado facilita la organización y búsqueda de datos acústicos [23], mediante técnicas DSP/PR es posible incorporar métodos de detección y localización de vocalizaciones de fauna vocalmente activa [4]. Estos métodos se pueden usar para contestar una amplia gama de preguntas de investigación que abordan diferentes escalas biológicas, desde individuos, poblaciones, comunidades y ecosistemas hasta paisajes [10],[1]. A pesar de que la clasificación automatizada no es un concepto nuevo, sí lo es para muchas aplicaciones; la mayoría de la investigación en clasificación de objetos se ha enfocado en computación y robótica [18]. Éste es un reto multidisciplinario, que requiere de la aplicación de áreas relacionadas con bioacústica, taxonomía, sistemática, ecología del paisaje sonoro, computación y electrónica.

## 2 Adquisición de información

El monitoreo acústico es un método no invasivo, lo que implica que no se requiere capturar individuos para contar con información biológica de interés. Además, se puede apoyar este tipo de monitoreo mediante el uso de herramientas tecnológicas disponibles actualmente [19], obteniendo una amplia cobertura espacial y temporal, e incluso es posible reducir los costos que implican otras técnicas. Los componentes básicos de hardware que se usan para la adquisición de datos son micrófonos, grabadoras, fuentes de energía, mecanismos de inicio y final de grabaciones y cubiertas para los equipos [24]. Los avances en tecnología bioacústica incluyen el desarrollo de sistemas de grabación cableados e inalámbricos, que permiten la recolección de datos en múltiples lugares a lo largo del tiempo [17]. Los micrófonos utilizados dependen del organismo que se requiera estudiar; hay micrófonos especializados para la captura de infrasonidos (elefantes, cetáceos) y de ultrasonidos (murciélagos, algunos insectos). Para la captura de vocalizaciones de aves, ranas y algunos insectos usualmente se utilizan micrófonos omnidireccionales, los cuales muestrean el sonido con la misma eficiencia en todas las direcciones [17]. En [17],[18],[19] se encuentran

consideraciones técnicas sobre tipos de micrófonos, recomendaciones para la ubicación de estos en los sitios de monitoreo, y especificaciones sobre equipos y protocolos de comunicación.

Los sistemas de monitoreo bioacústico automatizado requieren de sistemas de grabación digitales que se puedan programar, del desarrollo de un sistema computarizado de reconocimiento y clasificación de sonidos y de una alta capacidad de almacenamiento de las grabaciones [3],[20]; estos equipos de grabación se conocen como unidades de grabación automática (ARUs por sus siglas en inglés). Las ARUs pueden programarse para hacer grabaciones de audio en diferentes horas del día de manera automática o activarse desde un teléfono móvil [25].

Actualmente existe una variada oferta de tipos de equipos de grabación automatizada en el mercado, sin embargo, no todos los equipos ofrecen la misma calidad de grabación. Las diferencias entre los equipos de grabación están vinculadas con varias características como por ejemplo en la relación señal a ruido, en la calidad y diseño de los circuitos electrónicos, la configuración, la protección del equipo a las condiciones de intemperie, frecuencia de muestreo, filtros de reducción de ruido de fondo y cronogramas de programación [26]. Las ARUs más utilizadas para este fin son los *songmeters*, sistemas de grabación E3A, teléfonos móviles y accesorios.

- **Songmeter SM2-GPS** (Wildlife Acoustics Inc., Concord, MA, USA, <http://www.wildlifeacoustics.com>): consta de dos micrófonos omnidireccionales (respuesta de muestreo: 20-2000 Hz). Estas unidades de grabación pueden tener un cronograma de grabación de acuerdo a los requerimientos de cada investigación, además están diseñadas de tal forma que se pueden conectar hidrófonos, micrófonos sensibles al ultrasonido e infrasonido [17],[27]. Para estudios en donde se desee triangular el sonido con fines de localización de individuos, los SM2 tienen la opción de ser conectados a un GPS.
- **Sistemas de grabación E3A**: las series Earthsong E3A DSS Bio-Acoustic Monitor Kit están diseñadas para grabar ambientes sonoros con una alta resolución, en estéreo. Este sistema utiliza dos micrófonos de condensador electret CZM (*Compression Zone Microphone*), y está patentado por River Foks Research Corporation (Patente # US 6,681,023 B1/20 Jan 2004), el cual incorpora una novedosa tecnología que combina las propiedades acústicas de un micrófono de zona de presión y un transformador acústico, en donde se incrementa la sensibilidad porque el sonido se refleja desde una superficie dura, creando una zona de mayor presión sonora [28],[29]. El equipo es protegido de la intemperie por una maleta a prueba de agua [29].
- **Teléfonos móviles y accesorios**: existen sensores incorporados en teléfonos

móviles que se programan para la captura de audio, tales como micrófonos externos omnidireccionales. Los datos se transmiten a través de Internet o conexiones móviles a una central de almacenamiento de datos. La energía es suministrada por paneles solares o baterías, algunas de estas unidades también utilizan preamplificadores reguladores de voltaje y controladores de poder<sup>1</sup>.

Debido a las diferencias tecnológicas de los equipos disponibles en el mercado, es recomendable usar un sólo tipo de equipo para efectos comparativos en programas de monitoreo y de esta manera reducir sesgos en la detectabilidad. En [26] se recomienda que en estudios de corto a mediano plazo se utilicen los mismos equipos por un máximo de cinco años, por diferentes razones primero, debido a que la tecnología cambia rápidamente, es mejor reemplazar con nuevas versiones de equipos y así poder gozar de las mejoras que los fabricantes hayan realizado; segundo, las versiones antiguas de los equipos salen del mercado y esto hace que sea casi imposible mantenerlas.

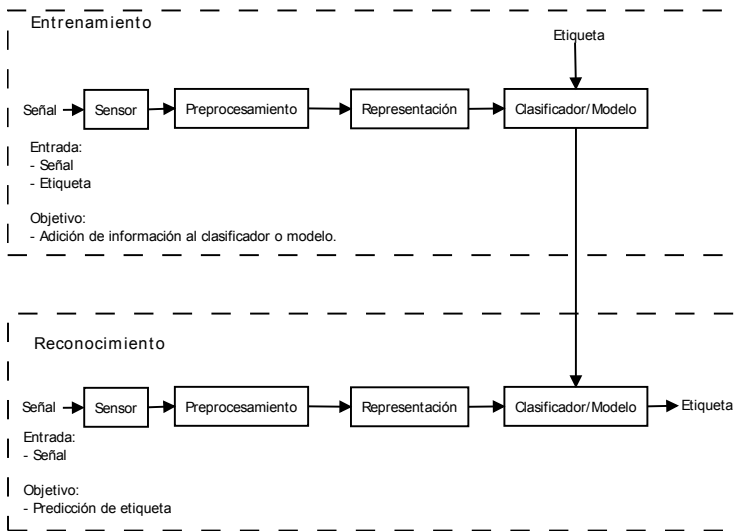
### 3 Herramientas para el reconocimiento automatizado de señales bioacústicas

Los procesos de análisis de señales acústicas se pueden llevar a cabo mediante algoritmos computacionales de reconocimiento. El objetivo de ello es tener la capacidad de analizar grandes cantidades de datos en lapsos cortos de tiempo y, además, permitir que ésto se haga de una forma sistematizada. La tecnología para el análisis automatizado de sonidos se ha enfocado casi completamente en reconocimiento del habla humana; sin embargo, hay una gran necesidad para la detección automatizada y clasificación de sonidos naturales no humanos [30]. En procesamiento de habla se han producido muchos avances en el análisis de series de tiempo, por esta razón, muchas de las técnicas que se usan en monitoreo bioacústico se apoyan en estas metodologías [4]. Los modelos de reconocimiento de voz generalmente se basan en la caracterización del vocabulario y la gramática. De manera similar, en las vocalizaciones de animales acústicamente activos, por ejemplo en aves, se pueden distinguir unidades básicas similares a los fonemas y, estructuras más complejas como cantos o frases [31], que son la base para crear los

---

<sup>1</sup><http://arbimon.com/arbimon/index.php/products-acoustics>

modelos de identificación automatizada [32]. En la Figura 1 se presenta un esquema de un sistema de monitoreo automatizado, se resalta que aunque este documento está orientado particularmente a los sistemas basados en señales acústicas, se puede adaptar fácilmente a un esquema que incluya datos de imágenes o datos del entorno tales como ubicación, condiciones atmosféricas, etcétera. En un sistema automatizado se distinguen dos partes: 1) entrenamiento y 2) reconocimiento; en el entrenamiento se diseña el mecanismo de inferencia a partir de información *a priori*; el reconocimiento consiste en el uso del sistema para la clasificación automatizada de la señal de entrada.



**Figura 1:** Etapas de un sistema de reconocimiento. El sistema se divide en entrenamiento (arriba) y reconocimiento (abajo). En cada una de estas partes se encuentran las etapas de adquisición, preprocesamiento, representación y clasificación.

Los sistemas de reconocimiento basados en bioacústica generalmente se diseñan para segmentar grabaciones, identificar la presencia de una especie objetivo o clasificar un segmento de canto de acuerdo a una baraja de opciones. Estos sistemas de reconocimiento se han aplicado en diferentes casos, por ejemplo en el enriquecimiento de la información de las coleccio-



nes de sonidos de tal manera que se facilita la navegación en bases de datos de audio, adicionando información de la localización en tiempo y frecuencia de las especies; esta necesidad surge porque aunque las grabaciones en archivos de sonidos de animales usualmente utilizan sistemas de catalogación mediante una serie de metadatos, es poco frecuente que las grabaciones tengan todas las especies de fondo catalogadas, bien sea por desconocimiento de las vocalizaciones de éstas o porque el colector no lo menciona [33].

La elaboración de algoritmos de clasificación de especies efectivos es un paso preliminar dirigido a la extracción de información útil de las grabaciones recolectadas en el campo [34]. Algunos de los desafíos relacionados con el desarrollo de software de reconocimiento son: los sonidos de animales se graban en ambientes altamente ruidosos [35], se pueden presentar distintas vocalizaciones al mismo tiempo (traslapadas), los individuos emisores de sonido pueden estar en movimiento (esto causa variaciones en intensidad); en el caso del reconocimiento automatizado en aves la dificultad depende, también, de la forma como las especies obtienen sus cantos. Las aves se dividen en dos grandes grupos, los sub-oscines que adquieren su repertorio de canto netamente a través de mecanismos genéticos, es decir son 100 % innatos, mientras que otro grupo los oscines o aves cantoras adquieren su repertorio mediante mecanismos genéticos y procesos de aprendizaje similares a la adquisición del habla en humanos [36]. Como consecuencia, la variabilidad en los cantos en individuos y poblaciones de una misma especie es alta, y es común que en estas especies se presenten dialectos y variaciones del canto debido a aislamientos geográficos [37]. Sin embargo, estas especies presentan llamados de corta duración que no varían sustancialmente entre poblaciones o notas dentro del canto que tampoco varían y que pueden utilizarse como las señales clave de reconocimiento especie-específico.

### **3.1 Estimación del desempeño en sistemas de reconocimiento automatizado**

La estimación del desempeño de un sistema de reconocimiento automatizado es fundamental, siendo necesario conocer la confiabilidad de los datos etiquetados por el sistema. Una forma básica de estimación del desempeño de clasificación o detección es la proporción de los objetos etiquetados correctamente, sin embargo, esto puede no ser muy ilustrativo en el caso

de problemas desbalanceados, es decir, cuando el número de objetos de cada clase difiere considerablemente, de manera que un sistema que indistintamente asigne la etiqueta de la clase mayoritaria a todas las muestras, podría reflejar un falso buen desempeño. En [38] se describen medidas de estimación de la fiabilidad de clasificadores biclase<sup>2</sup>, como la sensibilidad y la especificidad; éstas se pueden aplicar en tareas de reconocimiento multi-clase replanteando los problemas como en el caso de la técnica uno contra todos, es decir transformando el problema en varios problemas biclase, tomando en cada uno de ellos una clase como clase objetivo y las demás como no objetivo; o uno contra uno, en donde se evalúan los desempeños de cada par de clases [21],[39],[32]. En los problemas de clasificación de datos bioacústicos, frecuentemente se tiene una alta detección de falsos positivos [40]. En el caso de problemas multietiqueta —en donde un objeto puede pertenecer a varias clases— por ejemplo, en problemas de etiquetado de grabaciones, donde cada una de ellas puede contener vocalizaciones de diferentes especies, existen medidas especiales de estimación del desempeño, como *Hamming loss*, *Rank loss*, *One-error*, cobertura y *Micro-AUC* [41].

### 3.2 Preprocesamiento

El preprocesamiento es el conjunto de técnicas que se usan para mejorar la calidad de la señal o adaptarla a las condiciones requeridas para procesos posteriores como los de representación o clasificación. Algunos de los tratamientos que se suelen realizar son: filtrado (atenuación del ruido), normalización (estandarización de la intensidad de la señal de las grabaciones) y submuestreo (reducción de la cantidad de datos disminuyendo la frecuencia de muestreo). En los siguientes trabajos se describen algunos ejemplos de preprocesamiento: en [35] se submuestran las señales a 8kHz y se normalizan en amplitud; en [30] se usa un algoritmo de extracción de señales de ruido de fondo que consiste en la aplicación de filtros en donde el umbral se ajusta dependiendo de la banda de frecuencia, considerando que el nivel de ruido de fondo no es el mismo en cada una de ellas; en [41] se normalizan los valores en el espectrograma y se aplica un filtro para aumentar el

---

<sup>2</sup>Problema de clasificación con dos clases en donde los objetos pueden pertenecer a una clase denominada como clase positiva u objetivo o a otra denominada como negativa o no objetivo.

contraste de las vocalizaciones y reducir el ruido.

### 3.3 Segmentación

Las grabaciones contienen información acústica que no es relevante para los procesos de automatización del reconocimiento de especies, ya que las especies pueden presentar picos de actividad vocal o sus vocalizaciones están dispersas en el tiempo. En estos casos se requiere localizar los intervalos de tiempo que contienen vocalizaciones de animales en la grabación; este procedimiento se conoce como segmentación. Las vocalizaciones de algunos animales se han jerarquizado, por ejemplo en las aves, éstas se dividen en cuatro niveles: notas, sílabas, frases y cantos [42]; en la literatura se encuentran varios trabajos orientados a la clasificación de sílabas [4], [32],[35],[37], [42],[43]. Los métodos de segmentación más básicos están diseñados para detectar las regiones con mayor energía [44], sin embargo, en otros métodos más elaborados se busca detectar trayectorias características en las representaciones tiempo-frecuencia [37],[41],[45].

### 3.4 Métodos de representación

La representación de señales acústicas consiste en la selección de un conjunto de variables (llamadas también atributos, descriptores o características) que describan la información contenida en el sonido de una forma compacta, de tal forma que sea procesable computacionalmente. Existe un gran número de características que se pueden extraer de señales de audio y se usan para múltiples propósitos: reconocimiento automatizado de voz, evaluación de información musical, segmentación de audio y evaluación del sonido ambiental. En [17],[23],[24],[46],[47] se revisan diversos métodos para representar y procesar datos acústicos. A continuación se mencionan algunos descriptores tradicionales:

- **Representaciones tiempo-frecuencia:** Las señales de audio son series de tiempo. Para el análisis de patrones acústicos generalmente se usa la transformación de Fourier, el espacio transformado se conoce como dominio de la frecuencia cuya unidad de medida más común es el hercio (Hz). Se conoce como espectrograma la representación que se obtiene al realizar la transformada de Fourier a intervalos traslapados de la señal de audio.

El análisis tiempo-frecuencia de las señales bioacústicas en su gran mayoría se hace usando el espectrograma, además, éste tiene una gran acogida por los biólogos como herramienta de inspección visual. Sin embargo, en muchas aplicaciones de DSP se ha usado satisfactoriamente la transformada Wavelet, la cual es una modificación de la transformada de Fourier. Esta herramienta matemática ofrece información sobre la localización en tiempo y frecuencia de la señal y, al ser una representación bidimensional, se puede ilustrar de manera similar al espectrograma. Algunos trabajos en bioacústica resaltan sus propiedades en la representación de señales no estacionarias, con energía dispersa en un amplio rango de frecuencias, discontinuas y con picos de energía [10],[48],[49].

- **Los coeficientes cepstrales en las frecuencias de Mel:** Estos coeficientes, comúnmente conocidos por la sigla MFCCs (*Mel frequency cepstral coefficients*), son una de las representaciones más usadas para la clasificación de audio. A diferencia de la representación en el dominio de la frecuencia de Fourier, en los MFCCs la unidad de medida son los mels, en donde la escala de frecuencia se ubica de forma logarítmica, con lo cual se busca que la representación se acerque mejor a la percepción auditiva humana [34].
- **Codificación predictiva lineal:** La codificación predictiva lineal (*Linear predictive coding*, LPC) se basa en un análisis auto-regresivo, en donde se usa un predictor lineal para estimar el valor de cada muestra de una señal por una combinación lineal de valores anteriores. El objetivo de la LPC es estimar parámetros básicos, como las frecuencias formantes [47], que corresponden a las bandas de frecuencia donde se concentra la mayor parte de la energía de la señal acústica.

### 3.5 Métodos de clasificación

Gracias al desarrollo de técnicas de aprendizaje supervisado se pueden clasificar datos bioacústicos de forma rápida y precisa. La clasificación de vocalizaciones de animales tradicionalmente se ha relegado a enfoques cualitativos, estos requieren de un gran consumo de tiempo y frecuentemente están sesgados por la subjetividad [21]. En esta revisión se usa el término detección en los problemas cuyo objetivo es localizar la ubicación temporal de la emisión de la vocalización de una especie objetivo [4],[22],[23],[34],[35],[50]. El término clasificación se usa para etiquetar un canto [17],[21],[24],[30], para ello se usa un algoritmo conocido como clasificador, cuya entrada corresponde a un conjunto de valores relacionados con atributos o características del objeto; el clasificador requiere entrenarse previamente con un conjunto de datos estadísticamente representativos. Generalmente, en la etapa de preprocesamiento de los problemas de clasificación se requiere la aplicación de algoritmos de segmentación; en [4], por ejemplo, en donde se denomina

como unidad de reconocimiento a cada sílaba segmentada. Un desafío que no se ha estudiado a profundidad, teniendo en cuenta las variaciones temporales y geográficas de las vocalizaciones de animales, es el desarrollo de sistemas autónomos adaptativos que se acomoden a nuevas circunstancias con ninguna o poca intervención humana; un trabajo que tiene en cuenta esto es [23], en donde para la clasificación y detección se usa una metodología llamada sistema de soporte en línea de memoria perceptual (MESO: *Perceptual Memory to Support Online Learning in Adaptive Software*).

**3.5.1 Algoritmos de clasificación.** Los algoritmos de clasificación se usan para asignar etiquetas de clase a objetos ubicados en un espacio de representación o descritos por características [51],[52]. Dentro del área de reconocimiento de patrones, existen muchos algoritmos de clasificación establecidos, algunos de ellos se plantean como funciones matemáticas que mapean los valores de entrada a un conjunto con valores categóricos. Los siguientes clasificadores se han usado frecuentemente en reconocimiento de señales bioacústicas:  $k$  vecinos más cercanos ( $k$  nearest neighbors, kNN) [35], máquinas de vectores de soporte (*Support Vector Machine*, SVM) [3],[21],[22], bosque aleatorio [21],[22], *Boosting* adaptativo [22], funciones discriminantes [3],[21], redes neuronales [42], entre otros.

Frecuentemente, los problemas de clasificación se plantean en términos probabilísticos. Particularmente en reconocimiento de audio se usa la técnica conocida por la sigla en inglés HMMs (*hidden Markov models*). Estos modelos son un método probabilístico que ha sido usado amplia y exitosamente en diferentes escenarios de reconocimiento automatizado tales como la identificación de habla y hablante, la clasificación de señales sísmicas y la clasificación de secuencias de genes. Los HMMs resultan muy apropiados para el análisis de señales bioacústicas debido a su capacidad de tratar datos secuenciales con alta variabilidad (e, incluso, de diferentes longitudes) [34],[8],[32]. Estos cuentan con una clara y elegante interpretación bayesiana, además, se han desarrollado algoritmos rápidos y eficientes para realizar el entrenamiento y la prueba de los modelos [53].

## 4 Ejemplos de aplicación de monitoreo bioacústico automatizado

Actualmente, se están desarrollando redes de sensores inalámbricos que permiten adquirir paquetes de datos bioacústicos y datos ambientales (luz, velocidad del viento, humedad relativa, temperatura, precipitación, etcétera) simultáneamente, en diferentes escalas espaciales, en conjunto producen una recolección de datos ecológicos muy relevantes [54]. Una vez recolectados, los datos pueden ser agrupados y transmitidos a través de una red de computadores hacia un laboratorio de análisis bioacústico en donde se cuenta con la infraestructura necesaria para el almacenamiento, y capacidad técnica interdisciplinaria necesaria para el análisis de los procesos y patrones de la diversidad de la fauna vocalmente activa. El ciclo se cierra en la constante divulgación de la información ambiental y ecológica, lo cual es la base para comunicar a aquellos responsables de tomar decisiones sobre la conservación y el manejo del medio ambiente y para una divulgación transparente para el público en general. La clasificación automatizada de las vocalizaciones de diferentes especies (ya sea especies clave o un ensamble de especies) tiene el potencial de incrementar el conocimiento sobre los cambios de los ecosistemas debido a vectores de transformación (cambios de uso del suelo) y sus efectos sobre los patrones y procesos de la biodiversidad [23],[25],[55].

### 4.1 Reconocimiento bioacústico de poblaciones

Los estudios de poblaciones requieren de datos de alta calidad, y una forma de obtenerlos es mediante el reconocimiento de individuos [56]. Existen muchas técnicas que se han implementado para este fin, pero la mayoría utilizan capturas y marcaje, y a pesar de que a través de ellas se obtiene información valiosa, pueden causar cambios en el comportamiento de los individuos y hacerlos más visibles ante los predadores [57]. La individualidad en las vocalizaciones ha sido demostrada en diferentes especies de aves [58], así como en cetáceos [59] y elefantes. Las señales acústicas tienen el potencial de ser usadas como una alternativa al marcaje tradicional [57]. El uso de análisis canónicos discriminantes (DCA) fue efectivo para la correcta identificación de 12 individuos del ave Formicario mexicano (*Formicarius*

*monilinguer*), en bosques húmedos tropicales de México y esta técnica de monitoreo permitió mapear los territorios y monitorear la dinámica espacial y temporal de individuos previamente marcados y posteriormente reconocidos por sus vocalizaciones, gracias a lo cual se conocen aspectos comportamentales sexuales, sus movimientos en el tiempo, interacciones con el ambiente y dinámicas poblacionales [16].

## 4.2 Reconocimiento bioacústico de comunidades

El monitoreo acústico puede proveer datos confiables en riqueza y composición de especies. Esta técnica de monitoreo es un método eficiente y efectivo para evaluar la biodiversidad y sus cambios en escalas temporales y espaciales amplias. Las grabaciones con un único micrófono se han utilizado ampliamente para medir la riqueza y composición de especies, en aves, murciélagos, ranas e insectos [17]. La necesidad de implementar un sistema de monitoreo bioacústico automatizado nació en la década de los noventa, con estudios que buscaron evaluar la composición de especies de aves terrestres migratorias que hacen sus viajes durante la noche [60],[61], desde los sitios en donde se reproducen en el hemisferio norte hacia el hemisferio sur donde se resguardan de la época de invierno. Muchas de estas especies emiten cortos llamados mientras vuelan [62]. Estos llamados son especie-específicos y se cree que son utilizados por la dificultad de la comunicación visual; así, la bandada mantiene la ubicación espacial con el fin de evitar colisiones, mantiene también la cohesión durante el vuelo y puede estimar la dirección del viento. Con el fin de automatizar la clasificación de estos llamados por especies, Mills [60] creó un software para detectar esas vocalizaciones, en un rango de frecuencias entre 5 y 9 kHz, de duración entre los 50 y 100 ms. La detección utilizó los picos temporales de los llamados en las bandas de energía, para la clasificación se utilizaron redes neuronales. Otra aproximación al mismo problema utilizó los árboles de decisión. Este sistema de reconocimiento automatizado fue probado para 138 especies migratorias nocturnas australianas, 85 % de los llamados fueron clasificadas correctamente, 4 % incorrectamente y 8 % incierto.

## 5 Proyectos de monitoreo bioacústico automatizado

Los siguientes son proyectos de monitoreo bioacústico automatizado que se están desarrollando en la actualidad. En general ellos disponen de herramientas tecnológicas, hardware, software, equipos de comunicaciones y personal especializado:

- En Estados Unidos (EEUU), el programa de investigación en bioacústica *BRP* (Bioacoustic Research Program)<sup>3</sup> del Laboratorio de Ornitología de la Universidad Cornell, en Ithaca Nueva York, desarrolla equipos de grabación y software que usan científicos alrededor del mundo para estudiar la comunicación entre animales y monitorear la salud de las poblaciones de vida silvestre. Este laboratorio tiene dos ramas de investigación una de ellas en comunicación animal, en donde se llevan a cabo varios proyectos sobre monitoreo bioacústico y movimientos de varias especies, investigaciones que están en la vanguardia en la comprensión de patrones de migración y comunicación. La segunda rama de investigación es sobre algoritmos para desarrollo de software y métodos para el procesamiento de señales, con el fin de detectar, reconocer, localizar y hacer estudios comparativos de los sonidos de los animales.
- En Puerto Rico, el proyecto *ARBIMON* (Automated Remote Biodiversity Monitoring Network)<sup>4</sup> cuenta con estaciones temporales y permanentes de grabación y transmisión de archivos de audio, software para el manejo de base de datos y la clasificación automática de especies. Además, se desarrolla investigación orientada a la recolección de información ambiental mediante imágenes satelitales.
- En Europa, el proyecto *AMIBIO* (Automatic acoustic Monitoring and Inventorying of Biodiversity)<sup>5</sup> desarrolla estaciones autónomas con múltiples sensores en la región de Himeto (Grecia). Las estaciones de monitoreo se alimentan mediante paneles solares y se usan para grabar y transmitir señales de audio y datos climáticos. Los objetivos del proyecto son: evaluar la biodiversidad, estimar la densidad de animales, detectar la presencia o ausencia de especies raras o en peligro en áreas de difícil acceso, estimar la salud de ciertas especies, advertir la presencia de sonidos atípicos tales como la presencia de cazadores y detectar posibles desastres naturales inducidos por la actividad humana.
- El Laboratorio de Evaluación Ambiental Remota *REAL* (Remote Environmental Assessment Laboratory)<sup>6</sup> de la Universidad del Estado de Michigan estableció un observatorio ecológico a través de sensores para monitorear el ambiente de manera automatizada mediante una red de sensores que recolectan los datos de forma programada. Los paisajes sonoros son analizados en tiempo real con el

---

<sup>3</sup><http://www.birds.cornell.edu/brp/>

<sup>4</sup><http://arbimon.com/>

<sup>5</sup><http://www.amibio-project.eu/>

<sup>6</sup><http://www.real.msu.edu/>



fin de identificar las especies basándose en las vocalizaciones que emiten. Esta información es la base para las investigaciones bioacústicas e índices ecológicos. Toda esta información se almacena en una base de datos que se puede consultar por Internet.

## 6 Software comercial o libre

Los siguientes son programas computacionales que poseen herramientas de tratamiento de señales de audio y que se han usado para el análisis y reconocimiento bioacústico:

- Avisoft (<http://www.avisoft.com/>): herramienta para el análisis, edición y clasificación de sonidos. Software comercial, posee una versión “lite” para descargar.
- PAMGUARD (<http://www.pamguard.org/>): software especializado en la detección, localización y clasificación de mamíferos marinos. Código abierto.
- Raven (<http://www.birds.cornell.edu/brp/raven/RavenOverview.html>): software de análisis de sonidos del BRP de la Universidad Cornell. Software disponible en versiones comercial y libre.
- Song Scope (<http://www.wildlifeacoustics.com/products/song-scope-overview>): herramienta de análisis y reconocimiento de sonidos. Software comercial, tiene una versión de prueba de 15 días.
- SoundID (<http://www.soundid.net/>): herramienta de detección y reconocimiento de sonidos. Software comercial, tiene una versión de prueba de 30 días.
- XBAT (<http://www.birds.cornell.edu/brp/software>): herramienta de análisis de sonidos desarrollada con el software de análisis matemático MATLAB. Código abierto (MATLAB tiene licencia comercial).
- SonoBat (<http://www.sonobat.com/>): herramienta de análisis y reconocimiento de vocalizaciones de murciélagos. Software comercial y versión de prueba.

## 7 Retos y oportunidades para Colombia

Colombia es un país rico en biodiversidad y fauna vocalmente activa, cuenta con más de 1.860 especies de aves [5], 763 especies de anfibios, 479 especies de mamíferos, 2.000 especies de peces marinos y 1.435 especies de peces dulceacuícolas<sup>7</sup>. Las aves son un grupo que ha sido utilizado como indicador de cambios en la naturaleza, están distribuidas ampliamente, son

---

<sup>7</sup><https://www.siac.gov.co/contenido/contenido.aspx?conID=1252&catID=52>

fáciles de detectar por sus sonidos y, en comparación con otros grupos de animales, se cuenta con gran conocimiento de la biología de la mayoría de las especies. En Colombia existen cerca de 20 asociaciones de ornitología dispersas por el territorio nacional, las cuales pueden presentar interés en este tipo de monitoreo. A escala global, la contribución de asociaciones ornitológicas y de observadores de aves ha incrementado el conocimiento de las tendencias poblacionales de las aves, especialmente en Europa y Norte América. En Colombia, desde 1987, se han realizado censos anuales de aves, los cuales han contribuido con el incremento del conocimiento y la cohesión de las asociaciones ornitológicas; el método de muestreo empleado en estos censos anuales es puntos de conteo. Si las asociaciones ornitológicas implementaran un sistema de monitoreo acústico automatizado, bajo un diseño experimental sistematizado podrían operar no sólo por una vez al año, sino tener datos a escalas temporales y espaciales más amplias y de esta manera, contar con información a diferentes niveles, de comunidades y poblaciones.

Es cada vez más creciente el interés de diferentes entidades ante este tipo de aproximación a la biodiversidad, como por ejemplo fundaciones y organizaciones no gubernamentales (ONGs), instancias académicas en diferentes zonas de Colombia, quienes en conjunto podrían realizar nodos de monitoreo, bajo un sistema estandarizado de toma de datos en campo e identificación de especies en laboratorio. El Instituto Humboldt cuenta con dos soportes ideales para el desarrollo de un sistema de monitoreo bioacústico: la Colección de Sonidos Ambientales y el Laboratorio de Biogeografía Aplicada y Bioacústica. La Colección de Sonidos tiene la capacidad para ser el centro de almacenamiento de los especímenes bioacústicos; de esta manera, se garantiza su protección a largo plazo, y en conjunto con entidades académicas con experiencia en ingeniería electrónica, ciencias de la computación e informática y/o bioacústica nacionales e internacionales, podrían ofrecer las herramientas para la automatización del reconocimiento de las especies de fauna vocalmente activa, publicar esta información de manera gratuita y tener un sistema de información en línea sobre este tipo de monitoreo.

El monitoreo bioacústico automatizado es una herramienta llamativa en diferentes contextos, por ejemplo para el turismo ornitológico en Colombia, ya que es posible el desarrollo de aplicaciones para teléfonos móviles, y de

esta manera se facilita el reconocimiento de las especies en campo. El uso de sonidos de la fauna en sistemas educativos mediante textos escolares multimedia concientizaría a un público infantil sobre la alta biodiversidad del país. Adicionalmente, incrementar el conocimiento de este tipo de monitoreo en Colombia permitirá que empresas privadas tengan sistemas de monitoreo confiables y eficaces, con un soporte físico de la evidencia de la presencia de las especies en sus zonas de acción, ideal para estudios de impacto y planes de manejo ambiental que la Autoridad Nacional de Licencias Ambientales (ANLA) exige.

En cuanto al diseño e implementación de un sistema de reconocimiento automatizado, hay tres cuellos de botella: a) recolección de un número estadísticamente suficiente de ejemplos de vocalizaciones de las especies objetivo y no objetivo; b) segmentación de las grabaciones para extraer las vocalizaciones y descartar porciones de audio que no sean de interés; c) transformación de un sistema prototipo a uno definitivo, esto se refiere al desarrollo de interfaces de usuario, manuales, legalización de licencias y en general todo lo que se debe adicionar al sistema para ponerlo en funcionamiento. Las tareas a) y b) requieren más que experiencia técnica en informática y electrónica, la colaboración de expertos en ecología, esto se puede llevar a cabo a través de convenios de cooperación entre grupos o instituciones con personal capacitado en ciencia y tecnología. Una vez se cuenta con un conjunto de ejemplos para la generación de un prototipo de reconocimiento automatizado, la transformación a uno definitivo y la instalación son actividades que implican inversión en personal y, en la compra de equipos y software.

## 8 Conclusiones

Las herramientas tecnológicas disponibles actualmente, de adquisición, almacenamiento y procesamiento de información, se pueden usar como apoyo a actividades de recolección ambiental facilitando tareas como: conteo de especies, evaluación rápida de la biodiversidad con amplia cobertura espacial y temporal, enriquecimiento de la información en los estudios de ecología y conservación, desarrollo de actividades de monitoreo automatizado, monitoreo de especies raras o en peligro, detección de especies invasoras,

estimación del impacto del cambio climático, análisis de los efectos de la actividad antropogénica, entre otras [19].

Las señales acústicas ambientales no se pueden adquirir en condiciones controladas, de manera que los sistemas automatizados de reconocimiento deben contar con métodos robustos al ruido, a las variaciones en intensidad y deben ser rápidos (teniendo en cuenta la gran cantidad de datos que se requiere inspeccionar). Además, considerando la naturaleza aleatoria y no estacionaria de las señales bioacústicas (variaciones temporales y espaciales de sus parámetros), estos sistemas deberían permitir la inclusión periódica de correcciones sugeridas por expertos.

El monitoreo bioacústico tiene el potencial de suplir información en tiempo real de varios grupos taxonómicos simultáneamente, de manera sistematizada, abarcando amplias escalas tanto espaciales como temporales. Mediante este tipo de monitoreo se puede contar con información precisa sobre el comportamiento de las dinámicas de la naturaleza ante factores de cambio, sensores que están asociados al desarrollo de la sociedad, por ejemplo la minería, la agricultura, la infraestructura, así como factores derivados del desarrollo humano, como el cambio climático. El uso de esta novedosa técnica de monitoreo permite crear sistemas de alerta temprana con alta eficacia, de tal manera que da un soporte con base en datos contemporáneos, a los tomadores de decisiones ambientales.

En el mundo se han desarrollado exitosamente investigaciones y proyectos sobre monitoreo de la biodiversidad usando este tipo de herramientas tecnológicas. La cooperación entre instituciones facilitará la implementación y el uso de esta tecnología en el país; un esquema de monitoreo bioacústico automatizado estándar permite unificar esfuerzos aislados de investigaciones relacionadas con el tema. Siendo Colombia un país megadiverso, con alto número de especies, endémicas, amenazadas y ecosistemas estratégicos se justifican los esfuerzos por implementar este tipo de sistemas de monitoreo de nuestra biodiversidad.

## Agradecimientos

Esta revisión se realizó en el marco del convenio de cooperación no. 12-020 celebrado entre el Instituto de Investigación de Recursos Biológicos

“Alexander von Humboldt” y la Universidad Nacional de Colombia - Sede Manizales.

## Referencias

- [1] M. Depraetere, S. Pavoine, F. Jiguet, A. Gasc, S. Duvail, and J. Sueur, “Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland,” *Ecological Indicators*, vol. 13, no. 1, pp. 46 – 54, 2012. 172, 175
- [2] D. M. Marsh and P. C. Trenham, “Current trends in plant and animal population monitoring,” *Conservation Biology*, vol. 22, no. 3, pp. 647–655, 2008. 173
- [3] M. A. Acevedo, C. J. Corrada-Bravo, H. Corrada-Bravo, L. J. Villanueva-Rivera, and T. M. Aide, “Automated classification of bird and amphibian calls using machine learning: A comparison of methods,” *Ecological Informatics*, vol. 4, no. 4, pp. 206 – 214, 2009. 173, 176, 183
- [4] C.-H. Lee, C.-C. Han, and C.-C. Chuang, “Automatic classification of bird species from their sounds using two-dimensional cepstral coefficients,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1541 –1550, Nov. 2008. 173, 175, 177, 181, 182
- [5] J. E. Botero, D. Arbeláez, and G. M. Lentijo, “Métodos para estudiar las aves,” *Biocarta*, no. 8, pp. 1–4, Jul. 2005. [Online]. Available: <http://www.rnoa.org/sco/pdf/Biocarta008.pdf> 173, 187
- [6] W. J. Sutherland, *Ecological census techniques: a handbook*, 2nd ed. Cambridge University Press, 2006. 173
- [7] C. Ten Cate, “Birdsong and Evolution,” in *Nature’s music: the science of birdsong*, 2004, ch. 10, pp. 296–317. 173
- [8] W. Chu and D. T. Blumstein, “Noise robust bird song detection using syllable pattern-based hidden Markov models,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011*, may 2011, pp. 345 –348. 173, 183
- [9] B. R. Grant and P. R. Grant, “Hybridization and speciation in darwin’s finches: the role of sexual imprinting on a culturally transmitted trait,” *Endless Forms: Species and Speciation*, pp. 404–422, 1998. 173
- [10] R. Bardeli, D. Wolff, F. Kurth, M. Koch, K. H. Tauchert, and K.-H. Frommolt, “Detecting bird sounds in a complex acoustic environment and application to bio-acoustic monitoring,” *Pattern Recognition Letters*, vol. 31, no. 12, pp. 1524–1534, Sep. 2010. 174, 175, 182
- [11] D. C. Cugler, C. B. Medeiros, and L. F. Toledo, “Managing animal sounds-some challenges and research directions,” in *Proceedings V eScience Workshop-XXXI Brazilian Computer Society Conference*, 2011. 174

- [12] A. Farnsworth and R. W. Russell, “Monitoring flight calls of migrating birds from an oil platform in the northern gulf of Mexico,” *Journal of Field Ornithology*, vol. 78, no. 3, pp. 279–289, 2007. 174
- [13] D. Mennill and S. Vehrencamp, “Context-dependent functions of avian duets revealed through microphone array recordings and multi-speaker playback,” *Current Biology*, vol. 18, pp. 1314–1319, 2008. 174
- [14] K. J. Odom and D. J. Mennill, “A quantitative description of the vocalizations and vocal activity of the barred owl,” *The Condor*, vol. 112, no. 3, pp. 549–560, 2010. 174
- [15] S. Gaunt and A. McCallum, “Birdsong and Conservation,” in *Nature’s music: the science of birdsong*, 2004, ch. 12, pp. 343–362. 174
- [16] A. Kirschel, M. Cody, Z. Harlow, V. Promponas, E. Vallejo, and C. Taylor, “Territorial dynamics of Mexican ant-thrushes *formicarius moniliger* revealed by individual recognition of their songs,” *Ibis*, vol. 153, no. 2, pp. 255–268, 2011. 174, 185
- [17] D. T. Blumstein, D. J. Mennill, P. Clemins, L. Girod, K. Yao, G. Patricelli, J. L. Deppe, A. H. Krakauer, C. Clark, K. A. Cortopassi, S. F. Hanser, B. McCowan, A. M. Ali, and A. N. G. Kirschel, “Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus,” *Journal of Applied Ecology*, vol. 48, no. 3, pp. 758–767, 2011. 174, 175, 176, 181, 182, 185
- [18] D. Chesmore, “The automated identification of taxa: Concepts and applications,” in *Automated Taxon Identification in Systematics: Theory, Approaches and Applications*, ser. Systematics Association special volumes, N. MacLeod, Ed. Boca Raton, FL: CRC Press, 2008, vol. 74, ch. 6, pp. 83–100. 174, 175
- [19] D. Chesmore, K.-H. Frommolt, D. Wolff, R. Bardeli, and S. Huebner, “Computational bioacoustics: New tools for assessing biological diversity,” May 2008, side Event at the ninth meeting of the Conference of the Parties (COP 9). Bonn, Germany. 174, 175, 190
- [20] M. A. Acevedo and L. J. Villanueva-Rivera, “Using automated digital recording systems as effective tools for the monitoring of birds and amphibians,” *Wildlife Society Bulletin*, vol. 34, pp. 211 – 214, 2006. 174, 176
- [21] D. W. Armitage and H. K. Ober, “A comparison of supervised learning techniques in the classification of bat echolocation calls,” *Ecological Informatics*, vol. 5, no. 6, pp. 465 – 473, 2010. 174, 180, 182, 183
- [22] F. Briggs, X. Fern, and R. Raich, “Acoustic classification of bird species from syllables: an empirical study,” Oregon State University, Tech. Rep., 2009. 174, 182, 183
- [23] E. P. Kasten, P. K. McKinley, and S. H. Gage, “Ensemble extraction for classification and detection of bird species,” *Ecological Informatics*, vol. 5, no. 3, pp. 153 – 166, 2010. 175, 181, 182, 183, 184

- [24] T. S. Brandes, “Automated sound recording and analysis techniques for bird surveys and conservation,” *Bird Conservation International*, vol. 18, no. S1, pp. S163–S173, 2008. 175, 181, 182
- [25] R. Mason, P. Roe, M. Towsey, J. Zhang, J. Gibson, and S. Gage, “Towards an acoustic environmental observatory,” in *eScience, 2008. eScience '08. IEEE Fourth International Conference on*, 2008, pp. 135–142. 176, 184
- [26] R. S. Rempel, C. M. Francis, J. N. Robinson, and M. Campbell, “Comparison of audio recording system performance for detecting and monitoring songbirds,” *Journal of Field Ornithology*, vol. 84, no. 1, pp. 86–97, 2013. 176, 177
- [27] D. Mennill, M. Battiston, D. R. Wilson, J. R. Foote, and S. M. Doucet, “Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour,” *Methods in Ecology and Evolution*, vol. 3, pp. 704–712, 2012. 176
- [28] K. A. Hobson, R. S. Rempel, H. Greenwood, B. Turnbull, and S. L. Van Wilgenburg, “Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system,” *Wildlife Society Bulletin*, pp. 709–720, 2002. 176
- [29] L. A. Venier, S. B. Holmes, G. W. Holborn, K. A. Mcilwrick, and G. Brown, “Evaluation of an automated recording device for monitoring forest birds,” *Wildlife Society Bulletin*, vol. 36, no. 1, pp. 30–39, 2012. 176
- [30] T. S. Brandes, “Feature vector selection and use with hidden Markov models to identify frequency-modulated bioacoustic signals amidst noise,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 6, pp. 1173–1180, aug 2008. 177, 180, 182
- [31] P. Caycedo-Rosales, “Estudio comparativo del canto entre poblaciones del soterrey de apolinar (*cistothorus apolinaris*, troglodytidae) presentes en la cordillera oriental de los andes colombianos,” Universidad Nacional de Colombia, 2000, tesis de pregrado. 177
- [32] B. Lakshminarayanan, R. Raich, and X. Fern, “A syllable-level probabilistic framework for bird species identification,” in *Proceedings of the Fourth International Conference on Machine Learning and Applications*. Los Alamitos, CA, USA: IEEE Computer Society, Dec. 2009, pp. 53–59. 178, 180, 181, 183
- [33] R. Bardeli, “Similarity search in animal sound databases,” *IEEE Transactions on Multimedia*, vol. 11, no. 1, pp. 68–76, jan 2009. 179
- [34] F. Briggs, R. Raich, and X. Z. Fern, “Audio classification of bird species: A statistical manifold approach,” in *Proceedings of the Ninth IEEE International Conference on Data Mining*. Los Alamitos, CA, USA: IEEE Computer Society, Dec. 2009, pp. 51–60. 179, 182, 183
- [35] C.-J. Huang, Y.-J. Yang, D.-X. Yang, and Y.-J. Chen, “Frog classification using machine learning techniques,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 3737–3743, 2009. 179, 180, 181, 182, 183

- [36] D. Todt, "From birdsong to speech: a plea for comparative approaches," *Anais da Academia Brasileira de Ciências*, vol. 76, no. 2, pp. 201–208, 2004. 179
- [37] L. Ranjard and H. A. Ross, "Unsupervised bird song syllable classification using evolving neural networks," *Journal of the Acoustical Society of America*, vol. 123, no. 6, pp. 4358 – 4368, 2008. 179, 181
- [38] M. Bramer, *Principles of Data Mining*. Springer, 2007. 180
- [39] S. Fagerlund, "Bird species recognition using support vector machines," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, pp. 64–64, 2007. 180
- [40] I. Urazghildiiev, C. Clark, T. Krein, and S. Parks, "Detection and Recognition of North Atlantic Right Whale Contact Calls in the Presence of Ambient Noise," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 3, pp. 358–368, Jul. 2009. 180
- [41] F. Briggs, B. Lakshminarayanan, L. Neal, X. Z. Fern, R. Raich, S. J. K. Hadley, A. S. Hadley, and M. G. Betts, "Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach," *Journal of the Acoustical Society of America*, vol. 131, no. 6, pp. 4640–4650, Oct. 2012. 180, 181
- [42] C.-H. Chou, P.-H. Liu, and B. Cai, "On the studies of syllable segmentation and improving MFCCs for automatic birdsong recognition," in *Proceedings of the 2008 IEEE Asia-Pacific Services Computing Conference*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 745–750. 181, 183
- [43] A. Härmä, "Automatic identification of bird species based on sinusoidal modeling of syllables," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '03*, vol. 5, Apr. 2003, pp. 545–548. 181
- [44] D. Stowell and M. D. Plumbley, "Birdsong and C4DM: A survey of UK birdsong and machine recognition for music researchers," Centre for Digital Music, Queen Mary, University of London, London, UK, Tech. Rep. C4DM-TR-09-12, Jul. 2011. 181
- [45] L. Neal, F. Briggs, R. Raich, and X. Z. Fern, "Time-frequency segmentation of bird song in noisy acoustic environments," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011*, May 2011, pp. 2012 – 2015. 181
- [46] C.-H. Lee, S.-B. Hsu, J.-L. Shih, and C.-H. Chou, "Continuous Birdsong Recognition Using Gaussian Mixture Modeling of Image Shape Features," *IEEE Transactions on Multimedia*, vol. 15, pp. 454–464, 2013. 181
- [47] D. Mitrovic, M. Zeppelzauer, and C. Breiteneder, "Features for content-based audio retrieval," in *Advances in Computers: Improving the Web*, ser. Advances in Computers, M. V. Zelkowitz, Ed. Elsevier, 2010, vol. 78, pp. 71 – 150. 181, 182
- [48] S. Ntalampiras, I. Potamitis, and N. Fakotakis, "Probabilistic Novelty Detection for Acoustic Surveillance Under Real-World Conditions," *IEEE Transactions on Multimedia*, vol. 13, no. 4, pp. 713–719, Aug. 2011. 182
- [49] A. Selin, J. Turunen, and J. T. Tanttu, "Wavelets in recognition of bird sounds," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 141–141, Jan. 2007. 182



- [50] J. L. Goyette, R. W. Howe, A. T. Wolf, and W. D. Robinson, "Detecting tropical nocturnal birds using automated audio recordings," *Journal of Field Ornithology*, vol. 82, no. 3, pp. 279–287, 2011. 182
- [51] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, 1st ed. Springer, 2007. 183
- [52] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. Wiley, 2001. 183
- [53] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, May 1989. 183
- [54] D. Chesmore, "Automated bioacoustic identification of species," *Anais da Academia Brasileira de Ciências*, vol. 76, no. 2, pp. 436–440, 2004. 184
- [55] J. Wimmer, M. Towsey, B. Planitz, P. Roe, and I. Williamson, "Scaling acoustic data analysis through collaboration and automation," in *e-Science (e-Science), 2010 IEEE Sixth International Conference on*. IEEE, 2010, pp. 308–315. 184
- [56] W. J. Sutherland, *From individual behaviour to population ecology*. Oxford University Press Inc., 1996. 184
- [57] P. K. McGregor, T. M. Peake, and G. Gilbert, "Communication behaviour and conservation," in *Behaviour and Conservation*. Cambridge University Press, 2000. 184
- [58] H. Slabbekoorn, "Singing in the wild: the ecology of birdsong," in *Nature's music: the science of birdsong*, 2004, ch. 1, pp. 178–2005. 184
- [59] V. Deecke, J. Ford, and P. Spong, "Quantifying complex patterns of bioacoustic variation: use of a neural network to compare killer whale (*orcinus orca*) dialects," *The Journal of the Acoustical Society of America*, vol. 105, p. 2499, 1999. 184
- [60] H. Mills, "Automatic detection and classification of nocturnal migrant bird calls," *The Journal of the Acoustical Society of America*, vol. 97, p. 3370, 1995. 185
- [61] A. Taylor, "Bird flight call discrimination using machine learning," *The Journal of the Acoustical Society of America*, vol. 97, p. 3370, 1995. 185
- [62] W. M. Tyler, "The call-notes of some nocturnal migrating birds," *The Auk*, pp. 132–141, 1916. 185