

Clasificación de especies de aves en base a características del canto

A. Nassiff, M. Rolé, G. Tessi

Supervisores: M. Gerard, T. Molas Giménez, L. Vignolo

Resumen—El presente trabajo surge a partir de la necesidad de desarrollar una metodología que permita identificar la especie a la que pertenece un ave determinado, según sus características en el canto. A raíz de estudios previos sobre dicha problemática que demuestran la posibilidad de identificar las especies según su canto, se ha logrado un desarrollo innovador que permite alcanzar este objetivo. Este desarrollo se basa, principalmente, en la extracción y manipulación de características del canto, las cuales conformarán *índices*. Debido a que estos valores no son fijos para los cantos de varios ejemplares de una misma especie, se utilizará un enfoque estadístico tanto para la creación de la base de datos como para la identificación propiamente dicha. Para verificar la robustez de la metodología ante la presencia de ruido en un canto a identificar, se realizaron diferentes pruebas contaminando señales conocidas de forma tal que la relación señal-ruido varíe en un rango preestablecido de valores. Los resultados obtenidos al aplicar esta metodología han sido satisfactorios, alcanzando así el objetivo propuesto anteriormente.

Palabras clave—identificación, canto, aves, índices, sílabas

I. INTRODUCCIÓN

La Argentina, un país con una gran variedad de climas, presenta diferentes variedades de aves habiendo unas 979 especies diferentes, distribuidas en diferentes regiones del territorio como selvas tropicales, lagunas, ríos, montañas y territorios llanos. Debido a la gran variedad existente, resulta interesante identificarlas *a priori* con un buen grado de certeza.

Un ave puede ser caracterizada por diferentes aspectos ya sean visuales, sonoros, de su hábitat natural, comportamiento, entre otros. El presente trabajo se enfoca en desarrollar una metodología que permita identificar las aves a partir de características en el canto.

El canto de las aves es una estructura compleja a la hora de su análisis. El mismo puede descomponerse en cuatro niveles de jerarquía: notas, sílabas, frases y canción [1], [2]. Los cambios en frecuencia para estos niveles se producen a gran velocidad en comparación a los del ser humano, de varios órdenes de magnitud, y su espectro es mucho más amplio que el de estos últimos. La mayoría de la especies tienen cantos que pueden ser considerados como señales cuasi-periódicas [3], como se ilustra en las figuras 1 y 2.

Utilizando diferentes técnicas de procesamiento digital de señales, es posible aislar las características de un canto para su análisis. Una forma de modelar las mismas es utilizando *índices* [3], [4], los cuales representan de manera numérica rasgos particulares y distintivos del canto de las diferentes especies.

El enfoque del presente trabajo estará orientado en la utilización de *índices*, para lo cual se definirán seis tipos diferentes a partir de propiedades temporales y frecuenciales del canto.

Estudios anteriores han intentado identificar las diferentes especies de aves a partir de sus características sonoras,

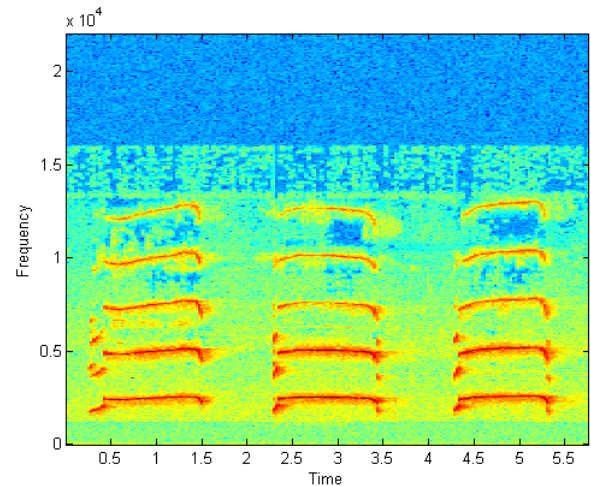


Fig. 1. Espectrograma del canto de un Águila Poma.

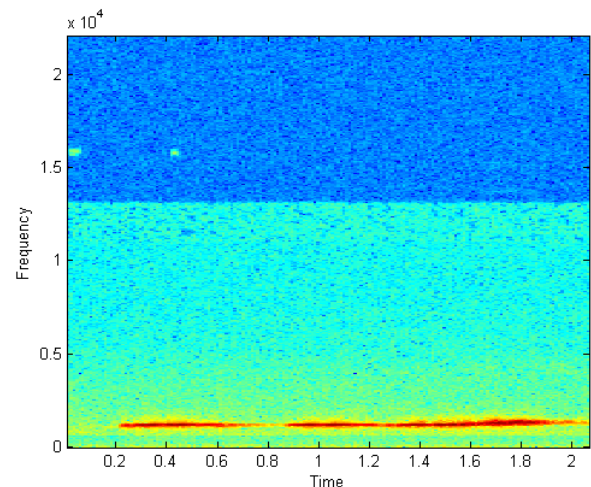


Fig. 2. Espectrograma del canto de un Tataupa Listado.

mediante técnicas como los modelos ocultos de Markov y la aproximación de sílabas del canto por ondas sinusoidales moduladas en amplitud y frecuencia [5].

A fin de poder comparar resultados y obtener conclusiones sobre el éxito de este trabajo, se considerarán los resultados obtenidos mediante el método de identificación de aves por medio de plantillas [6], y por ello se utilizará el mismo conjunto de especies de aves.

II. MATERIALES

Para la realización del presente trabajo, se requieren grabaciones de cantos de aves de diferentes especies, de forma que pueda ponerse a prueba el método a desarrollar.

Estas grabaciones han sido obtenidas del DVD "Sonidos

de Aves de Argentina y Uruguay”, de Bernabe Lopez-Lanus y facilitado por Lucas Genzelis.

III. METODOLOGÍA

El método consta de dos etapas, referidas a la creación de la base de datos que se utilizará para las comparaciones durante la identificación, y a la identificación propiamente dicha.

La primer etapa consiste en procesar los cantos de varias especies, tanto en los dominios temporal y frecuencial, para obtener sus características. Dichas características se utilizarán para crear *índices*, que serán los valores a comparar durante la identificación. Por último se calcula la media y el desvío estándar para cada conjunto de *índices* de cada especie.

Para poder manipular los *índices*, se optó por utilizar un vector con la siguiente estructura

$$\begin{bmatrix} \text{duracion de silabas} \\ \text{tiempo entre silabas} \\ \text{cantidad de silabas} \\ \text{frecuencia principal 1} \\ \text{frecuencia principal 2} \\ \text{frecuencia espectral media} \end{bmatrix}$$

La segunda etapa consiste en calcular los *índices* de la señal incógnita y compararlos con los de la base de datos, utilizando la función de distribución gaussiana como criterio estadístico.

Estas etapas se detallan a continuación.

A. Etapa 1 - Creación de la base de datos de índices

Las sílabas¹ son las características principales en el dominio temporal [7], las cuales son determinadas a partir de un proceso de ventaneo (utilizando ventanas de Hanning), con solapamiento de un tercio del tamaño de la ventana [4], [5]. A cada ventana se le calcula la energía y se obtiene la máxima de entre todas ellas. Luego, se toma el 10 % de dicha energía máxima como umbral mínimo y se lo compara con la energía de todas las ventanas. Si ésta es superior al umbral, la posición de inicio de la ventana indica el inicio de una sílaba; el final de la misma es determinado cuando la energía es inferior al umbral, tomando la posición de fin de la ventana. Una vez finalizado este proceso, se obtienen los valores en muestras que determinan el inicio y final de todas las sílabas en una señal, y con este conjunto de datos se pueden calcular los tres primeros *índices* de la siguiente manera:

- La duración de las sílabas se calcula como la diferencia entre el tiempo final e inicial de cada una.
- El tiempo entre sílabas se calcula como la diferencia entre el tiempo inicial de la sílaba siguiente con el tiempo final de la sílaba actual, siendo cero en caso de que haya una sola sílaba.
- La cantidad de sílabas se calcula contando la cantidad de intervalos de duración.

Una vez identificadas las sílabas para un canto, se buscan los *índices* en el dominio frecuencial. Se calcula la Transformada Discreta de Fourier de cada una de ellas, analizado siempre el rango de frecuencias positivas, y se buscan las dos primeras frecuencias principales (de mayor magnitud)

y la *frecuencia espectral media* (MSF) [8]. Las frecuencias se calculan a partir de sus magnitudes, considerando para la primera el valor máximo de magnitud, y para la segunda anulando un rango de valores alrededor de la primera y buscando nuevamente la magnitud máxima. La MSF se calcula como

$$MSF = \frac{\sum_i \omega_i \varepsilon_i}{E} \quad (1)$$

siendo ω la frecuencia, ε la magnitud asociada a dicha frecuencia y $E = \sum_i \varepsilon_i$ la energía total. Finalizado este proceso, se obtienen los *índices* restantes.

B. Etapa 2 - Identificación

Este proceso consiste en calcular el conjunto de *índices* de la especie incógnita y a partir de este, generar una matriz de probabilidades utilizando la función de distribución normal, teniendo en cuenta la media y desvío de cada especie que se encuentran en la base de datos.

Las probabilidades se calculan a partir de la función de distribución normal, la cual se define como

$$P(x) = e^{-\left(\frac{(x - \mu)}{\sigma}\right)^2} \quad (2)$$

donde x es el *índice* de la especie incógnita y μ y σ son la media y el desvío, respectivamente, para el mismo *índice* de la especie a comparar en la base de datos.

Como el objetivo es determinar la especie de nuestra señal incógnita, se calcula el promedio de las probabilidades para cada especie y se toma el valor más alto de entre todos los promedios. De esta manera, se obtiene como resultado la especie identificada.

IV. ANÁLISIS DE RESULTADOS

Para corroborar los resultados de la metodología se utilizó el mismo conjunto de señales con el que se creó la base de datos, se calcularon los *índices* de cada una de ellas y se los comparó con aquellos en la base de datos. En la Tabla I se presentan los resultados obtenidos al calcular la tasa de errores para cada especie.

TABLA I
RESULTADOS DE LA IDENTIFICACIÓN DE LAS ESPECIES UTILIZADAS
PARA LA CREACIÓN DE LA BASE DE DATOS.

Especies	Porcentaje de certeza %	Cantidad de errores
Águila Poma	100	0
Aguilucho Andino	80	1
Batará Negro	80	1
Chiflón	100	0
Chimango	100	0
Inambú Montaraz	100	0
Paloma Domestica	100	0
Pato Overo	100	0
Pingüino Patagónico	100	0
Tataupa Listado	100	0

Se puede apreciar como la metodología identifica correctamente las especies de las señales del conjunto utilizado para la creación de la base de datos. También se realizaron pruebas con señales que no pertenecían a dicho conjunto, cuyos resultados se muestran en la Tabla II.

Hay que destacar que estas señales de prueba contienen ruido, ya que han sido grabadas del ambiente natural; a

¹También se denominan *explosiones*.

TABLA II
RESULTADOS DE LA IDENTIFICACIÓN DE LAS ESPECIES CON EL
CONJUNTO DE PRUEBAS.

Especie Real	Especie Identificada	Probabilidad
Águila Poma 1	Águila Poma	95.61
Águila Poma 2	Águila Poma	95.60
Aguilucho Andino 1	Aguilucho Andino	79.18
Aguilucho Andino 2	Aguilucho Andino	79.17
Batará Negro 1	Batará Negro	96.69
Batará Negro 2	Batará Negro	94.43
Chiflón 1	Chiflón	95.40
Chiflón 2	Chiflón	95.80
Chimango 1	Chimango	90.78
Chimango 2	Chimango	86.60
Inambú Montaraz 1	Inambú Montaraz	96.80
Inambú Montaraz 2	Inambú Montaraz	97.52
Paloma Domestica 1	Paloma Domestica	76.72
Paloma Domestica 2	Paloma Domestica	73.20
Pato Overo 1	Pato Overo	92.47
Pato Overo 2	Pato Overo	92.00
Pingüino Patagónico 1	Pingüino Patagónico	91.92
Pingüino Patagónico 2	Pingüino Patagónico	84.95
Tataupa Listado 1	Tataupa Listado	94.50
Tataupa Listado 2	Tataupa Listado	96.47

pesar de ello, se logra un 100% de coincidencia. Por lo tanto, se puede afirmar que los resultados obtenidos mediante esta metodología son apropiados.

Para verificar cuánto influye el ruido en la identificación, se le agrega a la señal de prueba un ruido aleatorio, de forma tal que la relación señal-ruido de la señal varíe entre 1 y 50 dB. En la figura 3 se puede apreciar como la probabilidad de identificación para una señal que contiene el canto del Águila Poma disminuye a medida que el ruido aumenta, y que a valores de SNR inferiores a 18 dB el reconocimiento es erróneo ya que la probabilidad disminuye a menos del 65 %.

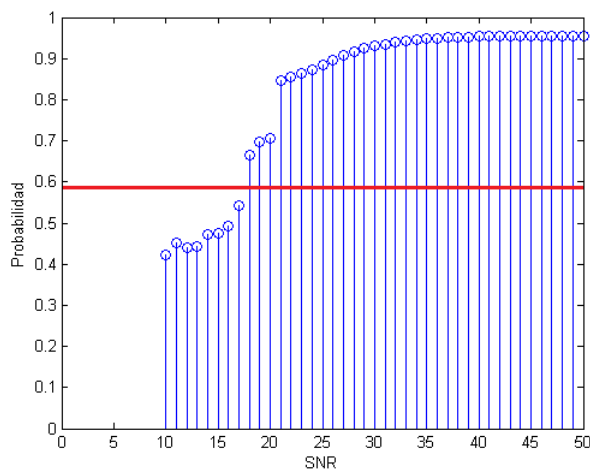


Fig. 3. Variación de la relación señal-ruido para el Águila Poma.

V. CONCLUSIONES Y TRABAJOS FUTUROS

A pesar de la incertidumbre a la hora de definir los *índices*, así como también su cantidad e importancia relativa, la metodología ha mostrado ser eficiente al presentar buenos resultados con muy poco error.

En contraste con el estudio de [6], la utilización de *índices* que representen fehaciemente las características propias del canto de las aves, otorga resultados muy similares para

el mismo conjunto de especies pero utilizando un número menor de señales para crear la base de datos de *índices*.

Luego de analizar los resultados obtenidos, se puede concluir que la metodología es eficiente y robusta, ya que detecta correctamente a la especie aún con un factor considerable de ruido.

Como trabajos futuros, la metodología queda abierta al estudio de los resultados que surjan ante la incorporación y/o redefinición de los *índices*, determinar si existe alguna influencia de un *índice* sobre otros, comprobar si determinados *índices* son más importantes que otros y estudiar si la misma puede distinguir a las aves que imitan el canto de otras. También podría aplicarse con fines prácticos a dispositivos que analicen, por ejemplo, la presencia de aves que puedan causar accidentes aéreos en los aeropuertos; discriminar si un ave determinada imita el canto de otra; determinar la zona o región en la que habita una especie; entre otras aplicaciones.

REFERENCIAS

- [1] C. K. Catchpole y P. J. B. Slater, *Bird Song: Biological Themes and Variations*. U.K. Cambridge Univ. Press, 1995.
- [2] P. Somervuo y A. Härmä, "Bird song recognition based on syllable pair histograms," *IEEE ICASSP 2004, Canada*, Mayo 2004.
- [3] A. Franzen y I. Gu, "Classification of bird species by using key song searching: A comparative study," *IEEE*, 2003.
- [4] P. Somervuo, A. Härmä, y S. Fagerlund, "Parametric representations of bird sounds for automatic species recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, Noviembre 2006.
- [5] A. Härmä, "Automatic identification of bird species based on sinusoidal modeling of syllables," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003.
- [6] L. Genzelis, E. R. Fredini, y C. Yones, "Identificación automática de especies de aves a través de su canto," *Facultad de Ingeniería y Ciencias Hídricas - UNL*, 2011.
- [7] A. Härmä y P. Somervuo, "Classification of the harmonic structure in bird vocalization," *IEEE ICASSP 2004, Canada*, Mayo 2004.
- [8] J. D. Sitt, "Biomimética vocal," *Facultad de Ciencias Exactas y Naturales - UBA*, pp. 35–36, 2009.