

REPRESENTACIÓN ÓPTIMA DEL ESPECTRO DE LLAMADAS DE ANUROS EN SISTEMAS DE CONTROL AMBIENTAL UTILIZANDO REDES DE SENSORES INALÁMBRICOS

Gómez Bellido, Jesús*; Luque Sendra, Amalia; Carrasco Muñoz, Alejandro

Grupo de investigación TEP022 y TIC150.

Departamentos de Ingeniería del Diseño y Tecnología Electrónica. Universidad de
Sevilla.

*E-mail: aluque2@us.es

RESUMEN

El análisis y la clasificación de los sonidos producidos por determinadas especies animales, en particular los anuros, han revelado que son un potente indicador de la existencia del cambio climático. Es por lo que los sistemas de monitorización ambiental que utilizan redes de sensores inalámbricos son de gran interés para obtener indicadores del calentamiento global. El presente documento se centra en este proceso de extracción de características explorando tres alternativas: el MPEG-7 estandarizado, los Filter Bank Energy (Energía de Bancos de Filtros, FBE) y los Mel Frequency Cepstral Coefficients (Coeficientes Cepstrales de Frecuencia de Mel, MFCC). Además, se consideran diferentes valores para cada opción en la extracción de las características del espectro.

PALABRAS CLAVE

Sensores, Sonidos, Ambiental, Control, Clasificación.

ABSTRACT

The analysis and classification of the sounds produced by certain animal species, in particular anurans, have revealed that they are a powerful indicator of the existence of climate change. Environmental monitoring systems using wireless sensor networks are of great interest for global warming indicators. This paper focuses on this feature extraction process by exploring three alternatives: the standardised MPEG-7, Filter Bank Energy (FBE) and Mel Frequency Cepstral Coefficients (MFCC). In addition, different values are considered for each option in the extraction of spectrum characteristics.

KEYWORDS

Sensors, Sounds, Environment, Control, Classification.

INTRODUCCIÓN Y OBJETIVOS

Según *Global Shapers Community*, uno de los problemas más serios que afectan al mundo hoy en día es el cambio climático. Una de las consecuencias conocidas del cambio climático es su impacto en el desarrollo de las funciones fisiológicas básicas de varias especies [1], [2], como el sonido producido en la llamada de apareamiento, que juega un papel fundamental en la selección y reproducción sexual de numerosas especies ectotérmicas [3], [4]. Sonidos, por lo tanto, críticos para la adaptación de las especies al medio ambiente.

La producción de sonido en animales ectotérmicos está fuertemente influenciada por la temperatura ambiente [5], [6], que también puede afectar a varias características del sistema de comunicación acústica. Como resultado, la temperatura puede afectar significativamente a los patrones de llamada, influyendo en la actividad reproductiva del anuro.

El análisis y la clasificación de los sonidos producidos por ciertas especies de anuros han revelado que son un fuerte indicador de las fluctuaciones de temperatura y, por lo tanto, de la existencia del cambio climático [7].

Sin embargo, estos estudios tienen que estar respaldados por un gran número de grabaciones de audio, que normalmente se recogen sobre el terreno y se analizan de forma individualizada en un momento posterior. Se estima que, en promedio, un experto necesita 2 minutos de escucha para identificar una especie en 1 minuto de audio [8], por lo que resulta poco práctico analizar manualmente los grandes volúmenes de datos acústicos proporcionados por las modernas redes de sensores. Por este motivo, es imprescindible desarrollar sistemas inteligentes que simplifiquen, automaticen y agilicen la tarea de análisis y etiquetado de las grabaciones sonoras.

TRABAJOS PREVIOS

Desde hace varios años, en colaboración con el Parque Nacional de Doñana, se está trabajando en el problema de la clasificación de los anuros como indicadores del calentamiento global, desplegando una Red de Sensores a través del parque para diversos fines.

En una primera contribución [9], se demostró que es posible realizar una clasificación automática de los sonidos de anuros en su hábitat natural (ambiente “ruidoso”). En [10] se amplió el estudio a más algoritmos de clasificación, para superar algunas dificultades encontradas en [9], continuando con la caracterización del sonido basado

en MPEG-7 y tratando cada *frame* como independiente. En [11] se propusieron seis métodos de clasificación, con la intención de explotar la información contenida en el orden de los *frames*. En [12], se estudió el tiempo necesario para calcular cada paso del proceso de clasificación: extracción de características; clasificador de la formación; y clasificación no secuencial y secuencial.

OBJETIVOS DE LA INVESTIGACIÓN

Basándonos en estos antecedentes, el objetivo principal de este trabajo es explorar el proceso de extracción de características, analizando la mejor manera de representar la información contenida en un *frame* de sonido. Además de las características MPEG-7 utilizadas en trabajos anteriores, se considerarán otras formas de representar cada espectro de trama, desde Filter Bank Energy (FBE) hasta los Coeficientes Cepstrales de Frecuencia de Mel (MFCC) comúnmente utilizados.

En los experimentos se ha empleado un amplio conjunto de datos, se cuenta con grabaciones del Museo Nacional de Ciencias Naturales, correspondiente a dos especies: sapo corredor (*epidalea calamita*) y sapo partero común (*alytes obstetricans*), con un total de 868 grabaciones que contienen 4 clases de sonidos. Dentro de este conjunto, se han seleccionado 13 grabaciones que para ser utilizadas como conjunto de entrenamiento. Estas grabaciones además de patrones de los cantos también contienen secciones de silencio y/o ruido. Para la realización de este estudio, se han utilizado 10 clasificadores y siete métricas de rendimiento diferentes.

RESULTADOS Y DISCUSIÓN

Extracción de características MPEG-7

Un primer enfoque para reducir significativamente el número de valores que representan un espectro es utilizar algún tipo de firma de éste. La idea no es almacenar la amplitud espectral en cada frecuencia, sino determinar ciertas características generales de la forma del espectro. Con este propósito, para la extracción de características se ha utilizado la norma ISO MPEG-7 [13]. De esta recomendación se ha seleccionado los parámetros más significativos para la clasificación, obteniendo del análisis 18 parámetros de 3 procesos diferentes.

Una de las principales ventajas del uso de las características MPEG7 es su estandarización, por lo que no es necesario ajustar ni seleccionar parámetros opcionales. Por lo tanto, utilizando los valores recomendados reflejados en la columna MPEG-7 de la Tabla 1, el rendimiento de la clasificación puede resumirse en la Figura

1. Como se puede observar, los mejores resultados lo obtienen el clasificador de Distancia Mínima con un resultado del 85%.

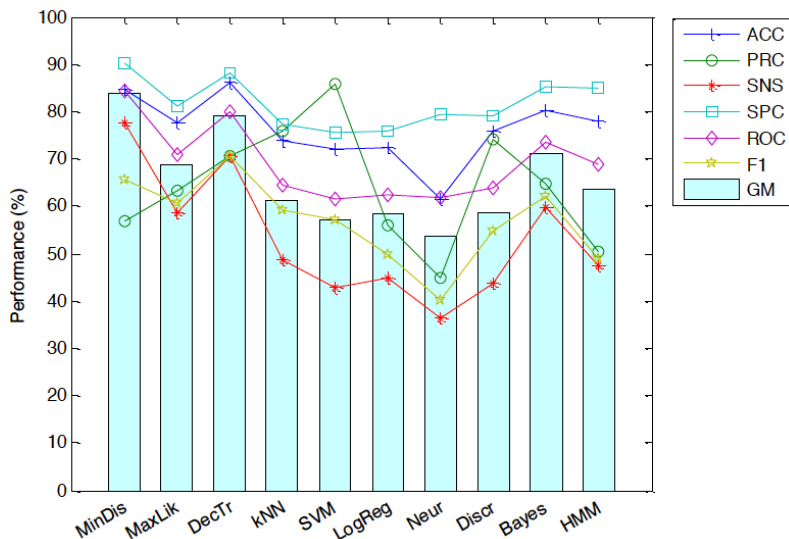


Figura 1. Representación de la clasificación usando características MPEG-7.

Banco de Filtros de Energía (FBE)

El segundo enfoque representa los sonidos usando el *Filter Bank Energy*, FBE. Como no existe una práctica estándar o común en la selección de opciones para FBE, con el objetivo de facilitar la comparación, hemos utilizado las mismas opciones que en MPEG-7. El desempeño en la clasificación obtenida usando esta estrategia puede resumirse en la Figura 2. Observamos que los mejores resultados se obtienen utilizando el clasificador de Máxima Probabilidad con un resultado del 92,69%.

Mel Frequency Cepstral Coefficients (MFCC)

Las características MFCC no están estandarizadas por lo que de forma inicial se han usado los valores por defecto de la implementación MFCC-HTK (Hidden Markov Model Toolkit), valores reflejados en la columna MFCC de la Tabla 1.

Los resultados de los distintos métodos de clasificación estudiados con este conjunto de características se resumen en la Figura 3, concluyendo el mejor comportamiento lo obtiene el clasificador Bayes con un resultado del 94,85%.

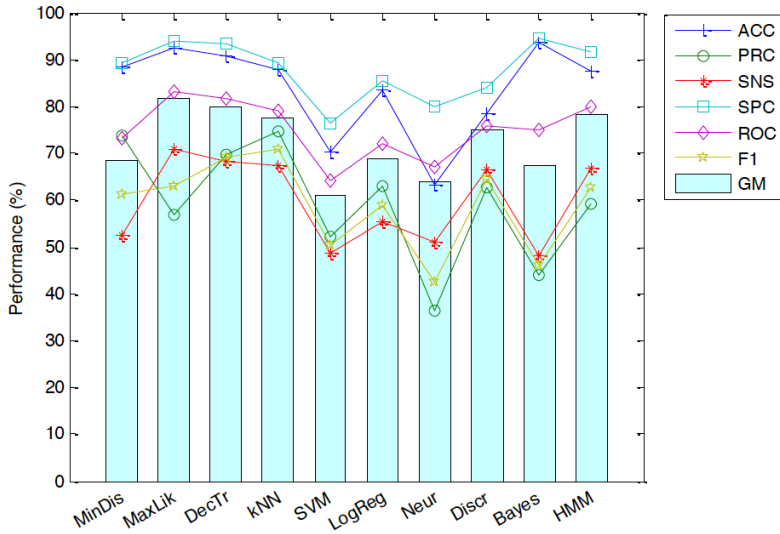


Figura 2. Representación de la clasificación usando características FBE.

MFCC Optimizado

El proceso de extracción de estas características del MFCC deja abierta muchas opciones, se va a estudiar si un conjunto diferente de valores para las opciones de extracción podría lograr mejores resultados de clasificación.

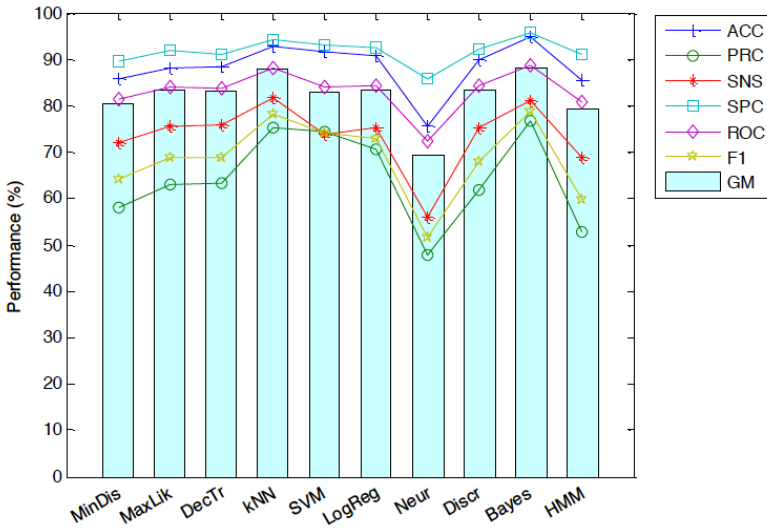


Figura 3. Representación de la clasificación usando características MFCC.

Realizar un estudio exhaustivo en la dimensión de los valores (R^{11}) se hace inviable en cuanto al tiempo de cálculo que conllevaría. Por esto, como alternativa a este planteamiento, se ha empleado un enfoque mas sencillo y rápido, pero aún así efectivo. Se considera un punto de partida en el espacio de opciones y cada vez que éste se mueva lo hará en una sola dimensión. La primera búsqueda tendrá en cuenta el punto de partida (1) y los valores $O_1 - 1$ restantes de la primera propiedad. La búsqueda en la segunda dimensión requerirá el cálculo de los valores de $O_2 - 1$.

En la tabla columna MFCC-Optimizado de la Tabla 1 se refleja un resumen de los parámetros óptimos finalmente seleccionados para la clasificación. En la Figura 4 se muestra el rendimiento alcanzado al hacer uso de esta estrategia, obteniendo los mejores resultados el clasificador Bayes, con un 96,37%.

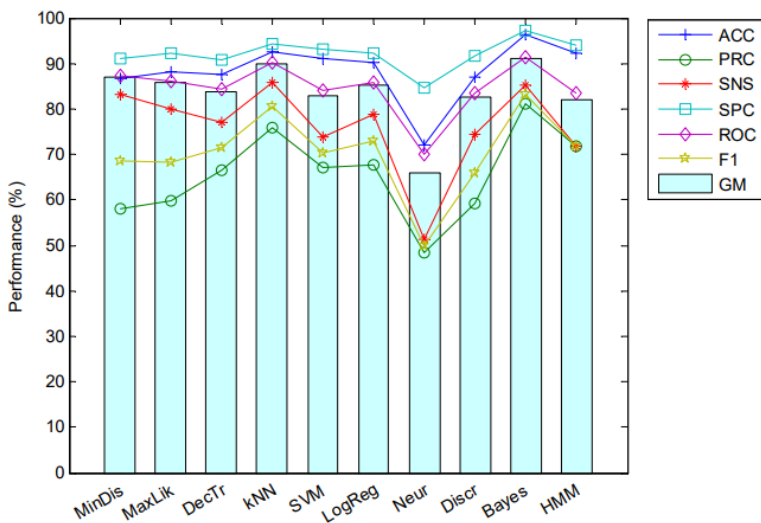


Figura 4. Representación de la clasificación usando características MFCC Optimizadas.

CONCLUSIONES

En conclusión, para lograr un buen compromiso entre la calidad de la clasificación y las consideraciones de implementación del ruido blanco gaussiano, se usará el método MFCC, manteniendo el valor por defecto de 13 características y haciendo uso de las opciones ajustadas de forma óptima. Si el compromiso que se requiere es una importante reducción del número de características, el MFCC supera claramente al FBE.

Tabla 1. Resumen valores para la extracción de características

Dominio	Función	Propiedad	MPEG-7	MFCC	MFCC-Opt
Tiempo	Preénfasis	α	-	0,97	-
	Framing	Ventana	Hamming	Hamming	Hamming
		T_w	30ms	25ms	20ms
		T_s	10ms	10ms	10ms
Frecuencia	FBE	L_f	64kHz	300Hz	1kHz
		H_f	16kHz	3700Hz	5kHz
		M	-	20	20
		Escala	-	Mel	Mel
Quefrecia	Cepstrum	Transformación	-	DCT	DCT
		C	-	13	-
	Liftering	L	-	22	20

AGRADECIMIENTOS

A los grupos de investigación TEP022 y TIC150 de la Universidad de Sevilla por los conocimientos adquiridos y los trabajos previos realizados.

A la Escuela Politécnica Superior de Sevilla por la organización de la Jornadas Doctorales que ha permitido la publicación del presente trabajo.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Pörtner, H.O. y Knust, R. (2007). Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science*, 315, pp. 95–97.
- [2] Huey, R.B., Deutsch, C.A., Tewksbury, J.J., Vitt, L.J., Hertz, P.E., Pérez, H.J.Á., y Garland, T. (2009). Why tropical forest lizards are vulnerable to climate warming. *Proceedings of the Royal Society of London. Series B, Biological sciences*
- [3] Bradbury, J.W. y Vehrencamp, S.L. (2011). *Principles of Animal Communication*, 2nd ed. Sinauer Associates: Sunderland, MA, USA.
- [4] Fay, R.R. y Popper, A.N. (Eds.) (2012). *Comparative Hearing: Fish and Amphibians*, Springer Science & Business Media: New York, USA, Volume 11.
- [5] Bellis, E.D. (1957). The effects of temperature on salientian breeding calls. *Copeia*, pp. 85–89.

- [6] Walker, T.J. (1962). Factors responsible for intraspecific variation in the calling songs of crickets. *Evolution*, 16, pp. 407–428.
- [7] Llusia, D., Márquez, R., Beltrán, J.F., Benitez, M., y Do Amaral, J.P. (2013). Calling behaviour under climate change: Geographical and seasonal variation of calling temperatures in ectotherms. *Global Change Biology*, 19, pp. 2655–2674.
- [8] Wimmer, J., Towsey, M., Roe, P., y Williamson, I. (2013). Sampling environmental acoustic recordings to determine bird species richness. *Ecological Applications*, 23, pp. 1419–1428.
- [9] Luque, J., Larios, D.F., Personal, E., Barbancho, J., y León, C. (2016). Evaluation of MPEG-7-Based Audio Descriptors for Animal Voice Recognition over Wireless Acoustic Sensor Networks. *Sensors*, 16, p. 717.
- [10] Luque, A., Romero-Lemos, J., Carrasco, A., y Barbancho, J. (2018). Non-sequential automatic classification of anuran sounds for the estimation of climate-change indicators. *Expert Systems with Applications*, 95, pp. 248–260.
- [11] Romero, J., Luque, A., y Carrasco, A. (2017). Animal Sound Classification using Sequential Classifiers. En *BIOSTEC 2017: 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, ScitePress Digital Library: Setubal, Portugal, pp. 242–274.
- [12] Luque, A., Gómez-Bellido, J., Carrasco, A., Personal, E., y Leon, C. (2017). *Evaluation of the Processing Times in Anuran Sound Classification*. *Wireless Communications and Mobile Computing*. 2017, 8079846.
- [13] ISO. ISO/IEC 15938-4:2001 (MPEG-7: Multimedia Content Description Interface), Part 4: Audio; ISO/IEC JTC, 1; ISO: Geneva, Switzerland, 2001.