

Detección y caracterización automáticas basadas en tratamiento digital de imágenes y redes neuronales de señales tonales emitidas por cetáceos

ANTONIO SÁNCHEZ¹, F. JAVIER RODRIGO, PEDRO J. GARCÍA²,
ANDRÉS BUENO³, PATRICIO MUÑOZ, JUAN MORALES²,
JORGE LARREY², MARÍA CONSUELO BASTIDA², Y JOSÉ LUIS SANCHO²

1. Área Técnica de Estudios Avanzados y Tratamiento Digital de la Señal.
Sociedad Anónima de Electrónica Submarina (SAES).
2. Departamento de Tecnologías de la Información y las Comunicaciones.
Universidad Politécnica de Cartagena.
3. Departamento de Informática de Sistemas.
Universidad Católica San Antonio.

a.sanchez@electronica-submarina.com;
f.rodrigo@electronica-subamarina.com;
pedroj.garcia@upct.es; abueno@pdi.ucam.edu;
p.muñoz@electronica-submarina.com; juan.morales@upct.es;
jorge.larrey@upct.es; mc.bastida@upct.es; jl.sancho@upct.es

Proyecto/Grupo de investigación: Grupo de Teoría y Tratamiento de Señal.

Líneas de investigación: *Procesado Digital de Imágenes Médicas; Reconocimiento Estadístico de Patrones; Procesado de Acústica Submarina.*

Resumen

Se presenta un nuevo método para la detección y caracterización de señales tonales emitidas por cetáceos. Las señales son automáticamente detectadas a partir del espectrograma, aisladas mediante un proceso de segmentación basado en regiones, extraídas tras un proceso de resolución de cruces y caracterizadas automáticamente mediante un número fijo de coeficientes derivados de la aplicación de redes neuronales de funciones de base radial. Estos coeficientes pueden ser empleados para una posterior

clasificación de las señales basada en sus características. Las prestaciones del método han sido extensivamente evaluadas a partir de señales simuladas y registros de grabaciones en la mar.

1. Introducción

En la actualidad existe un interés creciente por la protección de la fauna marina en general, y por las especies de mamíferos marinos en particular, ante los efectos adversos para su bienestar e incluso para su supervivencia derivados de diferentes actividades humanas en el entorno marino [1]. El paso previo a su protección es la determinación de la presencia de los animales marinos en una determinada área. En la actualidad se emplean diferentes técnicas para monitorizar la presencia de mamíferos marinos. La aproximación clásica consiste en su observación visual en la superficie marina, aunque este método presenta restricciones significativas derivadas de la necesidad de la presencia de los animales en la superficie y cerca de la plataforma de observación y de la existencia de condiciones de luminosidad suficientes. La monitorización acústica no se ve afectada por estas restricciones y permite alcances de detección significativos. Está restringida a la detección de animales vocalizantes, pero esta es una característica inherente a los mamíferos marinos que utilizan la acústica para la comunicación, el establecimiento de relaciones sociales, la navegación y la ecolocalización [1]. Otras técnicas basadas en radar, luz infrarroja e imagen térmica están siendo investigadas en la actualidad para mejorar la detección de mamíferos marinos en la superficie [2].

Los cetáceos comprenden las ballenas, los delfines y las marsopas y constituyen el orden de mamíferos marinos con el mayor número de especies, totalizando más de 80 [3], y con individuos más vocalizantes. Habitan todos los mares y océanos del planeta y su repertorio vocal está constituido por señales pulsadas de corta duración y banda ancha y por señales tonales de banda estrecha, moduladas en frecuencia y con un amplio rango de duraciones [4].

El presente artículo se centra en las señales tonales emitidas por los cetáceos y presenta un método automático para su detección y caracterización. Los métodos automáticos se configuran como un medio efectivo para llevar a cabo tareas de monitorización durante prolongados intervalos temporales y reducen de forma significativa las tareas de los operadores humanos. El artículo está estructurado en cuatro secciones. La primera de ellas es la presente introducción. La Sección 2 presenta una descripción del método empleado. La Sección 3 describe los resultados obtenidos de la aplicación del método a señales simuladas y registros de grabaciones en la mar. En la cuarta y última Sección se presentan las conclusiones del estudio.

2. Descripción del método

El método propuesto cubre las dos primeras etapas de un sistema de reconocimiento de patrones: detección (pre-procesado) y extracción de características o caracterización. Sus principales etapas se presentan en la siguiente figura:

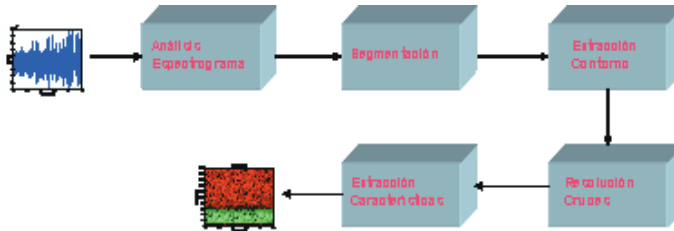


Figura 1: Diagrama de bloques del método propuesto para la detección y caracterización automática de señales de cetáceos.

Se describen a continuación las tareas llevadas a cabo en cada una de estas etapas:

- **Análisis Espectrograma.** La etapa inicial lleva a cabo el cálculo del espectrograma de la señal. Una FFT de tamaño 512, ponderada mediante una ventana Hanning y con un 50% de solapamiento, proporciona una resolución adecuada en tiempo (5.33 ms) y frecuencia (375 Hz) para permitir la detección y caracterización de las señales tonales inmersas en un entorno de ruido enmascarante. A continuación los datos son normalizados en relación al ruido utilizando un estimador estadístico sobre ventanas de observación en frecuencia. El siguiente paso es convertir la matriz resultante en una imagen de intensidad con valores en el rango entre 0 y 255. El resto de las etapas de la cadena de proceso operan sobre imágenes.
- **Segmentación.** Se aplica un proceso de segmentación basado en regiones. Se trata de un método que extrae regiones homogéneas a partir de un conjunto de píxeles iniciales que reciben la denominación genérica de ‘semillas’ [5]. Se selecciona un conjunto de semillas para la imagen completa, basado en la condición de exceder un determinado umbral. A partir de estos píxeles iniciales, las regiones correspondientes a señales tonales candidatas crecen incorporando los píxeles vecinos que presentan un valor superior al nivel de referencia de ruido. Este nivel se establece a partir de los datos de la imagen completa.
- **Extracción de Contorno.** Una vez finalizado el proceso de segmentación se obtiene un conjunto de señales tonales candidatas. Normalmente estos objetos extraídos se extienden a lo largo de más de un pixel en frecuencia en cada trama temporal. El objetivo es obtener pares tiempo-frecuencia

que permitan disponer de una curva bidimensional para cada señal tonal candidata. El proceso aplicado para conseguir este resultado es seleccionar el pixel en frecuencia de mayor valor para cada trama temporal. Las curvas obtenidas reciben la denominación de contorno de frecuencia.

- **Resolución de cruces.** Las señales candidatas pueden incluir los contornos de frecuencia de señales tonales individuales o de señales solapadas. En este último caso se implementa un algoritmo de resolución de cruces que permite deshacer el solapamiento entre señales. Este algoritmo extrae las señales a partir del seguimiento de su evolución desde su pixel de inicio.
- **Extracción de Características.** Es la última etapa y en ella se implementa inicialmente un proceso de selección entre las señales candidatas, de modo que sólo son consideradas como señales tonales validas aquellas que cumplen los requisitos establecidos de variación en tiempo y frecuencia. Estas señales son finalmente caracterizadas por un número fijo de coeficientes correspondientes a los pesos de la capa de salida de la red neuronal RBF utilizada para aproximar cada curva. Un total de 16 coeficientes han mostrado ser suficiente para capturar con precisión la información tiempo-frecuencia contenida en las señales. Estos coeficientes pueden ser empleados como entradas a una nueva red neuronal que puede llevar a cabo la última etapa del sistema de reconocimiento de patrones: la clasificación de las señales tonales detectadas y extraídas sobre la base de su vector de características. Un esquema de una red neuronal RBF se presenta en la Figura 2.

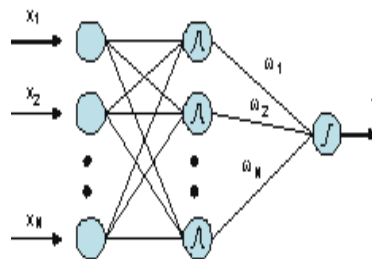


Figura 2: Estructura de tres capas de una red neuronal RBF, donde x_1, \dots, x_N denotan los vectores de datos de entrada a la red neuronal y w_1, \dots, w_M corresponden a los pesos aplicados a las salidas de los nodos de la capa oculta. Por último, el vector y es la salida de la red neuronal.

3. Aplicación del método y resultados

Las prestaciones del método propuesto se han evaluado a partir de señales simuladas y grabaciones de señales de cetáceos en la mar. Los resultados se

presentan a continuación.

3.1. Aplicación a señales simuladas

Para evaluar la efectividad del método se han generado ficheros que contienen señales simuladas que imitan la evolución en tiempo y frecuencia de las señales tonales emitidas por los cetáceos y se han insertado en un entorno de ruido *Gaussiano*. Uno de estos ficheros, cuyo espectrograma se presenta en la Figura 3, tiene una duración de cuatro segundos e incorpora cinco señales correspondientes a barridos lineales y cuadráticos en frecuencia, de las cuales cuatro se cruzan en dos grupos de dos, y dos señales espurias.

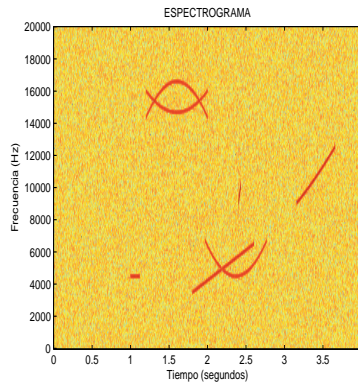


Figura 3: Espectrograma del fichero de señales simuladas de cuatro segundos de duración

Tras el cálculo del espectrograma las siguientes etapas de la cadena de procesamiento que implementa el método corresponden a la normalización, para obtener una referencia común de ruido, y la segmentación.

La salida de esta última etapa se presenta en la Figura 4(a)¹, en la que se puede apreciar que los objetos que se cruzan son segmentados como un único objeto. Tras la segmentación se llevan a cabo los procesos de extracción de contorno y resolución de cruces. Finalmente se seleccionan los objetos validos (señales tonales) y se extraen sus características mediante redes neuronales RBF. La salida del proceso de selección de objetos se presenta en la Figura 4(b), donde se presentan las imágenes individuales obtenidas correspondientes a cada objeto en una única imagen por simplicidad. Las Figuras 4(c) a 4(g) contienen la salida de la etapa de extracción de características donde cada señal individual queda caracterizada por un conjunto de 16 coeficientes.

¹En lo que sigue, las Figuras se presentarán siguiendo el orden alfabético de izquierda a derecha y de arriba hacia abajo.

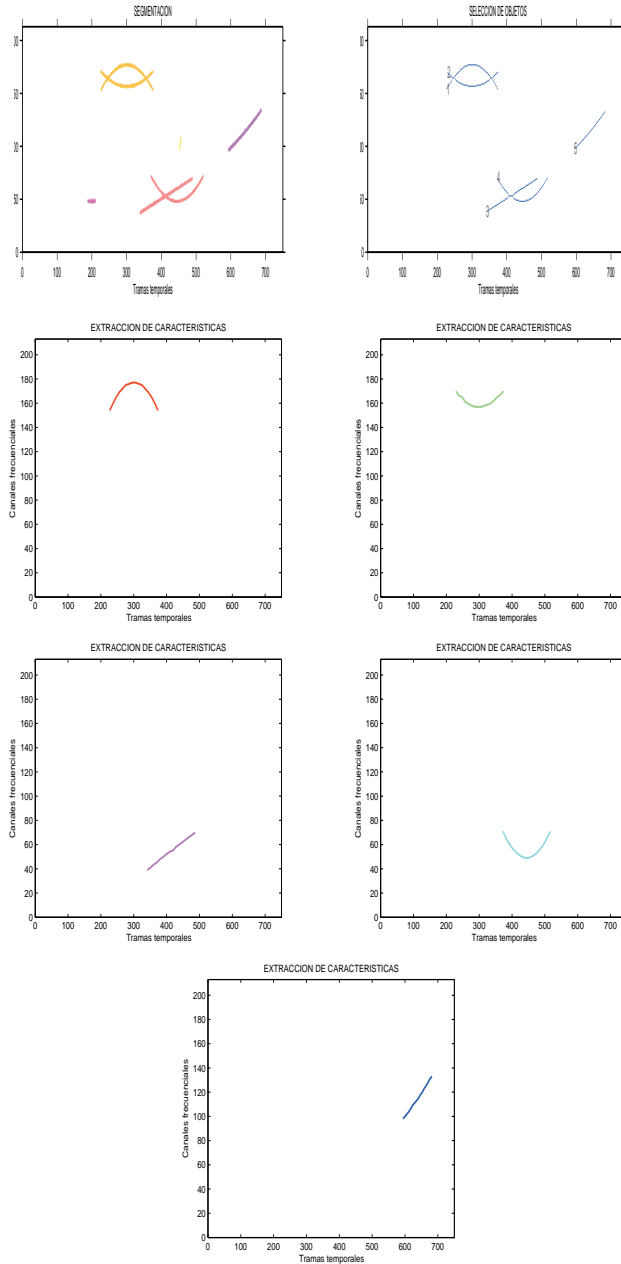


Figura 4: Salidas de diferentes etapas de la cadena de procesamiento para la señal simulada analizada. (a) Salida de la etapa de segmentación, (b) Salida de la etapa de selección de objetos donde los objetos se han presentado en una única imagen en lugar de imágenes individuales por simplicidad y (c) salida de la etapa de extracción de características.

3.2. Aplicación a señales grabadas en la mar

Una vez analizadas las prestaciones de detección sobre señales simuladas el siguiente paso consiste en su evaluación ante señales grabadas en entornos marinos. Para ello, el método se ha aplicado a un amplio conjunto de señales tonales de delfines (silbidos) obtenidas en un extenso rango de ambientes marinos que varían desde señales aisladas con alta relación señal-a-ruido (RSR) a grupos de señales solapadas con baja RSR. En los diferentes casos analizados el método propuesto ha demostrado su efectividad para la detección y caracterización automática de las señales tonales emitidas por los delfines. Como un ejemplo se muestran a continuación los resultados del procesamiento de dos registros de segmentos de grabaciones. En la Figura 5 se muestra el espectrograma de ambos registros.

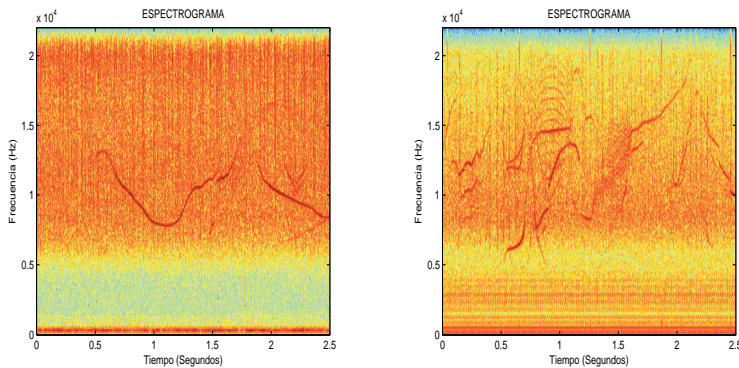


Figura 5: Espectrograma de los dos registros de segmentos de grabaciones analizados: (a) registro de complejidad baja y (b) registro de complejidad alta.

En las Figuras 6(a) y 6(b) se presentan las salidas de las etapas de segmentación. Las Figuras 6(c) y (d) corresponden a la extracción de características para los dos registros analizados.

4. Conclusiones

Se ha presentado un nuevo método para la detección y caracterización automática de señales tonales emitidas por cetáceos. Las prestaciones del método han sido evaluadas con señales simuladas y registros de señales tonales emitidas por delfines. Combina técnicas de tratamiento digital de imágenes y procesamiento de redes neuronales y proporciona resultados prometedores en la tarea de detección y clasificación de mamíferos marinos sobre la base de las señales acústicas emitidas. El método sigue el principio de modularidad en su diseño para facilitar el proceso de futuras modificaciones.

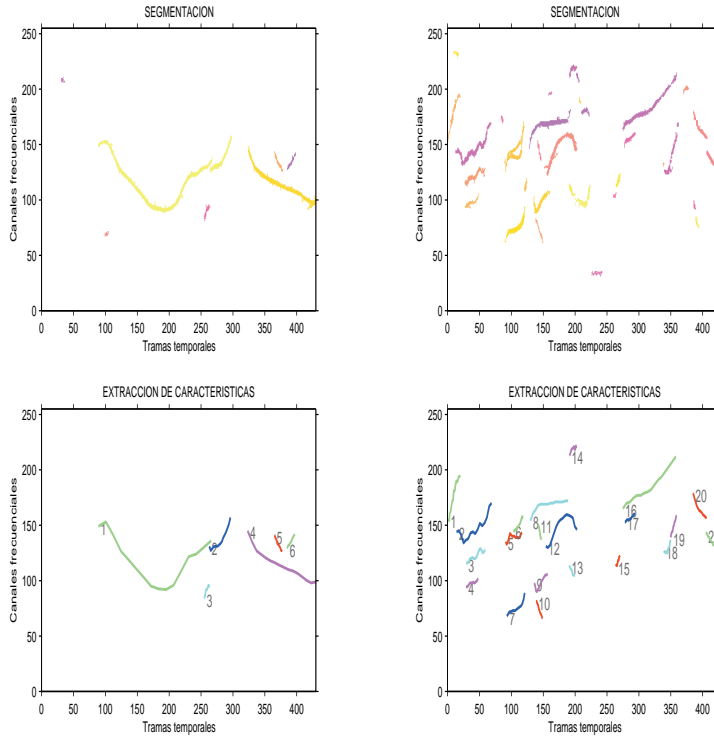


Figura 6: La Figura 6 (c) contiene seis señales tonales independientes y puede ser catalogado como de baja complejidad. La Figura 6 (d) incorpora cruces entre señales y alto nivel de ruido interferente por lo que puede ser catalogado como de complejidad alta.

Referencias

- [1] Richardson WJ, Green CR, Malme CI & Thomson DH. Marine mammals and Noise. Academic Press; 1995.
- [2] Pavan G. Acoustic risk mitigation in the Mediterranean Sea. Current situation and recommendations. En: Actas de la Conferencia Underwater Defence Technology (UDT), Nápoles (Italia); Junio 2007.
- [3] Reeves R, Stewart B, Clapham PJ & Powell JA. Guide to marine mammals of the world. Chanticleer Press, Inc.; 2002
- [4] Tyack P.L. & Clark C.W. Communication and acoustic behaviour of dolphin and whales. En W.W.L. Au, A.N. Popper & R.R. Fay eds. Hearing by Whales and Dolphins. Springer Verlag; 2000.
- [5] González RC & Woods RE. Digital Image Processing. Prentice Hall; 2001.