



Innovación Social Cuaderno Red de Cátedras Telefónica



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Entendiendo el lenguaje de los cetáceos

Cátedra Telefónica Universitat Politècnica de València
“Tecnologías para la innovación social y ambiental”

Algoritmo de detección y/o localización de vocalizaciones de belugas

Ramón Miralles Ricós
Diciembre 2010

Biografía



Ramón Miralles Ricós

Nació en Valencia (España) en 1971. Recibió el título de Ingeniero de Telecomunicación y el de Doctor Ingeniero en Telecomunicación por la Universitatd Politècnica de València en 1995 y 2000 respectivamente. En 1996 se incorporó como profesor en el Departamento de Comunicaciones en la Escuela Politécnica Superior de Gandía. Desde el año 2000 hasta la actualidad trabaja como profesor titular en la Escuela Técnica Superior de Ingenieros de Telecomunicación (Valencia). Es responsable del desarrollo de algoritmos y sistemas de inspección ultrasónica para el control de calidad en la industria. Es miembro del Instituto de Telecomunicación y Aplicaciones Multimedia (ITeam) en el cual ejerce también su labor como secretario del mismo.

Sus líneas de investigación prioritarias son: los estadísticos de orden superior, el procesado no lineal de las señales y las aplicaciones de procesado de señal en los sistemas de ultrasonidos. Ha publicado mas de 60 artículos en estas áreas (incluyendo artículos en revistas y conferencias nacionales e internacionales).

Índice

1. Introducción
2. Producción de sonidos en los cetáceos
3. Algoritmo automático de detección y clasificación
4. Bibliografía

1. Introducción

Poco se sabe de los mecanismos de producción de sonido por parte de los cetáceos. Si bien se sabe en algunos casos que órganos están implicados, la funcionalidad de cada uno de los órganos pasa por meras conjeturas en la mayoría de las especies. En este proyecto se pretende implementar un sistema de detección de vocalizaciones para animales en cautividad que funcione en tiempo real con dos objetivos. Por un lado este sistema podrá ser utilizado por los cuidadores del parque para controlar el grado de bienestar (altamente correlado con la tasa de vocalizaciones) y por otro el estudio y clasificación de estos sonidos podría aportar un nuevo enfoque que permitiera conocer mejor los mecanismos de producción de sonidos por parte de los cetáceos.

2. Producción de sonidos en los cetáceos

Los seres humanos producen sonido al expeler aire a través de la laringe. Las cuerdas vocales dentro de la laringe se abren o cierran según sea necesario para separar la corriente de aire en paquetes discretos. Estos paquetes son moldeados por la garganta, la lengua y los labios para formar el sonido deseado. La producción de sonido en los cetáceos difiere marcadamente de este mecanismo. El mecanismo preciso varía en los dos grupos de cetáceos: los odontocetos (belugas, cachalotes y delfines) y los misticetos (ballenas barbadas- incluyendo las más grandes, como la ballena azul). En el caso de las belugas los sonidos se generan al pasar aire por unas estructuras tridimensionales en su

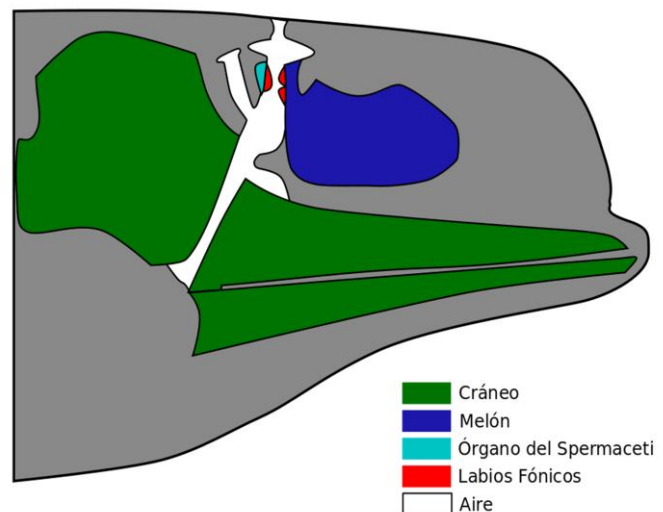


Figura 1. Órganos involucrados en la producción de sonidos de los odontocetos

cabeza (ver Figura 1) que podríamos asimilar con las fosas nasales humanas y que se conocen con la denominación inglesa de “phonic lips”. Las vibraciones generadas en los “phonic lips” son conducidas a la cabeza del animal. Aquí se forma el sonido es emitido en la dirección apropiada.

Al igual que sucede en los mamíferos terrestres existe una alta correlación, para el caso de los mamíferos acuáticos, entre la cantidad de sonidos emitidos y el grado de bienestar de estos. Ha sido ampliamente demostrado que en periodos de estrés como el producido por la manipulación veterinaria, la separación de crías o el transporte y cambio de ubicación, los animales modifican la tasa de emisión de sonidos. Investigadores del Oceanogràfic han verificado este comportamiento en las dos belugas (*Delphinapterus leucas*) que se encuentran en sus instalaciones [1, 2]. En esta misma línea, a lo largo de un proyecto de varios años de duración que comenzó en el año 2003, los investigadores del Oceanogràfic han sido capaces de aislar 32 tipos de vocalizaciones que configuran el repertorio acústico de las belugas. Algunas de estas vocalizaciones pudieron ser asociadas a categorías de comportamiento [3] (Agresividad, interacción con los cuidadores, interacción con los visitantes, etc).

Toda esta información es metódicamente recogida mediante unos hidrófonos y digitalmente grabada para su posterior análisis y escucha. Las características de estos sonidos hacen que el software habitualmente empleado en el análisis del habla humana no pueda ser usado directamente para la detección ni el reconocimiento automático de las vocalizaciones de las belugas. Es por ello que el análisis de toda esta información requiere del trabajo de un técnico con un cierto grado de entrenamiento y formación capaz de contabilizar los diferentes tipos de vocalizaciones grabadas.

Investigadores del iTEAM- Instituto de Telecomunicación y Aplicaciones Multimedia con su amplia experiencia en las técnicas de detección y clasificación mediante procesado digital de la señal están trabajando en la creación de una serie de algoritmos que en parte automatizaran las labores de los técnicos y permitirán un análisis continuo y en tiempo real de las vocalizaciones de los cetáceos en cautividad. Adicionalmente esta herramienta permitirá establecer parámetros estadísticos que podrán servir de ayuda para los cuidadores y la mejor comprensión de estos mamíferos.

En la sección siguiente se detalla de forma sencilla como las técnicas de procesado de señal pueden ayudar en esta labor a los técnicos del Oceanogràfic.

3. Algoritmo automático de detección y clasificación

La detección de vocalizaciones de belugas en un entorno como el del Oceanográfico no es una tarea fácil. La cúpula donde se encuentran los tanques que contienen las belugas y que es necesaria para una correcta climatización produce una gran cantidad de ecos y reverberaciones. De este modo los hidrófonos captan sonidos de los visitantes, así como de otras especies que puedan estar bajo la misma sala (morsas, focas, etc.). También hay que añadir a todo este repertorio de sonidos los sonidos producidos por las cuerdas empleadas para amarrar las plataformas flotantes empleadas por los cuidadores y el personal del parque. Dada la complejidad del problema, el algoritmo de detección en el que los investigadores del iTEAM están trabajando es un algoritmo de dos fases claramente diferenciadas. En una primera fase, y tras acondicionar la señal se realiza una detección de actividad por energía. Sólo aquellos fragmentos de la grabación donde se han producido detecciones son pasados a la fase de clasificación (ver Figura 2). Este esquema facilita la implementación del sistema en tiempo real con equipos de escasa potencia computacional y por lo tanto abarata los costes del sistema final.

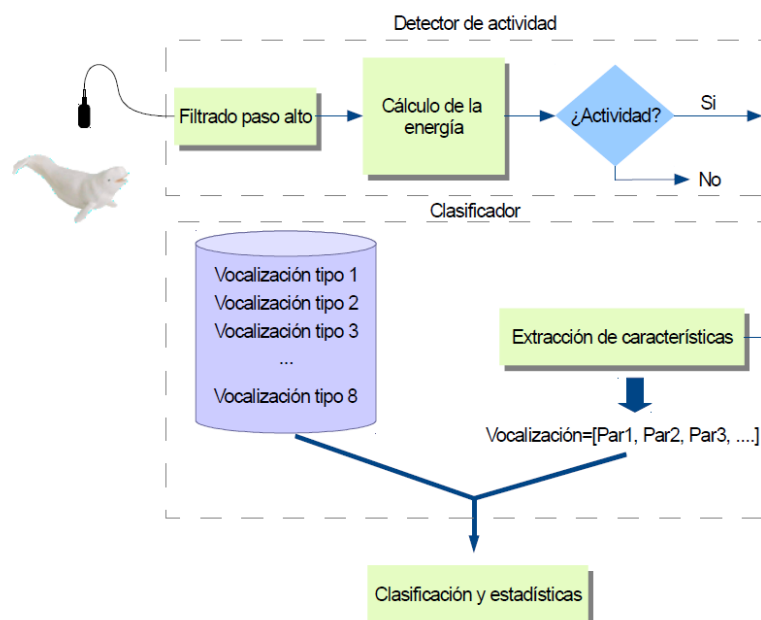


Figura 2. Algoritmo de detección/clasificación de vocalizaciones de belugas

La segunda etapa emplea el análisis discriminante cuadrático para realizar la clasificación y descartar aquellos sonidos que no provienen de vocalizaciones de las belugas.

Con el objeto de poder disponer de una base de datos lo suficientemente abundante de vocalizaciones (un algoritmo como el empleado necesita de más de una vocalización de cada tipo para poder ser entrenado apropiadamente) se redujo el número de vocalizaciones a 6, de la mezcla de las cuales se forman todos los sonidos que producen dichos cetáceos, vocalizaciones multitonales, trenes de pulsos, tonos, silbidos, palmadas con las mandíbulas y vocalizaciones de ecolocalización.

Adicionalmente se añadieron a estas vocalizaciones los sonidos procedentes del movimiento del agua y de las cuerdas anteriormente mencionadas dando el conjunto de 8 vocalizaciones o categorías en las que se clasificaron todos los sonidos producidos.

La observación detallada de los diagramas tiempo frecuencia (Figura 3, (a), (b) y (c)) y la manipulación digital de cada una de estos sonidos permitió identificar una serie de parámetros propios de cada vocalización que podían ser empleados para su clasificación (pitch, frecuencia fundamental, factor Q, etc.). La explicación de la buena capacidad para discriminar de algunos de estos parámetros podría encontrarse en el hecho que los sonidos que poseen características asociadas a la frecuencia de vibración (pitch) son creados a partir de vibraciones y por ello poseen una sonoridad mucho mayor que aquellos que son creados por otros órganos diferentes como por ejemplo los golpes de mandíbula, que producen sonidos sordos.

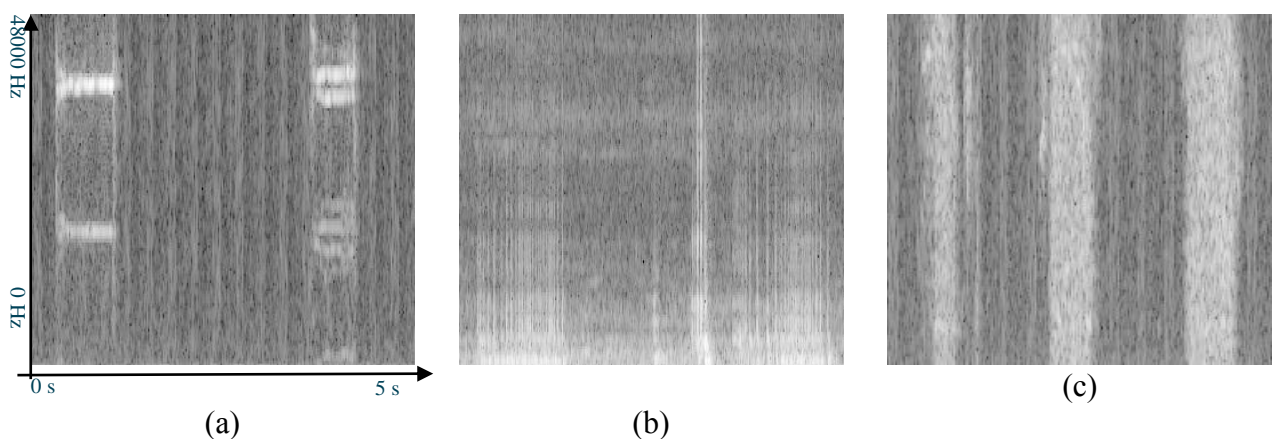


Figura 3. Diagramas tiempo-frecuencia de diferentes sonidos captados a la salida del detector de energía. (a) Vocalización multitonal, (b) Ruidos producidos por las cuerdas de amarre de las plataformas y (c) Trenes de pulsos.

A estos parámetros, con cierto sentido físico, se unieron otros parámetros típicamente empleados por el grupo en diversos algoritmos de clasificación (parámetros estadísticos, medida del grado de no linealidad, etc.). Con todo ello se conformo un vector de características, firma exclusiva de cada una de las vocalizaciones, que se empleo para realizar la clasificación. La tabla siguiente muestra algunos de estos parámetros.

Tabla 1. Detalle de algunos parámetros empleados en el vector de características para clasificación

	<i>Parámetro</i>	<i>Descripción</i>
1	Frecuencia fundamental	Frecuencia más baja del espectro de frecuencias tal que las frecuencias dominantes pueden expresarse como múltiplos de esta frecuencia fundamental.
2	Factor Q frecuencia fundamental	Relación entre la frecuencia fundamental y su ancho de banda
3	Pitch	Frecuencia percibida de los sonidos sonoros (producidos por la vibración de los labios fónicos).
4	Media de la sonoridad	Nivel de amplitud del pitch normalizado por la energía.
5	Autocovarianza de tercer orden de la señal temporal	Medida del grado de no linealidad de una señal.
6	Reversibilidad de la señal temporal	Otra medida del grado de no linealidad de una señal.

El estudio de los parámetros más significativos de los enumerados en la tabla 1 se realizó aplicando un test t-Student a cada uno de los parámetros y comparando las diferentes probabilidades.

En la actualidad el GTS una vez preparados los conjuntos de entrenamiento/ test y seleccionados los parámetros de los vectores de características, esta trabajando en el uso de clasificadores con análisis discriminante. Los primeros resultados son alentadores y permitirán una clasificación automática con una baja probabilidad de error. Igualmente algunos de los indicadores parecen aportar ciertos indicios de que la capacidad de

modulación y el grado de control en la producción de sonidos por parte de las belugas es muy alto.

4. Bibliografía

[1] Castellote M. Fossa F. and Esteban J.A. 2005 Comportamiento acústico de las belugas de L'Oceanogràfic. Proceedings del Congreso anual de la Sociedad Española de Sociedad Española para la conservación y estudio de los mamíferos (SECEM). Valencia 3-6 Diciembre de 2005.

[2] Castellote, M. Esteban, J.A. and Fossa F. 2007 The acoustic behaviour of L'Oceanogràfic beluga whales 1st International Workshop on Beluga Whale Research. Valencia Marzo 2007

[3] Castellote, M. , Fossa, F. and Esteban, J.A. 2006 Context-specific vocalizations in captive beluga whales ECS06 Annual Conference of the European Cetacean Society 2006