FEATURE-AIDED TRACKING FOR MARINE MAMMAL DETECTION AND CLASSIFICATION

Odile Gerard[1*], Craig Carthel[1], Stefano Coraluppi[1], and Peter Willett[2]
[1]NATO Undersea Research Centre, Viale S. Bartolomeo 400, 19126 La Spezia, Italy
[2]ECE Department, University of Connecticut, Storrs, CT 06269, USA
[*] Contact author. E-mail: odileg@free.fr.

ABSTRACT

This paper presents a method to detect and classify odontocete echolocation clicks as well as to estimate the number of animals that are vocalizing. A transient detector using the Page test [1-3] is used to extract the clicks: the click time, the click duration, the click amplitude and the spectral information of the clicks are extracted. A probability distribution over the species is assigned to each click, based on the spectral information of the click. The estimation of the number of animals is done using feature-aided multi-hypothesis tracking (MHT) algorithms. The association is based on the assumptions of slowly-varying click amplitude and intra-click timing [4-5]. This work has been done on the dataset provided by the organizers of the *3rd International Workshop on the Detection and Classification of Marine Mammals using Passive Acoustics*, Boston, July 2007. This dataset consists of training and test data; the training data includes vocalizations of three species: Blainville's beaked whale (*Mesoplodon densirostris*), Risso's dolphin (*Grampus griseus*) and short-finned pilot whale (*Globicephala macrorhynchus*).

SOMMAIRE

Cet article présente une méthode de détection et classification de clics d'écholocation d'odontocètes ainsi que d'estimation du nombre d'animaux vocalisant en même temps. Un détecteur de transitoires utilisant le test de Page [1-3] permet d'extraire les clics : leurs instants, durées et amplitudes ainsi que leurs spectres sont stockés. L'analyse du spectre d'un clic permet de lui affecter une probabilité de distribution parmi les différentes espèces. L'estimation du nombre d'animaux se fait à l'aide d'un algorithme de tracking (multi-hypothesis tracking MHT). L'association des clics est basée sur l'hypothèse que l'amplitude et l'intervalle entre deux clics varient lentement en fonction du temps. Ce travail a été réalisé sur le jeu de données mis à disposition par les organisateurs du *3rd International Workshop on the Detection and Classification of Marine Mammals using Passive Acoustics*, Boston, Juillet 2007. Ce dernier se compose de données d'entrainement sur trois espèces : Mésoplodon de Blainville (*Mesoplodon densirostris*), dauphins de Risso (*Grampus griseus*) et globicéphales (*Globicephala macrorhynchus*) et de fichiers test.

1 INTRODUCTION

The most reliable means to detect echolocating cetaceans is acoustic: one listens for "clicks". It is of interest to detect and classify the clicks automatically, and subsequently to determine how many animals are present.

The process observed from each animal is a sequence of clicks whose inter-event times and whose amplitudes vary slowly. From the observer's point of view there is the superposition of an unknown number of such processes, in addition to spurious measurements, hence both tracking and data association are helpful in determining the number of independent sources.

Numerous approaches exist to the tracking problem. Contact-based approaches are of interest here, since clicks provide contact-level measurement information. These techniques include sequential (scan-based), as well as batch processing techniques. In earlier work we documented our results in the analysis of hydrophone datasets with a variety of approaches; the most effective, at least at the time being, has proven to be the multi-hypothesis tracking (MHT) based approach. In this work, we further develop this tracker to include click feature information that allows to classify clicks originating from different species of vocalizing mammals.

Section 2 provides a description of the transient detection algorithm. Section 3 describes the assignment of probability over species to each click. Section 4 describes the feature-aided MHT algorithm. Section 5 provides some results. Conclusions are in section 6.

2 TRANSIENT DETECTION ALGORITHM

The transient detection algorithm is a slightly modified version of the algorithm described in [1]. The algorithm is summarized in figure 1.

First, the data is high-pass filtered to remove part of the noise and avoid detection of whistles (Butterworth order 8, cut-off frequency: 15 kHz). The squared time series of the filtered data is normalized (using an exponential averager) and then submitted to the Page test. The Page test is a sequential detector that provides robustness against unknown signal duration as it detects the start and the end of a signal. At this step, the time, the duration, the amplitude and the spectral information of the click are stored for the next processing steps.

3 PROBABILITY DISTRIBUTION OVER SPECIES

The first step is to have some criteria to distinguish the clicks of the various species.

For the dataset provided for the workshop the maximum sampling frequency is 96 kHz (that means the spectrum is limited to frequencies below 48 kHz). This sampling frequency is not high enough to characterize the entire click spectrum of the different species; because of this limitation the criteria to distinguish the species will be based on the lower portion of the spectrum.

Below we describe some characteristics of the clicks of the three species of interest in the dataset. Based on these characteristics, to each click we identify a normalized likelihood vector that quantifies the goodness of fit of the click spectrum to those of the species of interest. In particular, the four-dimensional likelihood vector includes one element for each species, and one for none of these.

The likelihood vector impacts the track-to-click association scores that are also based on amplitude and Inter-Click Interval (ICI) information, as discussed further in section 4. In particular, the track state includes a probability distribution over the four classes of interest; the inner product between this distribution and the click likelihood vector impacts the track-click association score.

3.1 Blainville's beaked whale

The energy of regular clicks of Blainville's beaked whale is distributed between the -10dB endpoints of about 26 and 51 kHz with a sharp cut-off below 25 kHz and a more gradual cut-off at the high end [6]. The spectra of a few Blainville's beaked whale clicks coming from the training data are illustrated in figure 2; they correspond to the description of [6].

The spectrum of the extracted clicks is not always as nice as the examples of figure 2, as it depends on the quality of the signal, the signal to noise ratio and the quality of the click extraction; what seems important to recognize these clicks is that they have their maximum frequency above 25 kHz and a very sharp cut-off frequency between 20 and 25 kHz. The buzz clicks are different [6] from the regular clicks but no specific criterion to classify them was used.

3.2 Risso's dolphin

For this species the clicks seem to have more spectral diversity; in all the training data files there are some clicks with narrow bursts in their spectrum. These bursts seem to be typical of the Risso's dolphin. They are not always at the same frequency, and there is not always the same number of bursts, but many of these bursts are around 22, 25 and 31 kHz. Some other clicks don't have these bursts at all and are more difficult to characterize. Figure 3 gives an example of some Risso's dolphin clicks spectrum coming from the training data.

3.3 Short-finned pilot whale

The clicks of the pilot whale vary significantly and are not easy to characterize. In the training data, what is often observed is a maximum between 15 and 20 kHz with energy until the end of the band (45 kHz). Some examples in the training data contain clicks with a maximum frequency above 25 kHz. Figure 4 gives an example of some pilot whale clicks spectrum coming from the training data.

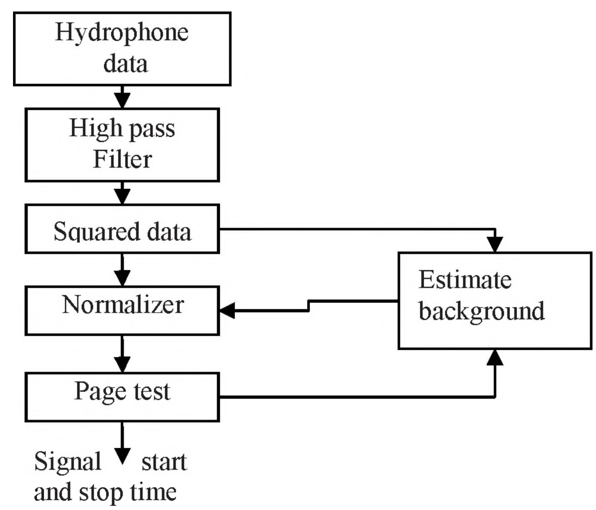

Figure 1: Block diagram of detection scheme.

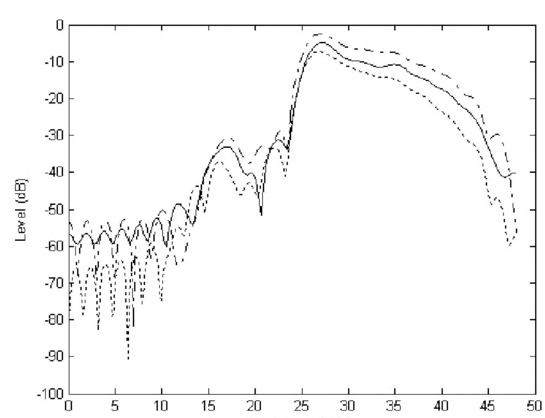

Figure 2: Examples of Blainville's beaked whale regular click spectra (from the training data).

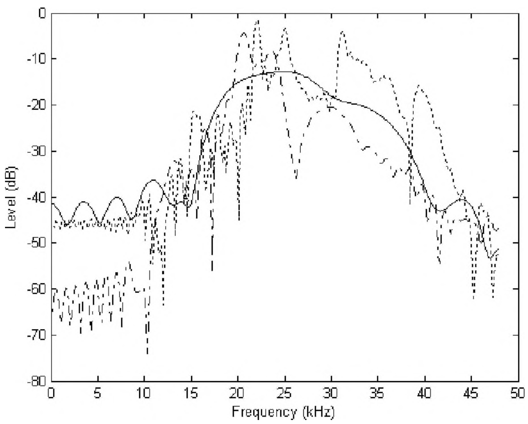

Figure 3: Examples of Risso's dolphin click spectra (from the training data).

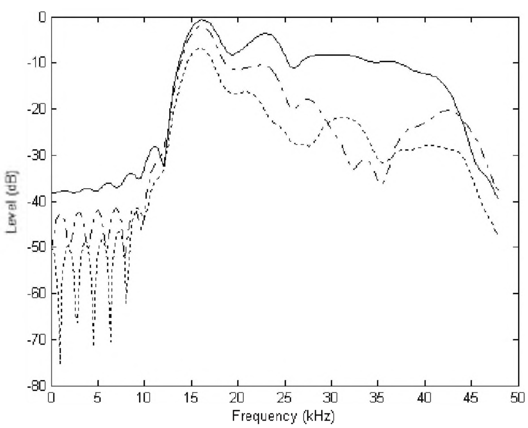

Figure 4: Examples of pilot whale click spectra (from the training data).

4 MHT APPROACH TO CLICK ASSOCIATION

The clicks of interest are echolocation clicks that are emitted by the animals to find prey. The clicks are regular with some pauses; the tracker associates these sequences of clicks. Thus, the track (or associated sequence of clicks) of a single animal will not be contiguous; rather, each animal may generate a number of click sequences separated by lengthy pauses. Our estimate for the number of animals is given by the largest number of tracks that coexist at any time.

Signal processing of hydrophone data results in a single time series of clicks. This time series includes sub-sequences that originate from an unknown number of vocalizing whales, as well as possible spurious clicks. For each marine-mammal originated sub-sequence, we assume the click amplitudes (dB) are slowly varying. Changes in amplitude and intra-click timing are due to animal motion, ambient disturbances, multi-path effects, etc. Animal feeding patterns are another source of change. Each sub-sequence may have missing detections. Our dynamical model for each sub-sequence is the following:

$$20\log x_{k+1} = 20\log x_k + w_k\,, \qquad (1)$$

$$(t_{k+1} - t_k) = (t_k - t_{k-1}) + v_k\,. \qquad (2)$$

x_k is the click amplitude of the click at time t_k; w_k and v_k are noise terms with variance $q_w(t_k - t_{k-1})$ and $q_v(t_k - t_{k-1})$, respectively; the time dependence results from integration of an underlying continuous-time dynamical model.

From equations (1-2), we see that the state of the sub-sequence at time t_k is given by $\left[x_k \quad t_k \quad t_{k-1}\right]'$. As noted above, the overall observed click sequence is given by the union of the marine-mammal originated sub-sequence, with an additional (unmodelled) spurious false click sequence. In the following we have $X_k = 20\log x_k$. Equation (1) becomes:

$$X_{k+1} = X_k + w_k\,, \qquad (3)$$

Neglecting transmission loss differences from one click to the next, the model applies to the received signal amplitude.

The identification of the model parameters q_w and q_v requires the use of clean datasets for which each vocalization sequence has few missed clicks and these originate from the same animal. The workshop dataset does not provide the possibility to estimate these parameters because there is not enough data with just one animal vocalizing.

Our past work in MHT tracking has focused on ground and undersea surveillance, the latter based on the use of active sonar; see [7-8] and references therein. Here, we have leveraged the same data association methodology and track management logic, with appropriate modification to kinematic and measurement modeling, recursive filtering, and measurement gating logic. Kinematic modeling is given by (2-3), with parameter settings as noted above. We assume perfect measurements of click times and amplitudes.

The tracker processes the click time series in sequential fashion. At each step, the set of tentative track hypotheses is updated with the current click. With a fixed latency, known as n-scan, track hypotheses are resolved; by this we mean that a single global hypothesis is maintained and all conflicting track hypotheses are pruned. The global hypothesis selection is based on maximization of the sum of track scores, where each track score is a log-likelihood value that includes a track initiation penalty term.

Track hypotheses are generated on the basis of track validation criteria: each click initiates a tentative track. A later click leads to a track update hypothesis if the resulting ICI is low enough, and if the click amplitudes are sufficiently close, based on a chi-squared criterion; a track coast hypothesis is also generated.

Under the hypothesis of two (or more) associated clicks, subsequent track updates require that a chi-squared criterion be met in both amplitude and ICI. Tentative tracks are confirmed with a minimum-click criterion. Tracks that fail to satisfy this criterion are discarded.

In our past work, the tracker did not exploit click feature information beyond click time and amplitude. In the present work, we exploit feature information in the form of the species type probability distribution described in section 3. Thus, the track state includes a species type vector, in addition to the current estimate of ICI and click amplitude.

As the time series click data is processed, the current click is compared against all active tracks, and only those clicks that satisfy the chi-squared criterion previously mentioned (for ICI and amplitude), in addition to a proximity test for species type, are considered as feasible track-click associations. Finally, it should be noted that elements of the species type vector are clipped at each track update: that is to say, each element is bounded away from 1, to avoid insensitivity over time to new data.

5 RESULTS

The results of our automatic-tracking formalism applied on sperm whale clicks can be found in [5-6]. In those efforts, the datasets supported identification of the relevant motion parameter estimates. In the present work, the dataset does not allow the identification of these parameters because it does not provide sequences of clicks long enough from any single animal. Without the estimation of these parameters, the method is challenged when many animals are present. Below we give the result of a very simple case for Risso's dolphin; for Risso's dolphin and pilot whales, we did not estimate the number of animals vocalizing at the same time.

The ICI of Blainvilles' beaked whales is typically between 0.2 s and 0.4 s [6, 9-10]. With this limitation, even in the presence of many animals, click association is possible even if the right values of model parameters are not known. Some illustrative examples are presented below.

Note that, in the sequence of figures that go with the examples, tracks are plotted alternatively in dashed or dotted line. We do not know definitively whether distinct tracks originate from the same animal. Nonetheless, we estimate the number of vocalizing mammals as the maximum number of co-existing tracks present in a given dataset.

5.1 Training data: Risso's dolphin.

The tracker was applied to a simple case of Risso's dolphin coming from the training data. Figures 5 and 6 give respectively the amplitude of the tracks and the ICI for each track. These results are obtained with the parameters $q_w = 30\,s^{-2}$, $q_v = 0.004$. In this case, clicks of one animal are associated leading to an ICI around 0.6 s. Some clicks are not associated into tracks (figure 5): they can come from an animal that is far and whose clicks are not detected regularly enough to be associated, or possibly they are echoes of the associated clicks.

5.2 Training data: Blainville's beaked whale

For the Blainville's beaked whale data, because of the directivity of the clicks and the fact that these animals have a neck and can move their heads, we have chosen a large value for q_w which allows for large variations in click amplitude. The following parameters are used: q_w = $50\,s^{-2}$, q_v = 0.01. Figures 7 and 8 give respectively the amplitude of the tracks and the ICI for each track for one file of the training data. In this case, almost all the clicks are associated. Figure 9 gives the estimated number of whales versus time. It seems that for this file there are often two whales vocalizing simultaneously.

5.3 Test Data

For the test data, only the number of beaked whales is estimated. The tracks are plotted only if the probability for a track to come from a beaked whale is more than 0.5. We will present two examples. In the first one, most of the clicks come from a sperm whale, but there are also some Blainville's beaked whale clicks. Figures 10 and 11 give respectively the amplitude of the tracks and the ICI for each track generated from this data. Many clicks are not associated (those coming from the sperm whales); nevertheless, some clicks are associated at various times (figure 10). Figure 11 is given for a short time window, so as to illustrate how many whales are vocalizing at the same time. In this example, it seems that a maximum of two whales are vocalizing at the same time.

In the second example, all clicks have been identified (by the data provider) as coming from the pantropical spotted dolphin. Nonetheless, at two points in the time series, the tracker associates clicks with tracks having a high probability to be from beaked whales. Figures 12 and 13 give respectively the amplitude of the tracks and the ICI for one of the two tracks generated from this data. The ICI of these clicks matches the expected ICI of the Blainville's beaked whale, and is consistent with the species type probability distribution of the track. Note that the ICI is not used in determining the species type probability distribution, nor is it used in identifying individual animals. Rather, as discussed previously, the species vector is determined through the spectral content information in the clicks.

Figure 14 gives an example of the spectrum of a click coming from a pantropical spotted dolphin (continuous line) and the spectra of two of the associated clicks (dashed and dotted lines). The spectra of the associated clicks really have the typical shape of the Blainville's beaked whale, which is quite different from the other clicks in the time series. Finally, figure 15 gives the temporal signal of these clicks: they too seem to have the typical shape of the Blainville's beaked whale [11]. For all these reasons, we conclude that these few clicks likely come from a Blainville's beaked whale.

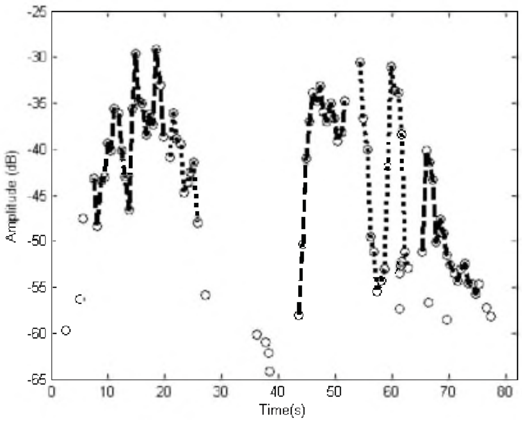

Figure 5: Risso's dolphin click amplitude data (circles) and MHT output (in dashed or dotted lines).

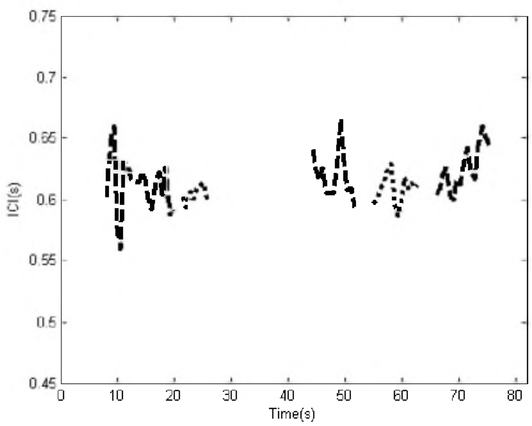

Figure 6: Sequences of Risso's dolphin ICIs for tracks generated by the MHT tracker.

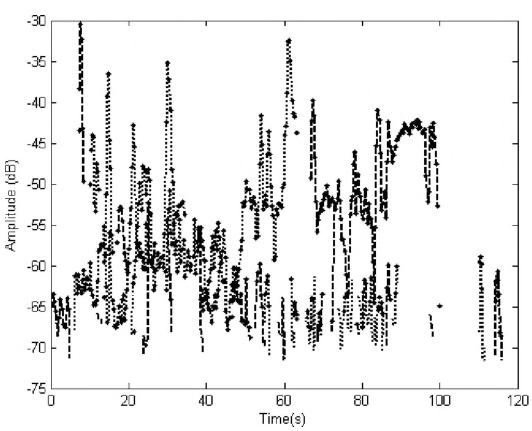

Figure 7: Blainville's beaked whale click amplitude data (dots) and MHT output (dashed or dotted lines).

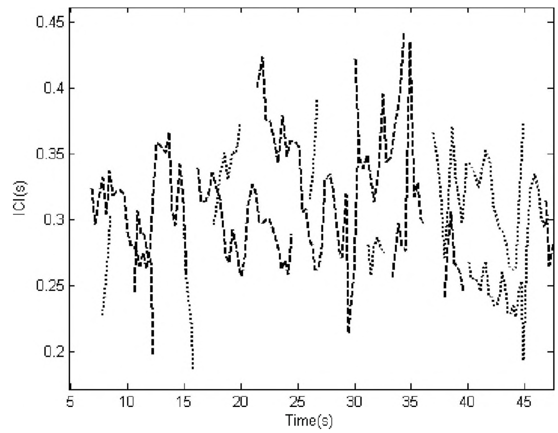

Figure 8: Sequences of Blainville's beaked whale ICIs for tracks generated by the MHT tracker.

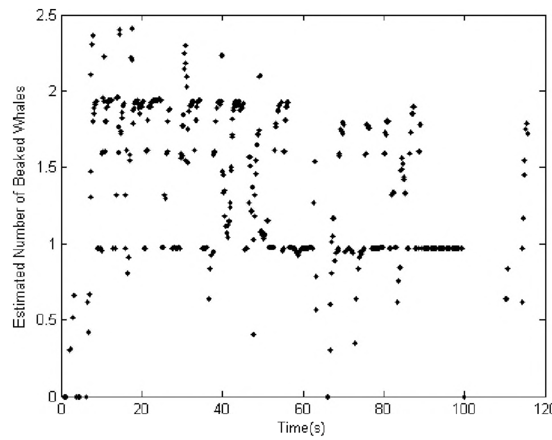

Figure 9: Estimated number of whales vocalizing versus time

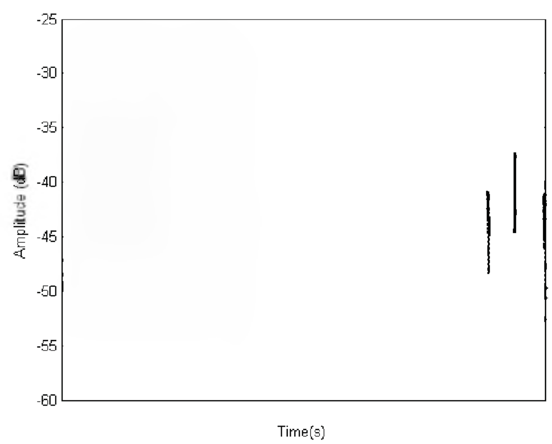

Figure 10: Click amplitude data (dots) and Blainville's beaked whale clicks MHT output (test data, first example).

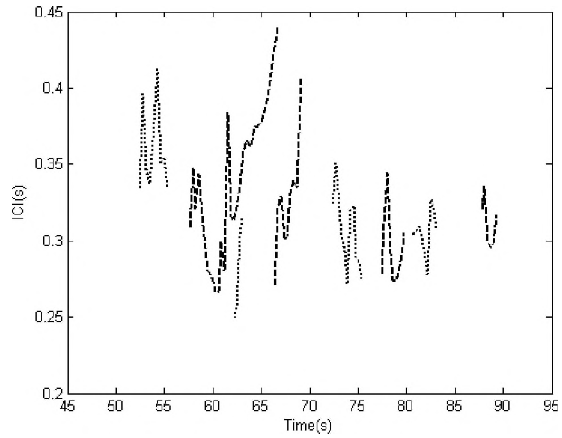

Figure 11: Sequences of ICIs for tracks generated by the MHT tracker.

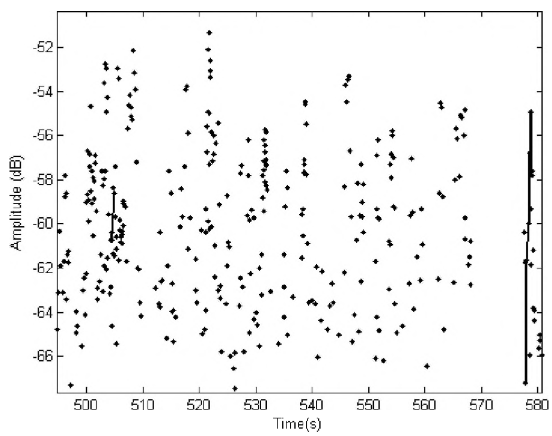

Figure 12: Click amplitude data (dots) and Blainville's beaked whale clicks MHT output (test data, second example).

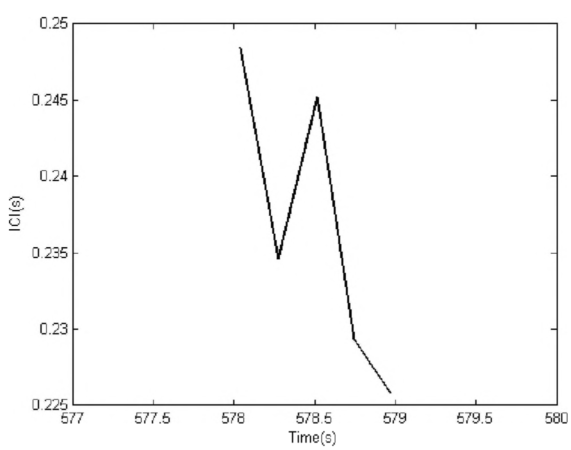

Figure 13: ICIs for a track generated by the MHT tracker.

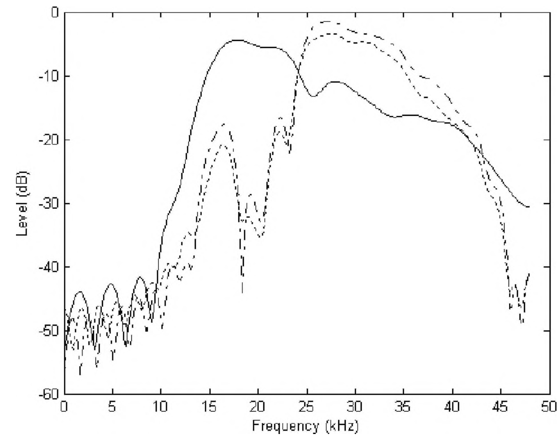

Figure14: Examples of pantropical spotted dolphin click spectrum (continuous line) and – probably – Blainville's beaked whale click spectrum (dashed and dotted line; test data, second example).

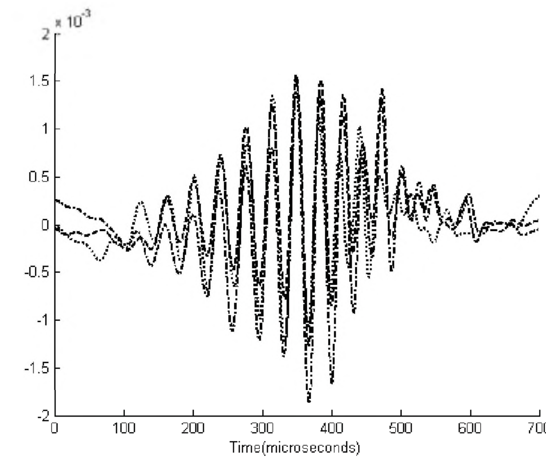

Figure15: Temporal signal of the associated clicks of – probably – the Blainville's beaked whale (test data, second example).

6 CONCLUSIONS

This paper presents a novel application and extension of target-tracking technology to marine-mammal detection and classification; the paper extends our past work to include feature-aided tracking. The results are promising, and help in classifying beaked whales' clicks as well as to estimate the number of animals present.

Our approach is challenged when many animals are present, especially if they are not beaked whales. To improve the results, it would be helpful to identify features or processing methods to distinguish animals within the same species, and, more generally, to determine an improved methodology to assign the click feature vector. Improved feature vector information would directly support improved tracking performance.

Finally, the parameters in our kinematic modelling of ICI and amplitude dynamics should be species dependent. Thus, a coupled approach to kinematic and classification filtering has the potential further to improve detection and classification performance.

7 ACKNOWLEDGMENTS

We thank the Woods Hole Oceanographic Institution (WHOI) and the Naval Undersea Warfare Center (NUWC) for providing datasets used in this study. We wish also to thank all the people involved in the organization of the *Third International Workshop on Detection and Localization of Marine Mammals using Passive Acoustics*, Boston, July 2007. And finally we thank the French Navy to have partly funded this study.

8 REFERENCES

1. D.A. Abraham, Passive acoustic detection of marine mammals, SACLANTCEN SM 351, 2000, NATO Undersea Research Centre, La Spezia, Italy.
2. R. Laterveer, S.P. Beerens, S.P. van Ijsselmuide, M. Snyder, J. Newcomb, A high frequency transient detector suitable for monitoring marine mammal activity, in Proceedings of UDT Europe, June 2004, Nice, France.
3. D.A. Abraham, P.K. Willett, Active Sonar Detection in Shallow Water using the Page Test, IEEE Journal of Oceanic Eng. Vol. 27 (1) January 2002.
4. O. Gerard, S. Coraluppi, W. Zimmer, and P. Willett, A Chorus of Whales: Evaluation of Sequential and Batch Approaches to Time-Series Tracking, in Proceedings of OCEANS 2006, September 2006, Boston MA, USA.
5. O. Gerard, C. Carthel, S. Coraluppi, and P. Willett, MHT and ML approaches to marine mammals detection, in Proceedings of OCEANS 2007, June 2007, Aberdeen, UK.
6. M. Johnson, P.T. Madsen, W.M.X. Zimmer, N.A. Aguilar de Soto and P.L. Tyack (2006), Foraging Blainville's beaked whales produce distinct click types matched to different phases of echolocation, The Journal of experimental biology 209 (24), 5038-5050.
7. S. Coraluppi and C. Carthel, Recursive Track Fusion for Multi-Sensor Surveillance, Information Fusion, vol. 5(1), March 2004.
8. S. Coraluppi and C. Carthel, "Distributed Tracking in Multistatic Sonar", IEEE Transactions on Aerospace and Electronic Systems, vol. 41(3), July 2005.
9. M. Johnson, P.T. Madsen, W.M.X. Zimmer, N.A. Aguilar de Soto and P.L. Tyack (2004), Beaked whales echolocate on prey, Proc. Roy. Soc. Lond. B 271, S383-S386
10. P.T. Madsen, M. Johnson, N.A. Aguilar de Soto, W.M.X. Zimmer, and P.L. Tyack (2005), Biosonar performance of foraging beaked whales (Mesoplodon densirostris), The Journal of experimental biology.,208, 181-194.
11. M.A. Roch, M.S. Soldevilla, J.C. Burtenshaw, R. Hoenigman, S.M. Wiggins, and J.A. Hildebrand., Comparison of machine learning techniques for the classification of echolocation clicks from three species of odontocetes, in Canadian Journal of Acoustics (current issue).

Photo Credit:Bahamas Marine Mammal Research Organisation