

A Chorus of Whales: Evaluation of Sequential and Batch Approaches to Time-Series Tracking

Odile Gerard, Stefano Coraluppi, Walter Zimmer
NATO Undersea Research Centre
La Spezia, Italy
{gerard, coraluppi, zimmer}@nurc.nato.int

Peter Willett
ECE Department
University of Connecticut
Storrs CT, USA
willett@engr.uconn.edu

Abstract- This paper applies target-tracking technology to a novel application: the processing of mammal vocalizations or clicks, with the goal of identifying the number of marine mammals in a surveillance region. This problem has direct application to marine mammal mitigation efforts in the context of active sonar operations.

I. INTRODUCTION

Although the exact mechanism is not clear, there is considerable evidence that some species of marine mammals can suffer significant harm from active sonar operations. In recent years, the NATO Undersea Research Centre has taken a lead role in international collaborative research efforts to study this problem. Further information is available online [1].

It is a matter of both environmental interest and common decency to ascertain that no marine mammals are nearby prior to active sonar operations. The most reliable means to detect echolocating cetaceans is acoustic: one listens for "clicks". The problem is complicated when several whales are in the vicinity: it is then of interest to know how many are present, such that one can be sure when all have left the area.

Some species of marine mammals exhibit regular click vocalization, especially sperm whales. In this case, the observations process from each animal is a sequence of events (actually: clicks) whose inter-event times and whose amplitudes, while not constant, vary slowly according to range and behavior (click repetition rate increases and amplitude decreases as the distance to prey closes). From the observer's point of view there is the superposition of several (an unknown number of) such processes, in addition to spurious (false) measurements, and hence both tracking and data association will be necessary to determine the number of independent data sources.

A number of approaches exist to the tracking problem. These include non-contact based approaches (track-before-detect and so-called "Bayesian tracking"), as well as contact-based approaches [2-3]. The latter class of methods is of interest here, since clicks provide contact-level measurement information. Contact-based approaches include sequential (scan-based) methodologies, as well as batch processing techniques.

Most scan-based tracking algorithms include track initiation, track maintenance, and track termination components. Numerous track-maintenance methodologies of varying complexity exist, including statistical nearest neighbor (NN) techniques, probabilistic data association (PDA) and extensions (e.g. the JPDA), and multi-hypothesis tracking (MHT) [2, 4]. MHT techniques provide improved performance at the cost of a short time latency. Recent developments in MHT technology applied to undersea surveillance are reported in [5].

Batch processing techniques are effective at identifying dim targets, with some target motion and target number assumptions. Recent applications of these techniques to undersea surveillance are reported in [6].

The observation process of interest here is a highly non-traditional one for multiple target tracking, whose algorithms usually expect scan-based "hits" of possible target locations. In this paper, we analyze the hydrophone data and develop baseline detection processing and a NN tracking solution. In the future, we plan to compare these results with a more sophisticated MHT-based sequential processing scheme, as well as a batch algorithm based on Streit & Luginbuhl's PMHT model [7]. Evaluation is on the basis of accurate determination of target number.

II. PROBLEM STATEMENT

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 subsequence, we assume that the click amplitude (in dB) and Inter-Click Interval (ICI) are slowly varying. Changes in amplitude and intra-click timing may be due to animal motion, ambient disturbances, multi-path effects, etc. A more significant source of changes will be due to animal feeding patterns; it remains to be seen how effectively the simple dynamical model described below handles these changes. Each subsequence may have some missing detections.

Our dynamical model for each subsequence is the following:

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 01 SEP 2006	2. REPORT TYPE N/A	3. DATES COVERED -	
4. TITLE AND SUBTITLE A Chorus of Whales: Evaluation of Sequential and Batch Approaches to Time-Series Tracking		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NATO Undersea Research Centre La Spezia, Italy		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)	
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited			
13. SUPPLEMENTARY NOTES See also ADM002006. Proceedings of the MTS/IEEE OCEANS 2006 Boston Conference and Exhibition Held in Boston, Massachusetts on September 15-21, 2006, The original document contains color images.			
14. ABSTRACT			
15. SUBJECT TERMS			
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	UU
			18. NUMBER OF PAGES 6
			19a. NAME OF RESPONSIBLE PERSON

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

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

In equations (1-2), x_k is the click amplitude of the click at time t_k , while w_k and v_k are process noise or disturbance terms with variance $q_w(t_k - t_{k-1})$ and $q_v(t_k - t_{k-1})$, respectively. (The time dependence of variances results from time-integration of an underlying continuous-time dynamical model.)

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

$$X_{k+1} = X_k + w_k. \quad (3)$$

As we neglect differences in transmission loss from the whale to the hydrophone from one click to the next, this model applies to the received signal amplitude.

III. SCAN-BASED PROCESSING: BASELINE ALGORITHM

We define a simple scan-based tracking algorithm that will provide a performance baseline for the batch processing approach as well as further scan-based tracking upgrades. The algorithm is given below.

Baseline Algorithm (summary)

- First contact initiates a track;
- For each subsequent contact:
 - If there exists at least one neighboring track, associate the contact to the “closest” track, and update the track accordingly;
 - Otherwise start a new track;
- Terminate all tracks after T sec with no update or after M missed detections;
- After processing, remove tracks with fewer than N associated clicks.

The algorithm above requires a non-negative T , positive integers M and N , and an association gating parameter G . For each track, the prediction, association gating, and update steps are defined as follows:

For tracks based on a single click ($k=1$):

- Prediction: $\xi(k+1|k) = X_k$; $P(k+1|k) = q_w(t_{k+1} - t_k)$;
- Innovations $\delta = (X_{k+1} - X_k)$;
- Gating test: $\delta P^{-1}(k+1|k)\delta \leq G$;

- Update: $\xi(k+1|k) = \begin{bmatrix} X_{k+1} \\ (t_{k+1} - t_k) \end{bmatrix}$.

For tracks based on multiple clicks ($k > 1$):

- Prediction: $\xi(k+1|k) = \begin{bmatrix} X_k \\ (t_k - t_{k-1}) \end{bmatrix}$;
- $P(k+1|k) = \begin{bmatrix} q_w(t_{k+1} - t_k) & 0 \\ 0 & q_v(t_{k+1} - t_k) \end{bmatrix}$;
- Innovations: $\delta = \begin{bmatrix} X_{k+1} \\ (t_{k+1} - t_k) \end{bmatrix} - \begin{bmatrix} X_k \\ (t_k - t_{k-1}) \end{bmatrix}$;
- Gating test: $\delta P^{-1}(k+1|k)\delta \leq G$;
- Update: $\xi(k+1|k) = \begin{bmatrix} X_{k+1} \\ (t_{k+1} - t_k) \end{bmatrix}$.

Note that the filtering equations above are the limiting form of the Kalman filter, as measurement noise tends to zero. (An alternative problem formulation in sec. 2 would have included both process and measurement noise.)

IV. DATASETS

For this study, we have chosen to work with sperm whale recorded hydrophone data. The sperm whale (*Physeter macrocephalus*) is the largest toothed whale, and may reach 18m in length and 50tons in weight. It is an active hunter, probably preying on giant squid. It can stay underwater for over an hour, at depths surpassing 1km. The vocalizations of sperm whales are made with brief pulsing sounds, called clicks. These clicks generally reach 30-35kHz in frequency, with high repetition rates. A short time after leaving the surface, the whale begins to click regularly, probably looking for food. A sequence of clicks is followed by brief periods of silence, or by sequences of clicks repeated at high repetition rates called creak or runs [8]. The amplitude of the creak is low and not often recorded. 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 will then be given by the largest number of tracks that coexist at any time.

The ocean medium is complex, and received acoustic data is correspondingly noisy. An additional source of amplitude variability is the directionality of the whale as an acoustic source: even if two clicks from the same whale are of the same power, the signal level at the hydrophone varies as a function depending whale orientation [9].

Three datasets will be used here. The first is from a *Dtag* acoustic recording tag [10] attached on a sperm whale in the Ligurian sea during the NURC sea trial Sirena'03 that was conducted in collaboration with the Woods Hole Institute. After the NURC research vessel (R/V Alliance) detected and localized a sperm whale, a small boat approached the whale to attach the *Dtag* on its dorsal surface. The signal is sampled at 96 kHz. An example of the signal recorded by the *Dtag* hydrophone and corresponding to a whale click is given Fig. 1.

The two other datasets were recorded by NUWC with a bottom-mounted hydrophone. The signals were sampled at 48 kHz. One is a 25min recording of either a single vocalizing sperm whale with reverberation, or two sperm whales. The other is a 20min recording of three or more whales. Fig. 2 and Fig. 3 illustrate a few seconds of hydrophone data for these two datasets.

The clicks were extracted by applying a simple threshold to the signal. The identification of the dynamical 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.

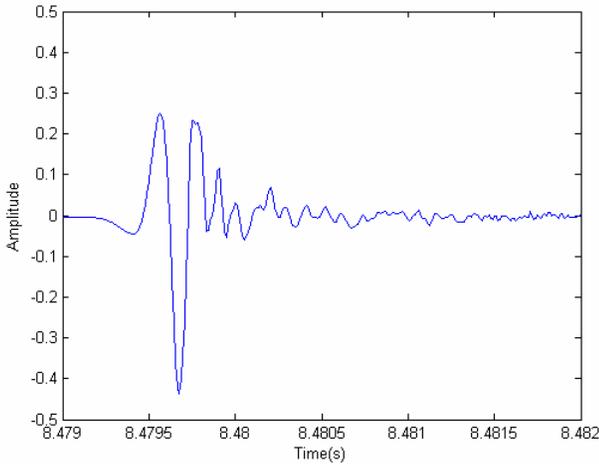


Figure 1. Click amplitude for the *Dtag* hydrophone (first dataset).

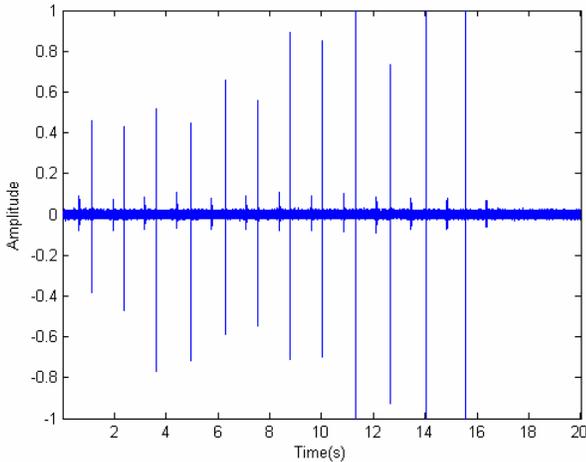


Figure 2. Sequence of clicks and click echoes (second dataset).

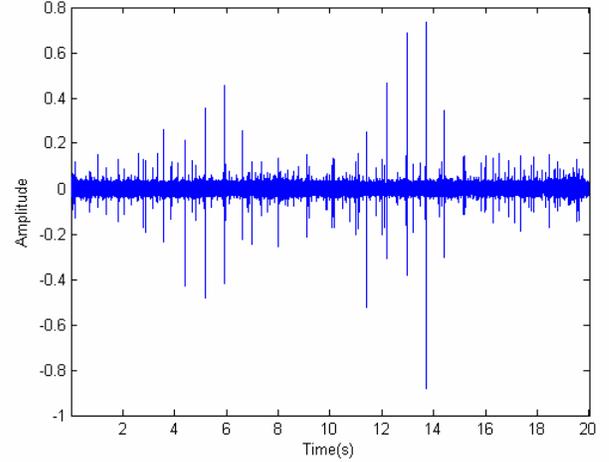


Figure 3. Sequence of clicks (third dataset).

V. PRELIMINARY EXPERIMENTAL RESULTS

The parameters q_w and q_v have been calculated for the first two datasets, where the signals are clean enough to isolate a single vocalizing sperm whale. For the first (*Dtag*) dataset, the parameter estimates were obtained from 3530 clicks and are as follows:

$$q_w = 14 s^{-2}, q_v = 0.072.$$

For the second dataset (NUWC bottom-mounted hydrophone) the parameter estimates were obtained from 950 clicks and are as follows:

$$q_w = 3.6 s^{-2}, q_v = 0.048.$$

No parameter estimation was performed using the third (noisy) dataset, which also originates from the NUWC bottom-mounted hydrophone. In this dataset, it is difficult to ascertain which clicks originate from the same animal, so that the dataset is not useful for model identification.

All the results that we illustrate in the following have been obtained with the following set of tracker parameters:

- Prediction: Track termination: $T=3, M=1$;
- Track validation: $N=9$;
- Click association: $G=9.2$, corresponding to a 99% data association gate.

If the gating test is not successful with any track, we double the click interval for all tracks and reapply the gating test. This accounts for the possibility of a missed click detection.

As previously mentioned, all the tracking results that we illustrate are for the baseline (NN) tracker.

A. First dataset: DTag data

We first apply the tracker to a time window of the first dataset that includes two vocalizing mammals, one with the tag and one without. (A different time window of the same dataset, where only the tagged whale was present, was used for model parameter estimation.)

Fig. 4 and Fig. 5 give respectively the amplitude of the tracks and the Inter-Click Interval for each track. The color changes each time a new track is plotted and 5 colors are used in total. (The color does not have any meaning.)

We can see that the two whales are distinctly tracked. Occasionally, there is fragmentation as we terminate track on the whale and initiate a new track. From an examination of Fig. 5, we can confirm that we indeed have a second whale and not an echo because the ICIs are different.

Note that starting at around 200s there are three tracks at the same time, two of which have the same ICI. It appears that the track on one whale terminates, and is replaced by a pair of interleaving tracks, each using every other click from the same animal. We confirm this by examining Fig. 6, which is a zoom of Fig. 4 on these two tracks. Note further that, as seen in Fig. 5 and as we would expect, the ICIs calculated for these two tracks are twice the true one.

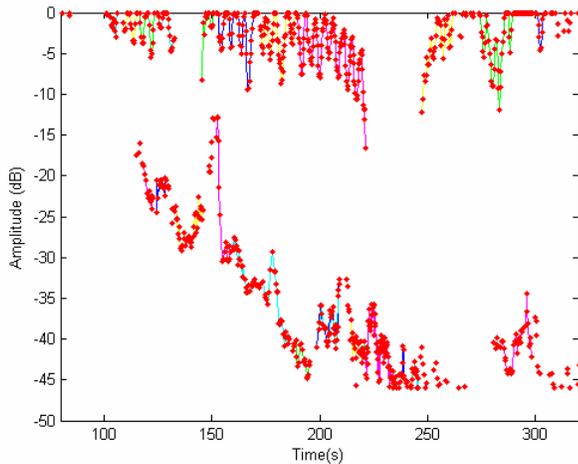


Figure 4. Click amplitude sequence (in red), and the resulting automatically-originated tracks (in different colors).

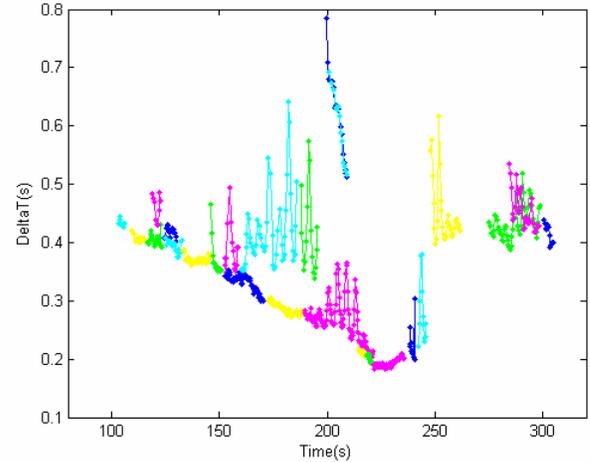


Figure 5. Sequences of Inter-Click Intervals for all tracks; note that the blue and cyan track with large ICIs are the duplicate tracks described above, and correspond to the same animal.

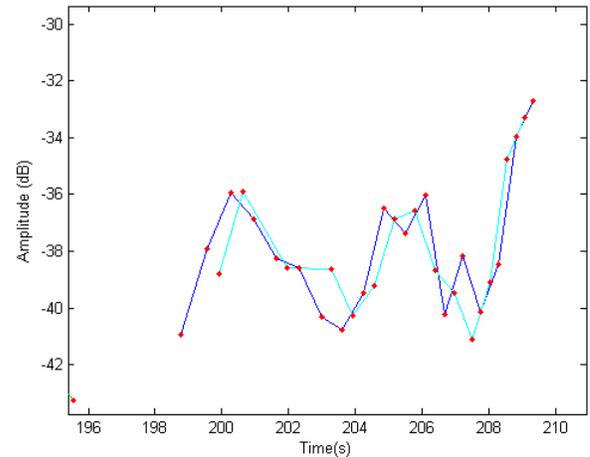


Figure 6. Zoom of figure 4, where we see the two interleaving tracks for clicks from the same animal.

B. Second dataset: Bottom hydrophone with one or two whales

We now apply the tracker to the first of the two NUWC datasets. Here we use second set of model parameter estimates. Recall that the estimates were based on the same dataset, though a much higher detection threshold was used so as to be sure to have only clicks from one animal.

Fig. 7 and Fig. 8 give, respectively, the click amplitude sequences for the tracks and the ICI sequences for a portion of the dataset. We note in Fig. 8 that the ICI is consistently the same for the two tracks that coexist at the same times; this demonstrates that what we observe is in fact an echo and not a second whale: the probability that two whales click with exactly the same ICI is low.

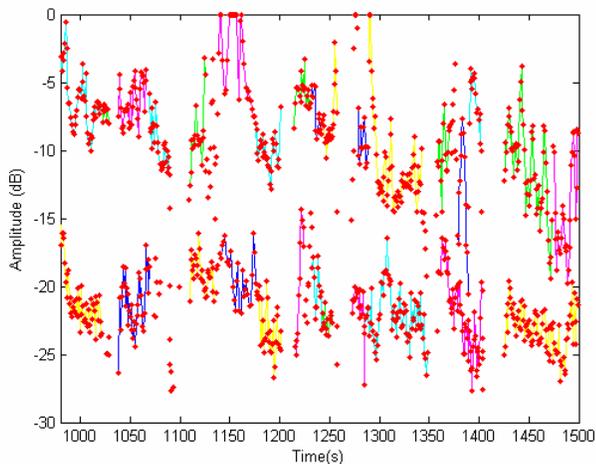


Figure 7: Click amplitude sequence (in red), and the resulting automatically-originated tracks (in different colors).

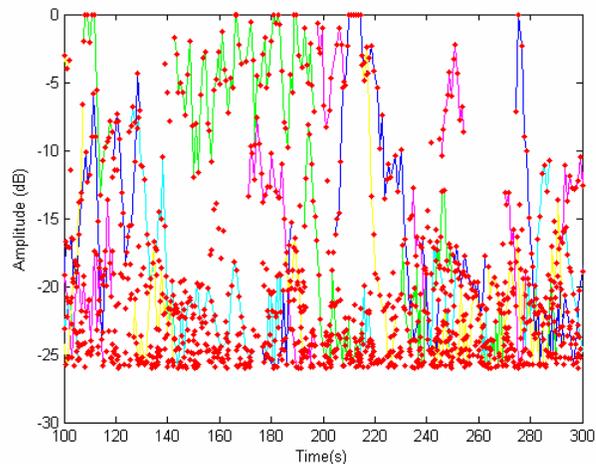


Figure 9: Click amplitude sequence (in red), and the resulting automatically-originated tracks (in different colors).

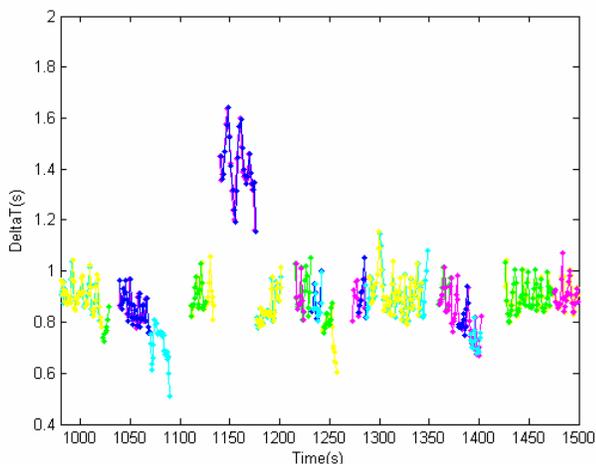


Figure 8. Sequences of Inter-Click Intervals for all tracks.

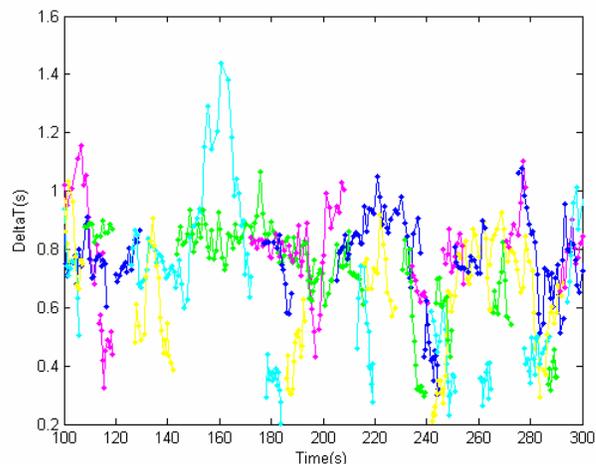


Figure 10: Sequences of Inter-Click Intervals for all tracks.

C. Third dataset: Bottom hydrophone with three or more whales

We now consider the third, and the most complex, of our datasets. This is the second NUWC dataset, and contains at least three vocalizing whales. Here we use model parameter estimates that are obtained by averaging the estimates that were based on the first two datasets. Fig. 9 and Fig. 10 give, respectively, the click amplitude sequences for all tracks that are generated, and the ICI sequences, for a portion of this dataset.

For this dataset, the tracker succeeds in associating clicks and generates many tracks that coexist at the same time (at least four). However, because of the complexity of the dataset (there are many clicks and they are not easily segmented into different amplitude and ICI ranges), the data association is not reliable. This suggests that further work is required: we need a better way to extract the clicks from the hydrophone data, since a simple threshold leads to too many spurious detections

for an acceptable detection probability. Alternatively, a higher threshold suppresses spurious detections, at the cost of a reduced detection probability.

Further, we need to improve upon the baseline tracking algorithm evaluated here. One possibility is to develop a more sophisticated scan-based approach (e.g. MHT). The strengths of this approach relative to a batch-processing approach remain to be evaluated.

VI. CONCLUSIONS

Marine mammal risk mitigation in the context of active sonar operations is a problem of increasing interest in the international community. This issue has been the focus of an ongoing research project at NURC. A key element of this work is the need for automated technologies to help sonar operators in monitoring the environment for possible marine-mammal presence.

As in other application domains, detection data is far too voluminous for an operator to contend with effectively. Thus,

there is a need for automated techniques to drastically reduce the data amount, while identifying key information contained in the data. In our case, the quantity of interest is an estimate of the number of mammals in the surveillance region. It should be noted that our work only addresses the detection of marine mammals that exhibit regular click vocalization.

This paper represents a first attempt to understand the nature of the click data that results from detection processing, as well as to study effective tracking techniques to identify the number of mammals. We have started simple: we have applied a simple threshold to the signal data, and have developed a straightforward nearest-neighbor approach to scan-based tracking.

Target tracking requires adequate kinematic modeling of targets. Thus, we have started our analysis of click data by first identifying relevant motion parameter estimates. Subsequently, we have applied our track to datasets of increasing complexity. The results are promising, but the complexity of the third datasets (as well as of other datasets for which results are not reported here), suggest that further work is required on both the detection and the tracking problems, in order to have an effective surveillance tool.

Continuing work on the tracking problem will include the following elements:

- Development and comparison of more sophisticated scan-based and batch processing approaches to data association;
- Exploitation of click frequency information;

ACKNOWLEDGMENT

We thank Peter Tyack and the Woods Hole Oceanographic Institute (WHOI) for collaborating with NURC in its sea trial

efforts, particularly for the expertise in tagging whales and assisting NURC in acquiring valuable datasets for further analysis.

We thank the Naval Undersea Warfare Center (NUWC) and the Atlantic Undersea Test and Evaluation Center (AUTC) for providing the second and third acoustic datasets used in this study. We wish also to thank Olivier Adam for making these datasets available to us through the *Second International Workshop on Detection and Localization of Marine Mammals using Passive Acoustics*, Monaco, Nov. 2005.

REFERENCES

- [1] <http://solmar.nurc.nato.int/>.
- [2] S. Blackman and R. Popoli, "Design and Analysis of Modern Tracking Systems", *Artech House*, 1999.
- [3] L. Stone, C. Barlow, and T. Corwin, *Bayesian "Multiple Target Tracking"*, *Artech House*, 1999.
- [4] Y. Bar-Shalom and X. Li, "Multitarget-Multisensor Tracking", *YBS Publishing*, 1995.
- [5] S. Coraluppi and C. Carthel, "Distributed Tracking in Multistatic Sonar", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41(3), July 2005.
- [6] P. Willett and S. Coraluppi, "MLPDA and MLPDHT Applied to MSTWG Data", to appear in *Proceedings of the 9th International Conference on Information Fusion*, July 2006, Florence, Italy.
- [7] R. Streit (ed.), "Studies in Probabilistic Multi-Hypothesis Tracking and Related Topics", *Scientific and Engineering Studies* Vol. SES-98-01, NUWC Division Newport, February 1998.
- [8] J. Mullins, H. Whitehead, and L.S. Weilgart. 1988. "Behaviour and vocalizations of two single sperm whales, *Physeter macrocephalus*, off Nova Scotia". *Can. J. Fish. Aquat. Sci.* 45(10): 1736-1743.
- [9] W. M. X. Zimmer, P.L. Tyack, M.P. Johnson, P.T. Madsen "Three-dimensional beam pattern of regular sperm whale clicks confirm bent-horn hypothesis", *The Journal of the Acoustical Society of America*, vol. 117(3): 1473-1485.
- [10] M.P. Johnson, P.L. Tyack "A digital acoustic recording tag for measuring the response of wild marine mammals to sound", *IEEE Journal of Oceanic Engineering*, vol. 28: 3-12.