

Blue Whale B and D Call Classification Using a Frequency Domain Based Robust Contour Extractor

Shyam Kumar Madhusudhana[†], Marie A. Roch[†], Erin M. Oleson[‡], Melissa S. Soldevilla[‡], John A. Hildebrand[‡]

[†] Department of Computer Science, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-7720

[‡] Scripps Institution of Oceanography, The University of California at San Diego, La Jolla, CA 92093-0205

Abstract-Passive acoustic monitoring of blue whales (*Balaenoptera musculus*) has been used to gain insight into their presence and behavior. Many of their calls have been shown to be detectable through spectrogram correlation due to the low variation in these stereotyped calls. In this work, we describe rule based classifiers for tonal B and D calls using the pitch/frequency contour information obtained from a contour extractor. B calls can be detected by spectrogram correlation, but the D calls are highly variable and are therefore difficult to detect using spectrogram kernel methods. Experiments on four hours of evaluation data from different field seasons show that 91.3% of B calls and 85.8% of D calls were correctly retrieved. For both types of calls, less than 2% of the retrieved calls were false positives.

I. INTRODUCTION

Recent uses of acoustic based methods for monitoring marine mammals have shown passive acoustic monitoring to be an attractive complement to visual methods. Passive acoustics have been used to obtain long-term information about their behavioral patterns such as migration, feeding and reproduction [1]. The temporal nature in blue whale B and D calls, coupled with high variability in frequency content, call duration and sweep rate, make detection/classification with existing systems (such as spectrogram correlation [2] and Entropic's fundamental frequency detector [3]) difficult. On average, B calls are ~20 s long and sweep gently downward with a fundamental frequency range from 18 Hz to 15 Hz and a series of higher frequency harmonics. The third harmonic is the most prominent harmonic which follows the fundamental between 55 Hz and 40 Hz. D calls are short (1-4 s) and are characterized by sharp but varying sweep rates between 95 Hz and 45 Hz. The presence of vocalizations by other marine mammals such as fin whales (*Balaenoptera Physalus*), call segmentation [4] and noise from ships, hardware, etc. add to the difficulty of the detection task.

The frequency domain based robust contour extractor described in [5] is used for obtaining time-frequency contours of calls. The

contour extractor performs two passes over per-frequency-bin mean normalized spectral data. In the first pass, close lying energy peaks across neighboring frames are identified. Such energy peak pairs exhibit continuity across a sequence of frames that contain a frequency-modulated signal. Sets of such energy peak sequences are seen as candidate contours. In the second pass, each candidate contour is either retained or rejected based upon its duration, sweep rate, etc. The output of the contour extractor consists of a sequence of frequency values (in Hz) representing sampled points on the contour of the identified call.

A run of the contour extractor on a segment of recording could output multiple time-separated frequency contours depending on the number of calls present in the segment and the apparent fragmentation of individual calls. In several cases, the contours obtained for individual calls may be fragmented as a result of-

- call fragmentation caused by drop-outs or calling behavior [6], or
- presence of high energy at frequencies other than the prominent harmonic during the call period.

Fig. 1 shows a sample segment containing a highly fragmented B call (from ~4s to ~20s). The fragmentation seen in the extracted contour is caused by both of the reasons mentioned above.

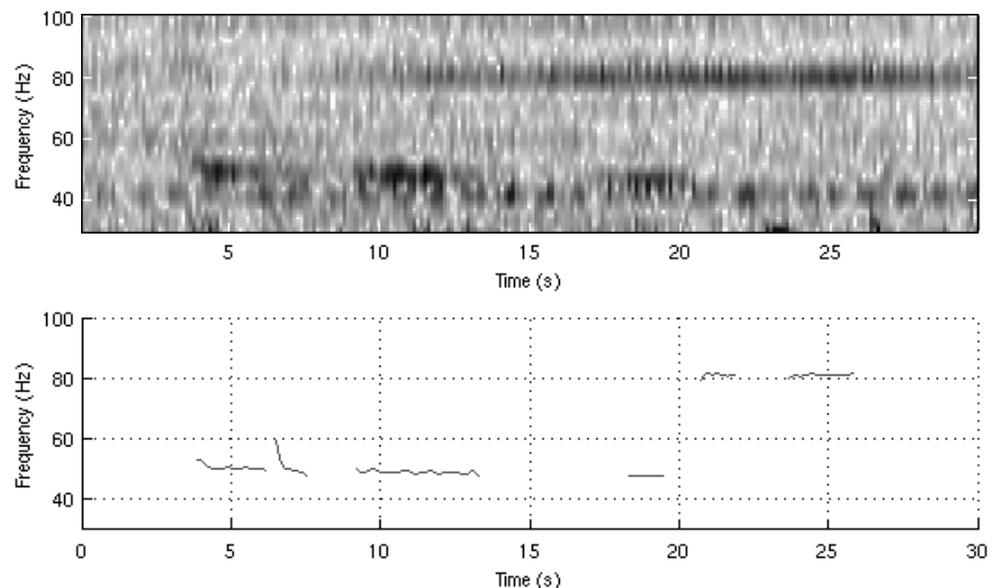


Fig 1. Spectrogram of a sample segment containing a B call and the extracted contours plotted on a time-frequency graph. The call exhibits high fragmentation. Spectrograms produced using 256 ms Hanning window and 1024 point FFT with an overlap of 50%.

The set of time-disjoint frequency contours provided by the contour extractor are used as features for the classification system, which consists of two rule-based expert systems which independently accept or reject sets of contours as B or D calls. Section II describes the general structure of the algorithm common to both the classifiers, with specifics presented in sections III and IV for B and D calls separately. Section V presents the results of using the classifiers on disjoint development and evaluation data from four field seasons.

II. OVERVIEW

The classifiers are expert systems whose rules were framed after examining the characteristics of a large number of samples of the two types of calls covering a majority of their variations. In the algorithm, we first make an assumption that each contour may represent only a fragment of a complete call. Based on this assumption, rules are developed to accept not just contours of complete calls but also other contours that exhibit a high correspondence to the general call structure and which are only shorter in duration than average call lengths. First, the classifier attempts to identify contours of call fragments (if any) and then relate them together to realize the complete call. The tests for identifying call fragments include checks on their duration, frequency range and sweep rate. Call fragmentation contributes to making classification a formidable challenge. Some short duration calls of blue whales or other species, and non-call sounds in the recordings may have characteristics that appear similar to either B or D call fragments and without careful consideration could easily be incorporated into a fragmented B or D call. The tests for relating successive contours include checks on their separation across both time and frequency, similarities in their sweep rates, total duration (fragment durations + separation), etc. The rules need to be developed such that they avoid the relating of fragments from separate calls (across both time and frequency).

The contour extractor produces a sequence of N contours $[c_1, c_2, \dots, c_N]$. Each detected call spans over a set of contours $\{c_i, c_{i+1}, \dots, c_{j-1}, c_j\}$, where $1 \leq i \leq j \leq N$. One or more contours in this set may not match the characteristics of the call being traced. Let R denote a subset of $\{c_i, c_{i+1}, \dots, c_{j-1}, c_j\}$ that contains only the contours that match the characteristics of the call being traced. While a call representation is being constructed, the start and end indices i and j are represented with labels $cBegin$ and $cEnd$. Contours c_{cBegin} and c_{cEnd} are the first and the last contours in R . At the start of the run, R is an empty set. The contours $[c_1, c_2, \dots, c_N]$ are examined successively. If a contour does not bear the characteristics of a call fragment, we skip to the next contour. Otherwise, checks are performed to see whether the initial assumption still holds ~~good~~, i.e. whether the current contour c_i is a portion of the larger call being traced. If the assumption holds ~~good~~, the current contour is considered to be an extension of the larger call being traced (add c_i to R and update $cEnd$). Where the

assumption fails, the call being traced (represented by contours in R) is considered to end at the last accepted contour (c_{cEnd}) and the current contour is considered to be the first fragment of a new call (clear R , add c_i to R , and update $cBegin$ and $cEnd$). An outline of the steps followed in the classifiers is shown in Fig. 2. At any point during the process, c_{cBegin} and c_{cEnd} always indicate the contours at the extremes of the call being traced and the set of contours that make up the traced call are available in $R = \{c_{cBegin}, \dots, c_{cEnd}\}$, regardless of whether the call is fragmented (where $cBegin < cEnd$ and $|R| > 1$) or not (where $cBegin = cEnd$ and $|R| = 1$).

Regardless of the cause of call fragmentation, it is important to treat such fragments as a single call ~~in order~~ to provide better call statistics for researchers studying the seasonal

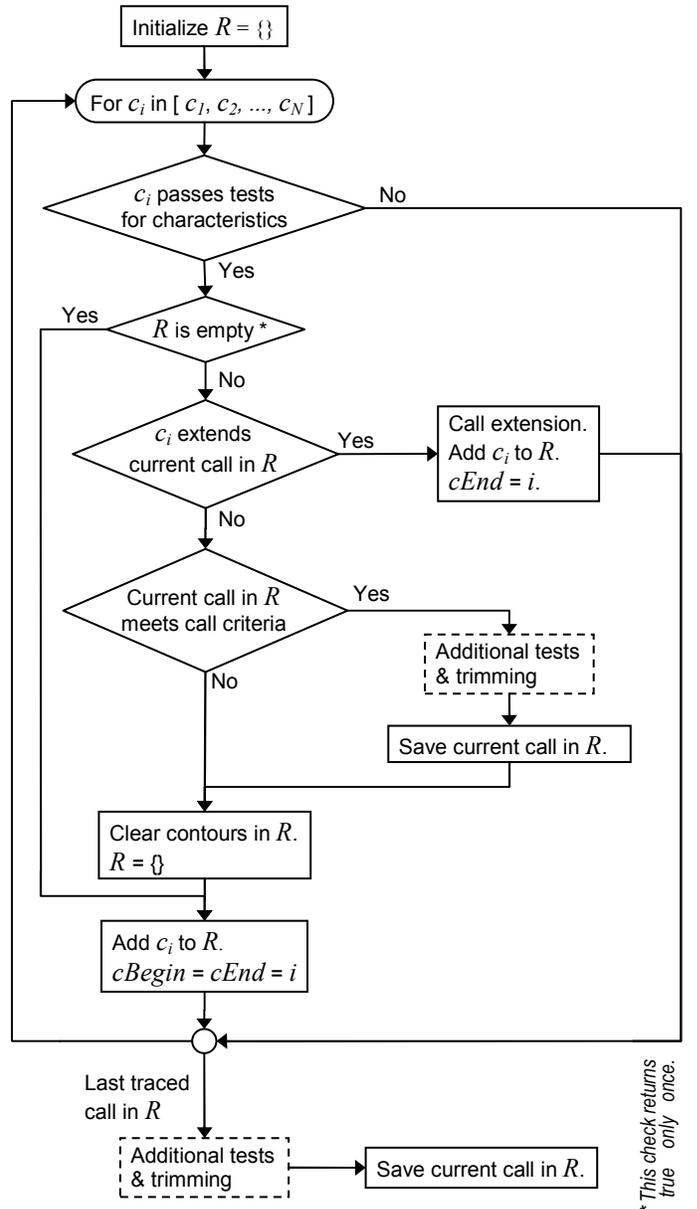


Fig 2. Overview of the classifiers. The step titled “Additional tests & trimming” is performed only in the D call classifier.

changes in the calling patterns. Let the times at which contour c_i began and ended be denoted by $tBegin_i$ and $tEnd_i$, respectively. Along with the call label, the classifier also reports fragmentation information which is one of the following:

- Complete call – when a single contour represents the entire call. This occurs when $(tEnd_{cEnd} - tBegin_{cBegin}) \geq$ minimum call length and $|R| = 1$.
- Composite call – when two or more successive contours, closely separated in time, together represent a single call. In this case, $(tEnd_{cEnd} - tBegin_{cBegin}) \geq$ minimum call length and $|R| > 1$.
- Partial call – when the contour or the set of successive contours in consideration has/have passed all tests for classification, but the total recognized length is small and hence is neither complete nor composite. Here, $(tEnd_{cEnd} - tBegin_{cBegin}) <$ minimum call length.

The minimum call lengths differ for B and D calls and the corresponding values used are provided in the appropriate sections.

In the case of D calls, occurrence of partial calls is a rarity and such detections are mostly due to sporadic noise and can be conveniently rejected. Hence the fragmentation information provided by the D call classifier only reports either complete or composite calls.

Since a B call has a more prominent characteristic of long call duration with a ~~xxx~~gentle downward sweep, the B call classifier is simpler in construction. On the other hand, since sporadic noise can also produce short and steeply down-sweeping contours that can be confused with D call contours, a D call classifier must consider more factors to produce accurate results. Additional tests are needed to address this issue. Also, overlapping B and D calls can occasionally cause the contour extractor to produce a single long contour that starts with a B call and ends in a D call or vice versa. As the B calls are several times longer than the D calls, an overlapping D call causes only a small error in the reported length of the B call and does not present significant problems for the B call detector. For the D call detector, failure to detect this situation would result in a large error. Hence a process to discard portions of the contours in this case is also needed. The “Additional testing & trimming” step in the D call classifier addresses these issues.

For each classifier, we first describe the call-specific tests for contour characteristics, then the tests to check inter-contour relationship and finally any call-

specific additional operations. Several constants are used during the process. They are represented with all uppercase names. Definitions of constants are not provided at the place of their first mention, in an attempt to keep the flow of the description of the classification process free from clutter caused by the frequent definitions of new constants. Instead, all the constants used for each classifier are summarized along with their values towards the end of the appropriate sections.

III. B CALL CLASSIFIER

The tests to determine if a contour c_i is a B call fragment are based upon statistics of c_i 's frequency. The unbiased mean (μ_i) and standard deviation (σ_i) are subjected to the following tests:

- Bounds test – check if the frequency interval within one standard deviation of the mean lies within the known frequency range of B calls.
Test: $(\mu_i - \sigma_i \geq L_FREQ_BOUND)$ and $(\mu_i + \sigma_i \leq U_FREQ_BOUND)$
- Duration test – check if the contour is long enough to be considered at least as a call fragment.
Test: $duration(c_i) \geq MIN_SEG_LEN$
- Dispersion test – check if the points in the contour are not too scattered around the mean frequency.
Test: $\sigma_i \leq SCATTER_THLD$

Tests to check if the contour sweeps downward are not made in order to nullify any errors in the estimates of the true contour obtained from the contour extractor. As an example, in Fig. 3 the extracted contour of the B call fragment (from ~4 s to ~11 s) does not have a downward sweep. Moreover, the test is not even necessary- partly due to the uniqueness in these characteristics of B call contours, and also since any contours caused by line noise never pass the bounds test. Some contours (like a D call fragment) may lie entirely within the lower and upper frequency bounds yet ~~are~~ too scattered to be a B call. The dispersion test handles this case.

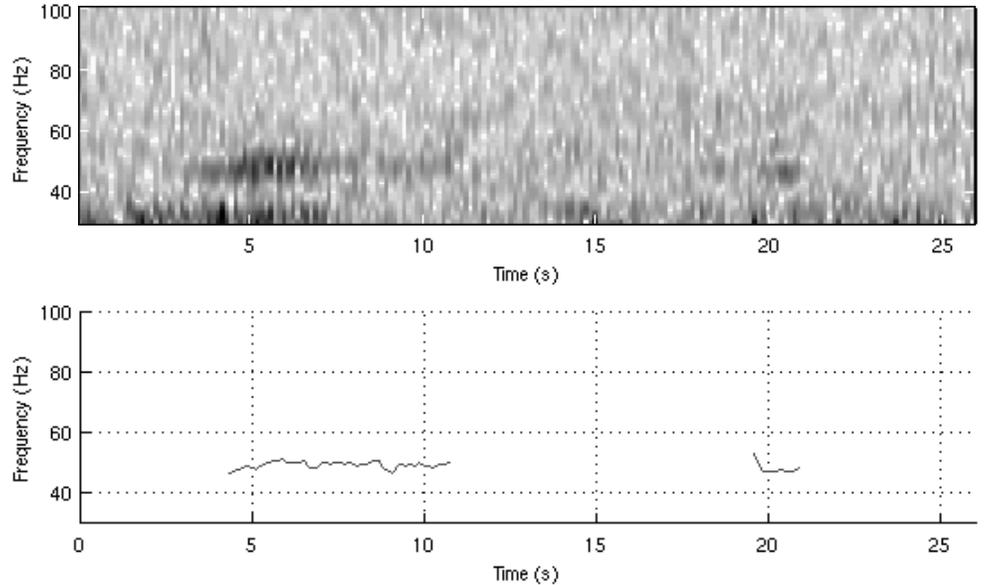


Fig 3. Spectrogram and the extracted contour of a sample segment containing a B call that shows a typical situation where a test for sweep direction could fail for B calls.

A contour that passes all of the above tests is accepted as a call fragment. If no calls were being traced, the current accepted contour (c_i) marks the beginning of tracing a new call. Contour c_i is added to R and $cBegin$ and $cEnd$ are set to i . Note that the contour that last passed the above tests is always c_{cEnd} .

If the contour c_i does not start the tracing of a new call, we check to see whether or not it extends the call currently being traced. Contour c_i is considered to extend a call if the separation between the contours c_{cEnd} and c_i (i.e., $tBegin_i - tEnd_{cEnd}$) is smaller than MAX_CALL_GAP , and if the following conditions hold true:

- $(tEnd_i - tBegin_{cBegin}) \leq MAX_CALL_LEN$: If the inclusion of c_i into the call being traced would not make its duration longer than MAX_CALL_LEN .
- $(\mu_{cEnd} + 0.5) > \mu_i$: If c_i is mostly at a lower frequency than c_{cEnd} . The 0.5 Hz correction factor is applied to compensate for any errors in the contour extractor estimates due to noise.

If the call being traced is extended by c_i , then c_i is added to R and $cEnd$ is set to i indicating the inclusion of c_i in the call which now spans from c_{cBegin} to the new c_{cEnd} .

If c_i does not extend the call being traced, then it marks the end of the call at c_{cEnd} . The traced call in R is reported, and then R is cleared in order to start tracing a new call. The start of a new call at the i^{th} contour is effected by adding c_i to R and setting $cBegin$ and $cEnd$ to i . Reported calls are marked as one of complete, composite or partial as described earlier.

After examining all contours, the last call being traced is considered to end at the last contour and again is marked as one of complete, composite or partial.

TABLE I
CONSTANTS USED IN THE B CALL CLASSIFIER

Constant	Meaning	Value
U_FREQ_BOUND	Frequency of a majority of points on the contour must be below this value (Hz)	60
L_FREQ_BOUND	Frequency of a majority of points on the contour must be above this value (Hz)	40
MIN_SEG_LEN	Contours shorter than this value are rejected (s)	1.5
MAX_CALL_GAP	Maximum separation between successive contours for considering them together as fragments of a single call (s)	8
MAX_CALL_LEN	A single contour or a group of successive contours is/are considered to represent one complete call when their total length exceeds this value (s)	22
MIN_CALL_LEN	Minimum call length. A single contour is considered to represent a complete call if its duration is longer than this value (s)	15
SCATTER_THLD	Threshold for the dispersion of contour points around the contour's mean frequency	4.5

IV. D CALL CLASSIFIER

In the D call classifier, two sets of rules are utilized for the test of characteristics. The first set leniently accepts contours

based on simple tests. Once a candidate D call has been completely detected, it is subjected to two additional operations. A truncation or trimming of the call may occur to eliminate any overlapping of B and D calls as previously described. After possible trimming, the slope is rechecked and then the remaining set of contours is subjected to additional tests for characteristics that are stricter. This second set of steps is shown as “Additional testing & trimming” in the flowchart of Fig. 2, and the flow of these operations is described in Fig. 4. Please note that in all further discussion, our definition of slope differs from its conventional definition in that we compute slope as $(y_1 - y_2)/(x_2 - x_1)$ instead of $(y_2 - y_1)/(x_2 - x_1)$, where y_1 and y_2 are the frequencies of the consecutive points at times x_1 and x_2 , on the contour. This alteration in definition is made in order to simplify discussion by helping the reader easily perceive comparisons of the magnitudes of downward slopes, as the change produces higher slope magnitude for steeper downward sweeps.

Let the maximum and minimum frequencies in c_i be denoted as $fmax_i$ and $fmin_i$, respectively. The contour characteristic test simply checks to see if $fmax_i$ occurred earlier than $fmin_i$ on the current contour. If no calls were being traced, the current contour (c_i) marks the beginning of tracing a new call. Contour c_i is added to R and $cBegin$ and $cEnd$ are set to i .

c_i is considered to extend a call if the separation between c_{cEnd} and c_i (i.e., $tBegin_i - tEnd_{cEnd}$) is smaller than MAX_CALL_GAP , and if the two conditions mentioned below hold true:

- The ending frequency of c_{cEnd} is not less than the starting frequency of c_i .
- The magnitude of the slope of the straight line joining the nearest ends of c_{cEnd} and c_i exceeds $SLOPE_L_BND$.

When the above is true, c_i is added to R and $cEnd$ is set to i indicating the inclusion of the i^{th} contour in the call which now spans from c_{cBegin} to the new c_{cEnd} .

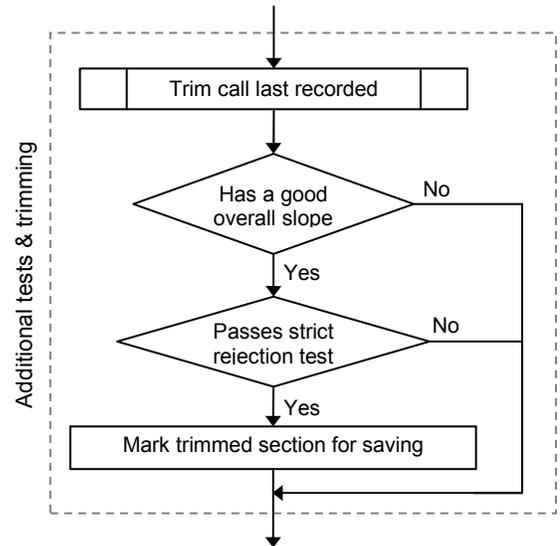


Fig 4. Description of the “Additional tests & trimming” step in the flowchart of Fig. 2, for the D call classifier.

Otherwise, if c_i were well separated from c_{cEnd} (i.e., $tBegin_i - tEnd_{cEnd} > MAX_CALL_GAP$) or if c_i started its sweep at a significantly high frequency ($fmax_i \geq U_FREQ_L_BND$), then it marks the end of the call being traced at c_{cEnd} . Before marking the start of a new call at the i^{th} contour by updating R , $cBegin$ and $cEnd$, contour trimming and additional tests are performed to confirm that the call contained in R is indeed a D call. The trimming process (described later in III A) diminishes the overall duration in the identified set of contours and sometimes even discards complete fragments. Hence it is necessary to test if whatever is retained after trimming still corresponds to the general D call structure. The trimmed set is examined to see if it has a steeply sweeping overall downward slope and then is subjected to the strict rejection test (processes described in III B).

Individual contours or a group of successive contours, corresponding to a single D call, that pass this rejection test are the ones accepted by the classifier. These are marked as complete or composite as described earlier. Finally, R is cleared in order to start tracing a new call. Then, c_i is added to R , and $cBegin$ and $cEnd$ are updated to mark the beginning of a new call at c_i .

After examining all contours, the last call being traced is considered to end at the last contour and again is marked as one of complete or composite as described before.

TABLE II
CONSTANTS USED IN THE D CALL CLASSIFIER

Constant	Meaning	Value
U_FREQ_L_BND	The lowest frequency above which the downward sweeps must start (Hz)	50
L_FREQ_U_BND	The highest frequency below which the downward sweeps must end (Hz)	50
MIN_CTR_LEN_T	Minimum duration that a contour (or a group of successive contours) must cover in order to be accepted (s)	0.5
MIN_CTR_LEN_F	Minimum frequency sweep length that a contour (or a group of successive contours) must cover in order to be accepted (Hz)	8
MAX_CALL_GAP	Maximum separation between successive contours for considering them together as fragments of a single call (s)	1
SLOPE_L_BND	Lower bound on the absolute value of the slope of the downward sweep (Hz/s)	10

A. Contour Trimming

Starting from the outer edges of the sets of points on the contours in R and proceeding inwards, the outer points are iteratively removed if the slopes with their neighboring points fall below 90% of $SLOPE_L_BND$, and the process continues until the total duration of the remaining points is greater than $MIN_CTR_LEN_T$. A threshold smaller than $SLOPE_L_BND$ was chosen here in order to avoid excessive trimming as the set of contours in consideration now has a higher probability of containing a D call. If, in the process, a contour is completely removed, the examining of slope

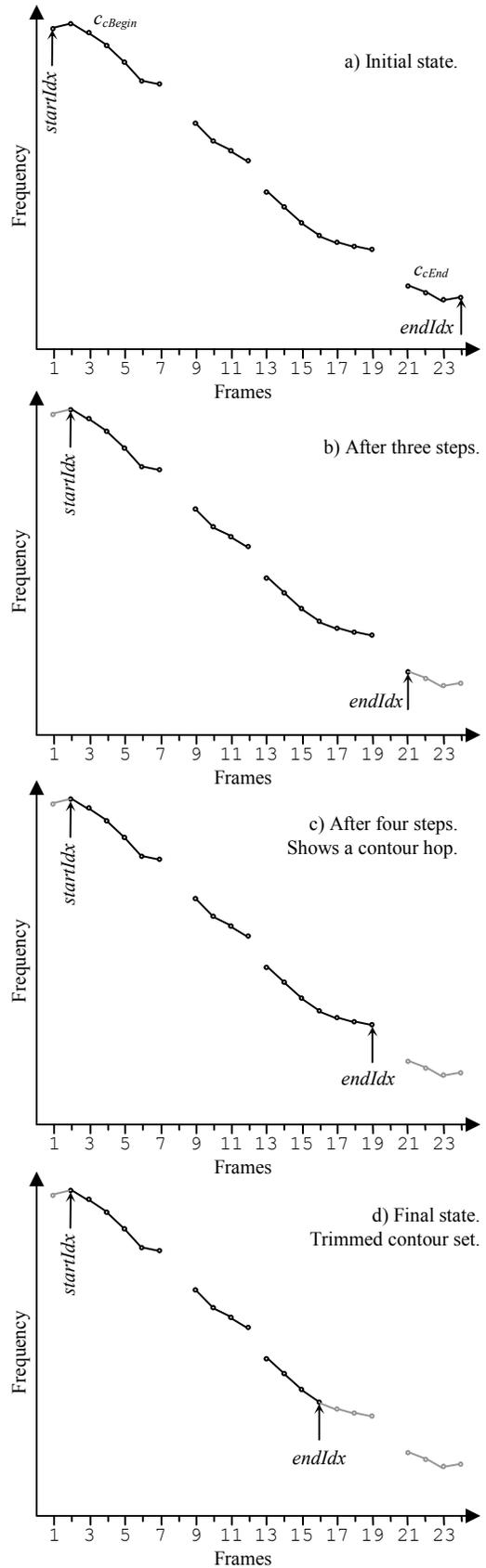


Fig 5. An example of the trimming process in the D call classifier, on a set of synthetic contours typical of a D call.

continues with points on the subsequent inner contour. When examination of slopes ceases in both directions, the portion of the contour (or the set of contours) remaining between the points at which the examinations stopped, is the trimmed result.

An example of the trimming process, on synthetic data, is shown in Fig. 5. Snapshots from various stages are provided. Fig. 5a shows the initial state before any processing has begun. The first and last contours in the considered set are identified as c_{cBegin} and c_{cEnd} in the figure. $startIdx$ and $endIdx$ identify the points at which slopes with their corresponding neighboring points are examined. Fig. 5b shows the state after three steps. Notice that progress in the forward direction at the left end had stopped after the first step at the frame 2. The progress in the backward direction at the right end has reached the end of a contour. The grayed portions on the outer extremes of the contours represent the cropped sections. In the following step, a hop is made to the next contour in the inward direction as shown in Fig. 5c. The progress in the backward direction ceases at frame 16, due to presence of steeper sweeps beyond that point, as shown in Fig. 5d. The portion remaining, at this stage, between $startIdx$ and $endIdx$ is the result of the trimming operation.

B. Overall Slope and Strict Rejection Test

The sub-routine for checking if the trimmed contour set has a good overall slope involves checking if, on an average, all the contours in the set have a good steeply sweeping downward slope. It is a straightforward test examining if the average of the slopes across consecutive pairs of points on all contours in the trimmed set exceeds 90% of $SLOPE_L_BND$. Again, a threshold slightly smaller than $SLOPE_L_BND$ is chosen for the same reason as in the trimming sub-routine. This test is necessary in those cases when

- the contours containing $startIdx$ and $endIdx$ have good slopes but the contours between them do not, and
- the contour containing $startIdx$ (or $endIdx$) has good slope only around the end point and it begins to change behavior as it proceeds inwards.

If the set of contours passes the above test, a final composite test is performed which confirms whether the call being examined is indeed a blue whale D call. The composite test (strict rejection test) includes checks to see if

- the call sweep terminates below $L_FREQ_U_BND$,
- the sweep length is at least $MIN_CTR_LEN_F$ on the frequency axis, and
- the sweep duration is at least $MIN_CTR_LEN_T$.

V. PERFORMANCE AND RESULTS

The contour extraction algorithm and both the classifiers have been implemented in Matlab. The contour extractor implementation runs at a real-time factor of over 40x on an AMD Athlon 1.19 GHz machine when processing 16 bit data

sampled at 1 kHz. The computational complexity of both B and D call classifiers is $\Theta(nm)$, where n is the number of contours in the considered segment and m is the average number of points in each contour. This means that the classifiers also have a very short turnaround time. Such high processing speeds are very favorable for the original intended application of recognizing B and D calls from extremely large recordings. Even with a non-optimized implementation, such high performance capability makes the system also suitable for use as a component in an on-site acoustic monitoring program.

Manual transcription of calls in large datasets is impractical. To provide a quality metric, 60 minute segments when blue whales were present and vocalizing were randomly selected from four different field seasons, between 2000 and 2004, and were manually transcribed. These are used as a ground truth table for evaluating performance. The classification results of the system for the chosen data are summarized in Table III.

Two significant metrics are used to quantify the performance of the individual functional entities (contour extractor, call classifier) in the system as well as the overall system itself. They are *recall* and *precision*. The units of interest at different levels in the system are the contours and the calls. The *recall* of a functional entity can be described as its accuracy, measured as the ratio of the number units of interest correctly identified to the actual number of such units present in the data considered. *Precision* can be defined as the ability of the entity to avoid reporting units falsely where they do not actually exist, and is measured as the ratio of correct identifications to the total number of units reported (including false identifications). The values of these metrics are represented as a percentage, and are computed using the average of the corresponding values from the four seasons.

When the units of interest considered are contours, the two metrics characterize the performance of the contour extractor. These values are listed under the column titled ‘‘Contour Extractor’’ at the bottom of Table III. The performance of a classifier in isolation identifies its ability to correctly classify calls (of the corresponding type) using the contour data obtained from the contour extractor. This is evaluated considering only valid contours of the corresponding call type from the output of the contour extractor and the number of such contours classified correctly. These results are listed under the column titled ‘‘Call Classifier’’ at the bottom of Table III. The values listed in the column titled ‘‘Overall System’’ provide a composite evaluation, as they depend on both the strength of the contour extractor and the accuracy of the classifier. They are computed as the fraction of the number of calls (of a type) present in the recordings that are correctly reported.

Using the results of performance evaluation, the various system parameters were extensively tuned in order to obtain optimal performance for the data considered. The values of the constants listed in Tables I and II were arrived at after successive cycles of performance evaluation and subsequent

TABLE III
PERFORMANCE EVALUATION USING DEVELOPMENT DATA

Year	2001		2002		2003		2004	
Contours in recordings	62		71		104		97	
Contours extracted	59		70		96		86	
Non-call contours wrongly extracted	3		9		2		4	
Call type	B	D	B	D	B	D	B	D
Calls present in recordings	22	28	51	13	89	12	36	36
Call contours extracted	22	26	46	10	82	10	33	29
Calls correctly identified	22	24	45	9	82	8	33	26
Falsely identified calls	1	0	1	1	0	0	2	0

Performance

	Contour Extractor	Call Classifier		Overall System	
		B Classifier	D Classifier	B Detector	D Detector
Recall	93.1%	99.5%	89.3%	91.9%	75.3%
Precision	94.5%	97.8%	98.5%	97.8%	98.5%

re-tuning of the parameters. The results reported in Table III were obtained using these final settings. In order to test if the final tuned system performed similarly on recordings with conditions different from those in the development set, it was re-evaluated using a different set of randomly selected 60 minute segments from each of the four different field seasons. Care was taken to make sure that these new segments were well separated (in time) from the segments from the same season that were used as development data. The performance of the system corresponding to these test data is summarized in Table IV.

As mentioned earlier, a D call is rather short and hence its fragments are even shorter. In several cases these fragments are shorter than a threshold and so are rejected in the first place, by the contour extractor. In certain other cases, the contour profiles of these short fragments closely match the profile of sporadic noise and hence these also get rejected, this time by the classifier. This results in the slightly lower recall values seen for a D call classifier/detector in comparison to their B call counterparts.

Comparing Tables III and IV, one can easily note that the system has displayed a good degree of stability across different sets of recordings. Extreme noisy conditions in the segment of recording considered in the development set, corresponding to the recording season of 2004, has caused the contour extractor to miss several D call contours. The apparent increase in performance seen with the D call detector for test data is due to this. The similarities in the recall and precision values for the D call classifier across the development and the

TABLE IV
PERFORMANCE EVALUATION USING TEST DATA

Year	2001		2002		2003		2004	
Contours in recordings	71		90		75		171	
Contours extracted	70		82		74		168	
Non-call contours wrongly extracted	2		1		1		2	
Call type	B	D	B	D	B	D	B	D
Calls present in recordings	41	15	60	22	45	20	50	119
Call contours extracted	40	15	56	20	44	20	50	117
Calls correctly identified	39	13	50	17	43	18	47	103
Falsely identified calls	1	2	0	0	2	1	6	0

Performance

	Contour Extractor	Call Classifier		Overall System	
		B Classifier	D Classifier	B Detector	D Detector
Recall	96.8%	94.2%	87.8%	91.3%	85.8%
Precision	98.5%	95.2%	98.1%	95.2%	98.1%

test sets, however, suggests that the D call classifier has been consistent. Overall, the other components have performed consistently.

ACKNOWLEDGMENT

Funding for data collection was provided by Frank Stone and Ernie Young from CNO-N45, Robert Holst from the Navy's Strategic Environmental Research and Development Program and Bob Gisiner and Ellen Livingston of ONR. Data was collected with the help of the SIO Whale Acoustic Lab, in particular, Sean Wiggins, Allan Sauter, Chris Garsha, and Graydon Armsworthy.

REFERENCES

- [1] E. M. Oleson, S. M. Wiggins, and J. A. Hildebrand, "Temporal separation of blue whale call types on a southern California feeding ground," *Animal Behaviour*, vol. 74, pp. 881-894, 2007.
- [2] D. K. Mellinger and C. W. Clark, "Recognizing transient low-frequency whale sounds by spectrogram correlation," *J. Acoust. Soc. Am.*, vol. 107, pp. 3518-3529, 2000.
- [3] D. Talkin, "A Robust Algorithm for Pitch Tracking (RAPT)," in *Speech Coding and Synthesis*, W. B. Kleijn and K. K. Paliwal, Eds. Amsterdam, The Netherlands: Elsevier, 1995, pp. 495-518.
- [4] C. L. Berchok, D. L. Bradley, and T. B. Gabrielson, "St. Lawrence blue whale vocalizations revisited: Characterization of calls detected from 1998 to 2001," *J. Acoust. Soc. Am.*, vol. 120, pp. 2340-2354, 2006.
- [5] Madhusudhana, S.K.; Oleson, E.M.; Soldevilla, M.S.; Roch, M.A.; Hildebrand, J.A., "Frequency based Algorithm for Robust Contour Extraction of Blue Whale B and D calls," *OCEANS 2008 - MTS/IEEE Kobe Techno-Ocean*, vol., no., pp.1-8, 2008.
- [6] E. M. Oleson, "Calling behavior of blue and fin whales off California," PhD dissertation, Scripps Institution of Oceanography, UC San Diego, California, USA, 2005.