THREE-DIMENSIONAL SINGLE-HYDROPHONE TRACKING OF A SPERM WHALE DEMONSTRATED USING WORKSHOP DATA FROM THE BAHAMAS

Christopher O. Tiemann
Applied Research Laboratories, University of Texas at Austin
P.O. Box 8029, Austin, TX 78713

ABSTRACT

A passive acoustic localization method for tracking the movement of a clicking sperm whale in three-dimensions using data from just one hydrophone is demonstrated using data made available for the 3rd International Workshop on Detection and Classification of Marine Mammals. One recording contains sperm whale clicks recorded on a bottom-mounted hydrophone on a steep slope of the Navy's AUTEC test range. When the direct-path acoustic ray arrivals from several clicks are time-aligned, persistent associated multipath arrivals of reflected ray paths can be identified for each click event and used for localization. Although the use of multipath arrival information is a standard procedure for range-depth tracking, a three-dimensional estimate of whale position can be obtained from the same multipath information with knowledge of an azimuthally-dependent environment relative to the receiver. In this case, azimuthal distinction arises from varied bathymetry. Multipath arrival patterns are matched to unique range-, depth-, and azimuth-dependent modeled arrival patterns to make an estimate of whale location. A three-dimensional whale track in range, depth, and bearing from the fixed hydrophone is presented.

SOMMAIRE

On démontre dans cet article une méthode de localisation acoustique passive pour suivre la trace d'un cachalot cliquant dans les trois dimensions utilisant les données d'un seul hydrophone, en utilisant les données disponibles à partir du Troisième Atelier Internationale sur la Découverte et la Classification des Mammifères Marins. On a enregistré les bruits secs du cachalot enregistré par un hydrophone monté au fond sur une pente raide du champ d'essai d'AUTEC de la Marine Américaine. Quand les arrivées des rayons acoustiques des plusieurs cliques avec des trajectoires directes sont alignés par intervalles, des arrivées multi-trajectoires associées et persistantes des trajectoires de rayons reflétés peuvent être identifiées pour chaque événement d'une clique et peuvent être utilisées pour la localisation. Bien que l'utilisation de l'information d'arrivée multi-trajectoire soit une procédure normale pour suivre la trace portée-profondeur, une estimation en trois dimensions de la position des cachalots peut être obtienne en utilisant les mêmes informations multi-trajectoires si l'on connait l'environnement azimuthalemment-dépendent relatif au récepteur. Dans ce cas, la distinction azimutale dérive de la bathymétrie variée. Les schémas des arrivées multi-trajectoires sont attribués aux schémas uniques qui dépendent de la portée, de la profondeur, et de l'azimut pour faire une estimation d'emplacement du cachalot. Une trajectoire en trois dimensions qui montre la portée, la profondeur, et le rapport du hydrophone fixe est donnée.

1. INTRODUCTION

Passive acoustic methods for monitoring marine mammal activity have been used for many years in censusing and behavioral studies, often in conjunction with visual surveys, because of the advantages they offer: they are unobtrusive and continue to work at times when animals are not visible (swimming underwater, nighttime, etc.) [1-3] Methods for not just detecting but also tracking the movement of animals underwater through analysis of their recorded vocalizations have advanced since early work that used geometric hyperbolic fixing techniques. [4-10] Techniques that exploit acoustic propagation models, multipath arrival information, or both can now provide alternate and possibly more accurate localization estimates. [11-15] Two-dimensional (2D) location solutions (range and depth) can be achieved using data from as few as one hydrophone, but a full three-dimensional (3D) estimate (range, depth, and unique bearing relative to a sensor) has previously required the use of multiple receivers. [12,16-20] A technique that uses range estimates from a single hydrophone to make a hydrophone-relative 3D track does not provide an absolute measurement of azimuth.[20]

In 2006, Tiemann et al. [21] demonstrated a model-based technique for using acoustic data from just one hydrophone to make three-dimensional estimates of sperm whale locations in the Gulf of Alaska. The method exploits multipath arrival information from recorded sperm whale clicks, yet it does not require specific ray path identification (i.e., direct-path, surface-reflected). The technique can not only account for waveguide propagation physics (ray

interaction with the sea surface and sea floor in particular), but in fact relies upon reflections to estimate bearing to the source (whale). While the single-hydrophone localization technique was demonstrated successfully in its first application, a dataset made available for the 3rd International Workshop on Detection and Classification of Marine Mammals provided an opportunity to further exercise the localization algorithm in another much deeper environment. This paper describes the application and results of that localization attempt.

Among the data that the Naval Undersea Warfare Center (NUWC) provided for the workshop was a 10-minute recording ("test data #9") containing numerous sperm whale vocalizations from what was assumed to be one animal. The data are from hydrophone #16 of the Navy's Atlantic Undersea Test and Evaluation Center (AUTEC) in the Bahamas, positioned about 5 m off the sea floor at 1386 m depth, from March 3, 2006, at 09:48 local time with a 96 kHz sample rate. NUWC also provided bathymetry information from a multibeam survey of the range; Figure 1 shows the topography around the receiver that provided data for this demonstration.

2. METHODOLOGY

Sperm whale vocalizations appear in the data as brief (~10 ms) broadband clicks with an inter-click interval of about 1 sec. The collection of all the multipath echoes from a single click event is an arrival pattern, and one step in the localization process is to identify these patterns in the data. Additionally, an acoustic propagation model is used to predict the arrival patterns expected at the receiver from hypothesized impulsive sources at many ranges, depths, and bearings around the receiver; these modeled arrival patterns are called the replica. After comparing the measured arrival patterns with the replica, the hypothesized source position for which they agree best is the best estimate of whale location. Repeating the localization for each click event creates a track of animal motion.

In cases where the environment is unique along radials at every bearing around the receiver, a source at every range/depth/bearing bin will have a unique multipath arrival pattern "fingerprint." It is azimuthal dependence in the environment that allows for bearing discrimination when using just one receiver, and the rough terrain around the hydrophone, as shown in Figure 1, ensures that the environment (bathymetry slice) looks different along radials in every direction.

2.1 Replica Generation

One input needed for the localization process is predictions of travel times for all ray paths between a source and the receiver for a suite of hypothesized sources on a grid of several ranges, depths, and bearings around the receiver. Specifically, the algorithm requires relative arrival times for all the multipath arrivals, not absolute travel times, so the time elapsed since the arrival of the direct ray path is calculated and saved for later use. The Gaussian beam acoustic propagation model BELLHOP [22] provides travel time predictions for the eigenrays assuming a grid of sources spaced 10 m in range out to 5 km and 5 m in depth to 2 km. The model uses a different bathymetry profile for radials spaced every 5° in azimuth around the receiver; the bathymetry data was that provided by NUWC as shown in Figure 1. The model assumes a source frequency of 5 kHz, a range-independent downward refracting soundspeed profile taken from the Levitus historical database, and geoacoustic properties of fine-grained sediments. [23]

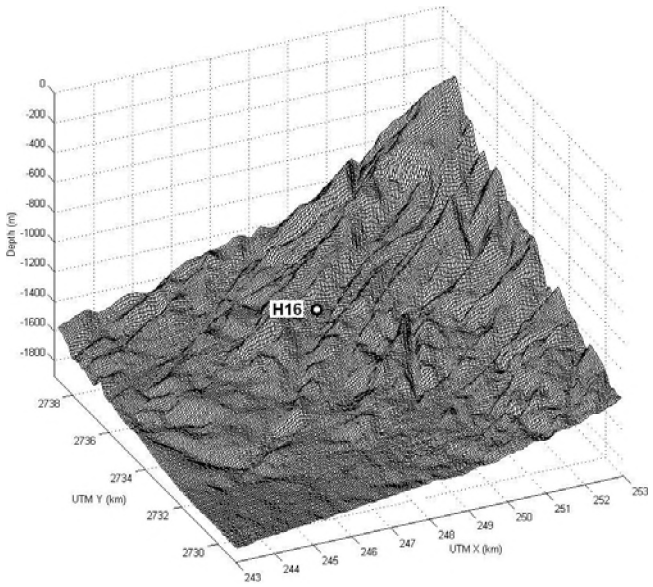

Figure 1. Bathymetry around AUTEC hydrophone #16. Coordinates are for UTM zone 18.

2.2 Arrival Pattern Extraction

A tool for automating the identification and extraction of arrival pattern information for click events in acoustic data is described in detail in Tiemann et al. [21], and it was used again here to extract information for 383 click events in the workshop data. A summary of its use follows.

Broadband sperm whale clicks are readily apparent when viewed as spectrograms like that of Figure 2a which shows four seconds of workshop data at the start of a click train; this spectrogram was made using 256-point fast Fourier transforms with 50% window overlap on data downsampled to 44 kHz (6 ms of data for each FFT). To make fainter click arrivals more apparent over background noise, spectrograms are summed over the frequency bins from 3 kHz to 22 kHz within each time bin, as shown in Figure 2b. Each peak in the spectral sum time series represents an arrival from either a direct or reflected ray path, with the direct paths typically having greater amplitudes than their associated echoes. In this example, the arrivals for five direct paths are spaced about 0.7 sec apart. Note that the inter-click interval between click events

was often less than the time separation between the two loudest arrivals of a given click event, as in this example. (The weaker reflected-path arrival at 393.9 sec is associated with the direct path arrival at 393.0 sec, not the one at 393.7 sec.)

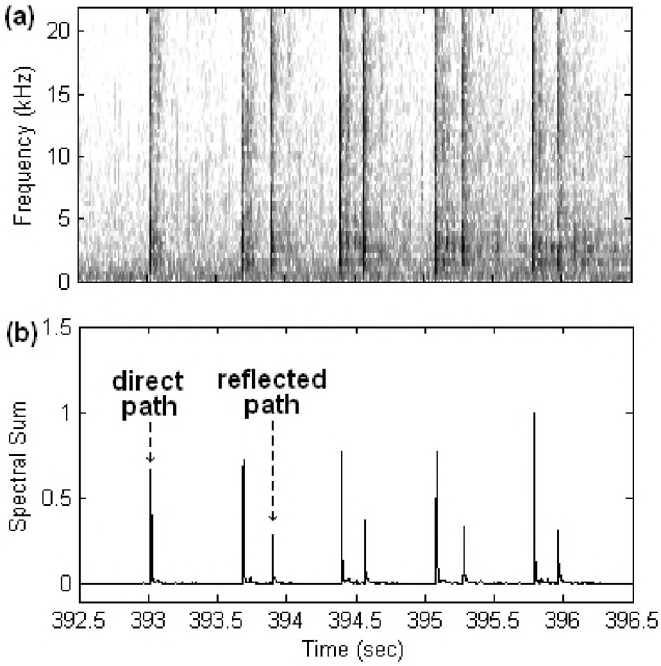

Figure 2. (a) Spectrogram of acoustic data from AUTEC; broadband sperm whale clicks appear as vertical stripes. (b) Spectrogram summed over frequency bins; broadband clicks appear as peaks.

The automated pattern matching tool assists in identifying the direct-path (earliest) arrivals for several click events in a click train. In order to assist in recognition of possibly faint multipath arrivals, a display tool time-aligns windows from the spectral sum time series which begin at the direct-path arrival. These aligned spectral sum excerpts can then be viewed as a two-dimensional color surface like that of Figure 3 where each horizontal slice conveys the relative amplitude and arrival time information of all arrivals occurring within 1.2 sec of a direct path arrival at relative time 0 sec. Note that the absolute time axis of Figure 3 indicates the time since the beginning of the data set; the relative time axis is the time elapsed since the direct path arrival of a given click event.

Persistent peaks in these surfaces that are time-aligned over every horizontal slice represent multipath arrivals that will be compared to the modeled arrival patterns, and another tool helps extract the arrival pattern information from these surfaces. Identifying the arrivals as either from surface-reflected paths, bottom-reflected paths, etc. is not required for the automated localization to follow, but such an interpretation of Figure 3 is provided as an example. The order of the arrivals in this example is direct path, bottom-bounce path, surface-bounce path, and bottom-surface-bounce path as labeled on Figure 3. Note that the isolated peaks on these surfaces that do not align in time with other peaks can be ignored during arrival pattern extraction as they are not associated with a given click event under consideration. For example, the isolated peaks around relative time 0.7 seconds are direct-path arrivals of the *next* click event; the ~0.7 sec inter-click interval is shorter than the length of the spectral sum excerpts used to make this figure.

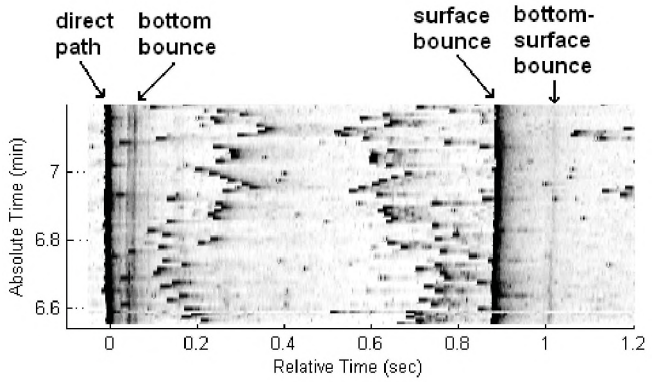

Figure 3. Time-aligned spectral sum excerpts, each starting with a direct-path arrival at relative time 0 sec, represented as a 2D surface.

A benefit of making the time-aligned spectral sum surfaces is that they allow the eye to integrate over multiple click events to recognize faint broadband arrivals that would not be obvious in any single spectral sum excerpt. For example, Figure 4 shows one of these surfaces made using data from most of the 10-minute workshop recording, 383 click events in all. Color scales were adjusted in this figure in attempts to make the faintest persistent arrivals visible, but doing so increases the visible clutter from non-aligned arrivals.

The same four persistent arrivals identified in Figure 3 are present in Figure 4 and are labeled there as well, but note how the relative spacing between the arrivals evolves with time. This is expected as the source is changing position relative to the fixed receiver, and it will have a new arrival pattern at every new location. The lengthened relative time axis of Figure 4 (now 3.5 sec) also shows two more arrivals around 2.5 sec relative time. The acoustic model predicts these should be present then and identifies them as a surface-bottom-surface reflected path and a bottom-surface-bottom-surface reflected path. To put that 2.5 sec delay into context, in the time between the direct-path arrival and its late echo, *three more* click events occurred, yet that late, faint arrival can still be associated with the correct direct-path arrival.

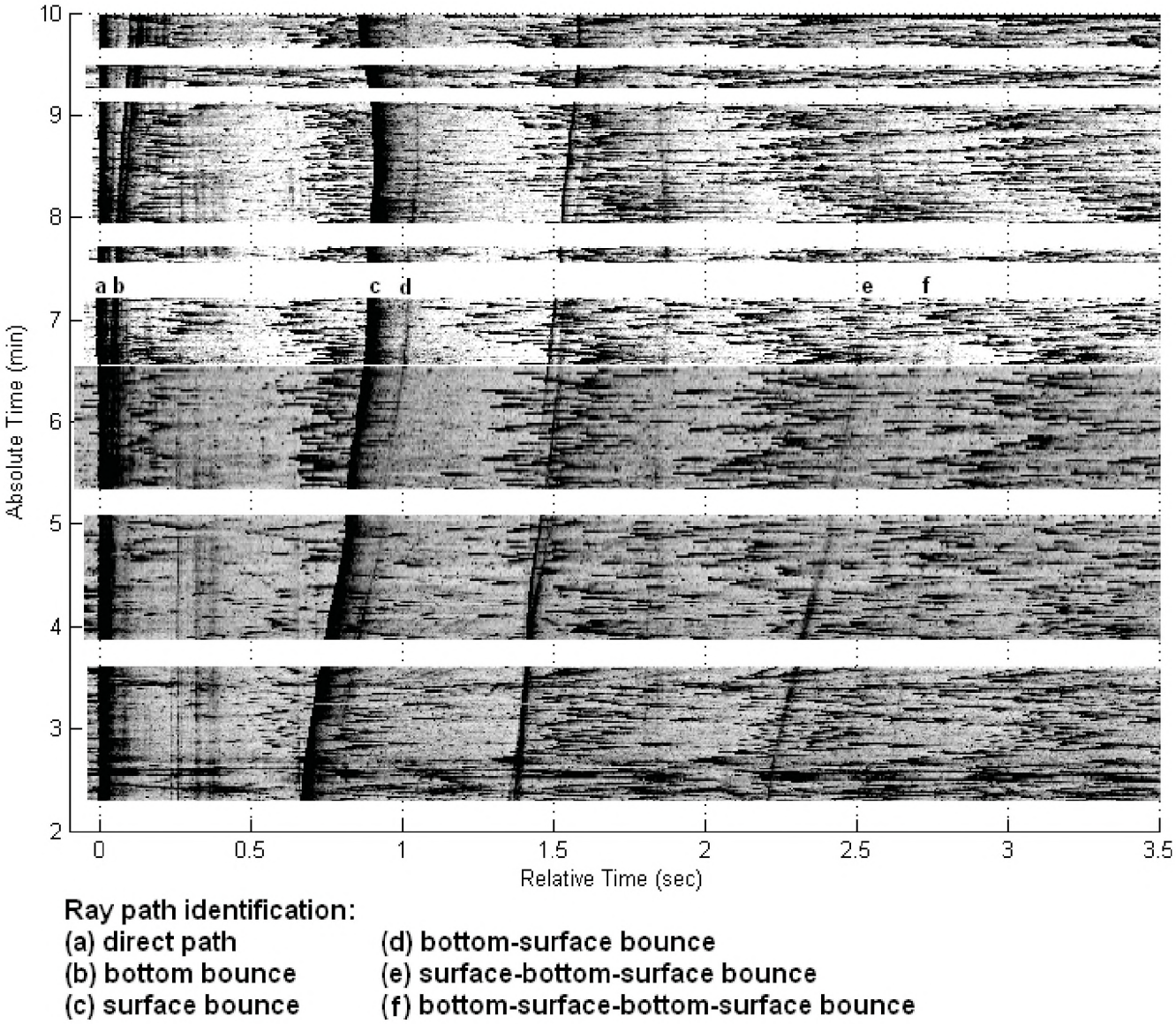

Figure 4. Time-aligned spectral sum excerpts, each starting with a direct-path arrival at relative time 0 sec, represented as a 2D surface. Persistent arrivals with identified ray path geometries are labeled.

Identification of the ray path geometries, though not required, can be challenging when the data indicates there are more ray arrivals than a model would predict, as was the case in this recording. For example, notice the arrivals present at about 1.5 sec relative time. These arrivals are curious not only because they split, merge, and even disappear but also because there is no match anywhere close to them in the modeled arrival patterns. One hypothesis to explain these arrivals is that they are from reflected paths outside of the vertical plane of propagation connecting the source and receiver. The acoustic propagation model used is limited to 2D problems (range/depth slices), yet the map of Figure 1 shows a rough 3D terrain with several ridges around the receiver which may allow for reflections of ray paths outside the vertical plane of modeling.

For another example of unusual arrivals, notice several that arrive at about 0.3 sec relative time, shortly after the direct path arrival. These arrivals are persistent throughout the entire record, and they are also locked in time to the arrival of the direct path, i.e. their relative arrival time never changes. Every other arrival shifts in time as the source moves except these. One hypothesis to explain these is that they are due to sound reflecting off the receiver housing, then off some nearby fixed terrain feature and back to the receiver. That scenario would cause stripes of constant relative arrival time like those seen in the figure regardless of the source location.

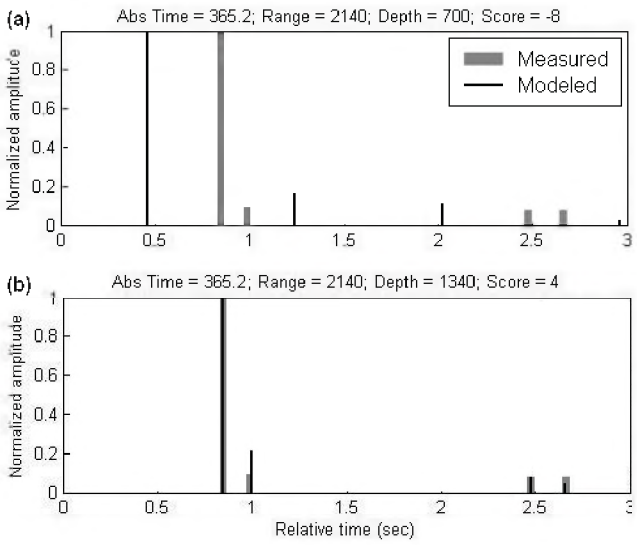

Figure 5. (a) Arrivals from a modeled source at 2140 m range, 700 m depth overlaid on measured arrivals from time 365.2 sec. No arrivals overlap, so it is given a low overlap score. (b) The same measured arrivals overlay modeled arrivals from 2140 m range, 1340 m depth. All measured and modeled arrivals overlap.

2.3 Ambiguity Surface Construction

The source location for each click event is estimated through the construction of bearing-dependent ambiguity surfaces. Each surface graphically conveys the likelihood that a whale was at a given range/depth bin when it vocalized. These surfaces have the same resolution as the replica, and a scoring mechanism assigns a score to each hypothesized source location based on how closely a measured arrival pattern matches the modeled arrival pattern for that location. The scoring technique and ambiguity surface construction are described in detail in Tiemann et al. [21]; a summary of its use here follows.

The score for every candidate source position is calculated by first counting the number of measured arrivals that have the same relative arrival times as those in the replica for that source position. From this score is subtracted the number of arrivals in both the data and replica that do not have a match in relative arrival time. A tolerance of 10 ms is used in defining a match in relative arrival times as that was a typical duration for a recorded click. To illustrate the scoring process, Figures 5a and 5b show the relative arrival time and amplitude information for a measured arrival pattern overlaid by modeled arrival patterns for two candidate sources at the same range but different depths. The first example has no overlapping arrivals among the 8 considered for a total score of -8; the second example shows all 4 arrivals from both the data and replica matching to result in a score of +4. Note that for early parts of the data set the direct-path and bottom-bounce arrivals were difficult to distinguish. Therefore, the bottom bounce path was not considered during the scoring process.

Scores are calculated for all candidate source positions and presented on an ambiguity surface like that of Figure 6, one surface for each bearing radial in the replica. Peaks on these surfaces indicate likely source positions, and the global maximum among all ranges, depths, and bearings is declared the best estimate of source location. Repeating the scoring process for every click event results in a track of whale motion.

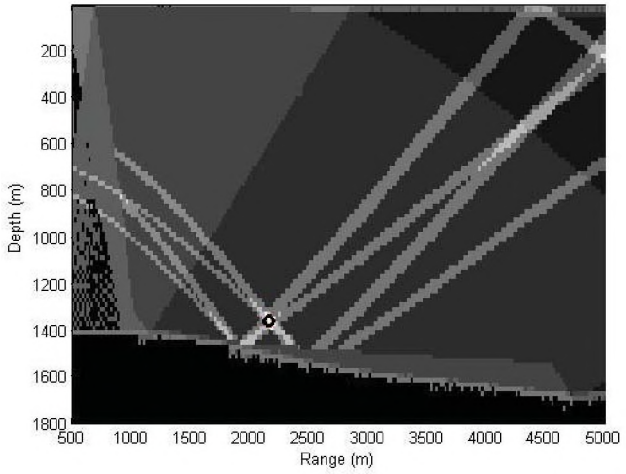

Figure 6. An ambiguity surface showing overlap scores on a vertical range/depth slice along bearing 200° from the receiver. The peak at 2140 m range, 1340 m depth is marked and indicates the best estimate of source location for the click event at time 365.2 sec.

3. RESULTS

The range, depth, and bearing estimates resulting from the localization process described above are presented in Figure 7. The whale track originates 3 km to the southwest of the receiver, ending within 2 km of the receiver and shallower. No independent ground truth data of whale motion was available for this workshop data, but the whale's average speed over this track was 2.9 m/s, comparable to other acoustically derived sperm whale swim speeds. [18,19] Figure 8 puts this track in context of the AUTEC environment, overlaying estimated positions on a plan view of the bathymetry contours. One can see that the whale's change in depth tracks the shallowing terrain. It is the 5° binning in the replica that causes a disjointed track when bearing changes; the whale did not suspend clicking.

4. DISCUSSION AND CONCLUSION

The ability to passively monitor the movement of marine mammals underwater benefits behavioral studies of such,

but the multi-sensor arrays typically required to make 3D tracks of animal motion increase the cost and complexity of those studies. An economical single-hydrophone solution for doing the same should be of benefit to the bioacoustics community. A previous effort demonstrated how to produce such a 3D estimate of sperm whale location by exploiting multipath arrival information from a single sensor, and this work used data provided by the 3rd International Marine Mammal Workshop to demonstrate that the same technique seems viable in deep water environments as well, at least in locations where there is some environmental variation in azimuth.

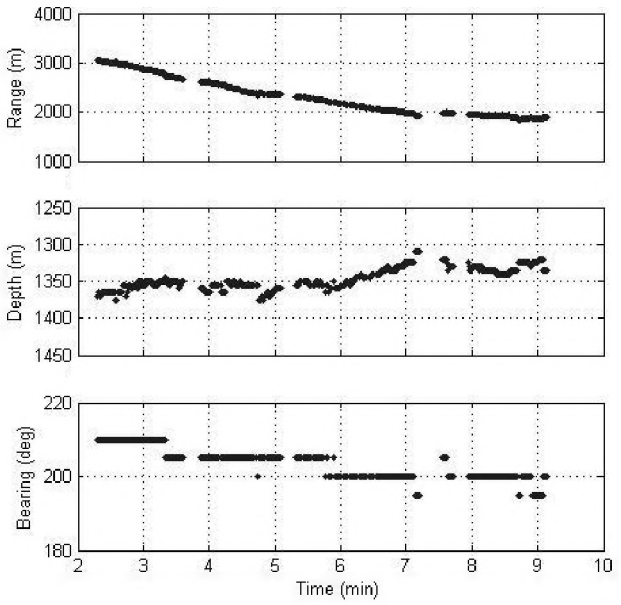

Figure 7. Range, depth, and bearing estimates relative to hydrophone #16 of the AUTEC range for a clicking sperm whale.

Azimuthal dependence in the environment is necessary in order to obtain a bearing estimate to the source. In locations where the topography is flat along all radials from the receiver, no unique bearing information can be obtained; however, the mechanics of the localization algorithm work exactly the same in providing a two-dimensional range/depth location estimate. Again, multipath arrival information can be exploited without knowledge of ray path geometry. A second recording provided for the marine mammal workshop ("test data #6") also contained sperm whale click trains, but it was taken from a hydrophone on a relatively flat part of the AUTEC range. Although not presented here, range/depth estimates of whale location were made from that data, but bearing information could not be resolved.

Questions of accuracy are warranted when demonstrating a localization technique, and there are two types of error that can negatively affect the method shown here: mismatch between the modeled and truth environment, and errors in measurements of the relative travel times between multipath arrivals of a given click event. Both contribute nonlinearly to an overall error. A thorough sensitivity analysis for this technique was described in Tiemann et al. [21] showing it to be reasonably robust against such errors: 14 m error due to environmental mismatch and 16 m error due to measurement inaccuracy in a simulated case.

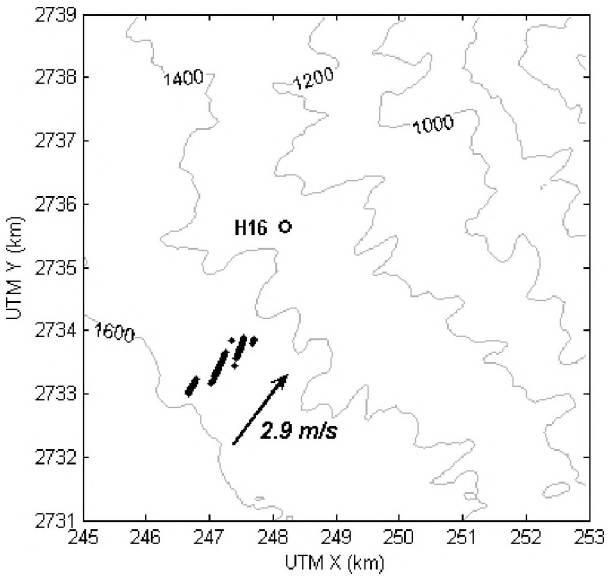

Figure 8. Estimates of sperm whale position overlaid on a plan view of bathymetry contours from the AUTEC range. Coordinates are for UTM zone 18; depths are in meters.

Lastly, although not implemented as such, the computational method for localization as presented here could be considered just a variation on the traditional hyperbolic fixing technique. In standard hyperbolic fixing, the time delay between an event recorded at two sensors defines a hyperbola of candidate source positions. The intersection of hyperbolas traced from multiple receiver pairs localizes the sound source. The same technique is basically being applied here (minus the isovelocity medium assumption), but in this single-hydrophone case, those additional receivers are *virtual*.

Recall that in ocean waveguide propagation, a ray path reflection off a boundary creates a virtual receiver vertically offset from the real receiver. Those real and virtual receivers are the foci in the definition of a hyperbola. The time of arrival difference between the direct and a reflected path, as measured during the arrival pattern extraction process, is like the time delay between an event arrival at the real and virtual receivers. Each additional multipath arrival time that can be measured contributes another virtual receiver and thus another hyperbolic path to the localization. To illustrate, consider the ambiguity surface of Figure 6. A direct ray path plus four late arrivals contributed to making this surface, so there are four paths of relatively high overlap scores. Those high-score areas even resemble hyperbolas for at least 2000 m away from the receiver, and they all

intersect at one range/depth bin to provide the localization estimate. With this understanding, a more elegant analytic solution for exploiting relative arrival times may become apparent as a substitute for the brute force computational method described here.

5. ACKNOWLEDGEMENTS

Thanks go to the mammal workshop organizers and to NUWC for providing not only the workshop dataset but also additional geometry and bathymetry information from AUTEC which made the localization possible. The Applied Research Laboratories at the University of Texas at Austin funded this analysis and algorithm development. Thanks also to Dr. Michael Revesz at ARL:UT for many helpful discussions and to Katherine Zaunbrecher and Jared LeBlanc for providing a translation of the abstract.

6. REFERENCES

1. R. Leaper, O. Chappell, and J. Gordon, "The development of practical techniques for surveying sperm whale populations acoustically," Rep. Int. Whal. Comm. **42**, 549–560 (1992).
2. J. Barlow and B.L. Taylor, "Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey," Marine Mammal Sci. **21**(3), 429–445 (2005).
3. C. Tiemann, S. Martin, J. Mobley Jr., "Aerial and acoustic marine mammal detection and localization on Navy ranges," IEEE J. Ocean Eng. **31**(1), 107-119 (2006).
4. S. Mitchell and J. Bower, "Localization of animal calls via hyperbolic methods," J. Acoust. Soc. Am. **97**, 3352-3353 (1995).
5. C.W. Clark, W.T. Ellison, and K. Beeman, "Acoustic tracking of migrating bowhead whales," IEEE Oceans Conference Proceedings **18**, 341-346 (1986).
6. J.L. Spiesberger and K.M. Fristrup, "Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography," Am. Nat. **135**, 107-153 (1990).
7. A.S. Frankel, C.W. Clark, L.M. Herman, and C.M. Gabriele, "Spatial distribution, habitat utilization, and social interactions of humpback whales, *Megaptera novaeangliae*, off Hawai'i determined using acoustic and visual techniques," Can. J. Zool. **73**, 1134-1146 (1995).
8. K.M. Stafford, C.G. Fox, and D.S. Clark, "Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean," J. Acoust. Soc. Am. **104**(6), 3616-3625 (1998).
9. V.M. Janik, S.M. Van Parijs, and P.M. Thompson, "A two-dimensional acoustic localization system for marine mammals," Mar. Mamm. Sci. **16**, 437-447 (2000).
10. C.W. Clark and W.T. Ellison, "Calibration and comparison of acoustic location methods used during the spring migration of the bowhead whale, *Balaena mysticetus*, off Pt. Barrow, Alaska, 1984-1993," J. Acoust. Soc. Am. **107**(6), 3509-3517 (2000).
11. C.O. Tiemann, M.B. Porter, and L.N. Frazer, "Localization of marine mammals near Hawaii using an acoustic propagation model," J. Acoust. Soc. Am. **115**, 2834–2843 (2004).
12. A. Thode, "Three-dimensional passive acoustic tracking of sperm whales (*Physeter macrocephalus*) in ray-refracting environments", J. Acoust. Soc. Am. **118** (6), 3575–3584 (2005).
13. A. Thode, D. K. Mellinger, S. Stienessen, A. Martinez, and K. Mullin, "Depth-dependent acoustic features of diving sperm whales *Physeter macrocephalus* in the Gulf of Mexico," J. Acoust. Soc. Am. **112**(1), 308–321 (2002).
14. W. M. X. Zimmer, M. P. Johnson, A. D'Amico, and P. L. Tyack, "Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale *Physeter macrocephalus*," IEEE J. Ocean. Eng. **28**(1), 13–28 (2003).
15. E.-M. Nosal and L.N. Frazer, "Sperm whale three-dimensional track, swim orientation, beam pattern, and click levels observed on bottom-mounted hydrophones," J. Acoust. Soc. Am. **122**(4), 1969-1978 (2007).
16. R. Aubauer, M.O. Lammers, and W.W.L. Au, "One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water," J. Acoust. Soc. Am. **107**, 2744–2749 (2000).
17. P.A. Lepper, K. Kaschner, P.R. Connelly, and A.D. Goodson, "Development of a simplified ray path model for estimating the range and depth of vocalising marine animals," in *Proc. Inst. Acoust.* (St. Albans, 1997), pp. 227–234.
18. M. Wahlberg, "The acoustic behaviour of diving sperm whales observed with a hydrophone array," J. Exp. Mar. Biol. Ecol. **281**, 53-62 (2002).
19. E.-M. Nosal and L.N. Frazer, "Track of a sperm whale from delays between direct and surface-reflected clicks," Appl. Acoust. **67**, 1187-1201 (2006).
20. C. Laplanche, O. Adam, M. Lopatka, and J. Motsch, "Male sperm whale acoustic behavior observed from multipaths at a single hydrophone," J. Acoust. Soc. Am. **118**, 2677-2687 (2005).
21. C. Tiemann, A. Thode, J. Straley, V. O'Connell, K. Folkert, "Three-dimensional localization of sperm whales using a single hydrophone," J. Acoust. Soc. Am. **120**(4), 2355-2365 (2006).
22. M.B. Porter and H.P. Bucker, "Gaussian beam tracing for computing ocean acoustic fields," J. Acoust. Soc. Am. **82**(4), 1349-1359 (1987).
23. F.A. Bowles, "A geoacoustic model for fine-grained, unconsolidated calcareous sediments," ARSRP Natural Laboratory, NRL/MR/7432-93-7082, March 9, 1994.