An Algorithm for the Localization of Multiple Interfering Sperm Whales Using Multi-Sensor Time Difference of Arrival

Paul M. Baggenstoss

Naval Undersea Warfare Center,

1176 Howell st,

Newport,

RI,

02841,

USA.

(Dated: May 5, 2011)
Abstract

In this paper an algorithm is described for the localization of individual sperm whales in situations where several near-by animals are simultaneously vocalizing. The algorithm operates on TDOA measurements observed at sensor pairs. The algorithm is able to resolve association “disputes” where a given TDOA measurement may fit to more than one position estimate. The algorithm also provides estimates of Cramer-Rao lower bound for the position estimates. We tested the algorithm with real data using TDOA estimates obtained by cross-correlating click-trains. The click-trains were generated by a separate algorithm that operated independently on each sensor to produce click-trains corresponding to a given whale and to reject click-trains from reflected propagation paths.

PACS numbers: 43.30.Sf, 43.30.Wi, 43.60.Jn
I. INTRODUCTION AND PROBLEM DEFINITION

There has been increased interest in the passive detection, localization, and classification of whales, mainly due to concerns over the impact of human activity such as active sonar on whale populations (Tiemann and Porter, 2004). Sperm whales (*physeter macrocephalus*) are best detected and localized by exploiting their loud echo location clicks. While some researchers have utilized the time difference between clicks received through different propagation paths (Nosal and Frazer, 2006), the predominant approach is to exploit the time difference of arrival (TDOA) of a given click at two or more sensors (Tiemann and Porter, 2004). Two general localization approaches exist. First, hyperbolic fixing (Mitchell and Bower, 1995) is a technique whereby a hyperbola is drawn on the geographic map corresponding to the locus of points consistent with the given TDOA at an assumed depth. The overlay of multiple hyperbolas from several TDOA measurements results in a convergence of hyperbolas at the correct horizontal whale location. Second, model-based methods create an ambiguity surface of the mean-square error between the measured time differences and the model time differences (Tiemann and Porter, 2004; White *et al.*, 2006). The methods are equivalent in principle, but the model-based methods lend themselves better to depth estimation and the use of more sophisticated propagation models that do not require the assumption of uniform sound-speed.

An important problem that arises with these methods is the existence of false or ambiguous TDOA values. There are three sources of these ambiguities.

1. **Cycle ambiguities.** Cycle ambiguities are introduced by the periodic nature of the clicking vocalizations (Tiemann *et al.*, 2006; White *et al.*, 2006). The repetition period of the clicks, known as inter-click interval (ICI), is relatively constant so one period can be mistaken for another. The predominant solution is to perform some kind of envelope correlation using a long time-window that contains many clicks (Morrissey

---

a) Electronic address: p.m.baggenstoss@ieee.org
et al., 2006; White et al., 2006; Bénard-Caudal and Glotin, 2009).

2. **Multi-path.** Multiple propagation paths or *multi-path* caused by reflections from the ocean surface or bottom cause additional ambiguities (Giraudet and Glotin, 2006; White et al., 2006). Attempts to eliminate multi-path on a single-click basis by exploiting the different energy decay-rates of direct and reflected paths have been used (White et al., 2006; Bénard-Caudal and Glotin, 2009). Multipath can also be eliminated by first estimating the echo delay at one sensor using autocorrelation, then eliminating multi-sensor time differences of arrivals which differ by the same estimated echo delay (Giraudet and Glotin, 2006).

3. **TDOA association.** Regardless of the multi-path and TDOA ambiguities of a given whale, there is the problem caused by the existence of multiple whales. Sperm whales often forage in groups, clicking simultaneously. To obtain a reliable localization solution, a given TDOA measurement must be associated with the correct whale, otherwise it can bias the location solution of another whale (Bénard-Caudal and Glotin, 2009). A method reported by (Giraudet and Glotin, 2006) associates TDOA estimates to a common whale by forming sensor quadruplets and insisting that the TDOAs satisfy simple arithmetic relationships. This method has the advantage that the TDOA estimates from the quadruplets are already associated with a given whale and can be used directly in a localization algorithm. The potential disadvantage is that the “hard” association decision is subject to error and requires simultaneous reception on at least 4 sensors.

Because our ultimate goal is the classification and/or identification of whales, we prefer a method that operates independently on each sensor to form the individual clicks into sequences of clicks, or *click-trains* that are from the same whale and propagation path (Bahl and Ura, 2003; Bahl et al., 2003; Baggenstoss, 2011). Classification is extremely difficult without performing click-train separation. The method of click-train formation and multipath elimination is described in a companion article (Baggenstoss, 2011).
Besides facilitating classification, the method has another advantage: although any 3-D localization method needs at least 4 TDOA measurements, our approach requires simultaneous click reception by pairs of sensors (instead of sensor quadruples by the method of Giraudet & Glotin - see last item above). This can be adapted to species with more directional emissions or wider-spaced hydrophones. This comes, however, at the cost of deferring the TDOA association problem to a later step - to be dealt with by the localization algorithm. The TDOA association problem is a classic chicken-and-egg problem: if the whale locations were known, the TDOA estimates could be better associated by comparison to the model TDOAs. But the whale location estimates can only be improved if the associations are correct. Such problems often lend themselves to the E-M algorithm (Redner and Walker, 1984), after which our algorithm is modeled.

In this paper, we describe an iterative model-based method that simultaneously solves the localization and TDOA association problems. Like some existing model-based approaches, it seeks a minimum of the weighted mean squared error between the measured and the model time differences (Tiemann and Porter, 2004; White et al., 2006). The primary innovation is the ability to simultaneously solve for multiple whales using “soft” association weights that assign each TDOA to the candidate solutions. The association weights and the position solutions are iteratively updated. A by-product of the algorithm is a solution error ellipse based on the Cramér-Rao lower bound.

II. TECHNICAL APPROACH

A. Main Assumptions

1. Simplifying Assumptions

We assumed a constant sound-speed of 1500 m/s in our localization experiments. This simplified many analytical results, however may be responsible for bias, particularly in the depth estimates. This does not limit the applicability of the method, which could be adapted
to more accurate sound-speed models. Also, to simplify the discussion, we do not track
whales. Thus, independent localization solutions are formed at each time update.

2. Inter-Sensor TDOA Estimation

We assume that sperm whale clicks are received at a number of bottom-mounted sensors
and that TDOA measurements from various combinations of sensor pairings have been mea-
sured. Although the algorithm can operate with any kind of TDOA measurement, the TDOA
measurements we used in the paper have been obtained using click-train formation. The
clicks received at each sensor have been grouped into separate click-trains and the reverber-
ation paths have been removed leaving only the direct-path click-trains. The click-trains are
formed using 12-second time windows which are updated each 6 seconds using 50 percent
time overlap. This process is described in the companion paper (Baggenstoss, 2011). By
relying on click-trains, the method lends itself better to the ultimate goal of classification.

The TDOA estimates are obtained by cross-correlating the click-trains from two sensors.
New inter-sensor TDOA measurements are available every 6 seconds. To eliminate potential
false measurements, TDOAs are subjected to a minimum correlation threshold of 0.9 and
a required minimum number of clicks (4 clicks). This results in a set of high-confidence
TDOAs from available sensor pairs. Each TDOA carries a weight corresponding to the
number of clicks involved in the correlation. The correlation process is described in the
companion paper (Baggenstoss, 2011).

TDOA measurements are represented by

$$\tau_i, 1 \leq i \leq M,$$

where $M$ is the total number of TDOA measurements collected from all the available sensors.
To identify the two sensor from which the measurement is made, we define the auxiliary
sensor indexes $s_i, r_i$. Thus, TDOA measurement $\tau_i$ is the time of difference for a click-train
received at sensors $s_i$ and $r_i$. Because sometimes we need to avoid the use of the auxiliary
indexes, we write $\tau_{AB}$ to indicate that this is a TDOA measured between sensors $A$ and $B$. 
3. **TDOA-whale association**

Once a TDOA measurement has been obtained by correlating two click-trains (from two different sensors) and if the correlation value is high, we can be reasonably certain that the two click-trains are from the same whale. We can expand this idea to form inter-associated groups. For example, if a TDOA $\tau_{AB}$ is obtained from correlation of click-train $A$ (at sensor $A$) and click-train $B$ (at sensor $B$), and TDOA $\tau_{BC}$ is obtained from correlation of click-train $B$ (at sensor $B$) and click-train $C$ (at sensor $C$), then clearly TDOA $\tau_{AB}$ and $\tau_{BC}$ are also associated. In this way, all of the available TDOA measurements can be organized into inter-related groups based on correlation.

Note that the chosen method has a certain chance of failure since the possibility exists that a set of TDOAs, say $\tau_{AB}$ and $\tau_{DE}$, exists that cannot be associated together. This happens if TDOAs $\tau_{AD}$, $\tau_{AE}$, $\tau_{BD}$ and $\tau_{BE}$ have been rejected because they do not exceed the cross-correlation threshold, so there are no common sensors. We did not use the TDOA association groups in the localization algorithm. Rather, we used the information as an aid to the manual initialization process by color-coding the hyperbolas according to groups. In the localization algorithm itself, we assumed no prior knowledge of which TDOA measurements were associated with which candidate solution (i.e. which whale).

B. **Localization Model**

We will introduce multiple whales and false TDOAs later. For now, assume for the moment that there is only one whale, all TDOA measurements are valid and associated with the given whale, and that we are considering TDOAs measured within a single time window of 12 seconds. Within this time window, we regard the whale as stationary. Let $z$ be an initial guess of the whale’s position. Let there be $M$ TDOA values available from click-train correlation between pairs of sensors. We assume that there are at least two TDOA values from three or more sensors for a two-dimensional solution and at least three TDOA values from four or more sensors for a three-dimensional solution. Let $n_i$, $1 \leq i \leq M$ be the
associated weights that represent the number of clicks that correlated to produce the given TDOA estimate. Let \( T(z, s, r) \) be the model propagation time difference between sensor \( s \) and sensor \( r \) assuming the whale is at location \( z \), where \( z = [x, y] \) for 2-D and \( z = [x, y, z] \) for 3-D. As an aside, note that

\[
T(z, s, r) = -T(z, r, s).
\]

We assume that

\[
\tau_i = T(z, s_i, r_i) + u_i,
\]

where \( u_i \) is zero-mean Gaussian independent noise with variance \( \sigma^2 \).

The assumption of independent and constant-variance TDOA estimation error is clearly a simplification. In reality, this error is not independent and the variance depends on many factors including the click detection method and click signal strength. However this seems like a reasonable simplifying assumption to make the problem tractable, and we do not anticipate this having a significant impact on the algorithm performance.

For the log-likelihood function we form a weighted sum of the log-likelihood functions for the independent measurements:

\[
Q(\tau_1, \ldots \tau_M; z, \sigma^2) = \sum_{i=1}^{M} n_i \left\{ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} [\tau_i - T(z, s_i, r_i)]^2 \right\},
\]

where \( n_i \) is the number of clicks involved in the TDOA estimate (the number of coincident clicks in the cross-correlated click-trains). Note that weighting the terms by \( n_i \) is equivalent to replicating the measurement \( n_i \) times. Our approach is to maximize \( Q(\tau_1, \ldots \tau_M; z, \sigma^2) \) over \( z \). Note that \( \sigma^2 \) can be estimated in parallel as the mean square TDOA error. The maximization of \( Q(\tau_1, \ldots \tau_M; z, \sigma^2) \) is more efficient if we analytically obtain the first and second derivatives with respect to \( z \).
1. First Derivatives

Let the 2 or 3-dimensional position vector be denoted by $\mathbf{z}$. Let the model TDOA error be

$$u_i = \tau_i - T(z, s_i, r_i).$$

Taking derivatives with respect to component $x$ of vector $\mathbf{z}$, we have

$$\frac{\partial Q(\tau_1, \ldots, \tau_M; \mathbf{z}, \sigma^2)}{\partial x} = \sum_{i=1}^{M} n_i \left\{ -\frac{u_i}{\sigma^2} (-T^x(z, s_i)) \right\},$$

where $T^x(z, s, r)$ is the partial derivative of model TDOA with respect to the $x$ component of $\mathbf{z}$. Note that we have removed the subscript $i$ and let $s, r$ represent any pair of sensor indexes. Let $D(z, z_s)$ be the distance from position $z$ to sensor $s$ (at position $z_s$). Then,

$$T^x(z, s, r) = \frac{\partial T(z, s, r)}{\partial x} = \frac{1}{C} \frac{\partial}{\partial x} \left\{ D(z, z_s) - D(z, z_r) \right\}$$

$$= \frac{1}{C} \frac{\partial}{\partial x} \left\{ \left[ (x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2 \right]^{1/2} - \left[ (x - x_r)^2 + (y - y_r)^2 + (z - z_r)^2 \right]^{1/2} \right\}$$

$$= \frac{1}{C} \left\{ D^{-1}(z, z_s) 2(x - x_s) - D^{-1}(z, z_r) 2(x - x_r) \right\}$$

$$= \frac{1}{C} \left\{ D^{-1}(z, z_s) (x - x_s) - D^{-1}(z, z_r) (x - x_r) \right\}$$

where $z_s = [x_s, y_s, z_s]$ in 3-D or $z_s = [x_s, y_s]$ in 2-D, is the sensor $s$ position and $C$ is the speed of sound. Partial derivatives with respect to position variables $y$ and $z$ are obtained similarly.

2. Second Derivatives

Taking second derivatives,

$$\frac{\partial^2 Q(\tau_1, \ldots, \tau_M; \mathbf{z}, \sigma^2)}{\partial x \partial y} = \sum_{i=1}^{M} n_i \left\{ -\frac{1}{\sigma^2} T^x(z, s_i, r_i) T^y(z, s_i, r_i) - \frac{1}{\sigma^4} u_i T^{xy}(z, s_i, r_i) \right\}.$$
C. Single Whale Algorithm

We present a simplified algorithm that assumes a single whale and that all TDOAs are associated with the whale. This simplified algorithm forms the basis of the multiple whale algorithm.

1. Algorithm Initialization

As there are generally many local maxima of (1), we will require a starting location in the vicinity of the global maximum. Later, when we discuss multiple whales, we address the problem of ambiguous solutions and automatic initialization. For now, we assume a single initial solution is available. One way to determine a likely starting point is based on intersections of hyperbolas (i.e. hyperbolic fixing). Each TDOA measurement corresponds to a locus of points forming a hyperbola in a two-dimensional geometric plot (see figure 1). Hyperbolic fixing can form the basis of manual or automatic initialization approaches. The depth of the whale must be assumed in order to create the hyperbolas. For the purpose of drawing an initial set of hyperbolas, we assumed a depth of 500 meters, which is a reasonable average depth for sperm whales. In figure 1, we show such a set of hyperbolas drawn using the TDOA obtained from a single 12-second data window. The initial position estimate would be placed at the convergence point of the hyperbolas.

2. Position Update

The parameters of the position solution include the TDOA error variance $\sigma^2$ and the position vector

$$z = [x, y, z]^T.$$

Let $z_l$ be the position vector at the $l$-th iteration. We apply the Newton-Raphson iteration (Kelley, 2003) to maximize equation (1). In place of the matrix of second derivatives (Hessian), we use the negative of the Fisher’s information matrix (FIM). The the expected value
FIG. 1. Hyperbola plots for a single 12-second dta window using initial depth estimate of 500m. There are 21 hyperbolas (21 TDOAs) obtained using 7 sensors. Data is from data set 1 (see section III).

of the Hessian equals the negative of the FIM. While using the true Hessian will result in the fastest convergence when near the local maximum, the FIM is guaranteed to be positive-definite and can avoid numerical problems when initialized far from the correct value. Refer to section V for additional details about the FIM. The algorithm update equation is

$$z_{l+1} = z_l + I^{-1}_{xyz}(z_l) \delta(z_l),$$

where $I_{xyz}(z_l)$ is the Fisher’s information matrix and $\delta$ is the vector of first derivatives (of $Q(\tau_1 \ldots \tau_M; z, \sigma^2)$ with respect to $x$, $y$, and $z$). Both are computed using the previous solution $z_l$. The algorithm converges when no further significant change in position or
objective function (1) is observed.

3. **Position Error Ellipse**

It is important not only to know the solution \( \mathbf{z} \) that maximizes (1), but also the error variance associated with this solution. This information is important when updating a recursive tracker. It is also critical in establishing, for example, whether the depth component can be trusted. It is very common for the \( x, y \) position to be accurate but the depth estimate to be inaccurate. This occurs typically when there are no nearby sensors. It is also of interest to know along which axis in the \( x, y \) plane the most error is expected to lie.

The \( x, y \) error ellipse is a region in space, drawn around the position estimate, that can be regarded as the region in the \( x, y \) plane where the true position must be, to some degree of certainty.

The Cramer-Rao lower bound (CRLB) matrix is the inverse of the Fisher’s information matrix (FIM) \( \mathbf{C} = \mathbf{I}_{x,y,z}^{-1} \) (Kay, 1988). The FIM for the position estimates based on the above model is derived in the appendix (section V). In our case, \( \mathbf{C} \) is the 3 by 3 error covariance matrix for the 3-dimensional parameter set \( \mathbf{z} = \{x, y, z\} \) for a location solution. This is an estimate of the asymptotic error covariance matrix. As the amount of data goes to infinity \((M \to \infty)\), we have

\[
\mathbb{E}((\mathbf{z} - \mathbf{z}_0)(\mathbf{z} - \mathbf{z}_0)') \to \mathbf{C},
\]

where \( \mathbb{E}(\cdot) \) is the expected value operator (computes the mean or statistical average). where \( \mathbf{z} \) is a solution estimate and \( \mathbf{z}_0 \) is the true location.

The solution error ellipse for solution \( \mathbf{z} \) are the locus of points \( \mathbf{z}^* \) where the inner product

\[
(\mathbf{z} - \mathbf{z}^*)'\mathbf{I}_{x,y,z}(\mathbf{z} - \mathbf{z}^*) = c,
\]

where \( c \) is a constant. For a Gaussian-distributed 2-D position error, with \( c = 1 \), 63 percent of the samples will lie within the ellipse. For a Gaussian-distributed 3-D position, 61 percent will lie within the ellipse. It is possible to use (1) to estimate the solution error ellipse based on the Cramer-Rao lower bound. These points can be drawn using eigen-analysis of
matrix $I_{x,y,z}$. A MATLAB code segment to do this is provided in a separate technical report (P.M.Baggenstoss, 2010).

D. Multiple Whales and False Time Delays

Not all of the inter-sensor TDOA measurements that are produced for a given time window are associated with a given whale position solution. Either they are false TDOAs altogether, or they may be associated with other whales. To handle the possibility that some TDOAs are not associated with the given solution, it is necessary to augment the weight $n_i$ in (1) with a solution-dependent membership probability, thus replacing the probabilistic weight $n_i$ with $n_i w_{k,i}$. Let there be $K$ candidate solutions with current localization solutions $z_k$, $1 \leq k \leq K$ and $K$ is equal to or greater than the number of whales (since we initialize the algorithm with many redundant solutions, this is virtually guaranteed). The membership probability,

$$w_{k,i}, \ 1 \leq k \leq K, \ 1 \leq i \leq M, \ \sum_{k=1}^{K} w_{k,i} = 1,$$

is the probability that measurement $\tau_i$ is associated with solution $k$. Thus a given TDOA is “soft-assigned” to the solutions. We consider two ways of calculating the membership probabilities $w_{k,i}$, “fixed-weight” (single-pass) and “iterative-weight”.

1. Fixed-Weight Membership Weights

For the fixed-weight method, we assign a fixed weight that depends only on the TDOA error. We used

$$w_{k,i} = \exp\left(-\frac{(\tau_i - T(z_k, s_i, r_i))^2}{2\sigma_m^2}\right),$$

where $\sigma_m$ is the fixed membership TDOA standard deviation of 0.1 seconds. The standard deviation $\sigma_m$ acts as a “soft” threshold for TDOAs to be included in the solution $k$. The value .1 was obtained empirically and offers a compromise between the desire to reject valid time-delays, but to reject invalid ones.
2. Iterative-Weight Membership Weights

The fixed-weight method fails if TDOAs are close matches to more than one solution. Take for example figure 2. There are three hyperbolas that intersect more than one solution.

![Figure 2](image)

**FIG. 2.** Hyperbolas from Data set 2 simultaneously showing four whales. The depth of 760 meters is assumed (this is the solution depth for the whale at $x = 13700$, $y = 6400$).

(there are four solutions indicated by the star-like convergence points for like-colored hyperbolas and labeled with “whale”). Because each of these TDOAs (hyperbolas) are consistent with more than one solution, the fixed-weight scheme would probably assign each of these TDOAs more or less equally to two solutions, biasing the localization positions. There is no easy way to be sure *a priori* which solution the TDOAs belong to. The color-coding seen in the figure was automatically-generated grouping based on correlation (see section II.A.3).
Although entirely correct in this instance, it cannot be relied on in every instance to make the correct associations. Thus, we must assume at the outset that any TDOA can associate with any solution.

The assignment problem is related to many problems in science that can be loosely called “clustering” algorithms - since assigning a TDOA to a whale can be considered the same as assigning it to a cluster. Algorithms that have probabilistic or “soft” assignments are iterative and are often called “fuzzy” clustering algorithms (Hoppner et al., 1999). These methods are also related to E-M algorithms (Redner and Walker, 1984). Soft membership functions (membership probabilities) allow an interactive “negotiation” to take place that attempts to resolve “disputes” happening when a given TDOA suits more than one solution. These “disputes” can be resolved if for a given solution, the TDOA errors for the correctly-associated TDOAs are smaller than the TDOA errors for the falsely associated TDOAs. Then, the incorrectly-associated TDOAs will have lower weight. This lower weight will reduce the bias, which in turn will drive the solution closer to the correct location, reducing the TDOA errors for the correctly-associated TDOAs even further. Eventually, the incorrectly-associated TDOAs will be driven out of the solution altogether.

To implement the iterative-weight approach, we will need a measure of the “importance” or “weight” of each candidate solution. This will allow us to use a large number of candidate solutions, and to simply ignore or delete low-weight solutions. Because there are a fixed number of TDOA values (M to be precise), and these TDOAs are soft-assigned to the K solutions, we can estimate the number of TDOAs that each position solution “owns”. Let \( \alpha_k \) be a relative measure of the number of clicks represented by solution k, normalized so that

\[
\sum_{k=1}^{K} \alpha_k = 1.
\]

The number of clicks is counted by summing the TDOAs owned by solution k weighted by \( n_i \), the number of clicks represented by the given TDOA. Clearly, \( \alpha_k \) is also a measure of the relative importance of solution k.

Let \( w_{k,i} \) be the assignment weights, equal to \( n_i \) times the probability that TDOA i is
assigned to solution $k$, therefore we must have

$$\sum_{k=1}^{K} w_{k,i} = 1.$$  

We estimate $w_{k,i}$ as follows. Let $\sigma_k^2$ be the solution-dependent time delay error variance. The un-normalized assignment weights are computed according to

$$\tilde{w}_{k,i} = \alpha_k (2\pi \sigma_k^2)^{-1/2} \exp \left( \frac{(\tau_i - T(z_k, s_i, r_i))^2}{2\sigma_k^2} \right).$$

We then normalize the weights to sum to 1 over $k$

$$w_{k,i} = \frac{\tilde{w}_{k,i}}{\sum_{l=1}^{K} \tilde{w}_{l,i}}.$$  

These weights are used in equation (1) by replacing $n_i$ with $n_i w_{k,i}$.

Note that each TDOA must have a non-zero assignment to at least one solution. In the case that some TDOA’s are spurious, we can assign them to a “dummy” solution. The “dummy” solution is included as one of the $K$ solutions, initialized to the geographic center of the sensor field and assigned a very large time-delay error variance. This way, any TDOA that does not match any of the whale positions will be assigned to this dummy solution by default.

In contrast to the fixed-weight method, these weights grow larger if either (a) $\sigma_k^2$ grows smaller (solution is tighter), or (b) if solution $k$ becomes “stronger” (larger $\alpha_k$). At each iteration, we update the position solutions $z_k$ using the new TDOA membership weights $w_{k,i}$. On the next iteration, we re-calculate the solution weights

$$\alpha_k = \frac{\sum_{i=1}^{M} n_i w_{k,i}}{\sum_{l=1}^{K} \sum_{i=1}^{M} n_i w_{l,i}},$$

3. Position Initialization

When using the iterative-weight algorithm, it is possible to initialize the algorithm with many redundant solutions. The “weaker” solutions will eventually be dropped when $\alpha_k$ becomes too small. These initial starting points can be determined by an automatic algorithm
that searches for converging hyperbolas (section II.C.1). If the convergence of hyperbolas is used for determining initial solutions, a good strategy would be to add a candidate solution at each point where three or more hyperbolas intersect (or nearly intersect). These weak convergence points often correspond to false local maxima of the target function (1). In addition, one can inject candidate solutions at multiple depths at each convergence point, say every 160 meters. Note that this initialization process assumes that no prior information is available. One could modify this process in a tracking application by using the solutions from the previous time step as starting points for the current time update.

4. Additional Initializations

We start the process by assuming the solutions are equally likely, setting \( \alpha_k = 1/K \). We initialized \( \sigma_k \) to 0.1 seconds.

5. Algorithm Summary

We summarize the algorithm as follows:

1. Obtain \( K \) potential location estimates where \( K \) can be much larger then the actual number of whales. Include a “dummy” solution located at the center of the sensor field.

2. Initialize \( \alpha_k, \sigma_k^2, 1 \leq k \leq K \) as described. Use a much larger \( \sigma_k^2 \) for the “dummy” solution (\( \sigma^2 = 1 \) is adequate).

3. Compute \( \tilde{w}_{k,i}, 1 \leq k \leq K, 1 \leq i \leq M \) as described.

4. Compute \( w_{k,i} \) from \( \tilde{w}_{k,i} \) as described.

5. Update \( \alpha_k \) as described.

6. Execute one iteration of the localization algorithm (section II.C.2), noting that \( \sigma^2 \) and \( n_i \) are replaced by the solution dependent \( \sigma_k^2 \) and \( n_i w_{k,i} \).
7. Go to step 2.

As the algorithm iterates, we prune weak or redundant solutions. We eliminate a solution if

1. \( \alpha_k \) falls below a threshold (problem dependent value, typically 0.01),

2. two solutions became physically closer than about 30 meters (weaker solution removed),

3. or the membership functions of two solutions became highly correlated indicating they represent the same set of TDOAs.

III. APPLICATION TO REAL DATA

A. Data Set Description

Details of the data sets we used in the examples are listed in table I.

B. Examples with One Whale

Whether a single or multiple whales were actually present, we always initialized the solution with multiple candidate solutions. Although an automated initialization is possible, we tested the algorithm using a graphical user interface by clicking at the convergence of the hyperbolas which were drawn at the initial depth of 500 meters. We injected candidate solutions at intervals of 160 meters depth. An example of an initial hyperbola plot is shown in figure 1. Because the main requirement in the initialization process is that at least one initial solution exists near the correct position, this manual initialization process can be replaced by an automatic one.

The final solution converged at 780 meters depth with a final CRLB that had a standard deviation of 4.3, 3.4, and 19.0 meters in x, y, and z. All other candidate solutions were automatically removed (see section II.D.5). The hyperbolas re-drawn for the 780 meters depth are shown in Figure 3. Notice that a much tighter convergence is seen in comparison
FIG. 3. Top: hyperbolas from Data set 1 showing one whale. Bottom: TDOA error values after convergence.

to Figure 1. The TDOA error values are also shown in the bottom of the figure. There is one error value for each TDOA (and each hyperbola). Error absolute values of 0.01 seconds and below are typical for valid TDOAs.

Only one whale was present in Data sets 1, 3, 4, and 5. Localization error ellipses for
data set 1 and the corresponding depth profile are shown in Figure 4. Ellipses are drawn for each time update at 6 second intervals. In the depth contour, the CRLB is plotted as an envelope surrounding the depth estimates. Similar to the error ellipses on the two-dimensional plane, this provides a reliable “warning” that solutions may have significant error. Notice that there are instances when the depth estimate exhibits outliers. This is caused by the unavailability of nearby sensors to provide adequate depth observability. Notice that at these times, the CRLB envelope expands, so is in fact a reliable indication of depth estimate variance. The position and depth results can be directly compared with figure 6 in (Giraudet and Glotin, 2006). The depth contour agrees in shape, but has an average of 800 meters instead of 700 meters in the reference. This may be due to our constant velocity assumption.
FIG. 4. Localization Error Ellipses and depth profile for Data set 1. Y-axis: depth. X-axis: time update number. The cyan “envelope” is the depth CR bound. The plots been color-coded so that the error ellipses and depth points can be associated. The results can be directly compared with figure 6 in (Giraudet and Glotin, 2006).
Results for data sets 3, 4, and 5 are presented in a separate technical report (P.M. Baggenstoss, 2010).

C. Examples with Multiple Whales

As many as four whales were present in Data set 2. An example of a hyperbola plot is shown in Figure 2 with clear evidence of four simultaneously vocalizing whales. The TDOAs have been automatically grouped into inter-related groups for display as described in section II.A.3. The hyperbolas have been color-coded according to group membership.

When initializing the algorithm for multiple whales, multiple candidate solutions are added at each convergence point (and at possible convergence points) at 160 meter depth separation.

As we discussed in section II.D.2 and illustrated in Figure 2, with multiple whales, TDOA estimates can overlap. Returning to Figure 2, consider the whale positioned farthest to the right indicated by the convergence of the hyperbolas at $x=18400, y=7500$. Notice that two of the hyperbolas emanating from this point pass very close to the centers of two other solutions. When solving for the positions of all of these whales using the fixed-weight membership function, positional errors will be introduced because those TDOAs belonging to other whales will not be rejected by the membership function (section II.D). This case was processed using the iterative-weight algorithm. Solutions were placed manually by clicking at the convergence point of each solution in figure 2 which was drawn using average depth of 500m. At each of the four convergence points, ten initial solutions were added in depth increments of 160 meters spanning the water column. As the algorithm iterated, solutions were removed according to pruning rules (section II.D.2). In ten attempts, each time clicking at random locations near the four convergence points, the algorithm correctly reported four surviving solutions with the same final locations. The results are shown in Figure 5. Blow-ups are shown in figure 6 in which it can be seen that there is extremely tight convergence at the solution points despite non-associated hyperbolas lying very close to the solutions. This
FIG. 5. Example of figure 2 after convergence of the iterative-weight algorithm. The hyperbolas are color-coded according to membership to the four remaining solutions. The final solution depth was used when drawing the hyperbolas of each solution independently. A cross (+) is drawn at the x-y location of each solution. Note that one hyperbola passing through the solution at (18300,7500) is two nearby hyperbolas. This was caused by the separate TDOA estimates from two parts of a click-train that was split in two - and for whatever reason (whale motion or separate estimation error) give slightly different TDOAs.

indicates that the membership weights have excluded these nearby hyperbolas (TDOAs).

The error ellipses for each of the four whales are shown in Figure 7. The tracks can be compared directly with figure 7 in (Bénard-Caudal and Glotin, 2009) where four whales are also reported.
FIG. 6. Expanded view of solution 1 through 4 in figure 5.
Depth profiles may be found in the technical report (P.M. Baggenstoss, 2010). In comparison to (Bénard-Caudal and Glotin, 2009), our depth estimates were more unreliable. This is likely due to the loss of time-delay measurements resulting from complications in click-train formation or due to thresholding of the click-train cross-correlation function.
IV. CONCLUSIONS

In this paper, we have presented a method of localizing sperm whales in environments containing several interfering whales using clicks detected on an array of bottom-mounted sensors. Like many model-based approaches, the method operates using the time difference of arrival (TDOA) of clicks received at pairs of sensors, comparing the TDOA measurements with a propagation model in order to minimize an error criterion. But, our method differs significantly from existing approaches. Unlike other methods, our approach can operate with multiple whales and time-delay estimates that have unknown associations (i.e. it is not known to which whale the time-delays belong). The TDOA association problem is handled simultaneously with the localization problem using iteratively-adapted “soft” association weights. We used TDOA measurements obtained by cross-correlating click-trains that are formed from clicks at individual sensors rather than from individual clicks or envelope correlation. This makes the method suitable for whale classification and identification. The method has as a by-product, the CRLB (error bound) that predicts the position estimation error ellipse in real-time. We demonstrated the method in examples containing up to four simultaneously vocalizing whales. We have demonstrated the ability of this method to resolve “disputed” associations - where a given TDOA is consistent with the positions of more than one whale.

V. APPENDIX 1: FISHER’S INFORMATION MATRIX FOR POSITION

The Fisher’s information matrix is a $P \times P$ matrix where $P$ is the total number of parameters (3 or 4 including $\sigma^2$, depending if depth $z$ is required). Let $x^*$ and $y^*$ be two arbitrary position parameters (they can represent any pair in the set $\{x, y, z\}$). Then the $x^*, y^*$ element of the Fisher’s Information matrix is

$$I_{x^*, y^*} = -\mathcal{E} \left\{ \frac{\partial^2 Q(\tau_1, \ldots, \tau_M; z, \sigma^2)}{\partial x^* \partial y^*} \right\} ,$$
where $\mathcal{E} \{ \} \text{ is the statistical expected value (mean value). Note that } \mathcal{E} \{ u_i \} = 0. \text{ Thus,}

$$I_{x^*,y^*} = \sum_{i=1}^{M} w_i \left\{ \frac{1}{\sigma^T} T_{x^*}^T (z, s_i, r_i) T_{y^*}^T (z, s_i, r_i) \right\},$$

where the superscript, for example $T_{x^*}$, represents the first derivative with respect to $x^*$.

**Acknowledgments**

This research was undertaken partly as a component of the DECAF project (Density Estimation for Cetaceans from passive Acoustic Fixed sensors), funded under the National Oceanographic Partnership Program jointly by the Joint Industry Programme and US National Marine Fisheries Service. I would like to especially thank DECAF project members, Len Thomas and especially Tiago Marques who provided very comprehensive and insightful comments and suggestions.

**References**


<table>
<thead>
<tr>
<th>Set</th>
<th>minutes</th>
<th>sensors</th>
<th>No. of whales</th>
<th>IDENTIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>5</td>
<td>1</td>
<td>Monaco 2005 nr. 2</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>6</td>
<td>4</td>
<td>Monaco 2005 nr. 1</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>7</td>
<td>1</td>
<td>AUTEC BRS 2007 Dive 1</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>8</td>
<td>1</td>
<td>AUTEC BRS 2007 Dive 2</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>7</td>
<td>1</td>
<td>AUTEC BRS 2007 Dive 3</td>
</tr>
</tbody>
</table>

List of Figures

FIG. 1  Hyperbola plots for a single 12-second data window using initial depth estimate of 500m. There are 21 hyperbolas (21 TDOAs) obtained using 7 sensors. Data is from data set 1 (see section III). 11

FIG. 2  Hyperbolas from Data set 2 simultaneously showing four whales. The depth of 760 meters is assumed (this is the solution depth for the whale at \( x = 13700, y = 6400 \)). 14

FIG. 3  Top: hyperbolas from Data set 1 showing one whale. Bottom: TDOA error values after convergence. 19

FIG. 4  Localization Error Ellipses and depth profile for Data set 1. Y-axis: depth. X-axis: time update number. The cyan “envelope” is the depth CR bound. The plots been color-coded so that the error ellipses and depth points can be associated. The results can be directly compared with figure 6 in (Giraudet and Glotin, 2006). 21

FIG. 5  Example of figure 2 after convergence of the iterative-weight algorithm. The hyperbolas are color-coded according to membership to the four remaining solutions. The final solution depth was used when drawing the hyperbolas of each solution independently. A cross (+) is drawn at the x-y location of each solution. Note that one hyperbola passing through the solution at (18300,7500) is two nearby hyperbolas. This was caused by the separate TDOA estimates from two parts of a click-train that was split in two - and for whatever reason (whale motion or separate estimation error) give slightly different TDOAs. 23

FIG. 6  Expanded view of solution 1 through 4 in figure 5. 24

FIG. 7  Localization Error Ellipses for Data set 2, whales 1-4. 25