Contact Localization and Motion Analysis in the Ocean Environment: A Perspective

JOSEPH C. HASSAB

(Invited Paper)

Abstract—This paper presents a perspective on contact localization and motion analysis (CLMA) in the ocean environment. Such well-studied cases as passive bearings-only contact motion analysis, multipath ranging, and localization/tracking via arrays are used to illustrate the concept. The performance of current CLMA systems is shown to depend on the type and characteristics of the basic measurements developed by their time delay estimators, the acoustic channels linking contact and observer, and the type and description of relative motion between contact and observer. General classes of CLMA schemes are presented. Complexity is shown to depend on the linearity/nonlinearity of their solution equations, the degree to which a contact is observable at each sample time, and the relative motion between contact and observer. Contributions to total system gain, biasing issues, and candidate solutions are discussed. Comprehensive references are provided.

I. INTRODUCTION

This paper discusses the problem of estimating the location and velocity of a sonar contact via observation and processing of acoustic data. Such data may be considered to include a desired signal and unwanted noise components. Embedded in the body of received data are differences in signal arrival times (time delays) as well as variations (Doppler shifts) in the signal. These differences and variations in signal are functionally dependent upon contact-observer geometry and environmental conditions.

Contact localization and motion analysis (CLMA) systems make use of a received signal's time delay and its variation in time to estimate a contact's location. These processing systems basically comprise a signal time delay estimator and a contact motion estimator (see Fig. 1). The time delay estimator maps the received acoustic data into recognizable and measurable clues (a dominant peak, valley, or slope on a curve, for example). These clues are further processed by the contact motion estimator so that estimates of time delays are smoothed and mapped into values for contact range, direction, depth, and velocity.

CLMA systems process data spatially as well as temporally. That is, they process data received simultaneously at spatially separated sensors, as well as data received during sequential observation intervals spread out in time. The total system gain results from both spatial and temporal gains. Spatial gain is influenced by such factors as size, number, placement, and configuration of sensors in the acoustic array. Temporal gain is influenced by the manner in which the received data are processed in time. In principle, joint optimization in both spatial and temporal dimensions is desirable for best system performance. In practice, considerations such as array stabilization, cost, spatial coherence, platform dimensions, etc., limit the achievable spatial gains, while nonstationarity, the requirement to yield an acceptable solution within a given time, etc., limit temporal gains.

Elements in a CLMA problem may be stationary (e.g., constant statistical measures on signal and noise, no relative motion between contact and observer, a homogeneous environment, etc.) or nonstationary. When the elements are stationary, processing is straightforward and is accomplished by a continual integration over the observed contact clues until desired accuracy in the solution is obtained. When the elements of the problem are nonstationary (e.g., moving contact/observer, a changing ray path channel, etc.) bias is introduced during a long contact observation interval due to smearing of the clue at the output of the time delay estimator. For instance, when the time delay is varying in time due to relative contact/observer motion, a correlator averaging time is kept short enough so that the time delay is quasi-stationary and the smearing effect of the time delay peak is avoided.

Thus observation of the contact must be limited to a brief time interval over which the process may be considered locally stationary. In this case, CLMA systems provide what may be considered "short-memory" or "snapshot" clues, which yield imprecise estimates of contact location and motion. However, with a succession of such brief time observation intervals, the system's temporal processor can extend the system memory and remove the biasing non-stationarity in the problem. It does this by superimposing the repeated short-memory estimates to enhance the invariant contact parameters in the problem, ultimately developing a well-defined estimate of the contact's location and motion. In Section VI, these concepts are developed further in the context of CLMA from a linear array.

In its totality, then, contact localization and motion estimation constitutes a process that is mathematically nonlinear and geometrically nonstationary in terms of contact/observer. It is a process not amenable to optimum global system syn-

 Manuscript received December 6, 1982; revised March 30, 1983.
 The author is with the Naval Underwater Systems Center, Newport Laboratory, Newport, R1 02841.

U.S. Government work not protected by U.S. copyright
thesis, as evidenced by most literature in the field which
generally deals with optimization of subsystems as realistic
conditions are introduced. Systems providing optimal per-
formance have been developed, but these are only for idealized
conditions such as stationary contact/observer and Gaussian
signal and noise, and extended contact observation times.

This paper aims to provide the reader who has had an
introduction to the CLMA problem with an overall perspective
developed in the context of the ocean environment. It pro-
vides numerous references for those interested in further study
of this subject. In Section II, six classes of CLMA problems are
identified and the difficulties in dealing with each of these
classes are mentioned. In Section III, illustrative examples
for these problem classes are introduced. In Section IV, the
types of errors encountered in the ocean environment, which
tend to make the various classes of problems more difficult
to solve, are categorized and techniques for reducing these
errors are highlighted. In Section V, the various points dis-
cussed in preceding sections are integrated into three elements
needed to formulate and solve CLMA problems in general.
In Section VI, three well-suited CLMA problems are reviewed
and used to illustrate the concepts developed in the preceding
sections.

II. GENERAL CLASSES OF CONTACT STATE
ESTIMATION PROBLEMS

Several types of contact state estimation problems are seen
in the literature available on this subject. These may be grouped
into general classes on the basis of source/observer motion,
linearity, and observability. That is, they can be grouped
according to the degree of relative contact/observer motion,
the complexity of their solution equations (linear or non-
linear), and the extent to which a contact is observable (i.e.,
the extent to which an observer can realize a unique solution
from the available data). For a moving observer and/or con-
tact, the problems increase in complexity as follows:

<table>
<thead>
<tr>
<th>Problem Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Linear solution problems with the contact’s state observable over each observation (sampling) interval.</td>
</tr>
<tr>
<td>B</td>
<td>Linear solution problems with the contact’s state observable only after multiple observation (sampling) intervals.</td>
</tr>
<tr>
<td>C</td>
<td>Linear solution problems with the contact’s state observable only after multiple observation (sampling) intervals, and only with motion constraints placed on contact and observer.</td>
</tr>
<tr>
<td>D</td>
<td>Nonlinear solution problems with the contact’s state observable over each observation (sampling) interval.</td>
</tr>
<tr>
<td>E</td>
<td>Nonlinear solution problems with the contact’s state observable only after multiple observation (sampling) intervals.</td>
</tr>
<tr>
<td>F</td>
<td>Nonlinear solution problems with the contact’s state observable only after multiple observation (sampling) intervals, and only with motion constraints placed on contact and observer.</td>
</tr>
</tbody>
</table>

Specific illustrations of the preceding classes of problems
are given in Section III. In the ocean context, the latter classes
are more prevalent than the former and are more difficult
to solve. These classes of problems are especially difficult
to solve when there is a mismatch between physical processes
and modeled processes (e.g., type of contact motion, raypath
model, etc.) or when there are large errors in the estimated
time delay parameters. An expanded discussion of these dif-
ficulties is given in Section IV. In analyzing the various CLMA
problems likely to be encountered, several general state-
ments hold true:

- Linear problems lend themselves readily to optimal esti-
mation with resulting minimum mean square estimation
error.
- Nonlinearity increases the complexity and the issues
involved in structuring an algorithmic estimator.
- Increased contact observability tends to improve the
quality of an estimate and speeds estimator convergence.
- Redundant observation (sampling) is required to reduce
the adverse effect of measurement errors.
- Constraints on observer/contact motion encumber the
estimation process by delaying estimator convergence,
lowering the quality of estimates, and degrading the abil-
ity of the estimator to adapt to mismatches between
modeled and physical processes.

Two types of estimator applications exist depending on
whether observation of the contact is by active [1], [2]
or passive [3] sonar. In the active case, the contact is ensoni-
fied by a signal emitted from the observer; estimation of
contact location and velocity is based on the observation and
processing of the backscattered returns from the contact. In
the passive case, the contact itself is an emitter whose signal
is received at the observer and processed for estimation of
contact location and velocity. For the active case, contact
localization and motion estimation falls into classes at the
beginning of the preceding list. For the passive case, the
problem falls into the classes predominantly at the end of the
list, which makes the estimation process a more difficult
one.

In either case, the observer is linked to the contact through
the intervening propagation medium. When analyzed, the
medium is seen to have distinguishable acoustic ray paths lying
within the usable beampatterns of both contact and observer.
Distinguishability here refers to the difference in path lengths
measured relative to a reference path or reference time. Each
difference in path length is reflected in the time delay incurred
by the signal as it propagates through the different paths.
In the active case, time delay refers to the difference in arrival
time between the reference emission time and reception of a
return. The term “time delay” in this paper also refers to the
time required for a sinusoid to repeat itself; i.e., its period.
These various time delays constitute the basic measurements that a time delay processor extracts from received signals. The desired contact state information is embedded within each time delay, which is characterized by the ray path structure within the sound channel.

III. REPRESENTATIVE PROBLEMS

Single Ray Path Channel

Representative problems are now described for a homogeneous and noiseless channel to illustrate in a simple context the various classes of problems. Consider first a problem involving a single path linking the contact to the observer (Fig. 2(a)). For the active case, the range to the contact and the rate at which the contact's range changes (range rate) are directly observable from the measured time delay through a linear relation. This represents a class A or B estimation problem (linear, with a high degree of observability) which is easily solvable [4]-[8].

The observer emits a pulse signal in a given direction and measures the time delay for its return. The range to the contact is related to the time delay by the equation

\[ R = \frac{ct}{2} \]  

where \( c \) is the in-water speed of sound. Subsequent time delay measurements yield the range rate \( \dot{R} \). As the measurements become noisy, more time delay measurements are required to yield the desired accuracy in range and range rate estimates.

A second measure of range and range rate between contact and observer is described by the Doppler shift in the signal, where

\[ R(t) = R(0) + ct - \frac{1}{f_0} \int_0^t f(t) \, dt \]

\[ \dot{R} = c \left[ 1 - f(t)/f_0 \right] \cdot \frac{(1)}{} \]

Here, \( f(t) \) is the received frequency and \( f_0 \) is the emitted frequency. In the active case, \( f_0 \) is known. In the passive case, other information such as contact direction may be used to estimate \( f_0 \).

Two-Path Channel

Another type of problem involves nonintersecting, two-path channels linking contact to observer (Fig. 2(b)). The measured time delay due to the different path lengths yields the direction to the contact in the plane containing sensors and contact. In the horizontal plane, the time delay \( \tau \) yields the far-field direction angle, \( \beta_a \) (known as bearing) where

\[ \beta_a = \sin^{-1} \left( \frac{ct}{L} \right) = \tan^{-1} \left( \frac{R_y}{R_x} \right) \]

Here, \( L \) is the separation between the sensors and \( R_x \) and \( R_y \) are the \( x \)- and \( y \)-components of the range \( R \) in the horizontal plane. In the vertical plane, the measured time delay gives the direction angle \( \beta_v \) (known as depression/elevation angle) where

\[ \beta_v = \tan^{-1} \left[ \frac{R_y}{(R_x^2 + R_y^2)^{1/2}} \right] \]

Here, \( R_z \) is the contact's \( z \)-coordinate in the vertical plane. When angular measurements fall in planes other than the vertical and horizontal planes, the direction angle is called a conical angle \( \beta_c \) and is related to \( \beta_a \) and \( \beta_v \) by

\[ \beta_c = \cos^{-1} \cos \beta_a \cos \beta_v = \cos^{-1} \frac{R_y}{R} \]

In this type of problem only directional information in the plane containing the ray paths is obtained directly from the time delay. Within each plane, the contact's state descriptors (location and velocity) may be obtained if motion constraints are placed on contact and observer; i.e., if the contact is presumed to maintain constant velocity while the observer makes at least one velocity change [9]-[12]. These constraints are used to develop contact observability that results in a consistent set of measurement equations from which the contact's state is estimated. In addition, the usual requirement for redundant data is needed to minimize the propagation of time delay measurement errors into contact state estimation errors. This situation leads to a class F type of estimation problem in which all of the encumbrances on the estimation process arise.

Three-Path Channel

The requirement for motion constraints on the estimation process may be relaxed in a type of problem which involves a third nonintersecting path linking contact to observer.
In this case, two time delays $\tau_1$ and $\tau_2$, are measured from which contact range $R$ and direction $\beta$ are obtained [13]

$$R = [L^2 - 0.5c^2(\tau_1^2 + \tau_2^2)]/[c(\tau_1 - \tau_2)]$$

$$\beta = \sin^{-1} [(c/2L)(\tau_1 + \tau_2) + (c^2/4LR)(\tau_1^2 - \tau_2^2)]. \quad (6)$$

Successive measurements yield contact velocity. This represents an estimation problem falling into classes D and E referred to earlier. Though the problem here is nonlinear, the relaxed requirement for observability yields a more stable estimator with faster convergence than the class F problem.

**Intersecting Ray Path Channel**

Thus far, nonintersecting sound ray paths have been considered. A fourth situation, commonly known as multipath problem, involves intersecting ray-paths [14]. For the simple two-path channel of Fig. 2(d), the time delay $\tau$ is related to range $R$ and depth $R_z$ of the contact relative to the receiver through the equation:

$$\tau = 1/c[(R^2 + 4H_0^2 - 4H_0 R_z)^{1/2} - R] \quad (7)$$

where $H_0$ is receiver or observer depth. This delay yields a contact direction in the vertical plane similar to the one in (4). As in the horizontal plane, the measurement equation contains two unknowns, $R$ and $R_z$. As with the nonintersecting three-path case, an additional path to the multipath situation (Fig. 2(e)) yields the desired observability or consistency of equations, leading to class D or E estimation problems. Unlike in the horizontal plane situation, however, it happens in practice that contact depth may be known (as would be the case with a surface contact). When contact depth is known, the presence of a third path is unnecessary to satisfy the observability condition. The third path would provide, in this situation, spatial redundancy which along with temporal redundancy would allow further opportunity to filter out errors in the contact’s state estimates. Depending on the situation, the multipath problem may belong to classes D, E, or F.

In practice, an integration of the preceding cases usually occurs with a merging of frequency and bearing data, active sonar time delay and bearing, bearing and depression/elevation angle, etc. The integration provides consistency to the measurement equations and improves the contact state estimation process. Differential Doppler between two paths, which provides an indication of time delay rate, is also considered. While time delay yields directional information when processed, time delay rate yields direct information on contact velocity. Recent research has addressed multisensor processing as well as the multicontact problem [15], [16].

**IV. MEASUREMENT AND MODELING ERRORS**

**Causes of Errors**

To convey the basic CLMA concepts, a homogeneous and noiseless ray path channel has been considered thus far. Such a channel yields a direct functional dependence between time delay vector $\tau$ and the contact state vector $x$

$$\tau = f(x) \quad (8)$$

where solution for the contact state $x = f^{-1}(\tau)$ is straightforward.

In practice, perfect or nearly perfect observations are rarely available. Vector errors $E_\tau$ are usually introduced due to the time delay measuring system, mismodeling of the environmental factors in the channel, mismodeling of the contact’s motion, or inaccurate monitoring of the observer’s own motion. Seldom is the spatial gain of a passive sonar system high enough to warrant neglecting these errors. Equation (8) must therefore be modified to account for the vector errors, such that

$$\tau = f(x) + E_\tau. \quad (9)$$

Regardless of the source of errors, their statistical characteristic influences selection of a particular contact localization and motion estimation process. In general, errors are characterized as either biased or unbiased, as discussed in the following text.

**Characterization of Errors: Biased or Unbiased**

This section delineates the sources of biased and unbiased errors in ocean-related CLMA problems. In the context of this paper, a biased error refers to the tendency of an estimated value to deviate from the true value in one direction. Biased errors may be constant or variable over a number of contact observation intervals. Constant bias may be due to differential dispersion in the channel paths (as between a volume and a bottom-reflected path), ray path curvature, or a non-Gaussian distribution of time delay estimates from the time delay processor for a low signal-to-noise ratio or low relative signal-to-noise bandwidth. Constant bias may be recognizable as a shift in the residual error between estimated and measured time delays; its effect may then be compensated for. Variable bias may be transient (as due to a contact maneuver) or persistent (as with a mismodel of the channel’s ray path curvature). Once a transient bias is recognized, adaptive control of the process noise may be successfully applied [17]. This amounts to effective reinitialization of the problem with some a priori information on the contact’s range. If a higher order motion model (one that allows estimation of a possible contact maneuver) is used, the estimator is more prone to instability, especially when only large unbiased errors are present. To deal with persistent variable bias requires a model of the process. If available, parameter estimation and process identification may be carried out with diminishing success if the estimation problem belongs to the later problem classer, i.e., classes E or F.

Unbiased errors may have Gaussian or non-Gaussian distributions. A Gaussian fluctuation of time delays may be due, for instance, to such effects as small perturbations in ocean sound speed profiles, or due to the ocean surface [18] or to the processing of time delays in the presence of limited noise. Even with a Gaussian error distribution on the time
delay estimates, their direct mapping into the desired contact states can result in increasingly non-Gaussian distributions as a function of the contact's range and off-broadside direction to the observer's array. From the estimation point of view, it is preferable to maintain an unbiased Gaussian distribution of errors, since this leads to manageable difficulties in the contact state estimation process. Many of the existing contact localization and motion estimator structures are designed on the basis of best unbiased mean square error reduction criteria.

Minimizing Errors

Minimizing errors in the time delay estimator and contact location estimator has relied principally on the assumption of unbiased Gaussian error distribution. This type of distribution is highly desirable since, as mentioned earlier, it makes problem analysis and solution implementation easier to accomplish than with a biased error distribution. However, biased errors do exist, as in the case of differential dispersion in a channel which may lead to "smearing" of often-used peak detectors. Such a situation would result in a biased estimate of time delay even if signal-to-noise ratio were good. When this type of bias results in large errors in the contact's state estimate, techniques such as variants on the complex demodulation technique must be used to remove the bias [19]; otherwise, the system is not useful as an estimator in that instance.

- Even for ideal channels with additive Gaussian noise, the distribution of time delay estimates becomes nonsymmetric as a function of decreasing signal spectra to noise spectra ratio [20], [21]. The resulting skewed distribution of errors is undesirable and has given impetus to the use of windowing and gating techniques to remedy the situation. Frequency windowing is incorporated into a basic time delay estimator to lower its threshold, while time delay gating is added to limit the search for the clue to the most probable region in the time delay estimator output. The induced stabilization of time delay estimates allow the usage of statistical estimators, such as linear weighted least square filters, to improve and assess the quality of the contact's state estimates.

Windowing

Windowing has been applied to the various types of time delay estimators that may be encountered in a CLMA problem. The specific time delay estimator used in a given problem depends on the number of sensors available and on the number of signal arrivals at each point [19], [22]–[32]. In one situation (Fig. 2(d), (e)), multiple acoustic propagation paths lead to intersection at a single sensing point. In this case, generalized cepstrum, autocorrelation, or complex demodulation techniques may be applied to interpret the resulting composite received data and measure the time delays. In another situation, propagation paths do not intersect at the sensing points (Fig. 2(a)–(c)). Here, spectral estimation or comparative signal analysis may be carried out at each sensing point, and generalized cross correlation, complex demodulation, or least square techniques carried out across sensing points.

Frequency windowing of time delay estimators has been the subject of extensive studies for both horizontal and vertical channels [28], [33], [34]. Addition of a properly designed window extends the region of satisfactory performance of a given conventional time delay estimator by lowering the estimators' operational threshold. An average improvement of 4 to 6 dB may be accrued. The windows are designed to remedy or compensate for physical conditions that affect unfavorably the performance of the estimator. They are dependent on signal spectra, noise spectra, and channel parameters. It should be stressed that windows must be designed to suit the estimator at hand and the situation under consideration, since improper windowing will deteriorate performance instead of improving it [33]–[35].

Gating and Filtering

The simplistic scheme of independently selecting the dominant clue for each time delay estimator output can deliver erratic time delay estimates whenever adverse but temporary conditions exist at the input [36]. For estimator initialization, some ensemble average over a number of successive time delay processor outputs can be taken to enhance the clue against mean background noise.

Where the clue is identified as having sufficient signal power over noise power, a time delay gate is centered at the corresponding output region, and clue estimation is executed over the gate output for each observation interval. The characteristics of the gate may be provided by a Kalman filter operating on the raw time delay estimates [20]. Such gating enhances the robustness of the processor against signal fades and limits the clue search to the most probable region in the processor output. A successful stabilization process of the estimates allows for automatic and quasi-optimal processing of the data to estimate contact location and motion. Furthermore, the linear Kalman filter for the time delays can detect easily contact maneuvers that yield a jump in the time delay rate, and can pass this information on to a Kalman filter that is estimating linearized contact state dynamics. The relation between time delay gates and spatial gates on the contact increases in complexity in line with the observability of the contact's state. For class A estimation problems, the two gates are directly proportional.

Statistical Smoothing

Even when time delays are estimated with unbiased Gaussian errors as would occur with high signal-to-noise spectra and long observation times (or as may occur following stabilization through windowing, gating, and filtering) direct mapping of the time delays into the contact's state can lead to biases in the estimation process. Reduction of this bias (and variance in contact state estimates) can be accomplished by judicious use of statistical estimation techniques over sequential and finite observations of the contact signal [13]. The contact state estimator is an expanding memory filter that maps imperfect time delay estimates into the invariant contact trajectory parameters (i.e., constant velocity, initial range) over which smoothing is performed. The smoothing reduces, jointly, the variance and the bias in the estimate of contact kinematic parameters. Such a scheme improves substantially on techniques that process inappropriately mapped time delays, or tech-
niques that directly transform the best time delays available into contact motion estimates. The latter approach is optimum only when stationariness of all elements in the problem can be assumed. For this limiting case, the approach using statistical smoothing converges automatically to the optimum estimates. Yet for generalized cases, it remains a viable approach for moving contacts at long ranges, for contact directions off the array’s broadside, and for high time delay variances.

Implicit in this discussion is a requirement for correct statistical descriptions of the processes at hand. The recovery from an incorrect statistical description in digital systems is aided by use of coupling loops for detecting such an event. The ensuing divergence is bypassed and the processes are routed in a degraded mode until the system recovers. When the traditionally separated signal and data processing stages are interactive [3], further improvement can take place because system deterioration is usually local and not total. Statistical smoothing as a means of minimizing errors is discussed further in Sections V and VI.

V. ELEMENTS IN THE FORMULATION AND SOLUTION OF CLMA PROBLEMS

Three elements need definition in the formulation of a contact’s state estimation process [17], [37]-[41]. These are encountered regardless of the class that the CLMA problem belongs to, and regardless of the application at hand. The three elements are:

1) a model of the relation between the contact’s state and the observables (i.e., time delays) as given in Section III,
2) a model of the contact’s state (e.g., stationary, constant velocity),
3) a criterion to filter out errors (discussed in Section IV) from the observables and models.

Of the various errors that are encountered, some are due to the time delay estimation process, some are due to the modeling of the channel, some are due to the presumed motion of the contact or the observer or both, and some are generated by the form of the data processing structure. Regardless of the error sources, filtering of unbiased errors has been dealt with collectively using varied estimation techniques. These include linear minimum variance, least squares, weighted least squares, maximum likelihood, and Bayes estimators. Performance of the resulting estimation procedures varies depending upon the available statistical descriptors. For Gaussian error distributions, the linear minimum variance estimates results agree with many of the others. In addition, nonlinear problems can be fitted through linearization, and minimum variance estimators can accommodate such cases with little or no knowledge of the probability density function of the errors. This latter characteristic explains the widespread use of linearization techniques since, more often than not, a probability distribution is merely conjectured.

As noted in element 1, the estimation problem begins by hypothesizing the functional relationship between received time delays and the contact’s state descriptors. Sensor diversity and channel diversity must be taken into account, since the number of measured time delays depends on the number of spatially separated sensors and on the number of intersecting ray paths in the sound channel (Fig. 2). Two time delays at each observation interval are needed to provide positional information on the contact. Synthetic diversity must also be considered; this refers to the orderly assembly of time delays estimated over successive observation intervals to enhance the available estimates and to provide otherwise unavailable estimates. With time delays that yield at each instant a single contact’s direction, the ranging relationship between moving contact/observer is quite circuitous and requires a series of time delay measurements combined with an observer velocity change.

Notwithstanding, the relational complexity, alignment of snapshot estimates of contact localization and motion requires a modeling of the nominal underlying processes. This calls for hypothesizing a dynamic model of the contact. Mismatches between real and modeled phenomena lead to biased errors, and estimates of these errors must be made along with estimates of the contact’s motion. Bias estimation remains a difficult problem, and bias due to the contact’s presumed motion has been most studied. In the underwater environment, the contact’s nominal motion is presumed to be predominantly constant in velocity interspersed with arbitrary maneuvers. The modeling presumes this type of motion with added unbiased perturbations to account for deviations on that motion. The perturbation input levels are varied to reflect the credibility in the evolution of the motion models. This control process is used in relation to the functional dependence of contact’s states upon observed time delays. Even when the contact’s motion model is inadequate, the evolution of the time delays has been modeled locally through nominal low-order polynomial expansions that prove helpful over a limited number of time delay estimates.

Given the contact’s dynamic model and the functional dependence of its state on the measured time delays, a criterion for “best” estimation of the contact’s states is chosen which yields the estimator structure. If a choice is made to minimize the average mean square error between estimated and true contact states, the procedure is a straightforward mathematical one applicable to varied situations. Other means to minimize errors, such as the maximum likelihood technique, can lead to insurmountable analytical difficulties for non-Gaussian statistics. The characteristics of the residual error between estimated and measured time delays is applied to weigh the adjustments on the contact’s states estimates until satisfactory minimization of the error is obtained. The residual error contains the cumulative error characteristics (biased or unbiased) which are sifted, either by an operator or automatically, so that the estimation process is conducted only on the data error characteristics that the estimator is designed to handle. Residual error characterization remains an active area of research; one in which detection of the bias has been stressed. Much of the attention has centered on adaptation to biasing caused by contact maneuvers [17], [42]-[45]. However, increasing attention is being paid to biasing due to sensor positioning [46] and environmental effects [14], and also on the effects of certain types of random errors [47], [48].
VI. BASIC CLMA ESTIMATION SOLUTIONS

Estimation of contact motion has been performed in a variety of specific applications based on observations of some indirect aspects of contact motion. In the ocean environment, perhaps the most familiar is the two-dimensional tracking of a noisy contact by using bearings-only observations. In this case, an observer monitors sequential bearings to a contact as it proceeds at a constant velocity. From these bearings, the observer estimates the contact’s range, course, and speed [12], [49]-[53]. Both contact and observer motions are presumed to be in the horizontal plane where the sequential bearing observations are gathered. Estimator convergence occurs only after a well-chosen velocity change by the observer. This requirement can be unwieldy, and may result in lengthy convergence time and unacceptable errors. The presence of another spatially separated sensor enhances the convergence process and has given impetus to contact localization and tracking by means of sensor arrays.

For three-dimensional tracking, it is well known in radar-sonar work that serious degradation of depression/elevation measurements is caused by multipath propagation. This occurs when the contact is at near-horizontal grazing angles or occurs within a beamwidth or so of a bounding surface. Several techniques have been investigated to reduce multipath errors, but these are generally ineffective, especially for contacts within one beamwidth of the bounding surface [54]. In the ocean environment, the difficulty with antimultipath techniques is compounded by the presence of two bounding surfaces and the focusing effect of volume inhomogeneity. Instead of aiming to overcome the multipath effect, one alternative is to capitalize on the resulting consistency in the system of equations relating the measurements to the contact’s position. More importantly, no limit on vertical beamwidth is set, thus relaxing the beamwidth constraint in many techniques. In the following sections, the contact motion estimators for the preceding problems are considered.

Noisy Bearings—Only CLMA

Bearings-only contact location and motion analysis represents a class F problem. It is a fundamental and well studied estimation problem in the underwater environment (Fig. 2(b)), and is the most often encountered and the most difficult to solve. Generally, contact velocity is assumed constant, observer motion is unrestricted, and contact and observer are assumed to be moving in the horizontal plane. The problem is inherently nonlinear because of the bearing measurements. Only three elements in the contact’s states are observable prior to an observer maneuver, and neither spatial nor channel diversities are available to develop a contact location estimation. Therefore, a synthetic sensor diversity must be developed to provide the observability and redundancy needed to filter out bearing errors.

Initial solutions to the bearings-only CLMA problem relied primarily on geometric constructions. With the introduction of the computer, it became possible for an operator to hypothesize a contact’s range, course, and speed, and then test each hypothesis until the resulting bearing from the hypothesized states fit the measured bearings in some approximate way. This technique remains viable when biased variants on the statistical description of the errors arise so that they are observable by the operator. Such manual techniques pervade many of the CLMA problems in the ocean environment. With statistical descriptors, automatic estimation algorithms may be applied. Recursive and batch processing algorithms are often used. Kalman filtering [17], [37]-[41] has found widespread application since it accommodates nonstationary process noise and more general types of contact/observer motion.

In the bearings-only problem, most of the estimation difficulties that could be expected to arise do so. To minimize filter divergence the most observable contact states must be identified and isolated from those whose observabilities are developed synthetically through a motion constraint. Also, the noisy bearings must not be submitted to nonlinear functions operations followed by expectation operators; this would lead to residual biases. Finally, mapping and smoothing must be performed over the contact’s state parameters with the longest time invariance. With these factors having been mentioned, some additional comments on the CLMA process, with references to the pertinent literature, are now presented.

The extended Kalman filter, formulated in a Cartesian state-space, can develop divergence problems caused by a premature convergence of the covariance matrix prior to the observer’s maneuver. Remedies for the divergence problem have been initially heuristic and call for rotation of the covariance matrix to align with the estimated bearing [11], or for a gating on the range estimates. Such techniques have yielded erratic results. Another technique calls for the use of a pseudo-measurement made up of the component of the correct range perpendicular to the measured bearing line. This measurement is linearly related to the contact state. This approach avoids the covariance collapse and ensuing divergence problem but produces biased state estimates that may not be negligible [49], [51], [55]-[67].

Recent approaches have considered the effect of the coordinate system and the location of the nonlinearity [50], [58]-[61] on the bearings only CLMA problem. It has been found that modified polar (MP) coordinates yield stable and unbiased estimates. The state vectors are bearing rate, range rate divided by range, bearing, and the reciprocal of range. The first three states are observable, while the fourth remains unobservable until an observer maneuver occurs. The degree of observability in the MP formulation is the reason for the resulting stability. The estimated range is separated from the covariance computation until the observer’s maneuver occurs. Unlike for linear filters, an appropriate choice of coordinate system is fundamental to the good performance of nonlinear filters that estimate contact states subject to observer motion constraints. It has been found that an indirect stability measure based on a bound for the decay rate of a Lyapunov function [61] yields, for the bearings-only Cartesian extended Kalman filter, the worst possible value for the stability criterion.

Finally, analysis of a two-sensor, omnidirectional array yields a contact tracking problem not in the horizontal plane. This problem is unobservable [62] prior to the observer’s
first maneuver. In addition, there is a sign ambiguity on the estimate of contact depth. An iterative, least squares algorithm was proposed to generate the contact estimate that uses the Householder transformation to solve the Gauss-Newton equations. For this tracking problem, other algorithm structures utilizing spherical [63] and MP polar coordinates have also been employed. Such algorithms have been used to process noisy conical angles only. As expected, their behavior is similar to those processing bearings-only measurements.

**CLMA from a Linear Array**

CLMA from a linear array, a class D problem, deals with the location and motion of a contact in the plane containing a linear array and contact [13], [64]–[68]. For the sake of simplicity, consider an array having three spatially separated elements (Fig. 2(c)). Two noisy time delays or two bearings to the contact are available at each observation of the contact over a short enough observation interval to permit the assumption of local stationarity. Bearing ambiguity as to contact position (to the right or left of the array) is considered resolved.

In contrast to the bearings-only problem where the observer monitors at each observation interval a single angular direction to the contact, in this instance the observer simultaneously monitors the direction from two spatially separated positions. Spatial diversity of the observer’s sensors yields a contact range estimate at each observation interval. Though the estimation problem remains nonlinear, the troublesome issues of contact observability are minimized. Those issues resurface as the contact range increases relative to the effective separation of sensors (sensor baseline) in the noisy ocean environment. As range increases, sensor baseline is, in effect, reduced so that reliance on synthetic aperture techniques again is required.

Time delay measurements are usually imperfect; this causes fluctuations in range and direction values, and subsequent errors in velocity estimates. When a Taylor expansion is carried out on the range and only the linear term is relevant, the mean values of contact range and direction are considered unbiased and their variance is a linear function of the time delay variance. For an effective sensor baseline, minimization of contact location variance leads to minimization of time delay variance. To effect this minimization, different windows are added to the basic time delays estimators with varying effectiveness [20], [28], [35], [69]–[73]. Such a ranging approach presumes stationary contact and sensor positions, as well as stationary signal and noise statistics.

This linear analysis is physically relevant at ranges close to the expansion point in the Taylor series and/or at small variances of the time delays. Bias in range becomes significant as the contact range increases, as the contact moves away from sensor array broadside, and as the time delay variance deteriorates with signal and noise conditions. When bias in range is not negligible, the relation between the variances of range to time delays becomes quite nonlinear. It has been found that the problems of range bias and variance with the limited observation intervals in the cross-correlator become intertwined. They must be minimized simultaneously through sequential smoothing of the time delays over successive observation intervals [74]. Otherwise, the bias can be substantial in various practical contact locations relative to the receiving array. Recently, this bias has been calculated in various forms [13], [74]–[76].

For a zero mean Gaussian noise on the time delays and a homogeneous channel, a lower bound on the range bias \( \langle R_b \rangle \) and corresponding variance \( \sigma_R^2 \) are [13]

\[
\langle R_b \rangle = (2a_T^2R^3/(L^4 \cos \beta), \quad \sigma_R^2 = (R_b)(R + 8 \langle R_b \rangle)
\]

(10)

where \( a_T^2 \) is the time delay noise variance. Equation (10) shows the explicit dependence of the range bias on time delay variance, contact range, and effective array length; it also shows that the range variance is inherently dependent on the residual bias. For the favorable conditions of Gaussianity and homogeneity, Fig. 3 illustrates a rapid deterioration in the contact range estimation process as a function of increasing contact range, off-broadside direction, and time delay estimator errors. Again, an improvement in performance requires that appropriate temporal processing be applied to develop a synthetic array aperture. Though the biasing issue has been explained in the context of CLMA from a linear array, it is relevant to both the noisy bearings-only CLMA and multipath CLMA problems.

Minimizing the error in the preceding range estimates has been accomplished for a single observation interval by increasing array length \( L \) and/or by minimizing the time delay variances. Practical considerations such as array dynamics, available space, and signal coherence eventually impose limitations on the permissible array size. There is therefore an interest in pursuing the alternative of extending the usefulness and effectiveness of an existing array by increasing the temporal processing gain. For variance reduction of a stationary contact, several maximum likelihood localization estimators have been developed [48], [77]–[81]. These techniques are optimal for negligible bias and sufficiently long observation times. Their results yield the most optimistic performance of the system and provide measure bounds for improvement possibilities.

In practice, signal and noise characteristics can slowly vary. Also, time delays from a moving contact may be considered only quasi-stationary over a finite observation interval. These constraints limit the observation time of the time delay estimators, hence deteriorating their performance from the optimal condition. Consideration has been given [82]–[86] to contact/observer induced nonstationarity on the time delays. Estimation of the Doppler effect has allowed some increase in the observation interval [87]. However, the interval must remain short enough so that the time delays vary according to a low order polynomial form. Complications arise due to the presence of noise and the unknown order \( K \) of the polynomial. The order \( K \) is not known a priori since it is a function of the relative range and the number of observation intervals. Over a limited number of observation intervals, however, \( x(t) \) is likely to vary in a linear or parabolic fashion and parameter estimation may be carried out with a short-memory filter [20], [36], [78]. This filter has other bene-
fits, since it can aid in the estimation of time delays through the design of a gating mechanism [20] or through a peak search in the correlator output. In addition, the resulting decrease in time delay variance allows an extended region of operation away from a given array before the need arises to precede the triangulation scheme by spatial gating. Such gating can be helpful in the estimation process when independent information is available to define the most probable region of contact location.

For further smoothing beyond the few observation intervals in the short-memory filter, the assumption of constant contact velocity is essential. There are only four unknowns to estimate over all the successive observation intervals. Then the noisy time delays are constrained within a processor to point to an estimate $\hat{R}_x(0), \hat{R}_y(0), \hat{V}_x, \hat{V}_y$ with a minimum mean square error. The highly expanded memory system provides the desired redundancy to yield an enhanced estimation of the unknown parameters. In such estimation problems, it is desirable to ultimately map the time delay observations onto the invariant and unbiased contact motion parameters over which smoothing is performed to reduce both the variance and the bias in estimating contact location [13]. This mapping imparts stationarity to the problem, thus allowing an effective increase in the averaging time of the localization system beyond that allowed in the time delay estimator.

When variants develop on the otherwise constant contact motion parameters, adaptive filtering techniques have been applied to transition the state estimates to the newly evolving invariant parameters. Accordingly, total contact motion is described piecewise, i.e., as consisting of nonmaneuvering portions and maneuvering portions. The maneuvering portions have been modeled as random velocity perturbations resulting in the use of adaptive filtering techniques [88]-[90], or as unknown but deterministic inputs resulting in the use of estimation/identification techniques [91]. The observability of the contact's state allows the application of estimation techniques to identify the bias due to the contact's maneuver. For the bearings-only problem, detection of the contact maneuver and adaptive reinitialization of the estimation process have been used most often.

**Multipath and Multisensor CLMA**

Tracking of a moving contact via noisy observation of multipath time delays has been made by a single sensor, as opposed to spatially separated sensors (Fig. 2(d), (e)). Depending on the details of the situation, this problem may belong to classes D, E, or F. Tracking in the vertical plane is affected by the ocean inhomogeneities. When the region of interest has a layered structure, precise ray path models must be derived from ray-path studies. When formulating the track-
ing problem, it is highly desirable to choose a simple mathematical model that describes the prominent characteristics of the propagation paths [92], [93]. In near-isospeed waters, one convenient alternative is to represent the actual sound speed profile by a series of constant gradients, and then to approximate the resultant by an effective sound speed; this replaces the curved ray paths with equivalent straight lines. However, it may happen that the chosen representation is not equivalent to the actual situation. This causes a bias in the time delay observation model, which translates into a steady state error or causes the filter to diverge [14].

One approach to compensate for the effects of medium inhomogeneity has been to generate real time propagation paths by means of modified, constant-gradient ray tracing algorithms. Ocean features can also be of help for time delay estimation and contact localization purposes. Certain features form patterns, and these can be used in conjunction with measured time delays to recognize the channel sub-space whose attached set of time delays matches those extracted by the time delay estimator [94], [95]. This leads to ray path identification and ultimately to an estimate of contact location and motion.

In multipath tracking, similar issues to those discussed previously arise as to observability and nonlinearity. A major difference, however, arises in the possibility of conducting, via single time delays measured recursively, contact range estimation without the requirement for an observer maneuver. This occurs when the contact’s depth happens to be known. In general, however, multisensor CLMA merges time delay measurements collected in various planes. Bearing and/or multipath time delay measurements are included in the estimation processors. The contact’s state estimation process is greatly aided by two multipath time delay measurements, which impart consistency to the observation equations at each instant of time. In contrast to the noisy, bearings-only case, tracking is performed in a three-dimensional frame with concomitant advantages [14], [89], [90], [95]–[97]. Principally, the depth parameter can be estimated, and the unwieldy requirement for a velocity change (needed for Kalman filter convergence) is eliminated. The convergence time for the filter is greatly diminished and, furthermore, the filter displays a low operational threshold. In addition, its stability is maintained when mismodeling exists in the observational or kinematical models.

With multipath-induced time delay measurements, an appropriate state vector can be defined such that an equivalent depression/elevation (D/E) angle can be related to the time delay measurement. Under this condition, algorithms for processing bearing and D/E angle have been developed. By employing bearing and D/E measurements, system observability requires either an appropriate azimuth maneuver or an observer depth excursion. CLMA performance has been assessed by examination of the Cramer–Rao bound for large range to array baseline geometries [9]. When compared to the bearings-only approach, the processing of both bearing and D/E measurements yields better range and range rate estimates. For the five-state problem (constant contact depth), the increased accuracy of the range estimate is due entirely to the improved range rate estimate. The degree of improvement in range and range rate estimation over bearings-only CLMA is proportional to the quality of D/E data or to the size of the D/E angle. When given perfect D/E measurements or large D/E angle with small error, the range variance for the five-state estimator approaches approximately 1/4 that of the bearings-only algorithm.

Recently, three-dimensional CLMA has been introduced using time delay measurements from an array that yields conical angle, D/E angle, and inverse range measurements. A modified spherical system arises that is applicable to the bearings-only maneuvering contact problem [63].

VII. SUMMARY AND CONCLUSIONS

This paper has provided a general perspective on contact localization and motion analysis (CLMA) in the ocean environment. It has presented representative CLMA problems, and in doing so illustrates that the difficulty in solving these problems increases as the geometric relationship between contact and observer becomes nonstationary, as the equations defining the problems become nonlinear, and as constraints are placed on observer and contact motion. Various sources of estimation errors are discussed, and the ability to identify, characterize, and control them is shown to be a significant part of the overall estimation process.

The total gain in CLMA systems is the result of spatial, environmental, and temporal factors. That is, gain depends on a system’s spatial aperture (array size, number and placement of sensors, etc.), on the sound ray paths in the propagation channel, and on the temporal processing techniques used to convert the sequentially received signal time delay information to contact state estimates. The spatial aperture and propagation channel provide, with temporal processing over each basic observation interval, rough estimates or snapshots of a contact’s location and motion. Although the quality of these snapshots is not necessarily enhanced by successive observation intervals, temporal processing aligns, superimposes, and filters them to enhance the contact’s state estimates. Such temporal processing techniques are seen to yield an increasingly greater percentage of overall system gain as contact observability decreases, as signal-to-noise spectra deteriorate, and as contact range to array size increases.

ACKNOWLEDGMENT

The author thanks Dr. S. Nardone, K. Gong, and W. Golembewski for their comments on the manuscript.

REFERENCES


—, “Properties of a nonlinear estimator for determining position and velocity from angle-of-arrival measurements,” in Proc. 14th


Dr. Hassab is a member of Eta Kappa Nu, Phi Kappa Phi, Tau Beta Pi, Pi Mu Epsilon, Sigma Pi Sigma, and Chi Epsilon. He is listed in the 1976 edition of Who’s Who in the United States. In 1979, he chaired the Office of Naval Research Conference on Time Delay Estimation and Applications and has been Session Chairman at several IEEE and NATO conferences.

Joseph C. Hassab was born in Jounieh, Lebanon, on January 1, 1941. He received the B.S. degrees in electrical engineering and in civil engineering, the M.S. degree in electrical engineering, and the Ph.D. degree from Drexel University, Philadelphia, PA, in 1966, 1967, 1968, and 1970, respectively.

From 1970 to 1971, he was an Assistant Professor of Physics at LaSalle College, Philadelphia, PA. In 1971, he joined the Naval Underwater Systems Center, Newport, RI. While at NUSC, he has conducted research and consulting resulting in approximately one hundred journal publications in the varied aspects of signal processing, target motion analysis, scattering, wave propagation, channel modeling, and numerical analysis. He has been, respectively, the Head of the Advanced Systems Analysis and Assessment Division, the Advanced Development Division, the Advanced Systems Division, Acting Department Head, and Chief Engineer. Presently, he heads the Systems Architecture and Targeting Division. Since 1973, he has been an Adjunct Professor at Roger Williams College, Bristol, RI, and since 1979 at Southeastern Massachusetts University, North Dartmouth, MA, teaching courses in radar, sonar, speech and seismic signal processing, electromagnetics, engineering instrumentation, digital signal processing, and control systems.