Application of Sequence Comparison Techniques
To Multisensor Data Fusion and Target Recognition

E. W. Libby1 and P. S. Maybeck
Air Force Institute of Technology
Wright-Patterson Air Force Base, OH 45433-7765

Abstract

A new class of techniques for multisensor fusion and target recognition is proposed using sequence comparison by dynamic programming and multiple model estimation. The objective is to fuse information on the kinematic state and "nonkinematic" signature of unclassified targets, assessing the joint likelihood of all observed events for recognition. Relationships are shown to previous efforts in pattern recognition and state estimation. This research applies "classical" speech processing-related and other sequence comparison methods to moving target recognition, extends the efforts of previous researchers through improved fusion with kinematic information, relates the proposed techniques to Bayesian theory, and applies parameter identification methods to target recognition for improved understanding of the subject in general. The proposed techniques are evaluated and compared to existing approaches using the method of generalised ambiguity functions, which leads to a form of Cramér-Rao lower bound for target recognition.

1 Introduction

Several authors in the multisensor fusion field have proposed using kinematic and nonkinematic information for "observation-to-track" assignment or target recognition [3,177-178] [7]. In [2,297-320], Mitis proposed the use of linear estimation techniques with target state vectors containing both kinematic and nonkinematic states, but his development treated these states as independent. In contrast, the multisensor research discussed here seeks to exploit the fact that nonkinematic or "signature" states, which are in general nearly direct functions of target-to-sensor aspect angle, are tightly coupled to target kinematics, in particular for turning targets. We are also interested in state-measurement relationships that are not well modelled with linear estimators.

The key step in this research was to consider the relationship between true and estimated kinematics and signatures for a track-propelled vehicle (e.g., a tank) in a turn. Unlike wheeled vehicles, tracked vehicles do not turn with a constant or even continuous radius of curvature, and in fact their motion can be described as nearly piecewise linear [12]. Consider the process of tracking a tank in a planar turn. Our sampled-data sensor provides a sequence of signature vectors, as measured at discrete times over the observation period. Concurrently, from conventional range/angle tracking and kinematic state estimation alone, for any feasible target class, we can hypothesise a sequence of anticipated signature vectors. Comparing the observed sequence to the kinematically-estimated sequence for the correct target, we see that their differences can be described fundamentally in terms of expansions and contractions of one sequence relative to the other. Can we compare these two sequences such that their origin target classes are seen to be identical, despite expansions and contractions?

This is a well-known problem in speech recognition, where the technique of "Dynamic Time Warping" (DTW) using dynamic programming (DP) [6, 12] for sequence comparison evolved during the 1970's [27, 28, 30]. The objective in that field was to quantify the similarity of two spoken words, despite pronunciation differences. For DTW comparison, words are represented by vector sequences, a line, but varied length, extracted from speech by any of several algorithms [27]. As in the "turning tank" example above, the differences between sequences describing the same word from different sources may be classified as expansions, contractions, and perhaps insertions and deletions [36].

As we began to develop DP-based sequence comparison for moving target recognition, the Target Recognition Technology Branch (WL/AARA) of the USAF's Wright Laboratory requested that the technique be applied to aircraft recognition using high-resolution resolution radar (RRRR). An exhaustive literature search was conducted, covering the areas of multisensor fusion, dynamic programming, and pattern recognition. The text by Shao and Kruskal [30] demonstrated the wide utility of what we will call "classical DP-based sequence comparison techniques, including DTW. Efforts by Barniv [235-154] illuminated a DP algorithm originally developed by Larson and Peschon [18] for state estimation. This "Larson and Peschon (L&P) algorithm" and classical DP sequence comparison methods to assign sequences of signatures to the correct target classes.

The second source was a classified paper [29] by Mieras et al. of Raytheon, whose approach appears to be fundamentally identical to Le Chevalier's, although independently conceived. It will be shown in this paper that the Le Chevalier and Mieras approaches can be posed as "suboptimal" applications of the L&P algorithm, and it appears that published applications of this algorithm provided the original inspiration for Mieras et al.

2 Classical Pattern / Target Recognition

Classical pattern recognition [14] generally follows two basic approaches — "decision theoretic" (i.e., parametric or statistical, and nonparametric or nearest neighbor concepts) and "syntactic". Syntactic methods consider the form of the symbols in a way that is independent of the data, whereas decision theoretic classifiers are dependent on the data. In practice, decision theoretic classifiers are often limited to aspect angle "windows" of given solid angle extent. Multiple signature observations can be used with classical techniques (e.g., Bayes' Rule, voting rules, etc.) to provide better estimates of class membership for unknown targets.

For any set of signature realisations, this matching process inherently defines a "maximum likelihood" (ML) estimate of target aspect angle (a peak estimate) over time for each library class. This ML aspect angle history for each candidate class contains much useful information, but evidently has never been used explicitly for classification purposes, although some have suggested [4]. Classical decision theoretic classifiers are ill-suited to use this information. It will be seen that the Le Chevalier and Mieras approaches use this information implicitly, and our research has explored other means of doing so.

3 Sequence Matching By DP

The Dynamic Time Warping (DTW) and Larson and Peschon (L&P) algorithms form the core of the moving target recognition approaches principally...
investigated by the authors to date. This section discusses them, their similarities and their difference.

3.1 Classical Sequence Comparison / DTW

In classical sequence comparison, generally all that we require of the sequence elements is that some distance metric exists by which one element can be compared to another. Each element will represent a discrete, sampled data representation from some feature space. The features generally represent observable quantities due to a physical (i.e., classically continuous) process or trajectory in some state space, where the true location in the state space at any time is unknown - the distinction between the state space of this trajectory and the feature space of the observables is an important one, and not always clear in the literature. Pausing to note that classical DP sequence comparison is a large class of algorithms including DTW, in the interest of space hereafter we will use the term DTW to refer to classical sequence comparison in general.

"Warping" or sequence comparison is the process of making associations between individual elements in the two sequences, computing the cost of each association according to the distance (measure of dissimilarity) between the element in one sequence and the element in the other, and finding the set of associations that gives the minimum total cost or distance. Associations are made subject to "continuity constraints," that limit, for example, the number of associations that can be made from one element of one sequence to elements of the other sequence, the number of elements that can be skipped, and so on. Continuity constraints prevent undesirable low cost associations between two sequences that really have significant differences.

This process can be posed as finding the minimum cost path through a space of associations, subject to transition constraints, and forward dynamic programming [12:10-11] provides a natural approach to determine this path. For the simplest form of "local" continuity constraint, the forward dynamic programming cost computation at each step can be written as:

\[ D(C_k) = d(c_k) + \min D(C_{k-1}) \]

where:
- \( c_k = (a_k, b_k) \) is the \( k \)-th element in a sequence of associations of elements from sequence \( A \) (\( m \) elements in number) with elements of sequence \( B \) (\( n \) elements in number), the association being between element \( a_k \) and element \( b_k \)
- \( C_k = (c_1, c_2, \ldots, c_k) \), the minimum cost sequence of associations leading to and including association \( c_k \)
- \( d(c_k) = \) the cost or distance of association \( c_k \), i.e., the distance in some sense between element \( a_k \) and element \( b_k \)
- \( D(C_k) = \) the total cost of reaching and accomplishing association \( c_k \) by the minimum cost sequence of allowable associations

For instance, if we consider representations, distance metrics, and path constraints have been the subject of much experiment, without identification of one particular "best" approach [27:297-303] [30:125-161]. A particular issue faced by DTW researchers is that of warping path length compensation - inherently, the algorithm above is biased toward solutions with a minimum number of associations. Where a substantial difference exists between \( m \) and \( n \), or where \( m \) and \( n \) are equal but the optimal association "paths" is highly nonlinear, classical DTW may be hard pressed to choose that association.

3.2 The Larson and Peschon (L&P) Algorithm.

Larson and Peschon proposed an algorithm [18] for estimating the sequence of \( n \) states or locations in some space with maximum a posteriori or MAP probability of producing an observed sequence of \( n \) measurements, conditioned on a priori information about transitions in the state space. They did not motivate their work as a tool for target recognition working on an aspect angle space, but we will apply it in this fashion.

Given a sequence of measurements \( X_n = \{x_1, x_2, \ldots, x_n\} \), Larson and Peschon wish to find the sequence of states \( X_n/b = \{s_1, s_2, s_3, \ldots, s_n\} \) that maximised the conditional probability density function:

\[ p(X_n/b | Z_n) = \max \{ p(x_1, x_2, \ldots, x_n | s_1, s_2, \ldots, s_n) \} \]

where the term "\( \max \)" refers to the operation of finding the maximum value of the indicated term, over all values of \( X_n \), representing the sequence of states \( \{x_1, x_2, \ldots, x_n\} \). Note, as do Larson and Peschon, that this is not to estimate the entire sequence up to the present, rather than simply the present state \( s_t \).

Next, Larson and Peschon were willing to assume independence of measurements \( x_t \) from states \( x_t \) and measurements \( x_t \) for \( t \neq t \), implying, for example, that the time interval required to take data for one measurement is less than and synchronized with the latter time in any one state, and that the measurement instrument is independent from event to event. With this assumption, then, they used Bayes' Rule to break the maximisation process into stages, making it suitable for solution by dynamic programming using the following equations (use of which will be discussed below):

\[ p(X_n/b | Z_n) = \begin{cases} \max \{ p(x_1 | s_1) \} \\ s_1 \end{cases} \]

which shows the final step in the process, a maximisation of \( p(x_t | b_t) \) over all possible final states \( s_t \), and:

\[ p(s_t) = \max \{ p(x_t | b_t) \} \]

Then, stepping theoretically to a hypothetical \( k \)-th level step:

\[ p(s_t | s_{t+1}, k+1) = \max \{ p(x_t | b_t) \} \]

or, equivalently, in the recursive form which is the heart of the algorithm:

\[ p(s_t | s_{t+1}) = \max \{ p(x_t | b_t) \} \]

The above equations are used in an iterative forward dynamic programming procedure which works as follows (from [18] with elaboration):

1. Quantise the state space \( [x_1] \) to obtain a grid consistent with the accuracy requirements of the problem.

2. Initialise the forward iterative procedure by defining \( I(x_0, 0) = p(x_0) \), the a priori density for each discrete \( x \) at time \( t_0 \).

3. For each quantised state \( x_1 \) (i.e., each possible discrete \( x \) at time \( t_1 \)), calculate \( I(x_1, 1) \), from \( x_1 \) and Eq. (6), with appropriate subscript changes for stage 1, rather than \( k+1 \).

4. Write \( s_0(x_1, 1) \) as the value of \( x_0 \) for which Eq. (6) is maximised in the previous calculation (establishing "pointers" which will be retraced to find the optimum state sequence).

5. Repeat steps (3) and (4) at each sampling instant until the \( k \)-th instant is reached. This repetition is one stage. This is the iterative forward dynamic programming procedure, moving forward through successive stages.

6. Determine the modal trajectory \( X_k/n \) by first using Eq. (3) to find \( s_k/n \) (i.e., the state with highest probability of being the terminus of the true state sequence) and then iteratively retracing the pointers set up in step (4), to find the optimal state sequence, i.e., \( s_k/n = X_k/n(s_k/n, k+1) \).

Since the factor \( p(s_k/n | X_k/n) \) is the same for all maximisations made at any time \( t_k \), the actual maximisation at any stage need not be done over the term shown in braces in Eq. (6), but rather only over the expression defined by computing this term without its denominator, denoted \( I^r(x_k/n, k+1) \).

3.3 Relating DTW and the L&P Algorithm.

Both DTW and the L&P algorithm are DP sequence comparison techniques. The fundamental difference between them is that DTW does not consider state transitions that occur off a single "one-dimensional" path in state space. In the usual DTW case, we have little knowledge of the underlying state space - only examples of the feature sequences produced by typical state trajectories. Observations from one state trajectory are simply compared to observations from another trajectory, and "warped" to allow for an optimal match. DTW generally attempts to associate an element of one sequence with more than one element of the other sequence, leading toward a bias for solutions that minimise the total number of associations.

On the other hand, the L&P algorithm can use information known a priori, or aside from the feature observations, about the likelihood of transitions in the state space. This allows the L&P algorithm to "investigate" more than one state trajectory. The L&P algorithm does not attempt to match more than one state space point with a given element in the feature sequence, and thus has no "arithmetic" bias toward short paths in the state space.

The drawback to the L&P algorithm is its "maximum likelihood" (ML) nature, in the sense that, given a set of \( n \) observations, it finds the set of \( n \) discrete states most likely to have generated the observations, subject to a
priori constraints: \( p(x_0) \) and \( p(x_{t+1} | x_t) \). It may be, however, that a state space region exists which has a higher overall probability of producing the given observations, when all possible trajectories over time through that region are considered. By comparison with a region chosen by the L&P algorithm, this "better" region might have many points which are rather likely to have originated the given observations, while the "L&P" region has a few well-positioned points which are very likely origin, but many that are quite unlikely. The use of DTW in such a case, forcing each point along a likely state trajectory to associate with an observation, could select the "better" region instead of that selected by the L&P algorithm.

Unfortunately, while the L&P algorithm can use the (relative) computational economy of DP to find the ML sequence of states in a state space of arbitrary dimension, the state space region with highest probability of generating the observed features can in general be found only by exhaustive search. A set of nominal or a priori likely trajectories through the state space would provide a starting point for such a search with DTW methods. The next section develops these ideas formally.

4 Moving Target Recognition: Theory

Using the classical Bayesien approach, we really desire to produce a pattern recognition system that estimates the a posteriori probability \( p(\omega_t | \mathbf{z}_t, \mathbf{z}_{<t}) \) that to be equal to \( \mathbf{z}_t \) is a target class \( \omega_t \), given a set of \( K \) trajectories \( \mathbf{z}_t = \{ \mathbf{z}_1^t, \ldots, \mathbf{z}_K^t \} \), a kinematic measurements \( \mathbf{z}_t^\text{kin} = \{ \mathbf{z}_1^t, \ldots, \mathbf{z}_K^t \} \) (note use of superscript \( d \) as in "dynamic", since \( d = 1 \) is a counting index in the L&P form), and a priori target class probability \( \pi(\omega_t) \) for each of \( J \) known target classes.

This ideal, but in practice unobtainable, system would consist of \( J \) functions, one for each target class, having a domain of the space of all measurements \( \mathbf{z}_t \) and time \( t \) and a range of the interval on the real line from zero to one, with the sum of the \( J \) function values equal to one (or less than one, if we wish to allow for unknown classes). Following Rao [28-33], however, in the absence of \( p(\omega_t | \mathbf{z}_t, \mathbf{z}_{<t}) \) (equivalently, the joint probibility density of target class and measurements), we are content to find a set of "generalised" likelihood functions such that the maximum value for each function is attained for the correct combination of target class and measurements. Note that the target signature libraries and aspect angle "words" used in classical ATR are likelihood functions.

What can we gain by considering the kinematics of the unknown target? Consider an abstract space \( \Theta \) of all possible target models and aspect angles over time as part of the domain of a matching function, \( \mathcal{L} \). A classical ATR likelihood function for a given target class \( \omega_t \) is defined by restricting the domain of \( \mathcal{L} \) to produce an \( L_t \) with domain \( S_t \subseteq \Theta \) corresponding to \( \omega_t \). Typically, we match sets of noise-corrupted feature observations (from another abstract space \( \Theta' \) which forms the remainder of the matching function domain) over time to elements in the first space. Unfortunately, these classical decision theoretic functions may give higher likelihoods than ideal for the wrong target class, in part because kinematically unlikely aspect angles and aspect angle transitions over time are allowed.

The key to the proposed approach is to restrict the domain of each function \( L_t \) further, requiring the target aspect angle over time to be consistent with the observed kinematics, since this restriction (correctly executed) should not adversely affect function values for measurements from the correct target class, but may lower the values for measurements from incorrect target classes. Note that for any likelihood function corresponding to target class \( \omega_t \), with the kinematically unrestricted and restricted matching domains denoted respectively by \( S_t \) and \( T_t \), we can show by contradiction that:

\[
\text{for } T_t \subseteq S_t, \quad (7) \sup_{\mathbf{z}_t} \mathcal{L}_t(T_t, \mathbf{z}_t) \leq \sup_{\mathbf{z}_t} \mathcal{L}_t(S_t, \mathbf{z}_t) \leq (8)
\]

If the kinematic restriction is done properly, and the measurements \( \mathbf{z}_t \in \mathcal{Z} \) do arise from class \( \omega_t \), the restricted domain should still include the region of highest original likelihood, and the restricted likelihood function values should tend to be equal to the unrestricted function values. If the measurements come from some other class \( \omega_t \), however, the restricted likelihood function values are more likely to be reduced. Thus, restricted likelihood functions promise better recognition. As we will show, the Le Chevalier and Miese approach moves in this direction by restricting the matching function domain to be consistent with feasible kinematics, or, in a suboptimal fashion, consistent with observed kinematics. By further, optimal restriction using observed kinematics, we will achieve a more highly "tuned" likelihood function (by analogy with a matched filter in the frequency domain). Restricting the matching function of the likelihood function according to kinematics is the under conditioning \( p(x_{t+1} | x_t) \), and may be known, on the observed information given by kinematic measurements \( \mathbf{z}_t^\text{kin} \).

5 Target Aspect Angle From Kinematics

The relationship between aspect angle and kinematics is strong for many target classes and has been exploited in sensor-augmented target trackers [17, 1]. In conventional aircraft in particular, aspect angle changes generally precede observable kinematic changes, as an aircraft rolls and creates an angle of attack for lateral acceleration to change its flight path. For this reason, researchers usually seek to improve kinematic state estimates using aspect angle information from pose estimates. In the sense of L&P, however, we will use kinematic information to define the "a priori" probabilities \( p(x_{t+1}^\text{kin} | x_t) \) for transitions in an aspect angle state space (hence the superscript \( a \) on \( x_t \)), or in the DTW sense, to define nominal trajectories through that state space.

As implied in the previous paragraph, aspect angle estimates from standard kinematic trackers using prior and current measurements may well be in error by 90 or more degrees at the start of a maneuver. On the other hand, most conventional aircraft (and other target classes also), once committed to a turn, will generally hold that turn for several seconds to gain the desired trajectory change. Thus, to determine the state of an aircraft at some time \( t \), we can use kinematic information not only prior to \( t \), but for several seconds afterward as well. Since our intention is to gain all possible information about the state of a turning aircraft, generally using information from a conventional extended Kalman filter (EKF) tracker, a natural step is to process the EKF outputs through an optimal fixed lag smoother (FLS) [28-18-17]. This was the basic approach in our research, using outputs from a tracking radar simulation performed using the "Multimode Simulation for Optimal Filter Evaluation" (MOSFE) software [9], postprocessed by a FLS with a fixed lag of two to three seconds.

Fig. 1 shows EKF and FLS mean performance in estimating one inertial component of target acceleration over 20 runs, where the true target acceleration is two \( g \) (64 ft/sec\(^2\)) during the period from three to eleven seconds and zero elsewhere. The upper solid curve is mean EKF error, while the lower solid curve is mean FLS error. In practice, although the FLS did an excellent job of correcting EKF state estimates as shown above, acceleration estimates were still too noisy to provide smooth aspect angle estimates, particularly in state directions where insufficient true acceleration made use of the optimal smoother pointless [23-11]. Thus, second-order polynomials were fitted to the filter/smoothing position estimates, and differentiated twice to obtain an acceleration estimate with error magnitudes that closely follow the mean FLS value in Fig. 1.

![Figure 1: Kinematic State Estimation Improvement by Smoothing](image-url)

The disadvantages of smoothing are added processing and the fact that our target information is no longer real-time. In general, we found that a high (e.g., 70%) estimate of the target acceleration was obtained with a four-second delay -- two seconds for the fixed lag smoother and two seconds for polynomial curve fitting. Following onset of a major maneuver, 2-3 more seconds of delay are desirable to identify steady state conditions (note how the FLS mean error curve in Fig. 1 begins to level out near the five second point). In any case, as shown in Fig. 1, for a 2-g turn lasting as little as eight seconds, the target acceleration can be estimated with high confidence for approximately five seconds. As we will show below, the advantage accrued in better position and velocity state information can be well worth the wait and processing, particularly for turning accelerations in excess of 1 g.

Once the target velocity and acceleration states are known and assumed to be in steady state relative to the target body frame, calculation of target-
sensors aspect angle and aspect angle rate is straightforward for any set of assumptions on target control parameters. In our research to date, we have assumed a conventional aircraft's coordinated turn motion— for any control method in which the plane of the wings is essentially normal to the lift vector—, we found deviations from the coordinated turn dynamics result only in an aspect angle position bias error which is ignored by our algorithms.

The kinematic state covariance estimate from the EKF/smoothers allows one to estimate the covariance of the kinematic aspect angle and aspect angle rate estimates, by use of the quadratic form:

$$P_{x} = E P E^{T}$$  \hspace{1cm} (9)

where:

- $P_{x}$ is a $4 \times 4$ matrix, a first order (linearized) covariance estimate for the error in the angular position and angular rate of the nominal aspect angle, in the direction of and normal to the nominal aspect angle path.
- $E$ is a $4 \times 6$ (row) by $6 \times 6$ (column) Jacobian matrix of partial derivatives, defined by determining the partial derivative of the angular position and rate along and normal to the nominal aspect angle path with respect to the target velocity and acceleration components along each inertial frame axis.
- $P$ is a $6 \times 6$ matrix, the filter/smoothers-estimated covariance of the target inertial velocity and acceleration estimates.

Straightforward extensions of this technique allow for calculation of angular state error “covariances” due to other variables. Due to unmodelled factors, the quantity in Eq. (9) can be treated as simply a lower bound on the true error covariance and a departure point for tuning.

6 The L&P Approach and $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$

The purpose of this section is to apply Bayes’ Rule [22], the L&P methodology (Sect. 5.2 and [18]) and aspect angle state transition information given by kinematic state estimates, to provide a posteriori probability $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$ for each class $w_{i}$, or a reasonable replacement.

Consider a set of $J$ a priori known target classes $w_{i}$, each represented by a target model having appropriate signature distributions associated with each aspect angle value. Given some discretization of continuous aspect angle on the targets (assumed the same for all target classes), each discrete aspect angle value is considered a state $x^{s}$. Now, for any given problem (i.e., any given measurement set $X_{m}^{k}$, $Z_{m}^{n}$) over some time interval, we can restrict our concern to a given aspect angle “window” or region— that is, we assume a negligible probability that the class presented aspect angles outside this region over the duration of the time interval corresponding to measurements $Z_{m}^{k}$.

Due to the smoothing process, the time interval corresponding to the kinematic measurements will generally contain the time interval of the signature measurements. The regions or windows may not be identical from class to class.

Now, define the super-region $X^{s}$ as the superset of all aspect angle cells or states that belong to the region of consideration for at least one target class, a total of $N_{c}$ cells or states in number. Any set of $k+1$ aspect angle cells, or aspect angle state history, corresponding for analysis purposes to discrete locations at signature measurement times along an aspect angle sequence which yields the $k$ aspect angle measurements $Z_{m}^{k} = [z_{1}, z_{2}, z_{3}, \ldots, z_{k}]$, will be denoted $X_{m}^{k} = [x_{1}, x_{2}, x_{3}, \ldots, x_{k}]$ (where $x_{k}$ is an a priori or starting state and the other $k$ states correspond one-for-one to the signature measurements $Z_{m}^{k}$). The (finite) number $N_{k}$ of possible such aspect angle sequences through $X^{s}$ is given by the number of permutations of $N_{c}$ things taken $k+1$ at a time, with replacement, or $(N_{c})^{k+1}$ sequences. We will denote the set of all such sequences as $X_{m}^{k}$. Henceforth, this development will refer to a particular “true” sequence of $k+1$ states as $X_{m}^{k}$ (consistent with the notation $w_{i}$ referring to an $i$th target class). Clearly, from the definition of $X^{s}$, some of these sequences $X_{m}^{k}$ are of negligible probability for one or more target classes, because they fall outside the subsets of $X^{s}$ appropriate for those classes. Other sequences are of negligible probability for all classes because they are kinematically unlikely.

We will see that applying the L&P equations (with reasonable modifications for target recognition) for any one model $w_{i}$ gives the particular state history $X_{m}^{k}$, any $X_{m}^{k}$ (i.e., the L&P estimate of the aspect angle sequence for model $w_{i}$), which maximizes the conditional probability $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$. With appropriate modifications, we will be able to find the joint conditional probability $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$, which we will sum over all possible $X_{m}^{k}$ to find the quantity that we desire, $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$. The object here is to understand the relationship between (1) the information given by the L&P approach, i.e., $X_{m}^{k}$, for a particular $w_{i}$, and a joint conditional probability associated with that state sequence and (2) the information that we want, $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$.

Further assumptions are:

1. Following L&P, assume that $s_{f}$ is independent of $x^{s}$ and $s_{f}$ for $t_{j} \neq t_{f}$.
2. If it should be clear, and kept in mind during this development that, for any $t_{i}$, $p(Z_{m}^{k} | X_{m}^{k}) = p(s_{f} | Z_{m}^{k})$, and analogously that $p(X_{m}^{k} | Z_{m}^{k}) = p(s_{f} | X_{m}^{k})$.
3. We assume that the "a priori" aspect angle state (cell) transition probability $p(X_{m}^{k} | X_{m}^{k-1}, Z_{m}^{k}, Z_{m}^{n})$ and probability of starting cell location $p(x_{1} | Z_{m}^{k}, Z_{m}^{n})$ are given by the EKF/smoothers determination of mean aspect angle, aspect angle rate, and associated covariance estimates as discussed in Sect. 5 (i.e., from kinematic information). Alternate approaches are discussed in [20].
4. We assume that $p(\omega_{i})$ (a priors) is known for each target class $w_{i}$, and furthermore that $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$, that is, that the kinematic measurements and derived kinematic state history provide no information as to the nature of the target. This last assumption is clearly neither true nor desirable if characteristic trajectories for various target types are classified probabilistically, and in that event another application of Bayes’ Rule will incorporate this information into the ATR decision. In this development, however, we wish to assess recognition improvement due to DP sequence comparison methods only, so all targets are considered equally likely to have executed the observed measures.
5. For the high range resolution radar signature case, we assume that uncertainties in range bin alignment and scale factor uncertainty are handled by finding $MAX(p(s_{f} | x^{s}, w_{i}))$ for any combination of signature measurement $s_{f}$ and trial aspect angle state $x^{s}$ on any model $w_{i}$, essentially following the “maximum likelihood” method discussed in [33], and used in [2, 25].

6.1 Relationship of $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$ to $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$

Recall that Larson and Posech sought the state history or sequence $X_{m}^{k}$ to maximize $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$ in some general state space. Analogously in our case, trying to find a "best" state sequence in aspect angle space over some model $w_{i}$, we might seek a state history to maximize $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$. Following the L&P approach with the above assumptions, we find (where the denominator term is given in the usual fashion by summing the numerator expression over all $X_{m}^{k}$):

$$p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n}) = \sum_{X_{m}^{k}} p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$$

$$p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n}) = \sum_{X_{m}^{k}} p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$$

Thus for any given target model $w_{i}$, we can conceptually use the L&P approach to find the set of states $X_{m}^{k}$ which maximizes $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$. So far, we have defined L&P-like conditional probabilities for the aspect angle state corresponding to one target class $w_{i}$. Now, we consider the a priori probability of class membership $p(\omega_{i})$, and define the desired a posteriori probabilities $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$. Assuming (with reservations as discussed above) that $p(\omega_{i}) = p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$, we start with the a priori probability $p(\omega_{i})$ and then multiply by $p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$, and continue as in the L&P development, to obtain in an analogous fashion (where the denominator is obtained by summing the numerator over all target classes and state sequences):

$$p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n}) = \sum_{X_{m}^{k}} p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$$

Now sum Eq. (11) over all possible $X_{m}^{k}$ for any given $\omega_{i}$ to obtain:

$$p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n}) = \sum_{X_{m}^{k}} p(X_{m}^{k} | Z_{m}^{k}, Z_{m}^{n})$$

Thus, the desired $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$ can be found rigorously only by keeping track of, and performing appropriate calculations for, all possible aspect angle sequences over all possible target models— that is, all $X_{m}^{k}$ over all $w_{i}$ (an exhaustive computation).

6.2 Relationship of $I^{*}(s_{f}, k)$ to $p(\omega_{i} | Z_{m}^{k}, Z_{m}^{n})$

Recall that the L&P method as implemented on different target models $w_{i}$ would find the $X_{m}^{k}$, say $X_{m}^{k}$, for each $w_{i}$, that maximizes
would have attempted, and iuggat that sukcguent prhng a university, to note that the oppdte direction to that conaieknt with the ability to describe the opposite direction to that conaieknt with the appropriate neighborhoods around the points in the "hat" over where, using the appropriate limiting process, we would have to do to find the denominator in Eq. (10). Examining \( I(x; \omega, k) \) closely, note that the preceding equation is equivalent to:

\[
\max I'(x; \omega, k) = \max p(x_{t+1} | z(t+1), z(t), \omega)
\]

Maximizing this quantity rather than the conditional probability is desirable because we avoid having to compute values for all \( X_t \in X_t \), which we would have to do to find the denominator in Eq. (10). Thus, it seems clear that one may need to accept delays of up to a few seconds and some form of smoothing to provide any reliable aspect angle estimates. If the target is determined to be turning during this period, that kinematic information can be used.

Recalling our comments in Sect. 3.3, it is also clear that the potential shortcoming of a L&P-type approach as applied to target recognition (e.g., Eq. (15), making decisions based on one set of \( \theta \) aspect angle states per target model) is that, assuming equal a priori probabilities for each target class \( \omega_i \), the sequence of states which yields the highest \( I(x; \omega, k) \) over all \( \omega_i \) may not fall on the particular target class \( \omega_i \) which has the highest \( \text{true} \) \( p(\omega_i | z(t), z(t+1), X_t) \). It may be that one model has a particular set of aspect angle states with associated signatures, such that the "best" sequence \( z_{t+1}^{n} \) traverses these points and gives this class the highest \( p(z_{t+1}^{n} | z(t), z(t+1), \omega_i) \), but if all possible \( X_t \) are considered, this class is less likely to have been the origin of the observed signatures \( z_{t+1}^{n} \) than some other. Hence the desire, as implied above, to investigate the contributions from multiple paths.

An alternate approach, then, is to construct sets of trajectories through the state space over the time frame of interest, using the same information on aspect angle from kinematics used to provide \( p(z_{t+1}^{n} | X_t, \omega = \omega_i) \), the a priori information for L&P-type approaches. These trajectories then imply sequences of feature observations, which can be compared to the observed sequences using DTW techniques. Further, in a departure from usual DTW, we can allow the "best path" to move from one trajectory to another. This defines a "two-dimensional" form of DTW.

7 Implementing Motion Fusion

Our research to date has involved seven basic algorithms, five of which use forms of DP sequence comparison for motion fusion, with the other two being expected to provide upper and lower bounds on performance. For the high range resolution radar (HRRR) domain in which this concept was evaluated, we used a Mahalanobis distance metric, which breaks the HRRR measurement or "sweep" as a vector of Gaussian components – this is by far the most popular HRRR metric to date [15, 25, 26, 33].

Consistent with the Mahalanobis metric approach, HRRR signatures were treated as 128-element (range "bin") vectors, downsampled from higher dimensioned vectors (choosing maximum range bins) by the signature generator [8] in dBm for desired aspect angles, polarisations, center frequencies, and bandwidths. Independent Gaussian noise realisations of constant variance were then added to each bin, using statistic gathered by analysis of actual HRRR tests [4] with a modified version of an existing program [26]. The assumption of noise independence from bin to bin was simply for convenience, since our test data showed significant cross-bin correlation (standard deviation was reasonably constant at 2-7 dBm along range and aspect angle extents). Since a maximum likelihood approach to range registration and constant noise variance were used, the signature comparison process was effectively a weighted correlation. In passing, we must note that the statistics of HRRR signatures deserve further research.

7.1 Independent Look (IL) Algorithm

This is a conventional decision theoretic target recogniser as discussed in Sect. 2. No restriction is placed on the "pose estimate" so generated – we find \( \max \{ I(x; \omega, 1) | \omega \} \) within the specified aspect angle window at each measurement time \( t \).

7.2 Perfect Knowledge of Aspect (PKA) Algorithm

This algorithm provides an upper bound on recognition performance in that it assumes that the recogniser knows perfectly the true aspect angle over time for each target class, matching the observed maneuver. For the Mahalanobis metric, this figure is simply the log joint maximum (classical) likelihood for observed or true aspect angles \( x_{t+1} \) over time, \( \log \sum_{\omega} p(z_{t+1}^{n} | x_{t+1}^{n}, \omega) \).

7.3 Fixed Bound (FB) Algorithm

This algorithm is an implementation of the Le Chevalier algorithm, as that approach is believed to work: the algorithm has no information from kinematic data.
matrices on the expected direction of aspect angle change, but knows that
the change is bounded. No subsequent processing is applied.

7.4 Full Larson and Peschon (L&P) Algorithm
This algorithm finds the natural log of the term in Eq. (14). Note that
the numerical values of this algorithm include contributions due to the
"a priori" aspect angle transition probabilities $p_{\alpha_{1\alpha_{2}}}$, $p_{\theta_{1\theta_{2}}}$, $p_{\phi_{1\phi_{2}}}$. Contributions
associated with the L&P "a priori" estimates $\bar{\alpha}$, $\bar{\theta}$, $\bar{\phi}$ were not included, to provide for
unbiased comparison of the curves.

7.5 One- and Two-Dimensional Path Warping
The one-dimensional path warping algorithm defines continuous one-
dimensional paths in aspect angle, parallel to and including the "nominal"
path given by the EKF/smother kinematic estimate. One-dimensional, un-
restricted endpoint [27, 28] DTW is performed along each trajectory. The
basic form requires continuous matching (no deletions). Each local path
cost is normalized by the total number of associations along that path.
This is a departure from usual DTW practice, and can lead to violations of
the "Principle of Optimality" [5], but worked well in our tests, since for the
proper match of measurements to target class, and constant measurement
noise statistics across the target length, the local average matching cost is
expected to be near the global average (see discussion in [20]).

Two-dimensional path warping uses the same set of trajectories defined
for the one-dimensional case, but local continuity constraints allow the op-
timum path to move from one trajectory to another. Other factors are as
for the one-dimensional case.

8 Performance Evaluation Approach
All of the algorithms discussed thus far are in fact interpretable in terms
of likelihood functions as defined by Rao [29]. Following the development
in [23-29, 90-101], we now introduce the use of generalised ambiguity functions
for assessing maximum likelihood estimator performance, as discussed by
Schweppe [31:376-381] and Maybeck [21].

8.1 The Generalised Ambiguity Function (GAF)
The generalised ambiguity function (GAF) is defined by the equation:

$$A_{\Omega}(\Omega_{1}, \Omega_{2}) \equiv \sum_{n} \int_{-\infty}^{\infty} L(\Omega, Z_{n}) r_{\Omega}(\Omega_{1}, \Omega_{2}) dZ$$

where:

- $\Omega_{1}$ is the particular combination of states $x(t)$ and parameters $y$ (the
  latter generally constant over the time interval of interest) for the
  truth system which generates the set of all possible measurement histories $Z_{n}$
  over which the integral is taken.
- $\Omega_{2}$ is a likelihood function defined for $\Omega_{1}$, operating on an
  element of this measurement history set, will ideally generate a higher
  value than any $L$ defined for some other value of $\Omega_{1}$, operating on an
  element of this measurement history set (the ambiguity function evaluates
  the extent to which this is true in the mean).
- $\Omega_{1}$ is the state/parameter value for which the likelihood function is defined,
  for evaluation against measurements generated by a truth model with state
  and parameter values $\Omega_{1}$.
- $L(\Omega, Z_{n})$ is the generalised ambiguity function, a function of $\Omega$ for a
  given $\Omega_{1}$ and likelihood function $L$.
- $r_{\Omega}(\Omega_{1}, \Omega_{2})$ is the probability density function of the
  measurements, given that the true states and parameters have the value $\Omega_{1}$.

Thus, the GAF is the expected value of likelihood functions defined for
combinations of states and parameters, conditioned on the true states and
parameters having particular values. For any particular value of $\Omega_{1}$ defining
the likelihood function, there is in fact a distribution of likelihood function
values produced, due to the different realisations of measurements produced
by a system with true states and parameters $\Omega_{1}$. Ambiguity functions can be
developed analytically for some likelihood functions [21, 22], or in any
case empirically by experiment or Monte Carlo simulation (as in our case).

Examining the ambiguity function for each realisable value of $\Omega_{1}$ and
for each such value of $\Omega_{2}$, a range of $\Omega_{1}$ encompassing reasonable state and
parameter values expected other than at $\Omega_{1}$, we desire that the function
have an easily discernible global maximum at $\Omega_{1}$ - i.e., that local maxima,
if present, are "widely'' separated from the global maximum at $\Omega_{1}$.

The curvature of the GAF at $\Omega_{1}$ can be related to the Cramér-Rao lower
bound (CRLB) [29] of the covariance for a state/parameter estimate ob-
tained by the use of that likelihood function [23, 24]. Recent interest has
been directed toward Cramér-Rao-like lower bounds for multisensor fusion-based
target tracking [10], but evidently no analogous bound has been de-
finied for multisensor target recognition [7].

8.2 Applying the Generalised Ambiguity Function
In general, to obtain the GAF in a Monte Carlo fashion, we define one like-
lihood function for each point of interest in state/parameter space. Each
likelihood function then operates on measurements from a system at some
"true" state/parameter point, unknown to the likelihood functions a priori.
The mean value of the likelihood function over a large number of measure-
ment sets defines the GAF. In classical ATR, the states (kinematic, tempera-
ture, etc.) of the candidate targets is readily defined, but real targets define
only discrete points in some infinite-dimensional, generally continuous "par-
arameter" space defined by their physical shape, materials, etc. Use of the
GAF in target recognition, then, requires the definition of pseudo-targets in
some sense "in between" real targets of interest.

To do this, pseudo-targets, we first define parent targets, or "points" in the (abstract) parameter space used by our target signature generator [8]. Each parent was defined by the same number of shapes and
surfaces, but the locations occupied in 3-D space by these objects differed
according to the size and shape of the respective target. Borrowing from the
language of computer graphics [5], then, 3-D linear interpolation "morphs"
(morphological, or shape, transformations) were performed to obtain new
"points" in target parameter space, or new targets in some sense "between"
the two parents. Fig. 2 shows an F-4 Phantom II and a MiG-21 as parent
targets, and a pseudo-target defined by 50% interpolation between the
parents. It must be emphasised that this linear interpolation was never
expected to translate into linear changes of the likelihood function outputs,
and it did not.

9 Results and Discussion
Parent target classes like those in Fig. 2 are readily separable in the feature
space and metric used here with any of the algorithms shown. More ambigu-
ous scenarios which demonstrate the power of the proposed approach were
generated by defining similar parents, high noise, and small morph fractions.
Typical outputs obtained in this way are shown in Figs. 3 and 4, for which
the parent target classes were a MiG-21 (the true target) and an SU-22
augmented with scatterers. The dotted vertical lines in the first figure indi-
cate parameter (target) interpolation values for which likelihood functions
were defined, and splits curve the connect the generalised ambiguity func-
tion (GAF) values to provide the curves shown. The second figure shows
percent correct recognition performance for likelihood functions tuned for
morph (parameter) values other than the true (unknown target) value (a
correct recognition is taken as one for which the properly tuned function
outputs a higher likelihood than the improperly tuned function). Relevant
target trajectory parameters are shown.

Note that the Independent Look (IL) algorithm defines the lower bound
on performance (worst), and the Perfect Knowledge of Aspect (PKA) algo-
rithm defines the upper bound (best). The Fixed Bound (FB) algorithms
provide significantly improved separation from the IL result, but the algo-
rithms which fuse filter/smother-provided kinematic information generally
provide equal or better separation in each case. Performance of the 1-D
and 2-D Warp algorithms is somewhat degraded in this simulation because
these methods force contiguous matches for measurements taken artificially.

Figure 2: F-4, 50% Morph, and MiG-21.
from particular discrete aspect angles. Results and anomalies are discussed in detail in [20].

Fig. 5 shows how progressive domain restriction provides better separation when measurements from one target class are matched to the library for another (i.e., wrong) class. This figure represents a region of solid angle in target aspect defined by the union of six "windows" or aspect angle bounds for any one measurement. The ML aspect angle locations identified by several algorithms over this angular extent for six measurements are shown — the true aspect angle is shown as well. Note the erratic aspect sequence selected by the IL processor and the still rather unlikely sequence selected by the FB algorithm. The DTW and L&P-based algorithms select more likely (linear) aspect angle paths, but their prediction to follow kinematically-reasonable paths forces a higher matching cost (lower likelihood) for this incorrect model-to-target association. In contrast, for measurements matched to their true origin target class, the different algorithms were much more likely to associate with the same aspect angle region.

The improvement from kinematic information fusion increases with the mean aspect angle rate or a g level of the target's turn. As turn rate increases, physics limits the number of possible aspect angle states $X^a$ (and therefore state sequences $X^a_t$), and we can limit the remaining matching domains even more severely to (fewer) sequences of expected length and direction. For the FB algorithm (with a fixed sampling rate), however, we must open the aspect angle bounds to give it any chance of tracking the nominal aspect rate on the true target. This increases dimensionality and gives it a greater chance of finding an improperly-high likelihood match on an incorrect target model. Other approaches for identifying infeasible aspect angle sequences may mitigate this problem, but may not effectively use the information available in observed kinematics.

Conversely, as turn rate decreases, the small mean aspect angle rate available to motion fusion algorithms tends to produce the same results as the FB algorithm, which assumes no mean rate, and can use small bounds when a small mean rate exits. For a zero-mean turn rate estimate, FB algorithms provide an effective approach — this is simply the limiting case of the L&P algorithm for a zero-mean, uniform $p(x_{k-1} | x_k, z_{k-1})$. For an aspect angle rate known to be zero, conventional decision theoretic recognition for a fixed aspect angle is most effective — this is in turn the limiting case of the FB algorithm for a bound of zero degrees.

Changes in the parameter space due to the morphing process can create apparently anomalous results, e.g., cases where the measurements from an F-4 were closer in Mahalanobis metric sense to sweeps from the MIG than they were to sweeps yielded by an interpolated target only 25% removed from the F-4. These cases resulted from the relative motion of scatterers during the morphing process, and were found to be physically reasonable after investigation. Modified morphing rules can resolve these anomalies.

Likelihood function differences for targets of interest (i.e., points of interest...
est in parameter space) are the key design criterion, but quick rolloff around the design point of each likelihood function should be of high secondary interest. The advantage to evaluating MI target recognition systems with GAFs is clearly that the method allows us to evaluate the curvature of the likelihood function away from its design point. This rolloff is directly related to the Crawford–Rao lower bound for the estimator used to develop the GAF [23]: practical evaluation of this bound using our approach requires one to generate target “morphs” or interpolations arbitrarily close to the design point, and evaluate the behavior of the GAF in this region.

The limiting value of the CRLB for these estimators is evidently given by the CRLB found in this fashion for the PKA algorithm (i.e., joint maximum likelihood for known aspect angle over time). In any case, the figures indicate that the separability of any two target classes may depend on factors other than behavior of the GAF near the true target parameter point. Therefore, this CRLB concept is perhaps not of greatest interest where we simply wish to identify a set of measurements as belonging to one of several a priori known points in some parameter space. The CRLB may be most useful where we wish to perform classical parameter estimation: for example, quantifying the extent to which we can estimate the optimum location in some finite-dimensional, model-based target parameter space to represent a previously unclassified real target, known only by measurements.

10 Further Directions and Conclusion

The dynamic programming-based approaches discussed here for exploiting the joint likelihood of kinematic and nonkinematic information in object recognition are just part of a class of techniques discussed more fully in [20]. Generically, these techniques are Bayesian multiple model parameter estimators [23:129–136] using linear and nonlinear models that exploit the unique coupling between states and parameters for different classes; an expansion both of (1) the efforts of a previous student of the second author [17] and (2) independent observations by Daum [3:177–178] made subsequent to the definition of this research.

The research described here has illuminated significant new directions for research in multisensor fusion and target recognition. Multisensor fusion of target kinematic and signature information is an exceptionally promising field. This fusion process can be viewed as exploiting the syntactic of physical processes, the joint likelihood of observable events, or restricting the domain of likelihood functions—in any case, it is clear that proper implementations of such fusion must provide recognition performance equal to or better than that of conventional “independent look” techniques.

References