3.2 Level 1 Processing

Level 1 processing is the low-level processing that results in target state estimation and target discrimination.[9] The term discrimination includes a hierarchy of processes, which from lowest to highest, encompass detection, orientation, classification (also called recognition in the older literature), and identification. The interpretation of these terms is shown in Table 3.1.[10–12] The ability to achieve a given level of discrimination depends on the resolution of the sensor and the SNR at the input to the sensor. These parameters may be traded off against each other to satisfy detection, classification, and identification requirements.[11–13]

Sensor outputs are combined through data association to produce the desired object or target discrimination level and target state estimate. The fusion algorithm used for target detection and classification process need not be the same as that used for state estimation and prediction. For example, a fusion algorithm that accepts highly processed data containing each sensor's best target-discrimination estimate can be the optimal one to use for the detection and classification problem when each sensor responds to independent signature-generation phenomena. But another fusion algorithm that accepts minimally processed data from more than one sensor and then analyzes and associates these data to form tracks may be optimal for obtaining the most accurate state estimates.

An overview of some 100 articles dealing with applications of information fusion, goals, system architectures, and mathematical tools has been compiled by Valet, Mauris, and Bolon.[14] Their literature survey addresses the selection of data and sensors that provide inputs to fusion systems, mathematical representation of the data and methods to combine them in an optimal way, and choice of output data format to enable easy interpretation of results and their further treatment.

Table 3.1 Object discrimination categories.

Category	Interpretation
Detection	Object is present
Orientation	Object is discerned as approximately symmetric or asymmetric and its orientation is determined
Classification	Class to which object belongs is discerned (e.g., building, truck, tank, man, trees, field)
Identification	Object is described to limit of an observer's knowledge (e.g., motel, pickup truck, M-1A1 tank, M-105 howitzer, soldier)

3.2.1 Detection, classification, and identification algorithms for data fusion

A taxonomy for detection, classification, and identification algorithms used in Level 1 processing is shown in Figure 3.4.[2,3,6,15–16] The major algorithm categories are physical models, feature-based inference techniques, and cognitive-based models. Other mathematical approaches for data fusion, not shown in the figure, are also utilized. These include random set theory, conditional algebra, and relational event algebra.[17] Random set theory deals with random variables that are sets rather than points. Goodman et al. use random set theory to reformulate multi-sensor, multi-target estimation problems into single-sensor, single-target problems.[17] They also apply the theory to incorporate ambiguous evidence (e.g., natural language reports and rules) into multi-sensor, multi-target estimation, and to incorporate various expert system methods (e.g., fuzzy logic and rule-based inference) into multi-sensor, multi-target estimation. Conditional-event algebra is a type of probabilistic calculus suited for contingency problems such as knowledge-based rules and contingent decision making. Relational-event algebra is a generalization of conditional-event algebra that provides a systematic basis for solving problems involving pooling of evidence. Still other data fusion approaches combine several of the illustrated methods, such as combinations of Dempster–Shafer with fuzzy logic and artificial neural networks with fuzzy logic.

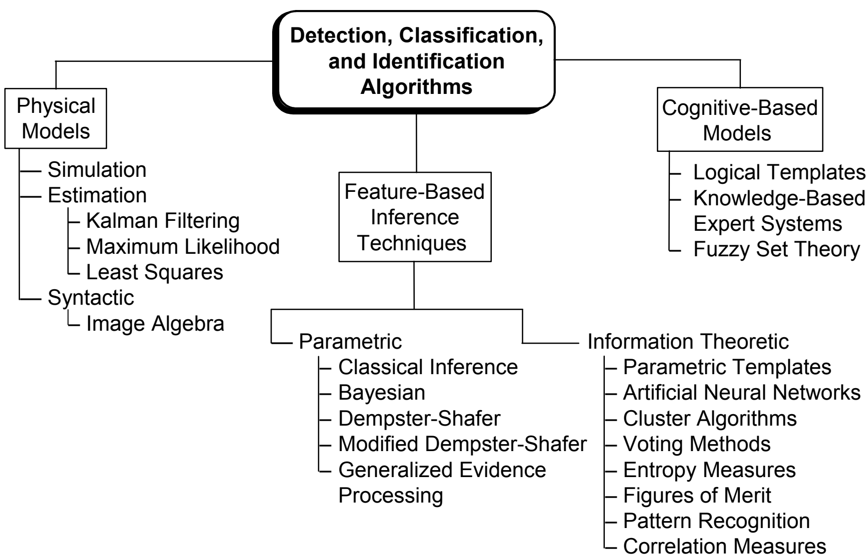

Figure 3.4 Taxonomy of detection, classification, and identification algorithms.[2,3,6,15–16]

3.2.1.1 Physical models

Physical models replicate object discriminators that are easily and accurately observable or calculable. Examples of discriminators are radar cross section as a function of aspect angle; infrared emissions as a function of vehicle type, engine temperature, or surface characteristics such as roughness, emissivity, and temperature; multi-spectral signatures; and height profile images. Table 3.2 lists feature categories used in developing physical models, and representative physical features and other attributes of the categories.[6]

Physical models estimate the classification and identity of an object by matching modeled or prestored target signatures to observed data as shown in Figure 3.5. The signature or imagery gathered by a sensor is analyzed for preidentified physical characteristics or attributes, which are input into an identity declaration process. Here, the characteristics identified by the analysis are compared with stored physical models or signatures of potential targets and other objects. The stored model or signature having the closest match to the real-time sensor data is declared to be the correct identity of the target or object.

Physical modeling techniques include simulation, estimation, and syntactic methods. Simulation is used when the physical characteristics to be measured can be accurately and predictably modeled. Estimation processes include Kalman filtering, maximum likelihood, and least squares approximation. The Kalman filter provides a general solution to the recursive, minimum mean-square estimation problem as long as the target dynamics and measurement noise are accurately modeled. Kalman filtering is discussed in Section 10.6, and maximum likelihood and least squares approximation in Sections 3.2.2 and 7.9. The syntactic methods, although listed under physical models, are described later as part of pattern recognition, a subset of information theoretic techniques.

An application of physical modeling based on laser-radar height-profile imagery is illustrated in Figure 3.6. The profile of a shrub and a tank are shown in the left image. The horizontal line passing through the turret of the tank identifies one scan or one profile slice through the image. The plot on the right represents the height of the features detected by the particular scan-line. If the scan-line were lowered to pass through the gun barrel of the tank, a height representing the barrel would be seen in the profile slice data.

When many height profiles produced by line scans through different regions of the laser imagery are compared, naturally occurring objects tend to have more random shapes than man-made objects. Thus, an object identification algorithm using shape as a classification criterion can be developed to differentiate between natural objects such as ground clutter (e.g., shrubs, boulder field, and trees) and man-made objects or potential targets having known height profiles.

Table 3.2 Feature categories and representative features used in developing physical models.

Feature Category	Representative Features	Other Attributes
Geometrical	Edges, lines, line widths, line relationships (e.g., parallel, perpendicular), arcs, circles, conic shapes, size of enclosed area	Represents the geometric size and shape of objects Man-made objects tend to exhibit regular geometric shapes with distinct boundaries
Structural	Surface area; relative orientation; orientation in vertical and horizontal ground plane; juxtaposition of planes, cylinders, cones	Develops a larger scale and contextual view of image segments
Statistical	Number of surfaces, area and perimeter, moments, Fourier descriptors, mean, variance, kurtosis, skewness, entropy	Used at local and global image levels to characterize image data
Spectral	Color coefficients, apparent blackbody temperature, spectral peaks and lines, general spectral signature	Man-made objects tend to possess distinct infrared spectral signatures
Time domain	Pulse characteristics (rise and fall times, amplitude), pulse width, pulse repetition interval, moments, ringing and overshoot, relationship of pulses to ambient noise floor	Selection of time-domain features versus frequency-domain features depends on transmitted waveform and received signal characteristics Less than 100-percent duty cycle signals favor time-domain analysis
Frequency domain	Fourier coefficients, Chebyschev coefficients, periodic structures in frequency domain, spectral lines and peaks, pulse shape and other characteristics, forced features (e.g., power spectral density of signal raised to N^{th} power)	Information is analogous to that from features in the time domain. 100-percent duty cycle signals favor frequency-domain analysis
Hybrid	Wavelets, Wigner–Ville distributions, cyclostationary representations	Useful for signals in which both time and frequency are important

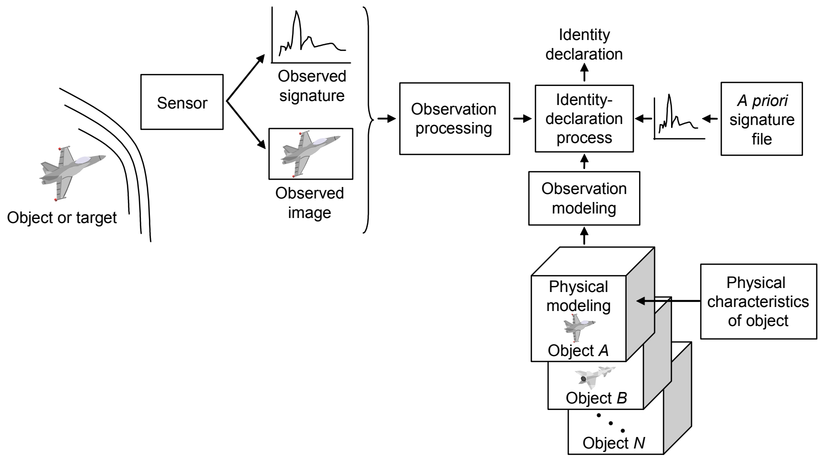

Figure 3.5 Physical model concept.

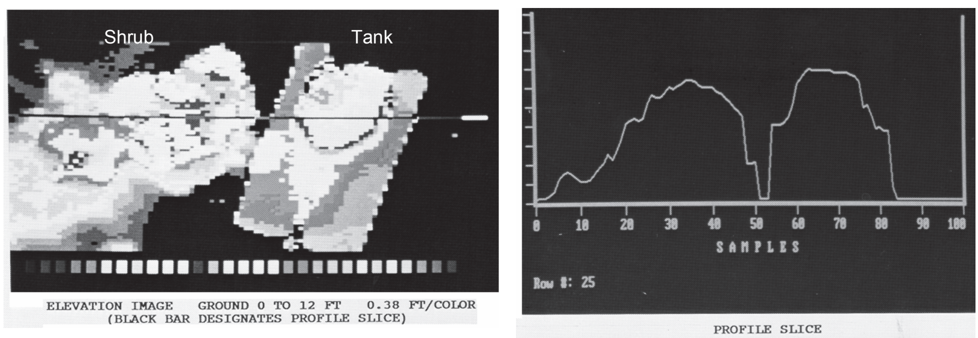

Figure 3.6 Laser radar imagery showing shapes of man-made and natural objects (photographs courtesy of Schwartz Electro-Optics, Orlando, FL).

3.2.1.2 Feature-based inference techniques

Feature-based inference techniques perform classification or identification by mapping data, such as statistical knowledge about an object or recognition of object features, into a declaration of identity. Feature-based algorithms may be further divided into parametric and information theoretic techniques (i.e., algorithms that have some commonality with information theory) as depicted in Figure 3.4.

Parametric techniques

Parametric classification directly maps parametric data (e.g., features) into a declaration of identity. Physical models are not used. Parametric techniques include classical inference, Bayesian inference, Dempster–Shafer evidential theory, modified Dempster–Shafer methods, and generalized evidence processing.

Classical inference gives the probability that an observation can be attributed to the presence of an object or event, given an assumed hypothesis. Its major disadvantages are: (1) difficulty in obtaining the density function that describes the observable used to classify the object, (2) complexities that arise when multivariate data are encountered, (3) its capability to assess only two hypotheses at a time, and (4) its inability to take direct advantage of *a priori* and likelihood probabilities.

Figure 3.7 illustrates a problem where classical inference is utilized to determine whether the detected radar illumination is from a Class 1 radar with low pulse repetition interval (PRI) or a Class 2 radar with higher PRI. A critical value of the PRI, designated as PRI_c, is selected based on acceptable Type 1 and Type 2 errors (defined in the figure). In this example, the null hypothesis H_0 (the statement being tested) is equated to "The observed PRI is less than PRI_c (i.e., it belongs to a Class 1 radar)" and the alternative hypothesis H_1 (the statement suspected of being true) to "The observed PRI is greater than or equal to PRI_c (i.e., it belongs to a Class 2 radar)." The probability that the observed PRI belongs to a Class 1 radar is calculated using a standardized random variable and the known probability density function that describes the PRI. The probability, computed assuming H_0 is true, that the standardized random variable assumes a value as extreme or more extreme than that actually observed is called the *P*-value of the test.

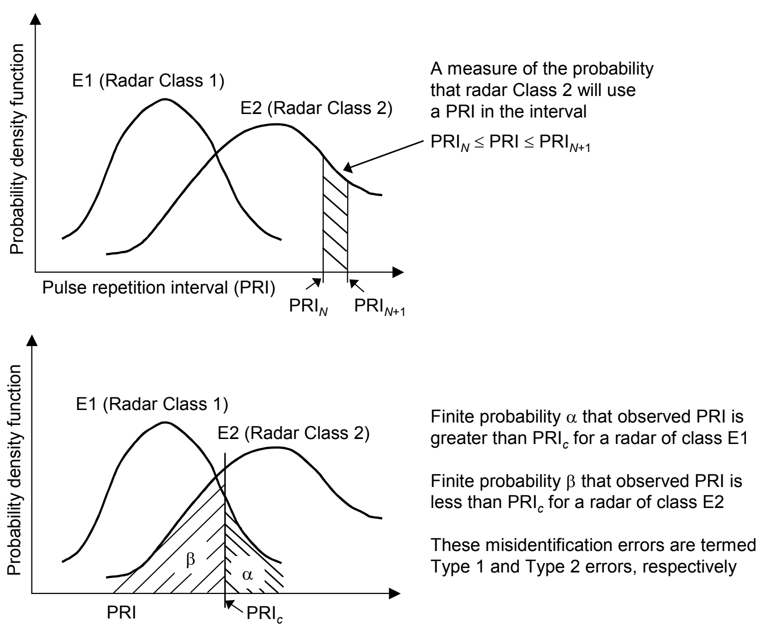

Figure 3.7 Classical inference concept [adapted from D.L. Hall, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, Norwood, MA (1992)].

The smaller the *P*-value, the stronger the evidence against H_0 provided by the data. If the *P*-value is as small as or smaller than α, the data are said to be statistically significant at level α. That is, the data give evidence against H_0 such that H_0 occurs no more than α percent of the time.

The significance-level α of any fixed level test is equal to the probability of the Type 1 error. Thus, α is the probability that the test will reject hypothesis H_0 when H_0 is in fact true. The probability that a fixed-level α significance test will reject H_0 when a particular alternative value of the parameter is true is called the power of the test against that alternative. Thus, the power is equal to 1 minus the probability of a Type 2 error for that alternative. These concepts are developed further in Chapter 4.

Bayesian inference resolves some of the difficulties with classical inference. It updates the *a priori* probability of a hypothesis given a previous likelihood estimate and additional observations and is applicable when more than two hypotheses are to be assessed.[6,18] The disadvantages of Bayesian inference include: (1) difficulty in defining the prior probabilities and likelihood functions, (2) complexities that arise when multiple potential hypotheses and multiple conditionally dependent events are evaluated, (3) mutual exclusivity required of competing hypotheses, and (4) inability to account for general uncertainty. Bayesian inference is discussed further in Chapter 5.

Dempster–Shafer evidential theory generalizes Bayesian inference to allow for uncertainty by distributing support for a proposition (e.g., that an object is of a particular type) not only to the proposition itself, but also to the union of propositions (disjunctions) that include it and to the negation of a proposition. Any support that cannot be directly assigned to a proposition or its negation is assigned to the set of all propositions in the hypothesis space (i.e., uncertainty). Support provided by multiple sensors for a proposition is combined using Dempster's rule. Bayesian and Dempster–Shafer produce identical results when all singleton propositions are mutually exclusive and there is no support assigned to uncertainty. A requirement of the Dempster–Shafer method is the need to define processes in each sensor that assign the degree of support for a proposition. Disadvantages of the method include the inability to make direct use of prior probabilities when they are known and the counterintuitive output sometimes produced when support for conflicting propositions is large. Several methods have been proposed to modify Dempster's rule through the use of probability transformations that better accommodate conflicting beliefs[19] and, in some cases, through the use of prior knowledge and spatial information.[20–26] Data fusion using Dempster–Shafer evidential theory and examples of its application are developed in more detail in Chapter 6.