A Tutorial on Multisensor Integration and Fusion

REN C. LUO AND MICHAEL G. KAY

Robotics and Intelligent Systems Laboratory
Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7906

***Abstract*—This paper presents a tutorial introduction to the subject of multisensor integration and fusion. The role of multisensor integration and fusion in the operation of intelligent systems is defined in terms of the unique type of information multiple sensors can provide. Multisensor integration is discussed in terms of basic integration functions and multisensor fusion in terms of the different levels at which fusion can take place. Numerical examples are given to illustrate a variety of different fusion methods. The paper concludes with speculations concerning possible research future directions and a guide to survey and review papers in the area of multisensor integration and fusion.**

I. INTRODUCTION

THE SYNERGISTIC use of multiple sensors by machines and systems is a major factor in enabling some measure of intelligence to be incorporated into their overall operation so that they can interact with and operate in an unstructured environment without the complete control of a human operator. The use of sensors in an intelligent system is an acknowledgement of the fact that it may not be possible or feasible for a system to know *a priori* the state of the outside world to a degree sufficient for its autonomous operation. The reasons a system may lack sufficient knowledge concerning the state of the outside world may be due either to the fact that the system is operating in a totally unknown environment, or, while partial knowledge is available and is stored in some form of a world model, it may not be feasible to store large amounts of this knowledge and it may not be possible in principle to know the state of the world a priori if it is dynamically changing and unforeseen events can occur. Sensors allow a system to learn the state of the world as needed and to continuously update its own model of the world. The motivation for using multiple sensors in a system is a response to the simple question: If a single sensor can increase the capability of a system, would the use of more sensors increase it even further? Over the past decade a number of researchers have been exploring this question from both a theoretical perspective and by actually building multisensor machines and systems for use in a variety of areas of application. Typical of the applications that can benefit from the use of multiple sensors are automatic target recognition, mobile robot navigation, industrial tasks like assembly, military command and control for battlefield management, target tracking, and aircraft navigation.

There are a number of different means of integrating the information provided by multiple sensors into the operation of a system. The most straightforward approach to multisensor integration is to let the information from each sensor can serve as a separate input to the system controller. This approach may be the most appropriate if each sensor is providing information concerning completely different aspects of the environment. The major benefit gained through this approach is the increase in the extent of the environment able to be sensed. The only interaction between the sensors is indirect and based on the individual effect each sensor has on the controller. If there is some degree of overlap between the sensors concerning some aspect of the environment that they are able to sense, it may be possible for a sensor to directly influence the operation of another sensor so that the value of the combined information that the sensors provide is greater than the sum of the value of the information provided by each sensor separately. This synergistic effect from the multisensor integration can be achieved either by using the information from one sensor to provide cues or guide the operation of other sensors, or by actually combining or fusing the information from multiple sensors. The information from the sensors can be fused at a variety of levels of representation depending upon the needs of the system and the degree of similarity between the sensors. The major benefit gained through multisensor fusion is that the system can be provided with information of higher quality concerning, possibly, certain aspects of the environment that can not be directly sensed by any individual sensor operating independently.

II. THE ROLE OF MULTISENSOR INTEGRATION AND FUSION IN INTELLIGENT SYSTEMS

This section describes the role of multisensor integration and fusion in the operation of intelligent machines and systems. The role of multisensor integration and fusion can best be understood with reference to the type of information that the integrated multiple sensors can uniquely provide the system. The potential advantages gained through the synergistic use of this multisensory information can be decomposed into a combination of four fundamental aspects: the redundancy, complementarity, timeliness, and cost of the information. Multisensor integration and the related notion of multisensor fusion are defined and distinguished. The different functional aspects of multisensor integration and fusion in the overall operation of a system are presented and serve to highlight the distinction between the different types of integration and the different types of fusion. The potential advantages in integrating multiple sensors are then discussed in terms of four fundamental aspects of the information provided by the sensors, and the problems associated with creating a general methodology for multisensor integration and fusion are discussed in terms of the methods used for handling the different sources of possible error or uncertainty.

Multisensor integration, as defined in this paper, refers to the synergistic use of the information provided by multiple sensory

087942-600-4/90/1100-0707$01.00

devices to assist in the accomplishment of a task by a system. An additional distinction is made between multisensor integration and the more restricted notion of multisensor fusion. *Multisensor fusion*, as defined in this paper, refers to any stage in the integration process where there is an actual combination (or fusion) of different sources of sensory information into one representational format. The information to be fused may come from multiple sensory devices during a single period of time or from a single sensory device over an extended time period. Although the distinction of fusion from integration is not standard in the literature, it serves to separate the general system-level issues involved in the integration of multiple sensory devices at the architecture and control level, from the more specific mathematical and statistical issues involved in the actual fusion of sensory information.

A. Potential Advantages in Integrating Multiple Sensors

The purpose of external sensors is to provide a system with useful information concerning some features of interest in the system's environment. The potential advantages in integrating and/or fusing information from multiple sensors are that the information can be obtained more accurately, concerning features that are impossible to perceive with individual sensors, in less time, and at a lesser cost. These advantages correspond, respectively, to the notions of the redundancy, complementarity, timeliness, and cost of the information provided the system.

- *Redundant* information is provided from a group of sensors (or a single sensor over time) when each sensor is perceiving, possibly with a different fidelity, the same features in the environment. The integration or fusion of redundant information can reduce overall uncertainty and thus serve to increase the accuracy with which the features are perceived by the system. Multiple sensors providing redundant information can also serve to increase reliability in the case of sensor error or failure.
- *Complementary* information from multiple sensors allows features in the environment to be perceived that are impossible to perceive using just the information from each individual sensor operating separately. If the features to be perceived are considered dimensions in a space of features, then complementary information is provided when each sensor is only able to provide information concerning a subset of features that form a subspace in the feature space, i.e., each sensor can be said to perceive features that are independent of the features perceived by the other sensors; conversely, the dependent features perceived by sensors providing redundant information would form a basis in the feature space.
- *More timely* information, as compared to the speed at which it could be provided by a single sensor, may be provided by multiple sensors due to either the actual speed of operation of each sensor, or the processing parallelism that may be possible to achieve as part of the integration process.
- *Less costly* information, in the context of a system with multiple sensors, is information obtained at a lesser cost when compared to the equivalent information that could be obtained from a single sensor. Unless the information provided by the single sensor is being used for additional functions in the system, the total cost of the single sensor should be compared to the total cost of the integrated multisensor system.

The role of multisensor integration and fusion in the overall operation of a system can be defined as the degree to which each of these four aspects is present in the information provided by the sensors to the system. Redundant information can usually be fused at a lower level of representation compared to complementary information because it can more easily be made commensurate. Complementary information is usually either fused at a symbolic level of representation, or provided directly to different parts of the system without being fused. While in most cases the advantages gained through the use of redundant, complementary, or more timely information in a system can be directly related to possible economic benefits, in multisensor target tracking fused information is sometimes used in a distributed network of target tracking sensors just to reduce the bandwidth required for communication between groups of sensors in the network.

B. Possible Problems

Many of the possible problems associated with creating a general methodology for multisensor integration and fusion, as well as developing the actual systems that use multiple sensors, center around the methods used for modeling the error or uncertainty in the integration and fusion process, the sensory information, and the operation of the overall system including the sensors. For the potential advantages in integrating multiple sensors to be realized, solutions to these problems will have to be found that are both practical and theoretically sound.

1) Error in the Integration and Fusion Process: The major problem in integrating and fusing redundant information from multiple sensors is that of "registration"—the determination that the information from each sensor is referring to the same features in the environment. The registration problem is termed the correspondence and data association problem in stereo vision and multitarget tracking research, respectively. Barniv and Casasent [5] have used the correlation coefficient between pixels in the grey level of images as a measure of the degree of registration of objects in the images from multiple sensors. Hsiao [26] has detailed the different geometric transformations needed for registration. Lee and Van Vleet [31] and Holm [25] have studied the registration errors between radar and infrared sensors. Lee and Van Vleet have presented an approach that is able to both estimate and minimize the registration error, and Holm has developed a method that is able to autonomously compensate for registration errors in both the total scene as perceived by each sensor ("macroregistration"), and the individual objects in the scene ("microregistration").

2) Error in Sensory Information: The error in sensory information is usually assumed to be caused by a random noise process that can be adequately modeled as a probability distribution. The noise is usually assumed not to be correlated in space or time (i.e., white), Gaussian, and independent. The major reasons that these assumptions are made is that they enable a variety of fusion techniques to be used that have tractable mathematics and yield useful results in many applications. If the noise is correlated in time (e.g., gyroscope error) it is still sometimes possible to retain the whiteness assumption through the use of a shaping filter [37]. The Gaussian assumption can only be justified if the noise is caused by a number of small

independent sources. In many fusion techniques the consistency of the sensor measurements is increased by first eliminating spurious sensor measurements so that they are not included in the fusion process. Many of the techniques of robust statistics (e.g., ε-contamination) can be used to eliminated spurious measurements. The independence assumption is usually reasonable so long as the noise sources do not originate from within the system.

3) Error in System Operation: When error occurs during operation due to possible coupling effects between components of a system, it may still be possible to make the assumption that the sensor measurements are independent if the error, after calibration, is incorporated into the system model through the addition of an extra state variable [37]. In well-known environments the calibration of multiple sensors will usually not be a difficult problem, but when multisensor systems are used in unknown environments, it may not be possible to calibrate the sensors. Possible solutions to this problem may require the creation of detailed knowledge bases for each type of sensor so that a system can autonomously calibrate itself. One other important feature required of any intelligent multisensor system is the ability to recognize and recover from sensor failure (cf. [8] and [27]).

MULTISENSOR INTEGRATION

The means by which multiple sensors are integrated into the operation of an intelligent machine or system are usually a major factor in the overall design of the system. The specific capabilities of the individual sensors and the particular form of the information they provide will have a major influence on the design of the overall architecture of the system. These factors, together with the requirements of the particular tasks the system is meant to perform, make it difficult to define any specific general-purpose methods and techniques that encompass all of the different aspects of multisensor integration. Instead, what has emerged from the work of many researchers is a number of different paradigms, frameworks, and control structures for integration that have proved to be particularly useful in the design of multisensor systems (see [33] for a review).

Many of the paradigms, frameworks, and control structures used for multisensor integration have been adapted with little or no modification from similar high-level constructs used in systems analysis, computer science, control theory, and artificial intelligence (AI). In fact, much of multisensor integration research can be viewed as the particular application of a wide range of fundamental systems design principles. Common themes among these constructs that have particular importance for multisensor integration are the notions of "modularity," "hierarchical structures," and "adaptability." In a manner similar to structured programming, modularity in the design of the functions needed for integration can reduce the complexity of the overall integration process and can increase its flexibility by allowing many of the integration functions to be designed to be independent of the particular sensors being used. Modularity in the operation of the integration functions enables much of the processing to be distributed across the system. The object-oriented programming paradigm and the distributed blackboard control structure are two constructs that are especially useful in promoting modularity for multisensor integration. Hierarchical structures are useful in allowing for the efficient representation of the different forms, levels, and resolutions of the information used for sensory processing and control; e.g., the NBS Sensory and Control Hierarchy [47] and logical sensor networks [22]. Adaptability in the integration process can be an efficient means of handling the error and uncertainty inherent in the integration of multiple sensors. The use of the artificial neural network formalism allows adaptability to be directly incorporated into the integration process.

A. The Basic Integration Functions

Although the process of multisensor integration can take many different forms depending on the particular needs and design of the overall system, certain basic functions are common to most implementations. The diagram shown in Fig. 1 represents multisensor integration as being a composite of these basic functions. A group of n sensors provide input to the integration process. In order for the data from each sensor to be used for integration it must first be effectively modeled. A *sensor model* represents the uncertainty and error in the data from each sensor and provides a measure of its quality that can be used by the subsequent integration functions. A common assumption is that the uncertainty in the sensory data can be adequately modeled as a Gaussian distribution. After the data from each sensor has been modeled it can be integrated into the operation of the system in accord with three different types of *sensory processing*: fusion, separate operation, and guiding or cueing. The data from Sensors 1 and 2 are shown in the figure as being fused. Prior to its fusion, the data from each sensor must be made commensurate. *Sensor registration* refers to any of the means (e.g., geometrical transformations) used to make the data from each sensor commensurate in both its spatial and temporal dimensions, i.e., that the data refer to the same location in the environment over the same time period. The different types of possible sensor data fusion are described in Section III. If the data provided by a sensor is significantly different from that provided by any other sensors in the system, its influence on the operation of the other senors may be indirect. The *separate operation* of of such a sensor will influence the other sensors indirectly through the effects the sensor has on the system controller and the world model. A *guiding or cueing* type of sensory processing refers to the situation where the data from one sensor is used to guide or cue the operation other sensors. A typical example of this type of multisensor integration is found in many robotics applications where visual information is used to guide the operation of a tactile array mounted on the end of a manipulator.

The results of the sensory processing function serve as inputs to the world model. A *world model* is used to store information concerning any possible state of the environment the system is expected to be operating in. A world model can include both *a priori* information and recently acquired sensory information. High-level reasoning processes can use the world model to make inferences that can be used to direct the subsequent processing of the sensory information and the operation of the system controller. Depending on the needs of a particular application,

information stored in the world model can take many different forms: In object recognition tasks the world model might contain just the representations of the objects the system is able to recognize, while in mobile robot navigation tasks the world model might contain the complete representation of the robot's local environment, e.g., the objects in the environment as well as local terrain features. The majority of the research related to the development of multisensor world models has been within the context of the development of suitable high-level representations for multisensor mobile robot navigation and control. Luo and Kay [33] describe a number of examples of world models used in mobile robots. The last multisensor integration function, *sensor selection*, refers to any means used to select the most appropriate configuration of sensors (or sensing strategy) from among the sensors available to the system. In order for selection to take place, some type of sensor performance criteria need to to established. In many cases the criteria require that the operation of the sensors be modeled adequately enough so that a cost value can be assigned to measure their performance. Two different approaches to the selection of the type, number, and configuration of sensors to be used in the system can be distinguished: "preselection" during design or initialization, and "real-time selection" in response to changing environmental or system conditions, e.g., sensor failure.

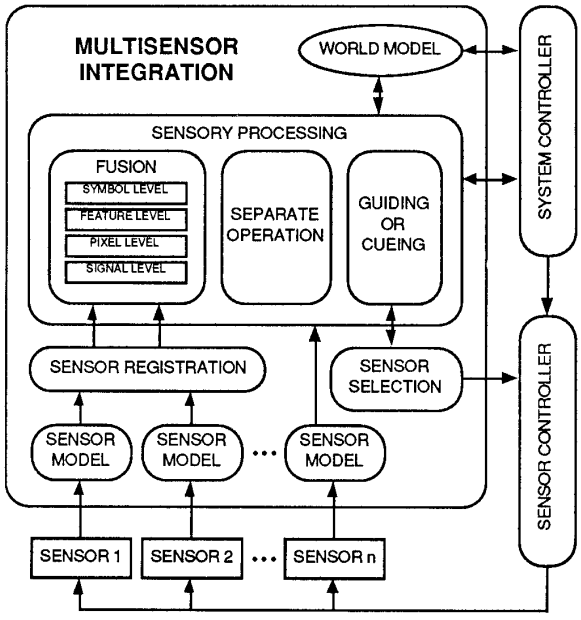


Fig. 1. Functional diagram of multisensor integration and fusion in the operation of a system.

B. *Networks and Rule-Based Systems*

Networks and rule-based systems are the most common forms of control structures used for multisensor integration. They can be used either individually or combined together as part of an overall control structure. They are especially useful when the sensors in a system are very dissimilar and the data they provide needs to be fused at multiple levels of representation, i.e., from signal- through symbol-level fusion. Because of the particular advantages of each structure, a rule-based system is most effective when it is used for top-level control and groups of network structures (e.g., Bayesian and neural networks) are used for lower-level control functions. Mitiche, Henderson, and Laganiere [40] have advocated the use of "decision networks" for multisensor integration. In a decision network, Bayesian and neural networks can be used as evaluating mechanisms at the nodes of a tree-structured production rule-based control network.

The use of network enable hierarchical structures of be efficiently represented (e.g., networks of logical sensors described in [22]) and allow the same formalism to be used to encode both the representational structure as well as the control structure; e.g., a hierarchical network can be used to both model an object and to control the decision process in multiple-hypothesis object recognition. The use of rule-based systems enable the implementation of many AI-based control schemes that offer extreme flexibility for integrating multiple sensors in complex systems because knowledge, in the form of production rules, can be added to the control structure in a modular and incremental fashion, and the production rules used in many systems can themselves be used for symbol-level fusion (see Section III-E-4).

The major problem with rule-based systems that limits their application to all levels of the control structure needed for multisensor integration is that, unless each rule in the system represents an inference that is independent of all the other rules in the system (i.e., the rule base forms a tree structure), improper conclusions can be drawn during the reasoning process; e.g., bidirectional inferences are not correctly handled, conclusions can not be easily retracted, and correlated sources of evidence are improperly treated [41]. A means of overcoming these difficulties, in portions of the control structure where individual rules can not be isolated from the effects of other related nodes, is through the use of the Bayesian formalism in which conditional probability expressions are used to represent factual or empirical information. A problem with the straightforward use of conditional probability expressions is that in order to assert a fact one must know its conditional dependencies with all of the other known facts. Bayesian networks [41] can be used to encode these dependencies as directed arcs between neighboring nodes in an acyclic graph so that they can be quickly identified. The network offers a complete inference mechanism that can identify, in polynomial time, every conditional dependence relationship.

III. Multisensor Fusion

The fusion of the data or information from multiple sensors or a single sensor over time can take place at different levels of representation (sensory information can be considered data from a sensor that has been given a semantic content through processing and/or the particular context in which it was acquired). As shown in Fig. 1, a useful categorization is to consider multisensor fusion as taking place at the *signal, pixel, feature*, and *symbol levels* of representation. Most of the sensors typically used in practice provide data that can be fused at one or more of these levels. Although the multisensor integration functions of sensor registration and sensor modeling are shown in Fig. 1 as being separate from multisensor fusion, most of the

TABLE I
COMPARISON OF FUSION LEVELS

Characteristics	Signal Level	Pixel Level	Feature Level	Symbol Level
Type of sensory information	single- or multi-dimensional signals	multiple images	features extracted from signals and images	symbol representing decision
Representation level of information	low	low to medium	medium	high
Model of sensory information	random variable corrupted by uncorrelated noise	stochastic process on image or pixels with multidimensional attributes	non-invariant geometrical form, orientation, position, and temporal extent of features	symbol with associated uncertainty measure
Degree of registration:				
spatial	high	high	medium	low
temporal	high	medium	medium	low
Means of registration:				
spatial	sensor coalignment	sensor coalignment or shared optics	geometrical transformations	spatial attributes of symbol if necessary
temporal	synchronization or estimation	synchronization	synchronization	temporal attributes of symbol if necessary
Fusion method	signal estimation	image estimation or pixel attribute combination	geometrical and temporal correspondence, and feature attribute combination	logical and statistical inference
Improvement due to fusion	reduction in expected variance	increase in performance of image processing tasks	reduced processing, increased feature measurement accuracy, and value of additional features	increase in truth or probability values

methods and techniques used for fusion make very strong assumptions, either explicitly or implied, concerning how the data from the different sensors is modeled and to what degree the data is in registration. A fusion method that may be very sound in theory can be difficult to apply in practice if the assumed sensor model does not adequately describe the data from a real sensor, e.g., the presence of outliers due to sensor failure in an assumed normal distribution of the sensory data can render the fused data useless, or the degree of assumed sensor registration may be impossible to achieve, e.g, due to the limited resolution or accuracy of the motors used to control the sensors.

The different levels of multisensor fusion can be used to provide information to a system that can be used for a variety of purposes; e.g., signal-level fusion can be used in real-time applications and can be considered as just an additional step in the overall processing of the signals, pixel-level fusion can be used to improve the performance of many image processing tasks like segmentation, and feature- and symbol-level fusion can be used to provide an object recognition system with additional features that can be used to increase its recognition capabilities. The different levels can be distinguished by the type of information they provide the system, how the sensory information is modeled, the degree of sensor registration required for fusion, the methods used for fusion, and the means by which the fusion process improves the "quality" of the information provided the system. A comparison of the different levels of fusion is given below and summarized in Table I.

Most methods of multisensor fusion make explicit assumptions concerning the nature of the sensory information. The most common assumptions include the use of a measurement model for each sensor that includes a statistically independent additive Gaussian error or noise term (i.e., location data), and an assumption of statistical independence between the error terms for each sensor. Many of the differences in the fusion methods included below center on their particular techniques (e.g., calibration, thresholding) used for transforming raw sensory data into a form so that the above assumptions become reasonable and a mathematically tractable fusion method can result. An excellent introduction to the conceptual problems inherent in any fusion method based on these common assumptions has been provided by [42]. Their paper provides a proof that the inclusion of additional redundant sensory information almost always improves the performance of any fusion method that is based on optimal estimation.

Clark and Yuille [11] have used a Bayesian formulation of sensory processing to provide a mathematical foundation upon which data fusion algorithms can be created and analyzed. The fusion algorithms are classified into two general types, distinguished by the manner in which information in the form of solution constraints is combined to obtain an overall solution to the multisensory processing problem. In "weakly coupled" data fusion algorithms the operation of the different sensory processing modules is not affected by the fusion process; in "strongly coupled" algorithms the output of a module does

interact with the other modules and effects their operation. Examples are given of data fusion applied to feature-level stereo, binocular and monocular depth cue, and shape from shading algorithms.

A. Fusion Levels in Automatic Target Recognition

Fig. 2 provides an example of how the different levels of multisensor fusion can be used in the task of automatic target recognition. In the figure, five sensors are being used by the system to recognize a tank: two millimeter-wave radars that could be operating at different frequencies, an infrared sensor (e.g., a forward-looking infrared sensor), a camera providing visual information, and a radio signal detector that can identify characteristic emissions originating from the tank. The complementary characteristics of the information provided by this suite of sensors can enable the system to detect and recognize targets under a variety of different operating conditions, e.g., the radars provide range information and their signals are less effected by atmospheric attenuation as compared to the infrared image, while the infrared sensor provides information of greater resolution than the radars and, unlike the camera, is able to operate at night.

The two radars are assumed to be synchronized and coaligned on a platform so that their data is in registration and can be fused at the signal level. The fused signal is shown in the figure as being sent both to the system, where it can be immediately used for the improved detection of targets, and as input to generate a range image of the target. The range image from the radars can then be fused at the pixel level with the intensity image provided by the infrared sensor located on the same platform. In most cases, an element from the range image can only be registered with a neighborhood of pixels from the infrared image because the differences in resolution between the millimeter-wave radars and the infrared sensor. An image from The fused image is sent both to the system, where it can be immediately used to improve target segmentation, and as input so that useful target features can be extracted from the image. The features from the pixel-level fusion can then be fused at the feature level with similar features extracted from visual image provided by the camera. The camera may be located on a different platform because the sensor registration requirements for feature-level fusion are less stringent than those for signal- and pixel-level fusion. The fused features are then sent both to the system, where they can be used to improve the accuracy in the measurement of the orientation or pose of the target, and as input features to an object program. The output of the program is a symbol, with an associated measure of its quality (0.7), indicating the presence of the tank. The symbol can then be fused at the symbol level with a similar symbol derived from the radio signal detector that also indicates the presence of the tank. The fused signal is then sent to the system for the final recognition of the tank. As shown in the figure, the measure of quality of the fused symbol (0.94) is greater than the measures of quality of either of the component symbols and represents the increase in the quality associated with the symbol as a result of the fusion, i.e., the increase in the likelihood that the target is a tank.

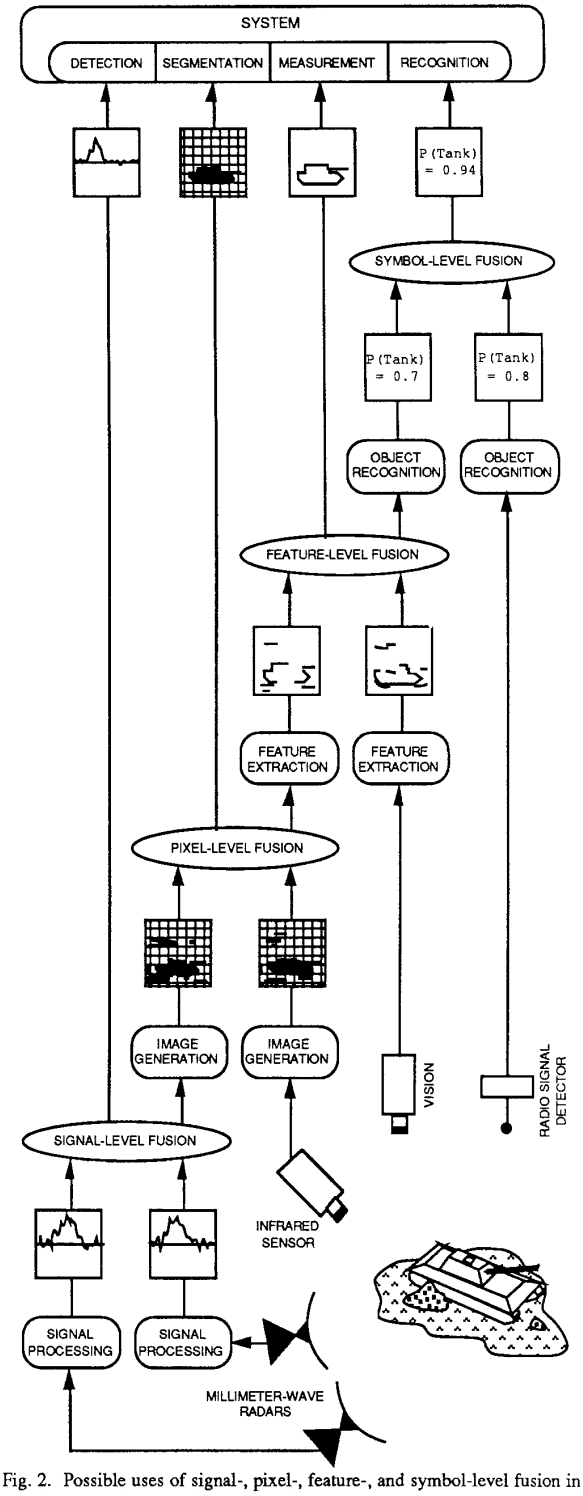


Fig. 2. Possible uses of signal-, pixel-, feature-, and symbol-level fusion in the automatic recognition of a tank.

The transformation from lower to higher levels of representation as the information moves up through the target recognition structure shown in Fig. 2 is common in most multisensor integration processes. At the lowest level, raw

sensory data are transformed into information in the form of a signal. As a result of a series of fusion steps, the signal is transformed into progressively more abstract numeric and symbolic representations. This "signals-to-symbols" phenomenon is also common in computational vision and AI.

B. Signal-Level Fusion

Signal-level fusion refers to the combination of signals of a group of sensors to provide a signal that is usually of the same form as the original signals but of greater quality. The signals from the sensors can be modeled as random variables corrupted by uncorrelated noise, with the fusion process considered as an estimation procedure. As compared to the other types of fusion, signal-level fusion requires the greatest degree of registration between the sensory information. If multiple sensors are used their signals must be in temporal as well as spatial registration. If the signals from the sensors are not synchronized they can be put into temporal registration by estimating their values at common points of time. The signals can be registered spatially by having the sensors coaligned on the same platform. Signal-level fusion is usually not feasible if the sensors are distributed on different platforms due to registration difficulties and bandwidth limitations involved in communicating the signals between the platforms. The most common means of measuring the improvement in quality is the reduction in the expected variance of the fused signal (see, e.g., Fig. 3(d)). One means of implementing signal level fusion is by taking a weighted average of the composite signals, where the weights are based on the estimated variances of the signals. If the signals are multidimensional the Kalman filter, for example, can be used for fusion.

1) Weighted Average: One of the simplest and most intuitive general methods of fusion is to first threshold redundant sensory information provided by a group of sensors to eliminate spurious measurements, and then take a weighted average of the information and use this as the fused value. While this method allows for the real-time processing of dynamic low-level data, in most cases the Kalman filter is preferred because it provides a method that is nearly equal in processing requirements and, in contrast to a weighted average, results in estimates for the fused data that are optimal in a statistical sense (a Kalman filter for a one-dimensional signal provides optimum weighting factors).

2) Kalman Filter: The Kalman filter [37] for a general introduction) is used in a number of multisensor systems when it is necessary to fuse dynamic low-level redundant data in real time. The filter uses the statistical characteristics of the measurement model to recursively determine estimates for the fused data that are optimal in a statistical sense. If the system can be described with a linear model and both the system and sensor error can be modeled as white Gaussian noise, the Kalman filter will provide unique statistically optimal estimates for the fused data. The recursive nature of the filter makes it appropriate for use in systems without large data storage capabilities. Examples of the use of the filter for multisensor fusion include: object recognition using sequences of images, robot navigation, multitarget tracking, inertial navigation, and remote sensing. In some of these applications the "U-D (unit upper triangular and diagonal matrix) covariance factorization filter" or the "extended Kalman filter" is used in place of the conventional Kalman filter if, respectively, numerical instability or the assumption of approximate linearity for the system model present potential problems. Durrant-Whyte, Rao, and Hu [15] have used the extended Kalman filter in a decentralized fusion architecture, and Ayache and Faugeras [3] have used it for building and updating the three-dimensional world model of a mobile robot.

3) Consensus Sensors: Luo, Lin, and Scherp [34] have developed a method for the fusion of redundant information from multiple sensors that can be used within a hierarchical phase-template paradigm for multisensor integration. The central idea behind the method is to first eliminate from consideration the sensor information that is likely to be in error and then use the information from the remaining "consensus sensors" to calculate a fused value. The information from each sensor is represented as a probability density function and the optimal fusion of the information is determined by finding the Bayesian estimator that maximizes the likelihood function of the consensus sensors.

C. Pixel-Level Fusion

Pixel-level fusion can be used to increase the information content associated with each pixel in an image formed through a combination of multiple images, e.g., the fusion of a range image with a two-dimensional intensity image adds depth information to each pixel in the intensity image that can be useful in the subsequent processing of the image. The different images to be fused can come from a single imaging sensor (e.g., a multispectral camera) or a group of sensors (e.g., stereo cameras). The fused image can be created either through the pixel-by-pixel fusion or through the fusion of associated local neighborhoods of pixels in each of the component images. The images to be fused can be modeled as a realization of a stochastic process defined across the image (e.g., a Markov random field), with the fusion process considered as an estimation procedure, or the information associated with each pixel in a component image can be considered as an additional dimension of the information associated with its corresponding pixel in the fused image (e.g., the two dimensions of depth and intensity associated with each pixel in a fused range-intensity image). Sensor registration is not a problem if either a single sensor is used or multiple sensors are used that provide images of the same resolution and share the same optics and mechanics (e.g., a laser radar operating at the same frequency as an infrared sensor and sharing the same optics and scanning mechanism). If the images to be fused are of different resolution, then a mapping needs to be specified between corresponding regions in the images. The sensors used for pixel-level fusion need to be accurately coaligned so that their images will be in spatial registration. This is usually achieved through locating the sensors on the same platform. The disparity between the locations of the sensors on the platform can be used as an important source of information in the fusion process, e.g., to determine a depth value for each pixel in binocular fusion. The improvement in quality associated with pixel-level fusion can most easily be assessed through the improvements noted in the performance of image processing tasks (e.g., segmentation, feature extraction, and restoration) when the fused image is being used as compared to the use of the individual component images.

The fusion of multisensor data at the pixel level can serve to increase the useful information content of an image so that more reliable segmentation can take place and more discriminating features can be extracted for further processing. Pixel-level fusion can take place at various levels of representation: the fusion of the raw signals from multiple sensors prior to their association with a specific pixel, the fusion of corresponding pixels in multiple registered images to form a composite or fused image, and the use of corresponding pixels or local groups of pixels in multiple registered images for segmentation and pixel-level feature extraction (e.g., an edge image). Fusion at the pixel level is useful in terms of total system processing requirements because use is made of the multisensor data prior to processing-intensive functions like feature matching, and can serve to increase overall performance in tasks like object recognition because the presence of certain substructures like edges in an image from one sensor usually indicates their presence in an image from another sufficiently similar sensor. Duane [14] has reported better object classification performance using features derived from the pixel-level fusion of TV and forward-looking infrared images as compared to the combined use of features derived independently from each separate image.

In order for pixel-level fusion to be feasible, the data provided by each sensor must be able to be registered at the pixel level and, in most cases, must be sufficiently similar in terms of its resolution and information content. The most obvious candidates for pixel-level fusion include sequences of images from a single sensor and images from a group of identical sensor (e.g., stereo vision). Other sensor combinations that make extensive use of pixel-level fusion include a coaligned laser radar and forward-looking infrared sensor. Although it is possible to use many of the general multisensor fusion methods for pixel-level fusion, e.g., Bayesian estimation, four methods are discussed in this section that are particularly useful for fusion at the pixel level: logical filters, mathematical morphology, image algebra, and simulated annealing. What makes these four methods useful for pixel-level fusion is that 1) each method facilitates highly parallel processing because, at most, only a local group of pixels are used to process each pixel, and 2) each method can easily be used to process a wide variety of images from different types of sensors because no problem or sensor specific probability distributions for pixel values are required, thus alleviating the need for either assuming a particular distribution or estimating a distribution through supervised training (only very general assumptions concerning pixel statistics are needed in simulated annealing to characterize a Markov random field for an image).

1) Logical Filters: One of the most intuitive methods of fusion the data from two pixels is to apply logical operators, e.g., if the values of both pixels are above particular thresholds the resulting AND filter is assumed to be true. Features derived from and image to which the AND filter was applied could then be expected to correspond to significant aspects of the environment. In a similar manner, an OR filter could be used to very reliably segment an image because all of the available information would be available. Ajjimarangsee and Huntsberger [2] have made use of some of the results concerning the means by which rattlesnakes fuse visual and infrared information to develop a set of logical filters that can be used for the unsupervised clustering of visual and infrared remote sensing information. Six logical filters are applied to remote sensing information that correspond to the six types of bimodal neurons found in the optic tectum of the rattlesnake: AND, OR, visible enhanced infrared, infrared enhanced visual, visible inhibited infrared, and infrared inhibited visible filters. The two inhibitory filters in effect implement an exclusive OR filter.

2) Mathematical Morphology: Mathematical morphology [20] is a method of image analysis that transforms each pixel of an image through the use of a set of morphological operators derived from the basic operations of set union, intersection, difference, and their conditional combinations; e.g., dilation and erosion operators are used to expand and shrink an image, respectively. Lee [20] has used binary morphological processing for the fusion of registered images from a pair of millimeter-wave radars operating at different frequencies. The fused image was found to improve road and terrain boundary detection. The binary morphological processing starts with two pixel-level feature sets extracted from the images of both radars. A high-confidence "core" feature set is derived from both feature sets through set intersection if both sets support each other, and through set difference if both sets are competing. A "potential" feature set is derived from both feature sets through set union if the sets are supporting, and through the union of one set with the complement of the other if both sets are competing. The morphological operations of conditional dilation and conditional erosion are used to fuse the core and potential feature sets. Conditional dilation extracts only those connected components of the potential feature set that have a non-empty intersection with the core feature set and is especially useful in rejecting clutter when the potential feature set includes both good feature and clutter components. Conditional erosion is useful for filling in missing segments of component boundaries in the core feature set.

3) Image Algebra: Image algebra [43] is a high-level algebraic language for describing image processing algorithms that is sufficiently complex to provide an environment for multisensor Pixel-level fusion. The four basic types of image algebra operands are coordinate sets, value sets, images, and templates. Coordinate sets can be defined as rectangular, hexagonal, or toroidal discrete arrays, or layers of rectangular arrays, and provide a coherent approach to the representation of sensor images that may have different tessellations or resolutions. If the images from the multiple sensors used for fusion have identical underlying coordinate systems the coordinate set is called homogeneous; otherwise it is termed heterogeneous. Value sets usually correspond to the set of integers, real or complex numbers, or binary numbers of a fixed length, and have the usual arithmetical and logical operations defined on them. A value set is called homogeneous if all of its values come from the same set of numbers; otherwise it is termed heterogeneous. Images are the most fundamental of image algebra's operands and are defined as the graph of a function from a coordinate set to a value set. Templates and template operations are the most powerful tool of image algebra and serve to unify and generalize into one mathematical entity the concepts of templates, masks, windows, the structuring elements of mathematical morphology, and other functions defined on pixel neighborhoods. Template operations

are used to transform images, and can be used to define a particular image processing algorithm in an implementation and machine independent way. The three basic template operations used to transform real-valued images are generalized convolution, multiplicative maximum, and additive maximum. Template operations can be used for local and global convolutions and for changing the dimensionality, or size and shape, of images.

Ritter, Wilson, and Davidson [44] have discussed the use of image algebra for multisensor pixel-level fusion. They have defined an image to be a "multisensor" image if its coordinate set is heterogeneous, and "multivalue" if its value set is homogeneous or "multidata" if it is heterogeneous. A data fusion function is any function that maps a value set of higher dimension to one of lower dimension. The most common data fusion operation is the "reduce" operation where, e.g., a vector-valued image is reduced to a real-valued image. A multilevel template can be thought of as a stack of templates that can operate differently on the different levels of a multivalue or multisensor image.

Chen [9] has extended the basic image algebra formalism to include incomplete and uncertain information. A stochastic image algebra is defined for image analysis through the use of a three-dimensional Markov random field as a model of an image. The use of a Markov random field to model images allows relaxation techniques like simulated annealing to be used for processing.

4) Simulated Annealing: Simulated annealing is a relaxation-based optimization technique that, when used in image processing applications [19], [50], amounts to viewing pixel values and the neighborhood in which they reside as states of atoms or molecules in a physical system. An energy function is assigned to the physical system and determines its Gibbs distribution. Due to the equivalence of the Gibbs distribution to a Markov random field, the energy function also determines an image model if the image can be represented as a Markov random field. Gradual temperature reductions in the energy function are used to relax or anneal the physical system towards a global minimum energy state which corresponds to the maximum *a posteriori* estimate of the true image given an initial image that is in some way corrupted. Local minima energy states are avoided during the relaxation process because changes in the system toward lower energy states are only favored and not strictly imposed.

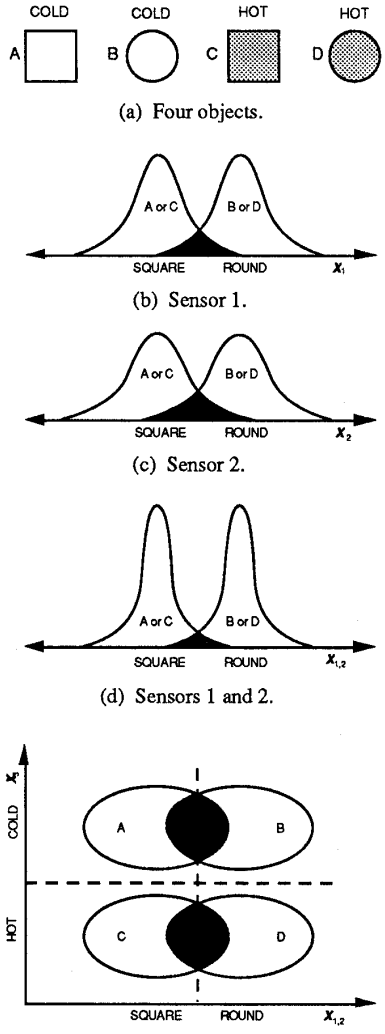
The use of simulated annealing for pixel-level fusion reduces to the problem of finding energy functions that can adequately describe appropriate constraints on the final fused image. Wright [51] has generalized the basic simulated annealing technique for image processing by creating a probability measure on a Markov random field which, in addition to modeling the information content of a single image, takes into consideration the information content of other registered images of the same view. Images from dissimilar sensors can be fused at the pixel level because the measure does not directly use the absolute values of the pixels in each image. Landa and Scheff [28] and Clifford and Nasrabadi [12] have used simulated annealing for the pixel-level binocular fusion of the images from two cameras in order to estimate depth. Clifford and Nasrabadi's fusion method uses intensity- and edge-based images together with optical flow data to compensate for partially occluded regions in images.

D. Feature-Level Fusion

Feature-level fusion can be used to both increase the likelihood that a feature extracted from the information provided by a sensor actually corresponds to an important aspect of the environment and as a means of creating additional composite features for use by the system. A feature provides for data abstraction and is created either through the attachment of some type of semantic meaning to the results of the processing of some spatial and/or temporal segment of the sensory data or, in the case of fusion, through a combination of existing features. Typical features extracted from an image and used for fusion include edges and regions of similar intensity or depth. When multiple sensors report similar features at the same location in the environment, the likelihood that the features are actually present can be increased and the accuracy with which they are measured can be improved; features that do not receive such support can be as spurious artifacts and eliminated. An additional feature, created as a result of the fusion process, may be either a composite of the component features (e.g., an edge that is composed of segments of edges detected by different sensors) or an entirely new type of feature that is composed of the attributes of its component features (e.g, a three-dimensional edge formed through the fusion of corresponding edges in the images provided by stereo cameras). The geometrical form, orientation, and position of a feature, together with its temporal extent, are the most important aspects of the feature that need to be represented so that it can be registered and fused with other features. In some cases, a feature can be made invariant to certain geometrical transformations (e.g., translation and rotation in an image plane) so that all of these aspects do not have to be explicitly represented. The sensor registration requirements for feature-level fusion are less stringent than those for signal- and pixel-level fusion, with the result that the sensors can distributed across different platforms. The geometric transformation of a feature can be used to bring it into registration with other features or with a world model. The improvement in quality associated with feature-level fusion can be measured through the reduction in processing requirements resulting from the elimination of spurious features, the increased accuracy in the measurement of a feature (used, e.g., to determine the pose of an object), and the increase in performance associated with the use additional features created through fusion (e.g., increased object recognition capabilities).

E. Symbol-Level Fusion

Symbol-level fusion allows the information from multiple sensors to be effectively used together at the highest level of abstraction. Symbol-level fusion may be the only means by which sensory information can be fused if the information provided by the sensors is very dissimilar or refers to different regions in the environment. The symbols used for fusion can originate either from the processing of the information provided by the sensors in the system, or through symbolic reasoning processes that may make use of *a priori* information from a world model or sources external to the system (e.g., intelligence reports indicating the likely presence of certain targets in the environment). A symbol derived from sensory information



(a) Four objects.

(b) Sensor 1.

(c) Sensor 2.

(d) Sensors 1 and 2.

(e) Sensors 1, 2, and 3.

Fig. 3. The discrimination of four different objects using redundant and complementary information from three sensors. (a) Four objects (A, B, C, and D) distinguished by the features "shape" (square vs. round) and "temperature" (hot vs. cold). (b) 2-D distributions from Sensor 1 (shape). (c) Sensor 2 (shape). (d) 2-D distributions resulting from fusion of redundant shape information from Sensors 1 and 2. (e) 3-D distributions resulting from fusion of complementary information from Sensors 1 and 2 (shape), and Sensor 3 (temperature).

represents a decision that has been made concerning some aspect of the environment (symbol-level fusion is sometimes termed "decision-level fusion"). The decision is usually made by matching features derived from the sensory information to a model. The symbols used for fusion typically have associated with them a measure of the degree to which the sensory information matches the model. A single uncertainty measure is used to represent both the degree of mismatch and any of the inherent uncertainty in the sensory information provided by the sensors. The measure can be used to indicate the relative weight that a particular symbol should be given in the fusion process. Sensor registration is usually not explicitly considered in symbol-level fusion because the spatial and temporal extent of the sensory information upon which a symbol is based has already been explicitly considered in the generation of the symbol, e.g., the underlying features upon which a group of symbols are based are already in registration. If the symbols to be fused are not in registration, spatial and temporal attributes can be associated with the symbols and used for their registration. Different forms of logical and statistical inference are used for symbol-level fusion. In logical inference the individual symbols to be fused represent terms in logical expressions and the uncertainty measures represent the truth values of the terms. If a logical expression represents a production rule, the uncertainty measures can be used to create certainty factors for the terms. In statistical inference the individual symbols to be fused are represented as conditional probability expressions and their uncertainty measures correspond to the probability measures associated with the expressions. The improvement in quality associated with symbol-level fusion is represented by the increase in the truth or probability values of the symbols created as a result of the inference process.

Henkind and Harrison [24] have analyzed and compared four of the uncertainty calculi used in many symbol-level fusion techniques: Bayesian estimation, Dempster-Shafer evidential reasoning, fuzzy set theory, and the confidence factors used in production rule-based systems. The computational complexity of these calculi are compared and their underlying assumptions are made explicit. Cheng and Kashyap [10] have compared the use of Bayesian estimation and Dempster-Shafer evidential reasoning for evidence combination.

1) An Object Recognition Example: Fig. 3 illustrates the distinction between complementary and redundant information in the task of object recognition. Four objects are shown in Fig. 3(a). They are distinguished by the two independent features shape and temperature. Sensors 1 and 2 provide redundant information concerning the shape of an object, and Sensor 3 provides information concerning its temperature. Fig. 3(b) and (c) show hypothetical frequency distributions for both "square" and "round" objects, representing each sensor's historical (i.e., tested) responses to such objects. The bottom axes of both figures represent the range of possible sensor readings. The output values x_1 and x_2 correspond to some numerical "degree of squareness or roundness" of the object as determined by each sensor, respectively. Because Sensors 1 and 2 are not able to detect the temperature of an object, objects A and C (as well as B and D) can not be distinguished. The dark portion of the axis in each figure corresponds to the range of output values where there is uncertainty as to the shape of the object being detected. The dashed line in each figure corresponds to the point at which, depending on the output value, objects can be distinguished in terms of a feature. Fig. 3(d) is the frequency distribution resulting from the fusion of x_1 and x_2. Without specifying a particular method of fusion, it is usually true that the distribution corresponding to the fusion of redundant information would have less dispersion than its component distributions. Under very general assumptions, a plausibility argument can be made that the relative probability of the fusion process not reducing the uncertainty is zero [42]. The uncertainty in Fig. 3(d) is shown as approximately half that of Fig. 3(b) and (c). In Fig. 3(e),

complementary information from Sensor 3 concerning the independent feature temperature is fused with the shape information from Sensors 1 and 2 shown in Fig. 3(d). As a result of the fusion of this additional feature, it is now possible to discriminate between all four objects. This increase in discrimination ability is one of the advantages resulting from the fusion of complementary information. As mentioned above, the information resulting from this second fusion could be at a higher representational level (e.g., the result of the first fusion, $x_{1,2}$, may still be a numerical value, while the result of the second, $x_{1,2,3}$, could be a symbol representing one of the four possible objects).

2) Bayesian Estimation: Bayesian estimation provides a formalism for multisensor fusion that allows sensory information to be combined according to the rules of probability theory. Uncertainty is represented in terms of conditional probabilities $P(Y \mid X)$, where $P(Y) = P(Y \mid X)$ if X remains constant. Each $P(Y \mid X)$ takes a value between 0 and 1, where 1 represents absolute belief in proposition Y given the information represented by proposition X and 0 represents absolute disbelief. Bayesian estimation is based on the theorem from basic probability theory known as "Bayes' rule":

$$P(Y \mid X) = \frac{P(X \mid Y)P(Y)}{P(X)},$$

where $P(Y \mid X)$, the "posterior probability," represents the belief accorded to the hypothesis Y given the information represented by X which is calculated by multiplying the "prior probability" associated with Y, $P(Y)$, by the "likelihood" $P(X \mid Y)$ of receiving X given that Y is true. The denominator $P(X)$ is a normalizing constant.

The redundant information from a group if n sensors, S_1 through S_n, can be fused together using the odds and likelihood ratio formulation of Bayes' rule. The information represented by X_i concerning Y from S_i is characterized by $P(X_i \mid Y)$ and the likelihood $P(X_i \mid \Re Y)$ given the negation of Y, or by the "likelihood ratio":

$$L(X_i \mid Y) = \frac{P(X_i \mid Y)}{P(X_i \mid \neg Y)}.$$

Defining the "prior odds" on Y as

$$O(Y) = \frac{P(Y)}{P(\neg Y)},$$

and assuming that the operation of each sensor is independent of the operation of the other sensors in the system, the "posterior odds" on Y given the information $X_1, ..., X_n$ from the n sensors are given by the product

$$O(Y \mid X_1,...,X_n) = O(Y)\prod_{i=1}^{n} L(X_i \mid Y).$$

The posterior odds are related to the posterior probability by

$$P(Y \mid X_1,...,X_n) = \frac{O(Y \mid X_1,...,X_n)}{1+O(Y \mid X_1,...,X_n)}.$$

The above formulation can also be used to fuse together a sequence of information from a single sensor provided that the uncertainty of the information can be assumed to be independent over time.

The application of Bayesian estimation for multisensor fusion can be illustrated using the object recognition example given above. Sensors 1 and 2, S_1 and S_2, provide redundant information relative to each other concerning the shape of the objects to be recognized. Let the propositions S and R represent the hypotheses that the object being sensed is square or round, respectively, and let S_1, R_1, S_2, and R_2 represent the shape indicated in the information provided by S_1 and S_2.

Given the information $P(S_1 \mid S) = 0.82$ and $P(S_2 \mid S) = 0.71$ from S_1 and S_2 concerning the hypothesis S, and assuming that square or round objects are equally likely to be encountered, i.e., $P(S) = P(R) = 0.5$, the posterior odds on S given the fusion of the information from both sensors are

$$O(S \mid S_1,S_2) = \frac{P(S)}{P(\neg S)} \cdot \frac{P(S_1 \mid S)}{P(S_1 \mid \neg S)} \cdot \frac{P(S_2 \mid S)}{P(S_2 \mid \neg S)}$$
$$= \frac{0.5}{0.5} \cdot \frac{0.82}{0.18} \cdot \frac{0.71}{0.29},$$

which corresponds to a posterior probability of

$$P(S \mid S_1,S_2) = \frac{O(S \mid S_1,S_2)}{1+O(S \mid S_1,S_2)} = \frac{11.15}{1+11.5} = 0.92.$$

In a similar manner, given the information $P(R_1|R) = 0.12$ and $P(R_2|R) = 0.14$, the posterior probability accorded the hypothesis R can be determined to be 0.02. The posterior probabilities of both hypotheses do not sum to unity in this example due to an assumed inherent uncertainty in the operation of S_1 and S_2 of 6 and 15 percent, respectively. If, for example, it is known *a priori* that only a third of the objects likely to be encountered are square, the posterior odds on S would be reduced by half and the odds on R would double.

3) Dempster-Shafer Evidential Reasoning: Garvey, Lowrance, and Fischler [18] introduced the possibility of using Dempster-Shafer evidential reasoning for multisensor fusion. The use of evidential reasoning for fusion allows each sensor to contribute information at its own level of detail, e.g., one sensor may be able to provide information that can be used to distinguish individual objects, while the information from another sensor may only be able to distinguish classes of objects; the Bayesian approach, in contrast, would not be able to fuse the information from both sensors. Dempster-Shafer evidential reasoning [46] is an extension to the Bayesian approach that makes explicit any lack of information concerning a proposition's probability by separating firm belief for the proposition from just its plausibility. In the Bayesian approach all propositions (e.g., objects in the environment) for which there is no information are assigned an equal *a priori* probability. When additional information from a sensor becomes available and the number of unknown propositions is large relative to the number of known propositions, an intuitively unsatisfying result of the Bayesian approach is that the probabilities of known propositions become unstable. In the Dempster-Shafer approach this is avoided by not assigning unknown propositions an *a priori* probability (unknown propositions are assigned instead to "ignorance"). Ignorance is reduced (i.e., probabilities are assigned to these propositions) only when supporting information becomes available.

In Dempster-Shafer evidential reasoning the set Θ, termed the "frame of discernment," is composed of mutually exclusive and exhaustive propositions termed "singletons." The level of detail represented by a singleton corresponds to the lowest level of information that is able to be discerned through the fusion of

information from a group of sensors or other information sources, e.g., a knowledge base. Given n singletons, the power set of Θ, denoted by 2^{Θ}, contains 2^n elements and is composed of all the subsets of Θ including Θ itself, the empty set ϕ, and each of the singletons. The elements of 2^{Θ} are termed propositions and each subset is composed of disjunction of singletons. The set of propositions $\{A_j \mid A_j \in 2^{\Theta}\}$ for which a sensor is able to provide direct information are termed its "focal elements." For each sensor S_i, the function

$$m_i\colon \{A_j \mid A_j \in 2^{\Theta}\} \rightarrow [0, 1],$$

termed a "basic probability assignment," maps a unit of probability mass or belief across the focal elements of S_i subject to the conditions

$$m_i(\phi) = 0,$$

and

$$\sum_{A_j \in 2^{\Theta}} m_i(A_j) = 1.$$

Any probability mass not assigned a proper subset of Θ is included in $m_i(\Theta)$ and is assumed to represent the residual uncertainty of S_i that is distributed in some unknown manner among its focal elements.

A "belief" or "support" function, defined for S_i as

$$bel_i(A) = \sum_{A_j \subseteq A} m_i(A_j),$$

is used to determine the lower probability or minimum likelihood of each proposition A. In a similar manner, "doubt," "plausibility," and "uncertainty" functions are defined as

$$\mathrm{dbt}_i(A) = \mathrm{bel}_i(A^c),$$

$$pls_i(A) = 1 - dbt_i(A),$$

and

$$u_i(A) = pls_i(A) - bel_i(A).$$

The degree of doubt in A is the degree of belief in the complement of A. The plausibility function determines the upper probability or maximum likelihood of A and represents the mass that is free to move to the belief of A as additional information becomes available. The uncertainty of A represents the mass that has not been assigned for or against belief in A. The Bayesian approach would correspond to the situation where $u_i(A) = 0$ for all $A \in 2^{\Theta}$. The Dempster-Shafer formalism allows the representation of total ignorance for a proposition A since $bel(A) = 0$ does not imply $dbt(A) > 0$, even though $dbt(A) = 1$ does imply $bel(A) = 0$. The interval $[bel(A), pls(A)]$ is termed a "belief interval" and represents, by its magnitude, now conclusive the information is for proposition A, e.g., total ignorance concerning A is represented as [0, 1], while [0, 0] and [1, 1] represent A as being false and true, respectively.

"Dempster's rule of combination" is used to fuse together the propositions X and Y from the two sensors S_i and S_j:

$$m_{i,j}(A) = \frac{\sum_{X \cap Y = A} m_i(X) m_j(Y)}{1 - \sum_{X \cap Y = \phi} m_i(X) m_j(Y)},$$

whenever $A \neq \phi$, and where $m_{i,j}$ is the orthogonal sum $m_i \oplus m_j$ and $X, Y \in 2^{\Theta}$. The denominator is a normalization factor that forces the new masses to sum to unity, and may be viewed as a measure of the degree of conflict or inconsistency in the information provided by S_i and S_j. If the factor is equal to 0 the sensors are completely inconsistent and the orthogonal sum operation is undefined. The combination rule narrows the set of propositions by distributing the total probability mass into smaller and smaller subsets, and can be used to find positive belief for singleton propositions that may be embedded in the complementary information (i.e., focal elements composed of disjunctions of singleton propositions) provided by a group of sensors.

The application of Dempster-Shafer evidential reasoning for multisensor fusion can be illustrated using the object recognition example given above. Θ is composed of the four singleton propositions A, B, C, and D, corresponding to the four objects to be recognized. Each of the three sensors used to recognize the objects is only able to provide information to distinguish a particular class of objects, e.g., square versus round objects. Sensors 1 and 2, S_1 and S_2, provide redundant information relative to each other concerning the shape of the objects, represented as the focal elements $A \vee C$ (square) and $B \vee D$ (round). The information from S_1 and S_2 is the same as that used to illustrate Bayesian estimation. Sensor 3, S_3, provides complementary information relative to S_1 and S_2 concerning the temperature of the objects, represented as the focal elements $A \vee B$(cold) and $C \vee D$ (hot).

The mass assignments resulting from the fusion of the information from S_1 and S_2 using Dempster's rule are shown in Table II. The probability mass assigned to each of the focal elements of the sensors reflects the difference in the sensors' accuracy indicated by the frequency distributions shown in Fig. 3(b) and (c); e.g., given that the object being sensed is most likely square, the greater mass attributed to $m_1(A \vee C)$ as compared to $m_2(A \vee C)$ reflects S_1's greater accuracy as compared to S_2. The difference in mass attributed to the object possibly being round reflects the amount of overlap in the distributions for each shape class. The mass attributed to $m(\Theta)$ for each sensor reflects the amount by which the focal element masses have been reduced to account for the inherent uncertainty in the information provided by each sensor. The normalization factor is calculated as 1 minus the sum of the two k's in the table, or $1 - 0.2 = 0.8$. As a result of the fusion, the belief attributed to the object being square has increased from $bel_1(A \vee C) = 0.82$ and $bel_2(A \vee C) = 0.71$ to $bel_{1,2}(A \vee C) = 0.93475$ (the sum of the $m_{1,2}(A \vee C)$ in the table). This increase is also indicated by the narrower distribution shown for the fused information in Fig 3(d) of Chapter 1.

Table III shows the mass assignments resulting from the fusion of the combined information from S_1 and S_2, $S_{1,2}$, with the focal elements of S_3. As a result of the fusion, positive belief can be attributed to the individual objects. The most likely object is A, as indicated by

$$bel_{1,2,3}(A) = m_{1,2,3}(A) = 0.85997,$$

and

$$dbt_{1,2,3}(A) = m_{1,2,3}(B) + m_{1,2,3}(C) + m_{1,2,3}(D) + m_{1,2,3}(B \vee D) + m_{1,2,3}(C \vee D)$$

TABLE II

FUSION USING SENSORS 1 AND 2.

		S_2		
		$m_2(A \vee C) = 0.71$	$m_2(B \vee D) = 0.14$	$m_2(\Theta) = 0.15$
S_1	$m_1(A \vee C) = 0.82$	$m_{1,2}(A \vee C) = 0.72775$	$k = 0.1148$	$m_{1,2}(A \vee C) = 0.15375$
	$m_1(B \vee D) = 0.12$	$k = 0.0852$	$m_{1,2}(B \vee D) = 0.021$	$m_1(B \vee D) = 0.0225$
	$m_1(\Theta) = 0.06$	$m_{1,2}(A \vee C) = 0.05325$	$m_1(B \vee D) = 0.0105$	$m_{1,2}(\Theta) = 0.01125$

TABLE III

FUSION USING SENSORS 1, 2, AND 3.

		S_3		
		$m_3(A \vee B) = 0.92$	$m_3(C \vee D) = 0.06$	$m_3(\Theta) = 0.02$
$S_{1,2}$	$m_{1,2}(A \vee C) = 0.93475$	$m_{1,2,3}(A) = 0.85997$	$m_{1,2,3}(C) = 0.056085$	$m_{1,2,3}(A \vee C) = 0.018695$
	$m_{1,2}(B \vee D) = 0.054$	$m_{1,2,3}(B) = 0.04968$	$m_{1,2,3}(D) = 0.00324$	$m_{1,2,3}(B \vee D) = 0.00108$
	$m_{1,2}(\Theta) = 0.01125$	$m_{1,2,3}(A \vee B) = 0.01035$	$m_{1,2,3}(C \vee D) = 0.000675$	$m_{1,2,3}(\Theta) = 0.000225$

$$= 0.11076.$$

The evidence for this conclusion is quite conclusive as indicated by a small uncertainty and a narrow belief interval for A:

$$\mathrm{u}1,2,3(\mathrm{A}) = \mathrm{pls}1,2,3(\mathrm{A}) - \mathrm{bel}1,2,3(\mathrm{A}) = 0.02927,$$

$$[bel_{1,2,3}(A), pls_{1,2,3}(A)] = [0.85997, 0.88924],$$

where the plausibility of A is

$$pls_{1,2,3}(A) = 1 - dbt_{1,2,3}(A) = 0.88924.$$

The least likely object is also quite conclusively D. If additional information becomes available, e.g., that the object was stored inside a refrigerated room, it can easily be combined with the previous evidence to possibly increase the conclusiveness of the conclusions.

4) Production Rules with Confidence Factors: Production rules can be used to symbolically represent the relation between sensory information and an attribute that can be inferred from the information. Production rules that are not directly based on sensory information can be easily combined with sensory information-based rules as part of an overall high-level reasoning system, e.g., expert systems. The use of production rules promotes modularity in the multisensor integration process because additional sensors can be added to the system without requiring the modification of existing rules.

The production rules used for multisensor fusion can be represented as the logical implication of a "conclusion" Y given a "premise" X, denoted as *if X then Y* or $X \rightarrow Y$. The premise X may be composed of a single proposition or the conjunction, disjunction, or negation of a group of propositions. The inference process can proceed in either a forward or backward chaining manner: in "forward-chaining" inference, a premise is given and its implied conclusions are derived; in "backward-chaining" inference, a proposition is given as a goal to be proven given the known information. In forward-chaining inference, the fusion of sensory information takes place both through the implication of the conclusion of a single rule whose premise is composed of a conjunction or disjunction of information from different sensors and through the assertion of a conclusion that is common to a group of rules.

Uncertainty is represented in a system using production rules through the association of a "certainty factor" (CF) with each proposition and rule. Each CF is a measure of belief or disbelief and takes a value $-1 \leq \mathrm{CF} \leq 1$, where CF = 1 corresponds to absolute belief, CF = −1 to absolute disbelief, and, for a proposition, CF = 0 corresponds to either a lack of information or an equal balance of belief and disbelief concerning the proposition. Uncertainty is propagated through the system using a "certainty factor calculus," e.g., the EMYCIN calculus [7].

Each proposition X and its associated CF is denoted as

$$X \; cf \; (\mathrm{CF}[X]),$$

where CF[X] is initially either known or assumed to be equal to 0. Given the set $\Re$ of rules in a system, each rule $r_i \in \Re$ and its associated CF is denoted as

$$r_i : X \rightarrow Y \; cf \; (\mathrm{CF}_i[X, Y]).$$

The CF of the premise X in r_i can be defined as

$$\mathrm{CF}_i[X] = \begin{cases} \mathrm{CF}[X] & \text{if } X = X_1 \\ \min(\mathrm{CF}[X_1], \ldots, \mathrm{CF}[X_n]) & \text{if } X = X_1 \wedge \ldots \wedge X_n \\ \max(\mathrm{CF}[X_1], \ldots, \mathrm{CF}[X_n]) & \text{if } X = X_1 \vee \ldots \vee X_n \\ -\mathrm{CF}[X] & \text{else.} \end{cases}$$

where each X_i is a proposition in X and $\neg X$ is the negation of X. The CF of the conclusion Y in r_i can be determined using

$$\mathrm{CF}_i[Y] = \begin{cases} -\mathrm{CF}_i[X] \cdot \mathrm{CF}_i[X, Y] & \text{if both CF's} < 0 \\ \mathrm{CF}_i[X] \cdot \mathrm{CF}_i[X, Y] & \text{else.} \end{cases}$$

The $\mathrm{CF}_i[X, Y]$ for r_i can be thought of as the $\mathrm{CF}_i[Y]$ that would result if r_i is implied and $\mathrm{CF}_i[X] = 1$.

If there is only one rule, r_Y, for which the unknown proposition Y is its conclusion, then $\mathrm{CF}[Y] = \mathrm{CF}_Y[Y]$. If there is more than one rule, then CF[Y] is determined by fusing together the $\mathrm{CF}_i[Y]$'s of all the r_i for which Y is their conclusion. Let

$$\Re_Y = \{r_i : x \rightarrow y \in \Re \mid \mathrm{CF}_i[x] \neq 0 \text{ and } y = Y\}$$

be the set of rules with known premises and Y as their conclusion. Given $N = |\Re_Y|$ such rules,

$$CF[Y] = CF[Y]_{j=N},$$

where, for every $r_i \in \Re_Y$, $CF[Y]_0 = 0$ and

$$CF[Y]_j = \begin{cases} CF[Y]_{j-1} + CF_i[Y]\cdot(1 - CF[Y]_{j-1}) & \text{both CF} > 0 \\ CF[Y]_{j-1} + CF_i[Y]\cdot(1 + CF[Y]_{j-1}) & \text{both CF} < 0 \\ \dfrac{CF[Y]_{j-1} + CF_i[Y]}{1 - \min(CF[Y]_{j-1}, CF_i[Y])} & \text{, else} \end{cases}$$

for $j = 1$ to N.

The application of production rules with certainty factors for multisensor fusion can be illustrated using the object recognition example given above. The information from the three sensors S_1, S_2, and S_3 is the same as that used in the illustrations of Bayesian estimation and Dempster-Shafer evidential reasoning. Let S_1 *cf* (0.87) and R_1 *cf* (–0.87) be the known propositions provided by S_1 concerning whether the objects being sensed are either square (S) or round (R), respectively. The two rules

$r_1 : S_1 \rightarrow S$ *cf* (0.94) and

$r_2 : R_1 \rightarrow R$ *cf* (0.94)

account for an inherent uncertainty of 6 percent in the information provided by S_1. Using only S_1, the certainty that the object being sensed is square is S *cf* (0.82) and that it is round is R *cf* (–0.82). The information S_2 *cf* (0.84) and R_2 *cf* (–0.84) from S_2, together with the additional rules

$r_3 : S_2 \rightarrow S$ *cf* (0.85) and

$r_4 : R_2 \rightarrow R$ *cf* (0.85)

can be fused with the redundant information from S_1 to increase the belief that the object is square to S *cf* (0.9478) and to increase the disbelief that it is round to R *cf* (–0.9478), where $CF[S] = 0.82 + 0.71(1 - 0.82)$ and $CF[R] = -0.82 - 0.71(1 - 0.82)$ corresponding to $\Re_S = \{r_1, r_3\}$ and $\Re_R = \{r_2, r_4\}$, respectively.

Let C_3 *cf* (0.94) and H_3 *cf* (–0.94) be the known propositions provided by S_3 concerning whether the objects are either cold (C) or hot (H), respectively. The two rules

$r_5 : C_3 \rightarrow C$ *cf* (0.98) and

$r_6 : H_3 \rightarrow H$ *cf* (0.98)

to account for the inherent uncertainty in S_3, together with the additional rules

$r_7 : S \wedge C \rightarrow A$ *cf* (1.0),

$r_8 : R \wedge C \rightarrow A$ *cf* (1.0),

$r_9 : S \wedge H \rightarrow A$ *cf* (1.0), and

$r_{10} : R \wedge H \rightarrow A$ *cf* (1.0),

enable the information from S_3 to be fused with the complementary information from S_1 and S_2 to determine the CF associated with the propositions A, B, C, and D, corresponding to the four possible types of objects. Having determined that C *cf* (0.92) and H *cf* (–0.92),

$$CF[A] = CF_7[S \wedge C] \cdot CF_7[S \wedge C, A]$$

$$= \min(CF[S] \wedge CF[R]) \cdot 1.0$$

$$= \min(0.9478, 0.92) = 0.92.$$

In a similar manner, $CF[B]$, $CF[C]$, and $CF[D]$ can be determined to be –0.9478, –0.92, and –0.9478, respectively.

The definition of a certainty factor calculus to use with production rules for multisensor fusion is *ad hoc* and will depend upon the particular application for which the system is being used. For example, the results of the object recognition example would more closely resemble the results found using Dempster-Shafer evidential reasoning if the definition of the CF of a conjunction of propositions in the premise of a rule was changed to correspond to the creation of a separate rule for each proposition, e.g., $S \rightarrow A$ and $C \rightarrow A$ instead of $S \wedge C \rightarrow A$ in r_7. Using this definition, the resulting CF's for A, B, C, and D would be 0.99, –0.014, 0.014, and –0.99, respectively (where a CF of 0 corresponds to a probability mass of 0.5).

IV. Conclusion

A. Future Research Directions

In addition to multisensor integration and fusion research directed at finding solutions to the problems already mentioned, research in the near future will likely be aimed at developing integration and fusion techniques that will allow multisensory systems to operate in unknown and dynamic environments. As currently envisioned, multisensor integration and fusion techniques will play an important part in the Strategic Defense Initiative in enabling enemy warheads to be distinguished from decoys [1]. Many integration and fusion techniques will be implemented on recently developed highly parallel computer architectures to take full advantage of the parallelism inherent in the techniques. The development of sensor modeling and interface standards would accelerate the design of practical multisensor systems [23]. Lyons and Arbib [35] have initiated the construction of a formal model of computation for sensory-based robotics that they term "robot schemas." Future extensions to their model will make it possible to reason about sensory interactions in a consistent and well-defined manner, and should facilitate the creation of the complex control programs required for multisensor robots. Continued research in the areas of artificial intelligence and neural networks will continue to provide both theoretical and practical insights. AI-based research may prove especially useful in areas like sensor selection, automatic task error detection and recovery, and the development of high-level representations; research based on neural networks may have a large impact in areas like object recognition through the development of distributed representations suitable for the associative recall of multisensory information, and in the development of robust multisensor systems that are able to self-organize and adapt to changing conditions (e.g., sensor failure).

The development of integrated solid-state chips containing multiple sensors has been the focus of much recent research [45]. As current progress in VLSI technology continues, it is likely that so-called "smart sensors" [38] will be developed that contain many of their low-level signal and fusion processing algorithms in circuits on the chip. In addition to a lower cost, a smart sensor

might provide a better signal-to-noise ratio, and abilities for self-testing and calibration. Currently, it is common to supply a multisensor system with just enough sensors for it to complete its assigned tasks; the availability of cheap integrated multisensors may enable some recent ideas concerning "highly redundant sensing" [49] to be incorporated into the design of intelligent multisensor systems—in some cases, high redundancy may imply the use of up to ten times the number of minimally necessary sensors to provide the system with a greater flexibility and insensitivity to sensor failure. In the more distant future, the development of micro or "gnat" [15] robots will necessarily entail the advancement of the state of the art in multisensor integration and fusion.

B. Guide to Survey and Review Papers

A number of recent papers have surveyed and reviewed different aspects of multisensor integration and fusion. An article on multisensor integration in the *Encyclopedia of Artificial Intelligence* has focused on the issues involved in object recognition [4]. Mitiche and Aggarwal [39] discuss some of the advantages and problems involved with the integration of different image processing sensors, and review recent work in that area. Garvey [17] has surveyed some of the different artificial intelligence approaches to the integration and fusion of information, emphasizing the fundamental role in artificial intelligence of the inference process for combining information. A number of the different knowledge representations, inference methods, and control strategies used in the inference process are discussed in his paper. Mann [36] provides a concise literature review as part of his paper concerning methods for integration and fusion that are based on the maintenance of consistent labels across different sensor domains. Luo and Kay [32, [33] and Blackman [6] have surveyed some of the issues of and different approaches to multisensor integration and fusion, with Blackman providing an especially detailed discussion of the data association problem, and Hackett and Shah [21] have surveyed a number of multisensor fusion papers and have classified them into the following six categories: scene segmentation, representation, three-dimensional shape, sensor modeling, autonomous robots, and object recognition. Recent research workshops have focused on the multisensor integration and fusion issues involved in manufacturing automation [23] and spatial reasoning [28]. Techniques for multisensor integration and fusion have been included in recent textbooks on artificial intelligent [13] and pattern recognition [48].

REFERENCES

[1] Adam, J.A. (1989). Star Wars in transition. *IEEE Spectrum*, 26(3), 32–38.

[2] Ajjimarangsee, P., and Huntsberger, T.L. (1988). Neural network model for fusion of visible and infrared sensor outputs. In P.S. Schenker (Ed.), *Proc. SPIE*, vol. 1003, *Sensor Fusion: Spatial Reasoning and Scene Interpretation* (pp. 153–160). Cambridge, MA.

[3] Ayache, N., and Faugeras, O. (1989). Maintaining representations of the environment of a mobile robot. *IEEE Trans. Robot. Automat.*, RA-5(6), 804–819.

[4] Bajcsy, R., and Allen, P. (1986). Multisensor integration. *Encyclopedia of Artificial Intelligence* (pp. 632–638). New York: Wiley.

[5] Barniv, Y., and Casasent, D. (1981). Multisensor image registration: Experimental verification. W.H. Carter (Ed.), *Proc. SPIE*, vol. 292, *Process. Images and Data from Optical Sensors* (pp.160–171). San Diego, CA.

[6] Blackman, S.S. (1988). Theoretical approaches to data association and fusion. In C.W. Weaver (Ed.), *Proc. SPIE*, vol. 931, *Sensor Fusion* (pp. 50–55). Orlando, FL.

[7] Buchanan, B.G., and Shortliffe, E.H., (Eds.). (1984). Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley.

[8] Bullock, T.E., Sangsuk-iam, S., Pietsch, R., and Boudreau, E.J. (1988). Sensor fusion applied to system performance under sensor failures. In C.W. Weaver (Ed.), *Proc. SPIE*, vol. 931, *Sensor Fusion* (pp. 131–138). Orlando, FL.

[9] Chen, S.S. (1989). Stochastic image algebra for multisensor fusion and spatial reasoning: A neural approach. In M.R. Weathersby (Ed.), *Proc. SPIE*, vol. 1098, *Aerospace Pattern Recog.* (pp. 146–154). Orlando, FL.

[10] Cheng, Y., and Kashyap, R.L. (1988). Comparison of Bayesian and Dempster's rules in evidence combination. In G.J. Erickson and C.R. Smith (Eds.), Maximum-Entropy and Bayesian Methods in Science and Engineering (vol. 2, pp. 427–433). Dordrecht, The Netherlands: Kluwer.

[11] Clark, J.J., and Yuille, A.L. (1990). *Data Fusion for Sensory Information Processing Systems*. Norwell, MA: Kluwer.

[12] Clifford, S.P., and Nasrabadi, N.M. (1988). Integration of stereo vision and optical flow using Markov random fields. *Proc. IEEE Int. Conf. Neural Networks* (pp. I-577–I-584). San Diego, CA.

[13] Dougherty, E.R., and Giardina, C.R. (1988). Mathematical Methods for Artificial Intelligence and Autonomous Systems. Englewood Cliffs, NJ: Prentice-Hall, pp. 273–277.

[14] Duane, G. (1988). Pixel-level sensor fusion for improved object recognition. In C.W. Weaver (Ed.), *Proc. SPIE*, vol. 931, *Sensor Fusion* (pp. 180-185). Orlando, FL.

[15] Durrant-Whyte, H.F., Rao, B.Y.S., and Hu, H. (1990). Toward a fully decentralized architecture for multi-sensor data fusion. *Proc. IEEE Int. Conf. Robotics and Automat.* (pp. 1331–1336). Cincinnati, OH.

[16] Flynn, A.M., and Brooks, R.A. (1988). MIT mobile robots—What's next? *Proc. IEEE Int. Conf. Robotics and Automat.* (pp. 611–617). Philadelphia, PA.

[17] Garvey, T.D. (1987). A survey of AI approaches to the integration of information. In R.G. Buser and F.B. Warren (Eds.), *Proc. SPIE*, vol. 782, *Infrared Sensors and Sensor Fusion* (pp. 68–82). Orlando, FL.

[18] Garvey, T.D., Lowrance, J.D., and Fischler, M.A. (1981). An inference technique for integrating knowledge from disparate sources. *Proc. 7th Int. Joint Conf. Artificial Intell.*, (pp. 319–325). Vancouver, BC, Canada.

[19] Geman, S., and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-6(6), 721–741.

[20] Giardina, C.R., and Dougherty, E.R. (1988). *Morphological Methods in Image and Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

[21] Hackett, J.K., and Shah, M. (1990). Multi-sensor fusion: A perspective. *Proc. IEEE Int. Conf. Robotics and Automat.* (pp. 1324–1330). Cincinnati, OH.

[22] Henderson, T., and Shilcrat, E. (1984). Logical sensor systems. *J. Robot. Syst.*, 1(2), 169–193.

[23] Henderson, T.C., Allen, P.K., Mitiche, A., Durrant-Whyte, H., and Snyder, W. (Eds.). (1987). *Workshop on multisensor integration in manufacturing automation* (Tech. Rep. UUCS–87–006). Univ. of Utah, Snowbird: Dept. of Comput. Sci.

[24] Henkind, S.J., and Harrison, M.C. (1988). An analysis of four uncertainty calculi. *IEEE Trans. Syst., Man Cybern.*, SMC-18(5), 700–714.

[25] Holm, W.A. (1987). Air-to-ground dual-mode MMW/IR sensor scene registration. In R.G. Buser and F.B. Warren (Eds.), *Proc. SPIE*, vol. 782, *Infrared Sensors and Sensor Fusion* (pp. 20–27). Orlando, FL.

[26] Hsiao, M. (1988). Geometric registration method for sensor fusion. In P.S. Schenker (Ed.), *Proc. SPIE*, vol. 1003, *Sensor Fusion: Spatial Reasoning and Scene Interpretation* (pp. 214–221). Cambridge, MA.

[27] Jakubowicz, O.G. (1988). Autonomous reconfiguration of sensor systems using neural nets. In C.W. Weaver (Ed.), *Proc. SPIE*, vol. 931, *Sensor Fusion* (pp. 197–203). Orlando, FL.

[28] Kak, A., and Chen, S. (Eds.). (1987). *Spatial Reasoning and Multi-Sensor Fusion: Proc. 1987 Workshop*. Los Altos, CA: Morgan Kaufmann.

[29] Landa, J., and Scheff, K. (1987). Binocular fusion using simulated annealing. *Proc. IEEE 1st Int. Conf. Neural Networks* (pp. IV-327–IV-334). San Diego, CA.

[30] Lee, J.S.J. (1988). Multiple sensor fusion based on morphological processing. In P.S. Schenker (Ed.), *Proc. SPIE*, vol. 1003, *Sensor Fusion: Spatial Reasoning and Scene Interpretation* (pp. 94–100). Cambridge, MA.

[31] Lee, R.H., and Van Vleet, W.B. (1988). Registration error analysis between dissimilar sensors. In C.W. Weaver (Ed.), *Proc. SPIE*, vol. 931, *Sensor Fusion* (pp. 109–114). Orlando, FL.

[32] Luo, R.C., and Kay, M.G. (1988). Multisensor integration and fusion: Issues and approaches. In C.W. Weaver, Ed, *Proc. SPIE*, vol. 931, *Sensor Fusion* (pp. 42–49). Orlando, FL.

[33] Luo, R.C., and Kay, M.G. (1989). Multisensor integration and fusion in intelligent systems. *IEEE Trans. Syst., Man Cybern.*, vol. SMC-19(5), 901–931.

[34] Luo, R.C., Lin, M., and Scherp, R.S. (1988). Dynamic multi-sensor data fusion system for intelligent robots. *IEEE J. Robot. Automat.*, RA-4(4), 386–396.

[35] Lyons, D.M., and Arbib, M.A. (1989). A formal model of computation for sensory-based robotics. *IEEE Trans. Robot. Automat.*, RA-5(3), 280–293.

[36] Mann, R.C. (1987). Multi-sensor integration using concurrent computing. In R.G. Buser and F.B. Warren (Eds.), *Proc. SPIE*, vol. 782, *Infrared Sensors and Sensor Fusion* (pp. 83–90). Orlando, FL.

[37] Maybeck, P.S. (1979, 1982). *Stochastic Models, Estimation, and Control* (Vols. 1 and 2). New York: Academic.

[38] Middelhoek, S., and Hoogerwerf, A.C. (1985). Smart sensors: When and where? *Sensors and Actuators*, 8, 39–48.

[39] Mitiche, A., and Aggarwal, J.K. (1986). Multiple sensor integration/fusion through image processing: A review. *Opt. Eng.*, 25(3), 380–386.

[40] Mitiche, A., Henderson, T.C., and Laganiere, R. (1988). Decision networks for multisensor integration in computer vision. In P.S. Schenker (Ed.), *Proc. SPIE*, vol. 1003, *Sensor Fusion: Spatial Reasoning and Scene Interpretation* (pp. 291–299). Cambridge, MA.

[41] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

[42] Richardson, J.M., and Marsh, K.A. (1988). Fusion of multisensor data. *Int. J. Robot. Res.*, 7(6), 78–96.

[43] Ritter, G.X., and Wilson, J.N. (1987). The image algebra in a nutshell. *Proc. First Int. Conf. Comp. Vision* (pp. 641–645). London.

[44] Ritter, G.X., Wilson, J.N., and Davidson, J.L. (1988). Image algebra application to multisensor and multidata image manipulation. In M.R. Weathersby (Ed.), *Proc. SPIE*, vol. 933, *Multispectral Image Processing and Enhancement* (pp. 2–7). Orlando, FL.

[45] Rojder, P., and Lauber, A. (1990). Using VLSI circuits for optimal signal handling in multisensorial robotic systems. In J.T. Tou and J.G. Balchen (Eds.), *Highly Redundant Sensing in Robotic Systems* (pp. 149–156). Berlin: Springer-Verlag.

[46] Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press.

[47] Swyt, D.A. (1989). AI in manufacturing: The NBS AMRF as an intelligent machine. *Robotics Autonomous Syst.*, 4, 327–332.

[48] Therrien, C.W. (1989). *Decision, Estimation and Classification: An Introduction to Pattern Recognition and Related Topics*. New York: Wiley, pp. 232–242.

[49] Tou, J.T., and Balchen, J.G. (Eds.). (1990). *Highly Redundant Sensing in Robotic Systems*. Berlin: Springer-Verlag.

[50] Wolberg, G., and Pavlidis, T. (1985). Restoration of binary images using stochastic relaxation with annealing. *Pattern Recog. Letters*, 3, 375–388.

[51] Wright, W.A. (1989). A Markov random field approach to data fusion and colour segmentation. *Image Vision Comp.*, 7(2), 144–150.