Multisensor Integration and Fusion in Intelligent Systems

REN C. LUO, SENIOR MEMBER, IEEE, AND MICHAEL G. KAY

Abstract — Interest has been growing in the use of multiple sensors to increase the capabilities of intelligent systems. The issues involved in integrating multiple sensors into the operation of a system are presented in the context of the type of information these sensors can uniquely provide. The advantages gained through the synergistic use of multisensory information can be decomposed into a combination of four fundamental aspects: the redundancy, complementarity, timeliness, and cost of the information. The role of multiple sensors in the operation of a particular system can then be defined as the degree to which each of these four aspects is present in the information provided by the sensors. A distinction is made between multisensor integration and the more restricted notion of multisensor fusion to separate the more general issues involved in the integration of multiple sensory devices at the system architecture and control level, from the more specific issues—possibly mathematical or statistical—involved in the actual combination (or fusion) of multisensory information. A survey is provided of the increasing number and variety of approaches to the problem of multisensor integration and fusion that have appeared in the literature in recent years—ranging from general paradigms, frameworks, and methods for integrating and fusing multisensory information, to existing multisensor systems used in different areas of application. General multisensor fusion methods, sensor selection strategies, and world models are surveyed, along with approaches to the integration and fusion of information from combinations of different types of sensors. Short descriptions of the role of multisensor integration and fusion in the operation of a number of existing mobile robots are provided, together with proposed high-level multisensory representations suitable for mobile robot navigation and control. Existing multisensor systems are surveyed in the following areas of application: industrial tasks like material handling, part fabrication (e.g., welding), inspection, and assembly; military command and control for battle management; space; target tracking; inertial navigation; and the remote sensing of coastal waters. A discussion is included of possible problems associated with creating a general methodology for multisensor integration and fusion—focusing on the methods used for modeling error or uncertainty in the integration and fusion process (e.g., the registration problem), the actual sensory information (i.e., the sensor model), and the operation of the overall system (e.g., multisensor calibration).

I. Introduction

In recent years interest has been growing in the synergistic use of multiple sensors to increase the capabilities of intelligent machines and systems. For these systems to use multiple sensors effectively, some method is needed for integrating the information provided by these sensors into the operation of the system. While in many multisensor systems the information from each sensor serves as a separate input to the system, the actual combination or fusion of information prior to its use in the system has been a particularly active area of research. Typical of the applications that can benefit from the use of multiple sensors are industrial tasks like assembly, military command and control for battlefield management, mobile robot navigation, multitarget tracking, and aircraft navigation. Common among all of these applications is the requirement that the system intelligently interact with and operate in an unstructured environment without the complete control of a human operator.

Manuscript received August 10, 1988; revised March 18, 1989. Preliminary versions of portions of this paper were presented at the SPIE Conference on Sensor Fusion, Orlando, FL, April 4–6, 1988, and published in the proceedings thereof, and the IEEE International Workshop on Intelligent Robots, Tokyo, Japan, October 31–November 2, 1988. This work was supported in part by the National Science Foundation under Grant DMC-8716126.

R. C. Luo is with the Robotics and Intelligent Systems Laboratory in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695-7911.

M. G. Kay is with the Department of Industrial Engineering, North Carolina State University, Raleigh, NC 27695-7906.

IEEE Log Number 8928399.

A. Biological Examples of the Synergistic Integration of Multisensor Information

Two of the major abilities that a human operator brings to the task of controlling a system are the use of a flexible body of knowledge and the ability to integrate synergistically information of different modality obtained through his or her senses. The increasing use of knowledge-based expert systems is an attempt to capture some aspects of this first ability; current research in multisensor integration is an attempt to capture, and possibly extend to additional modalities, aspects of this second ability. Thus a human's or other animal's ability to integrate multisensory information can provide an indication of what is ultimately achievable for intelligent systems (i.e., an existence proof) and insight into possible future research directions.

1) Ventriloquism: A well-known example of human multisensory integration is ventriloquism, in which the voice of the ventriloquist seems to an observer to come from the ventriloquist's dummy. The ability of visual information (the movement of the dummy's lips) to dominate the auditory information coming from the ventriloquist demonstrates the existence of some process of integration whereby information from one modality (audition) is interpreted solely in terms of information from another modality (vision). Howard [1] has reported research that found the discordance between visual and auditory information becomes noticeable only after the source of each has been separated beyond 30° relative to the observer (see Fig. 1). Notwithstanding ventriloquism, the use of information from these two modalities can increase the probability of

0018-9472/89/0900-0901$01.00

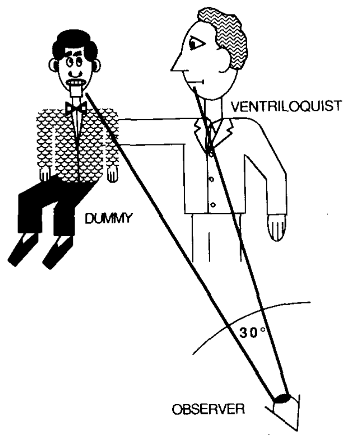

Fig. 1. Ventriloquism demonstrates existence of some process of human multisensory integration through ability of visual information (movement of dummy's lips) to dominate auditory information (from ventriloquist) for up to 30° separation of these information sources relative to observer.

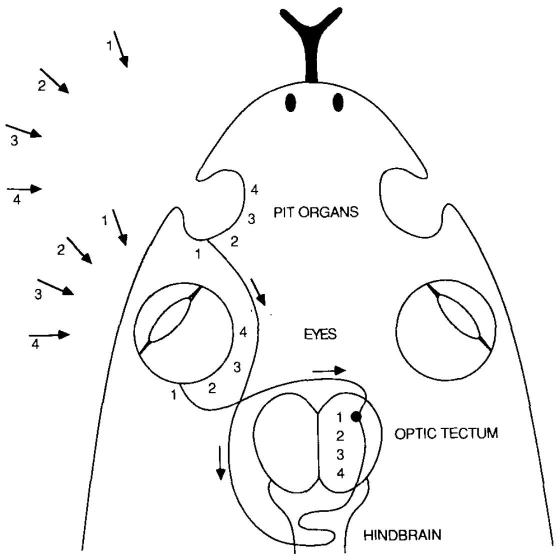

Fig. 2. Left eye and pit organ of rattlesnake are receiving information from Region 1 in environment. Information from both sources is represented on surface of optic tectum in similar spatial orientation. (Adapted from [2, p. 121].)

detecting an event in the environment when compared to the use of either modality alone.

2) Pit Vipers and Rattlesnakes: Although in humans the processes of multisensory integration have not yet been found, research on the less complex nervous systems of the pit viper and the rattlesnake has identified neurons in these snakes' optic tectums (a midbrain structure found in vertebrates) that are responsive to both visual and infrared information [2]. As shown in Fig. 2, both the left eye and pit organ of a rattlesnake are receiving information from Region 1 in the environment. Infrared information from the pit organ, together with visual information from the eye, are represented on the surface of the optic tectum in a similar spatial orientation so that each region of the optic tectum receives information from the same region of the environment. This allows certain "multimodal" neurons to respond to different combinations of visual and infrared information. Certain "or" neurons respond to information from either modality and could be used by the snake to detect the presence of prey in dim lighting conditions, while certain "and" neurons, which only respond to information from both modalities, could be used to recognize the difference between a warm-blooded mouse and a cool-skinned frog. The "and" neurons have been whimsically described as mouse detectors. In evolutionary terms, it seems likely that similar integration processes take place in the tectums of most other vertebrates—although at present only Newman and Hartline's [2] work on pit vipers and rattlesnakes has been reported.

B. Previous Surveys and Reviews

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 [3]. Mitiche and Aggarwal [4] discuss some of the advantages and problems involved with the integration of different image processing sensors, and review recent work in that area. Garvey [5] 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 [6] 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 [7], and Blackman [8] 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 (Section VI-D). Recent research workshops have focused on the multisensor integration and fusion issues involved in manufacturing automation [9] and spatial reasoning [10].

C. Overview of this Paper

Section II serves both to describe multisensor integration and fusion and to set the stage for the presentation in the four subsequent sections of a broad survey of current approaches to the problem. In each of these subsequent sections, increasingly more specific approaches are surveyed: from general paradigms and methods for integrating and fusing multisensory information, to existing multisensor systems used in different areas of application.

Section II describes the role of multisensor integration and fusion in the operation of an intelligent system. Multisensor integration and the related notion of multisensor fusion are defined and distinguished. A general pattern of multisensor integration and fusion is presented to highlight the distinction between the integration and the fusion of information in the overall operation of a system. The potential advantages in integrating multiple sensors are then discussed in terms of four fundamental aspects of the information provided by the sensors.

Section III presents approaches to different aspects of multisensor integration and fusion that are quite general in terms of their range of applicability. Initially a variety of paradigms, frameworks, and control structures are presented that have been proposed for the overall multisensor integration process. Work is then presented relating to two important integration functions: the preselection and real-time selection of sensors and the use of world models. The section concludes with a survey of general multisensor fusion methods.

Section IV surveys approaches to the integration and fusion of information from combinations of different types of sensors, with special emphasis given to vision and tactile sensor combinations because of the broad range of capabilities that this combination can provide an industrial robot.

Section V details the critical role played by multisensor integration and fusion in enabling mobile robots to operate in uncertain or unknown dynamic environments. A variety of proposed high-level representations for multisensory information are presented that are suitable for mobile robot control and navigation. The section concludes with a discussion of different sensor combinations that have been used in mobile robots, and short descriptions of the role of multisensor integration and fusion in the navigation and control of a number of existing mobile robots.

Section VI surveys a variety of multisensor systems in the following areas of application: industrial tasks like material handling, part fabrication (e.g., welding), inspection, and assembly; military tasks (e.g., command and control for battle management); space; target tracking; inertial navigation; and the remote sensing of coastal waters.

Section VII concludes with a discussion of possible problems and future research directions in the area of multisensor integration and fusion. Discussion of the possible problems centers around the methods used for modeling error or uncertainty in the integration and fusion process, the actual sensory information, and the operation of the overall system.

II. The Role of Multisensor Integration and Fusion in Intelligent Systems

In the operation of an intelligent system, 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. Prior to discussing these aspects, this section first provides a definition of the distinction between the notions of the integration and the fusion of multisensory information; secondly, a general pattern of multisensor integration and fusion is presented within the context of an overall system architecture to highlight some of the important functions in the integration process.

A. Multisensor Integration versus Fusion

Multisensor integration, as defined in this paper, refers to the synergistic use of the in formation provided by multiple sensory 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. (This definition would also apply to the fusion of information from a single sensory device acquired over an extended time period.) Although the distinction of fusion from integration is not standard in the literature, it serves to separate the more general issues involved in the integration of multiple sensory devices at the system architecture and control level, from the more specific issues involving the actual fusion of sensory information—e.g., in many integrated multisensor systems the information from one sensor may be used to guide the operation of other sensors in the system without ever actually fusing the sensors' information (e.g., Section IV-B).

B. A General Pattern

Fig. 3 is meant to represent a general pattern of multisensor integration and fusion in a system. While the fusion of information takes place at the nodes in the figure, the entire network structure, together with the integration functions, shown as part of the system, are part of the multisensor integration process. In the figure, n sensors are integrated to provide information to the system. The outputs x_1 and x_2 from the first two sensors are fused at the lower left-hand node into a new representation $x_{1,2}$. The output x_3 from the third sensor could then be fused with $x_{1,2}$ at the next node, resulting in the representation $x_{1,2,3}$, which might then be fused at nodes higher in the structure. In a similar manner the output from all n sensors could be integrated into an overall network structure. The dashed lines from the system to each node represent any of the possible signals sent from the integration functions within the system. The three functions shown in the figure are some of the functions typically used as part of the integration process. "Sensor selection" can select the most appropriate group of sensors to use in response to changing conditions, sensory information can be represented within the "world model," and the information from different sensors may need to be "transformed" before it can be fused or represented in the world

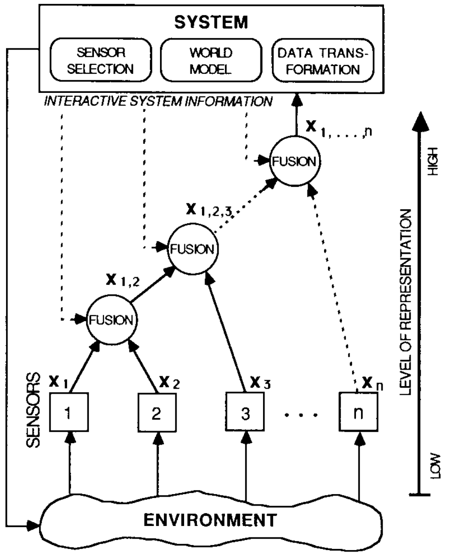

Fig. 3. General pattern of multisensor integration and fusion in system.

model. Shown along the right-side of the figure is a scale indicating the level of representation of the information at the corresponding level in the network structure. The transformation from lower to higher levels of representation as the information moves up through the structure 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 may be transformed into progressively more abstract numeric or symbolic representations. This "signals-to-symbols" paradigm is common in computational vision [11].

C. 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 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 one case fused information was used in a distributed network of target tracking sensors just to reduce the bandwidth required for communication between groups of sensors in the network (Section VI-D).

Fig. 4 illustrates the distinction between complementary and redundant information by using the network structure from Fig. 3 to perform, hypothetically, the task of object discrimination. Four objects are shown in Fig. 4(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. 4(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) cannot 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. 4(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 [12].) The uncertainty in Fig. 4(d) is shown as approximately half that of Fig. 4(b) and (c). In Fig. 4(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. 4(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 before, 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).

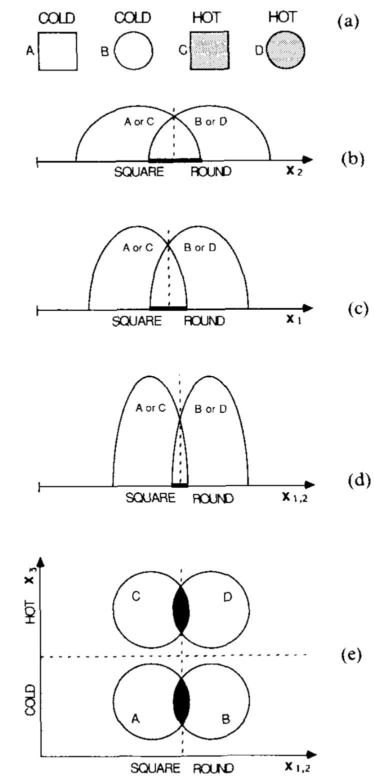

Fig. 4. Discrimination of four different objects using redundant and complementary information from three sensors. (a) Four objects (A, B, C, and D) distinguished by features "shape" (square vs. round) and "temperature" (hot versus cold). (b) Two-dimensional (2-D) distributions from Sensor 1 (shape). (c) Sensor 2 (shape). (d) 2-D distributions resulting from fusion of redundant shape information from Sensors 1 and 2. (e) Three-dimensional (3-D) distributions resulting from fusion of complementary information from Sensors 1 and 2 (shape), and Sensor 3 (temperature).

III. General Approaches to Multisensor Integration and Fusion

This section presents approaches to different aspects of the multisensor integration and fusion problem discussed in the previous section. Although some of the approaches were originally presented in terms of a specific application or combination of sensors, they are distinguished by their applicability to a broad range of systems in a number of possible applications.

A. *Paradigms and Frameworks for Integration*

1) Hierarchical Phase-Template Paradigm: Luo and Lin [13]–[17] have proposed a general paradigm for multisensor integration in robotic systems based upon four distinct temporal phases in the sensory information acquisition process (see Fig. 5). The four phases, "far away," "near to," "touching," and "manipulation," are distinguished at each phase by the range over which sensing will take place, the subset of sensors typically required, and, most importantly, the type of information desired. During the first phase, "far away," only global information concerning the environment is obtained. Typical information at this stage would be the detection, location, or identity of objects in a scene. The most likely types of sensors to be used during this phase would be noncontact sensors like vision cameras and range finding devices. If the scene is found to be of sufficient interest during the first phase, the manipulator

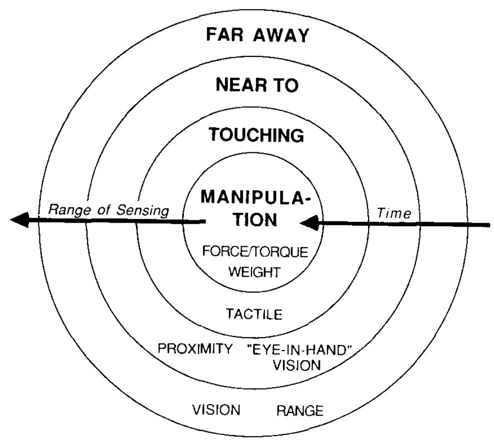

Fig. 5. Four phases of hierarchical phase-template paradigm.

can zoom in to obtain more detailed information. This leads to the second phase, "near to." Usually at this close range it is not possible to see the entire object, so noncontact sensors like proximity sensors or "eye-in-hand" vision systems, mounted on the gripper of the manipulator, are used. If it is desired to confirm or integrate the information from the previous two phases, one can proceed to the third phase, "touching." Contact sensors such as tactile sensors might be used at this phase. Finally if it is necessary to manipulate the object, one can proceed to the fourth phase, "manipulating." Sensors providing information concerning force/torque, slippage, and weight would typically be used during manipulation.

The information acquired at each phase is represented in the form of a distinct framelike template. Each template represents information that is both common to all phases (e.g., position and orientation of an object) and specific to the particular phase. During each phase of operation, the information acquired by each sensor is stored as an instance of that phase's template. The information from each sensor can then be fused into a single instance of the template (Section III-E-3 provides a description of the fusion method used).

2) Neural Networks: Current research in neural networks (e.g., [18]–[20]) is providing a common paradigm for the interchange of ideas between neuroscience and robotics; Pellionisz [21] has even introduced the term "neurobotics" to describe the possible use of brainlike control and representation in robotic systems. Although related to the adaptive learning control structure described in Section III-B-3, neural networks provide a fairly well-established formalism with which to model the multisensor integration process. Neurons can be trained to represent sensory information and, through "associative recall," complex combinations of the neurons can be activated in response to different sensory stimuli. "Simulated annealing" is one of many different techniques that can be used to find a global optimal state in a network based upon the local state of activation of each neuron in the network. Simulated annealing has been used to find optimal global paths for mobile robot navigation (Section V-B-2). "Self-organizing feature maps" as developed by Kohonen [20] can be used to reduce the dimensionality of the sensor signals while preserving their topological relationships.

Pearson *et al.* [22] have presented a neural network model for multisensor fusion based on the barn owl's use of visual and acoustic information for target localization. Separate visual and acoustic maps are fused into a single map (corresponding to the owl's optic tectum) which is then used for head orientation. Jakubowicz [23] has presented a neural network-based multisensor system that is able to reconfigure itself adaptively in response to sensor failure, and Dress [24] has explored the use of frequency-coded sensor information for fusion in neural networks.

3) Logical Sensors: A "logical sensor," as proposed by Henderson and Shilcrat [25], [26] and then extended in [27]–[33], is a specification for the abstract definition of a sensor that can be used to provide a uniform framework for multisensor integration. Through the use of an abstract definition of a sensor, the unnecessary details of the actual physical sensor are separated from their functional use in a system. In a manner similar to how an abstract data type separates the user from unnecessary algorithmic detail, the use of logical sensors can provide any multisensor system with both portability and the ability to adapt to technological changes in a manner transparent to the system.

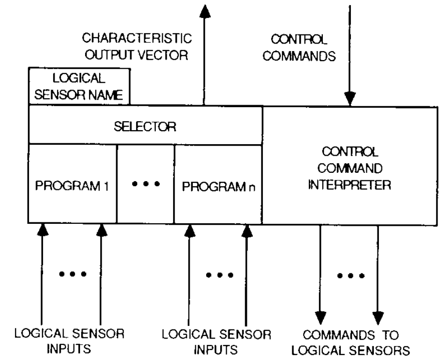

Fig. 6. Basic components of logical sensor. (Adapted from [27, fig. 1].)

Fig. 6 shows the essential elements of a logical sensor. Each logical sensor can serve as an element in a network of logical sensors, which itself can be viewed as a logical sensor. The "logical sensor name" uniquely defines a logical sensor. The "characteristic output vector" describes the data type of the stream of output vectors produced by the logical sensor. The "control commands" input to a logical sensor consist of both commands necessary to control the logical sensor and commands that are just passing through to other sensors lower in the network. The "control command interpreter" processes the incoming commands and sends appropriate commands to logical sensors lower in the network. The "selector" monitors the control commands issued to the logical sensor and the results of the various "program units"—acting as a "microexpert system" which knows the required function of the logical sensor. Each program unit serves to perform any required computation on the inputs to the unit. The logical sensor inputs are the output vectors of logical sensors lower in the network. When the logical sensor is an actual physical sensor, the raw data sensed from the environment can be considered as null inputs.

A hypothetical logical sensor-based range finder is shown in Fig. 7 that incorporates three physical sensors: an ultrasonic range finder and two cameras. Both cameras are used as input to a fast and a slow stereo logical sensor. Each of these logical sensors, which differ in terms of the speed and accuracy of their processing algorithms, are used as input to an overall stereo logical sensor which just serves a switching function based on control commands from the top-level logical sensor. The entire network of logical and physical sensors can provide for a range finder that is both robust in terms of the lighting conditions in which it can operate (i.e., the ultrasonic sensor for poor

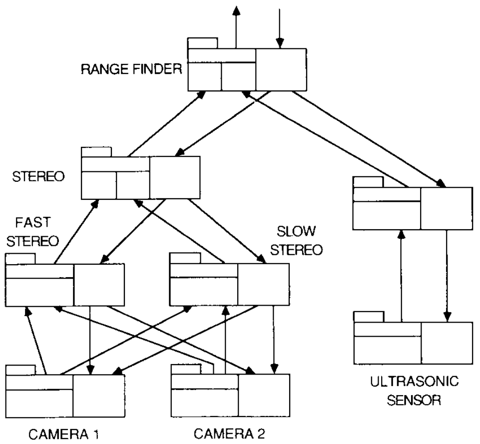

Fig. 7. Logical sensor network for range finder. (Adapted from [25, Fig. 6].)

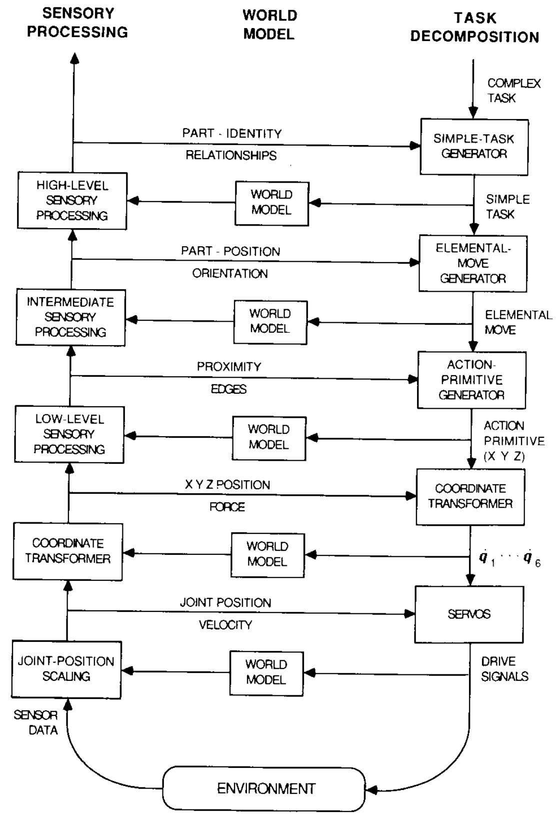

Fig. 8. NBS sensory and control hierarchy used to control multisensor robot. (Adapted from [38, figs. 5.24 and 9.6].)

lighting conditions) and, depending on time constraints, the speed at which it can operate. The range information could possibly be made more accurate if the redundant information from the stereo and ultrasonic sensors is fused at the top-level logical sensor in the network.

4) Object-Oriented Programming: In a similar manner to the logical sensors mentioned earlier, object-oriented programming is a methodology that can be used to develop a uniform framework for implementing multisensor tasks; Henderson and Weitz [29], [31] have, in fact, discussed the development of logical sensor specifications within an object-oriented programming context. In most object-oriented multisensor applications, each sensor is represented as an object. Objects communicate by passing messages that invoke specialized sensor processing procedures ("methods") based on the sensor's attributes and behavior. Each method is transparent to other objects, allowing possibly different physical sensors to be used interchangeably. Rodger and Browse [34] have used object-oriented programming for multisensor object recognition, and Allen [35] has developed an object-oriented framework for multisensor robotic tasks.

B. Control Structures

This section presents different structures that have been used to control the overall integration and fusion process. Control structures based on artificial intelligence (e.g., production systems) have not been included because a thorough discussion of their use in the context of multisensor integration and fusion can be found in [5].

1) The NBS Sensory and Control Hierarchy: The Center for Manufacturing Engineering at the National Bureau of Standards (NBS) is implementing an experimental factory called the Automated Manufacturing Research Facility (AMRF). As part of the AMRF, a multisensor interactive hierarchical robot control system [36]–[40] is being developed based, in part, on the mathematical formalism called the cerebellar model arithmetic computer [41], [42]. As shown in Fig. 8, the structure of the control system in AMRF consists of an ascending "sensory processing" hierarchy coupled to a descending "task-decomposition" control hierarchy via "world models" at each level. Input to the world model at each level comes from both the task unit at that level, and other unspecified locations in the system. The use of multiple levels is motivated by the observation that the complexity of a control program grows exponentially as the number of sensors and their associated processing increases. By isolating related portions of the required processing at one level, this complexity can be reduced. The large number of low-level processing tasks, which usually have to be done in real time, can be separated from the fewer, more complex, higher level processing tasks so that the required processing time at each level can become nearly equal. Assuming the required communication between processing levels will be much less than the communication within levels, complexity is reduced by requiring only a limited number of communication channels between levels. If the processing at each level can be done in parallel, the addition of more levels will not result in an exponential increase in complexity. The amount of processing at each level is further reduced by the use of *a priori* knowledge from the world model. The world models provide predictions to the sensory system concerning the incoming sensory information so that the amount of processing required can be reduced. The use of a world model promotes modularity because the specific information requirements of the sensory and control hierarchies are decoupled.

Fig. 8 provides an example of the control of a multisensor robot using the NBS hierarchy. Raw sensory data from the environment enter the system at the bottom. At this lowest level, most of the required sensory processing will be continuous monitoring of the robot's joint positions. Any deviation between the actual and expected data is sent as feedback information to the servos, and as summary information to the next level in the sensory processing hierarchy. More complex data, like that from vision sensors, is sent through to higher levels unmodified. At the very highest level in the system, the complex task and top-level world model filter down to lower levels in the hierarchy both expected and desired information values. It is at the intermediate levels where both of these information flows meet and interact. Based upon current sensory information, the world models are updated. The updated world models can then serve to modify the desired task control actions until, at the lowest level, the necessary drive signals are sent to the robot to initiate actions in the environment.

2) Distributed Blackboard: A blackboard architecture allows economical communication between distributed sensory subsystems in an integrated multisensor system. Each subsystem can send time-stamped summary output to a blackboard where it becomes available to any fusion process as well as the integration functions. The time stamp on the output in the blackboard allows for sensor information to be made commensurate before being fused. The blackboard can contain any system information needed by the integration functions. Any number of different fusion methods can be implemented using the output from the blackboard. Harmon *et al.* [43] have used a blackboard architecture to compare different methods of multisensor fusion, and Harmon (Section V-D-3) has used a blackboard architecture for autonomous vehicle control.

3) Adaptive Learning: Miller *et al.* [44]–[46] have applied an adaptive learning approach to the multisensor-based control of robotic manipulators. In experiments using this approach, the performance accuracy was limited by the resolution of the sensor feedback rather than by any limitations in the control structure. Adaptive learning is a method of control in which the system "discovers" the appropriate signals for control based on the output of the sensors. The system is taught a representative sample of correlated control signals and associated sensory outputs over the range of signals and sensory outputs encountered by the system. Based on the associations developed during this teaching phase, it is possible to have the system respond to any combination of sensory outputs with an appropriate control signal. The system requires no *a priori* knowledge of the relationship between the structural kinematics of the robot, or the desired control signals and their associated sensory outputs. It is this feature of the adaptive learning approach that makes it attractive when there are possibly multiple sensors interacting to produce complex output.

C. Sensor Selection Strategies

Sensor selection is one of the integration functions that can enable a multisensor system to select the most appropriate configuration of sensors (or sensing strategy) from among the sensors available to the system. 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.

1) Preselection: As an initial step towards a general methodology for optimal sensor design, Beni *et al.* [47] have derived a general relationship between the number and operating speed of available sensing elements as a function of their response and processing times. This relationship can be used to determine the optimal arrangement of the sensing elements in a multisensor system. In addition to the actual geometric arrangement of the sensing elements, consideration of the choice between adding sensing elements (static sensing) and moving the elements (dynamic sensing) is used in determining the optimal arrangement.

2) Real-Time Selection: Hutchinson *et al.* [48] have presented an approach to planning sensing strategies for object recognition in a robotic workcell. One sensor is used to form an initial set of object hypotheses and then subsequent sensors are chosen so as to disambiguate maximally the remaining object hypotheses. Grimson [49] has also considered the problem of recognizing objects in the workspace of a robot using minimal sets of sensory information, and Taylor and Taylor [50] have used "dynamic error probability vectors" to select the appropriate sensors necessary for the recovery from errors during an automatic assembly process.

D. World Models

The use of world models enables a multisensor system to both store and reason with previously acquired sensory information. World models are usually defined in terms of a high-level representation. Sensory information can be either added to a predefined model of the world (i.e., the system's environment), or used to create the model dynamically during operation. 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. Section V-B describes a number of examples of world models used in mobile robots. Included in this section are two representations that are more general in nature. These representations were developed within the context of a previously mentioned multisensor integration framework (logical sensors) and control structure (the NBS hierarchy).

TABLE I
GENERAL METHODS OF MULTISENSOR FUSION

Method	Operating Environment	Type of Sensory Information	Information Representation	Uncertainty	Measurement Consistency	Fusion Technique
Weighted average	dynamic	redundant	raw sensor readings	—	(thresholding possible)	weighted average
Kalman filter	dynamic	redundant	probability distribution	additive Gaussian noise	(thresholding, calibration)	filtering of system model
Bayesian estimate using consensus sensors	static	redundant	probability distribution	additive Gaussian noise	largest digraph in relation matrix	maximum Bayesian estimate of consensus sensor
Multi-Bayesian	static	redundant	probability distribution	additive Gaussian noise	ϵ-contamination	maximum Bayesian estimate
Statistical decision theory	static	redundant	probability distribution	additive noise	ϵ-contamination	robust minimax decision rules
Evidential reasoning	static	redundant and complementary	proposition	level of support versus ignorance	—	logical inference
Fuzzy logic	static	redundant and complementary	proposition	degree of truth	—	logical inference
Production rules	static	redundant and complementary	proposition	confidence factor	—	logical inference

1) The Multisensor Kernel System: Henderson *et al.* [51], [52] have presented the multisensor kernel system as a means of providing a representation for sensor information that is compatible with the specification of logical sensors (Section III-A-3). Object features are extracted from low-level sensory data and organized into a three-dimensional "spatial proximity graph" that makes explicit the neighborhood relations between features. Each feature is defined in terms of a logical sensor and is available to the system as the output of the logical sensor's characteristic vector. Subsequent sensory data can then either be matched in terms of the spatial proximity graph, or a "*k-d* tree" (a binary tree with k-dimensional keys that allows the nearest neighbors of one of k features to be found) is constructed, using the proximity graph, for faster processing.

2) The NBS Sensory System: Shneier *et al.* [53], [55], and Kent *et al.* [54] have described the kinds of processes involved in the higher levels of the sensory system of the NBS hierarchy (Section III-B-1). World models at each level in the hierarchy are used to create initial expectations about the form of the sensory information available at that level and then to generate predictions for the task control units in the hierarchy so that they do not have to wait for sensory processing to finish. Errors between the sensed information and the world model are used initially to register the model and later to maintain the consistency of the model during operation of the system.

E. General Fusion Methods

This section surveys different methods that have been proposed for general multisensor fusion (discussion of additional fusion methods relating to specific applications is included in Sections IV–VI). 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) 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 Richardson and Marsh [12]. Their paper provides a proof that the inclusion of additional redundant sensory information almost always improves the performance of any fusion method based on optimal estimation.

Pau [56], [57] describes a number of statistical pattern-recognition techniques that are appropriate for multisensor fusion. All of these techniques could be used to reduce the error in classifying objects through the use of multiple sensors to provide redundant information concerning features of the objects. To avoid an exponential increase in complexity as sensors are added to a system, a key requirement is that the number of features and levels in the recognition process increase at a slower rate than the number of sensors. To meet this requirement it becomes necessary to improve the overall methods of feature extraction and selection—two major areas of interest in pattern recognition. Thus multisensor fusion becomes a problem within the context of statistical pattern recognition. Pau describes a number of operators and techniques that can fuse the features perceived by the sensors to limit their growth as additional sensors are added [56] and introduces a representation for multisensor fusion that is based on "context truth maintenance" [57].

Table I summarizes for comparison the relevant aspects of each general multisensor fusion method presented in this section. The sequence in which the methods are presented corresponds roughly to the increasingly high levels of representation of the information being fused (see Fig. 3). The representations used extend from low-level probability distributions for statistical inference to high-level logical propositions used in production rules for logical inference. In addition to the level of representation of the multisensory information, distinctions can be made as to whether the method is appropriate when information is assumed to come from *static* or *dynamic* sources in the operating environment, and as to whether the information is redundant or complementary (Section II-C). Included in the table are the means used to represent uncertainty in the measurement and fusion processes, possible methods used to determine the consistency of sensor measurements (e.g., elimination of any spurious sensor measurements), and the actual techniques used for fusion.

1) Weighted Average: One of the simplest and most intuitive general methods of fusion is to take a weighted average of redundant information provided by a group of sensors 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 weighted average has been used for multisensor fusion in the mobile robot HILARE (Section V-D-1), after first thresholding the sensory information to eliminate spurious measurements.

2) Kalman Filter: The Kalman filter (see [66] 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 determine estimates recursively 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 (Section IV-A), robot navigation (Section V-D-4), multitarget tracking (Section VI-D-2), inertial navigation (Section VI-E), and remote sensing (Section VI-F). 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.

3) Bayesian Estimation using Consensus Sensors: Luo and Lin [13]–[17] have developed a method for the fusion of redundant information from multiple sensors that can be used within their hierarchical phase-template paradigm (Section III-A-1). The central idea behind the method is first to eliminate from consideration the sensor information that is likely to be in error and then to use the information from the remaining "consensus sensors" to calculate a fused value.

Fig. 9 shows a functional block diagram of the method. The information from each sensor is represented as a probability density function. Given readings from n sensors in the system, the resulting information is first made commensurate through preprocessing. An n by n distance matrix is created by calculating for each element (i, j) in the matrix the "confidence distance measure" between the information from sensors i and j. A confidence distance measure is defined to be equal to twice the area under the density function of sensor i between the readings from sensor i and sensor j. Use of this measure assumes that the domains of each sensor's density function are commensurate. If the density functions are assumed to be Gaussian, the distance can be computed by use of the error function. The distance matrix determined by the use of this measure will not be symmetric unless the density functions of all the sensors are identical. Threshold values, based on the required sensing accuracy, are then applied to the elements in the matrix. Elements not exceeding their threshold are represented by a one in a binary-valued n by n relation matrix. The largest connected digraph formed from this matrix will determine the group of consensus sensors most likely not to be in error. The optimal fusion of the information is determined by finding the Bayesian estimator that maximizes the likelihood function of the consensus sensors.

4) Multi-Bayesian: Durrant-Whyte [58]–[61] has developed a model of a multisensor system that represents the task environment as a collection of uncertain geometric objects. Each sensor in the system is described by its ability to extract useful static descriptions of these objects. An "ϵ-contaminated" (see paragraph 5, below) Gaussian distribution is used to represent the geometric objects. The sensors in the system are considered as a team of decision-makers. Together, the sensors must determine a team-consensus view of the environment. A multi-Bayesian approach, with each sensor considered a Bayesian estimator, is used to combine the associated probability distributions of each respective object into a joint posterior distribution function. A likelihood function of this joint distribution is then maximized to provide the final fusion of the sensory information. The fused information, together with an *a priori* model of the environment, can then be used to direct the robotic system during the execution of different tasks.

5) Statistical Decision Theory: McKendall and Mintz [62] and Zeytinoglu and Mintz [63], [64] have used statistical decision theory to develop a general two-step method for the fusion of redundant location data from multiple sensors. (Location data refer to sensor measurements that are modeled as additive sensor noise translated by the parameter of interest being sensed.) Sensor noise is modeled as the ϵ-contamination of a variety of possible proba-

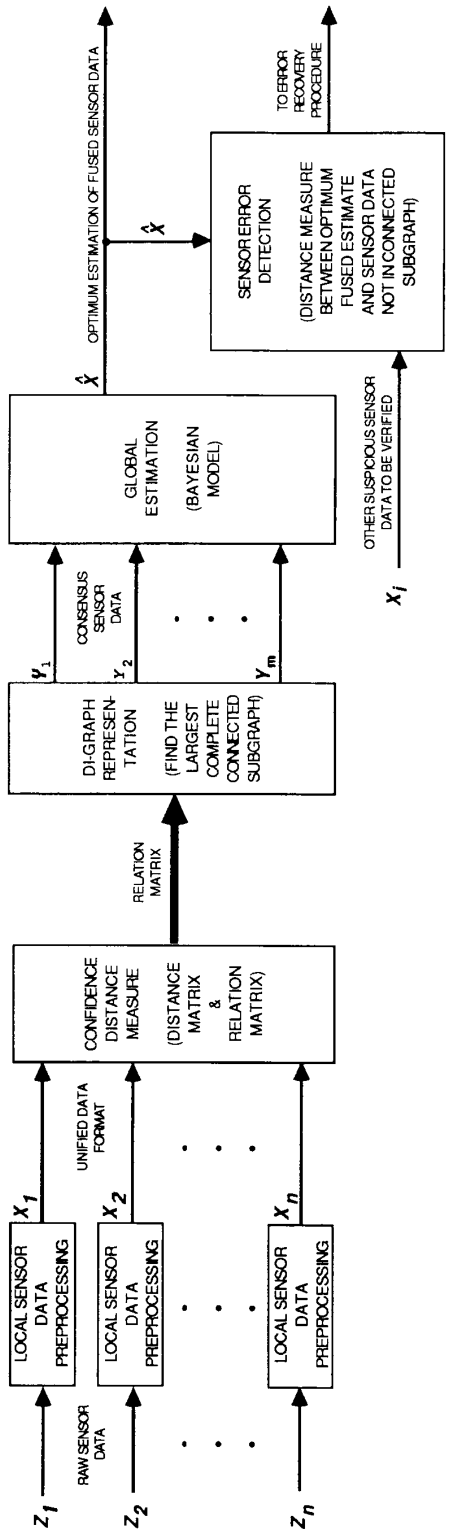

Fig. 9. Functional block diagram of consensus sensor fusion method. (Adapted from [17, fig. 2.)

bility distributions. The use of ϵ-contamination in the sensor model serves to increase the robustness of the decision procedure by removing a certain outlying fraction ϵ of the distribution to account for heavy-tailed deviations from the assumed noise distribution that may have been caused by spurious sensor readings. Initially, the data from different sensors are subject to a robust hypothesis test as to its consistency (see [65] for an introduction to "robust statistics"). Data that passes this preliminary test are then fused using a class of robust minimax decision rules.

6) Shafer–Dempster Evidential Reasoning: Garvey *et al.* [67] introduced the possibility of using Shafer–Dempster evidential reasoning in multisensor fusion. Bogler [68] and Waltz and Buede [69] have explored its possible application in, respectively, multisensor target identification, and military command and control. Shafer–Dempster evidential reasoning [70] is an extension to the Bayesian approach that makes explicit any lack of information concerning a proposition's probability by separating firm support for the proposition from just its plausibility. In the Bayesian approach all propositions (e.g., features 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 Shafer–Dempster 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.

7) Fuzzy Logic: Huntsberger and Jayaramamurthy [71] have used fuzzy logic to fuse information for scene analysis and object recognition. Fuzzy logic [72], a type of multiple-valued logic, allows the uncertainty in multisensor fusion to be directly represented in the inference (i.e., fusion) process by allowing each proposition, as well as the actual implication operator, to be assigned a real number from 0.0 to 1.0 to indicate its degree of truth. Consistent logical inference can take place if the uncertainty of the fusion process is modeled in some systematic fashion.

8) Production Rules with Confidence Factors: Kamat [73], Belknap *et al.* [74], and Hanson *et al.* [75] have used production rule-based systems for object recognition using multisensor fusion. Production rules are used to represent symbolically the relation between an object feature and the corresponding sensory information. A confidence factor is associated with each rule to indicate its degree of uncertainty. Fusion takes place when two or more rules, referring to the same object, are combined during logical inference to form one rule. The major problem in using production rule-based methods for fusion is that the confidence factor of each rule is defined in relation to the confidence factors of the other rules in the system, making it difficult

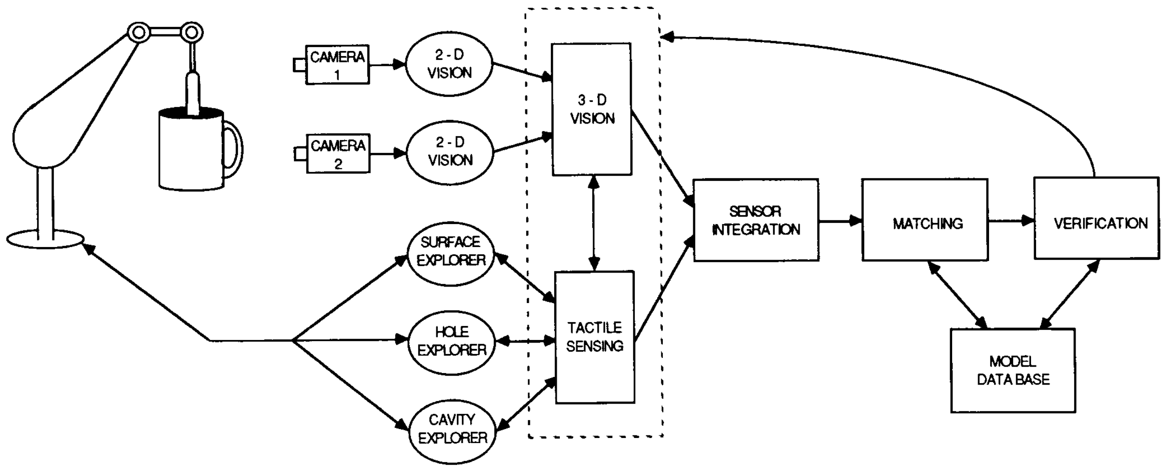

Fig. 10. Overview of Allen's robotic object recognition system. Mug is shown being recognized through vision and active exploratory tactile sensing. (Adapted from [91, Figs. 1.1 and 1.2].)

to alter the system when, for example, new sensors are added that require additional rules.

IV. Integration and Fusion using Specific Sensor Combinations

This section surveys a variety of approaches to the integration and fusion of information from combinations of different types of sensors. An effort is made to present approaches that exploit the specific characteristics of the information provided by each type of sensor and have an area of possible application that is quite general (e.g., object recognition). Sections V and VI present a wide variety of additional sensor combinations that have been used in mobile robot and other applications, respectively.

A. Vision

Visual information is the most powerful single source of sensory information available to a system. Many different types of nonvisual sensors are used in combination with vision sensors to compensate for some of the difficulties encountered in the machine processing of visual information. Tasks such as object recognition can sometimes require the aid of additional types of sensors to approach the capabilities of a human using just visual information. This section will not describe in detail the many integration and fusion techniques that use only vision sensors because there already is an extensive literature, including many reviews and book-length treatments (e.g., [76], [77]) covering the various aspects of computational vision.

Magee and Aggarwal [78] provide a review of recent research efforts aimed at combining intensity and range features derived from visual images to determine the structure of three-dimensional objects. In some of the work reviewed, information from one feature is used to guide the acquisition of information concerning another feature when the second feature requires a much longer processing time (e.g., intensity guided range sensing for object recognition [79] and the determination of motion parameters [80]). Research related to the fusion of sequences of images has used the "optical flow" of the images to determine the motion of objects in the image (see [81] for a recent review), and both a Bayesian [82] and extended Kalman filter [84], [85] to establish the surfaces of three-dimensional objects. Flachs *et al.* [83] have used a "complexity metric" as a mathematical basis for multisensor fusion in vision systems. Much of the research related to the use of multiple visual sensors has used the stereoscopic effect from the sensors to determine range information (see [86] for a recent stereo vision review, and [87] for an approach to binocular fusion that uses simulated annealing). Porrill [88] has used Gauss–Markov estimation together with geometric constraints to fuse multiple stereo views of a wire frame model. In robotics, overhead and "eye-in-hand" vision sensors have been combined for use in three-dimensional object recognition [89]. Many of these techniques, originally developed to fuse both sequences of images from a single vision sensor and images derived simultaneously from multiple sensors, have significantly influenced subsequent work in nonvisual fusion.

B. Vision and Tactile

As can be seen in Fig. 5 with reference to the hierarchical phase-template paradigm (Section III-A-1), the integration of vision and tactile information, together with a robot's own manipulation capabilities, gives that robot a wide range over which to receive sensory information. Combined vision and tactile sensors have been used to a great extent by industrial robots to perform both assembly and inspection tasks (Section VI-A).

1) Allen: Allen [3], [90]–[95] has developed a robotic object-recognition system that uses three-dimensional vision together with active exploratory tactile sensing (see Fig. 10). The system was developed to recognize common kitchen items like mugs, plates, pitchers, and bowls that had no discernible textures and were homogeneous in color. Recognition of objects like these pose a serious problem for many vision-only recognition systems because of the lack of features that can be used for matching and depth analysis. Through the integration of tactile information with vision, Allen's system is able to obtain information concerning any holes, cavities, or curved surfaces that can be used to identify a particular object.

The model data base represents each object in a hierarchical manner that allows the sensory devices to match the models at different levels of detail. The models are independent of viewpoint and scale and contain relational information that can be used to reduce the searching of the data base. Each object is modeled as a collection of components and features: the components correspond to the discernible differences in the surface of the objects (e.g., the body, bottom, and handle of the mug shown in the figure), and the features correspond to the object's holes or cavities. At the lowest level, each surface component is modeled as a grid of bicubic spline curves that form a patch. Both holes and cavities are modeled as right cylinders of constant cross-sectional area—cavities having the additional attribute of depth. The operation of the system can be summarized in a five-step recognition cycle [91], as follows.

Step 1: Initially, two-dimensional vision processing routines are applied to the image to determine bounded regions. The centroid of each region is calculated by using the matched three-dimensional stereo points of its boundary so that it can be used as the starting point for tactile exploration of the region. The depth and surface orientation of each region is determined using binocular stereo. Isolated edge pixels, which could be possible noise points, are excluded from consideration through thresholding. In a system relying only on vision to determine depth and orientation, this elimination of data to reduce error could result in a surface description that is not dense enough for recognition purposes. By allowing tactile sensing to explore further any uncertain regions, the regions that are identified from the remaining data will have a greater accuracy and can be used with higher confidence in later steps of the recognition process.

Step 2: The tactile sensing system explores each region identified by the vision system to determine if it is a surface, hole, or cavity. The tactile sensor used in this system is an octagonal cylinder covered with conducting surfaces mounted perpendicular to the mounting plate of the robot (the tactile sensor in the figure is exploring the cavity formed by the well of the mug). The sensor approaches a region orientated in a direction normal to the centroid of the region until either it contacts a surface, travels beyond a specified threshold used to indicate the presence of a cavity, or, if the sensor is able to travel its full length into the region without contact, the region is assumed to be a hole. If the region is a surface, a surface patch is constructed by integrating vision and touch. If the region is either a hole or cavity, the sensor moves in a sawtooth manner around the region's boundary to determine its shape.

Step 3: For regions that are surfaces, the vision and tactile sensing results of the previous steps, together with additional tactile sensing, are integrated to create three-dimensional surface patches that can be matched with the model data base. Starting from the location of contact with the surface, the tactile sensor uses knot points to determine the directions along which traces of the surface will be made. The points reported along each trace are combined into cubic least squares polynomial curves which can then be used to fill in areas of the surface that still lacked detail after stereo vision processing.

Step 4: The surfaces patches and closed curves (corresponding to holes or cavities) are matched against the model data base to find an object that is consistent with the sensory information. If more than one object is found to be consistent, a probabilistic measure is used to order the objects for verification.

Step 5: Once a consistent object is found the verification procedure is used for further active exploratory sensing to verify components and features of an object's model that have not been sensed. Visually occluded holes and cavities are verified using the tactile sensor. Verification of visually occluded surfaces using only the tactile sensor is difficult because vision is needed both to guide the sensor during traces of the surface and, most importantly, to establish that the region of interest is indeed a smooth surface that can be approximated by a patch. The system provides for robust viewer-independent object recognition because no *a priori* viewpoint or orientation of the object is assumed, e.g., any identifiable part of the model of an object can be used to invoke a search to verify the remaining surfaces or features of the model needed for recognition.

2) Integration using a Decision Tree: Luo and Tsai [96] have developed a system that uses two-dimensional vision, together with two tactile sensing arrays mounted on a gripper of a robot, to recognize objects. Moment invariants of an object's shape are used as features for recognition and to calculate the centroid of a region of the object needed for determining the proper grasping position for the gripper. During the initial learning phase the system creates a decision tree by first presenting all of the objects to be recognized to the vision sensor so that their top-view silhouette boundaries can be determined. If there still exist groups of objects that are indistinguishable in terms of this visual information, tactile information concerning the objects' lateral shape is obtained to make the objects in each group more distinguishable. Different predetermined lateral directions are used until all of the objects can be distinguished. The final result of this clustering process is a hierarchical decision tree with each leaf representing a single discriminable object, each nonterminal node corresponding to a group of objects that are indistinguishable at that level of sensory processing, and each arc associated with the effective lateral direction used by the tactile sensor to distinguish the child from the parent node. The first level below the root node in the tree corresponds to the initial version sensing; levels below the first represent successive stages of tactile sensing. Finally, the recognition phase proceeds by traversing the decision tree in the same direction as the tree was created until the object is able to be discriminated.

3) Object Apprehension: Stansfield [97]–[100] has presented a system which uses vision and tactile sensors for object "apprehension." Apprehension is defined as the determination of the properties of an object and the relationships among these properties without, as in recognition, going on to attach a label to the object as a whole. The system is structured as a modularized hierarchy of knowledge-based experts, each responsible for either the execution of an exploratory procedure or for the further processing of information from other exploratory procedures. An exploratory procedure extracts information concerning a predefined general-purpose primitive (e.g., compliance, elasticity, texture, etc.) or feature (e.g., edges, surface patches, holes, cavities, etc.) that is related to some aspect of the object's form, substance, or function [99]. Each feature is composed of one or more primitives or features. Modules in the lower portion of the hierarchy are responsible for processing the information from each sensor system, while the upper level modules integrate the information coming from the different sensor systems. Many of the higher level modules are able to function with information from a variety of different lower level modules. Spatial polyhedrons have been proposed recently by Stansfield [100] as a generic object representation that can extend the capabilities of the system to include recognition.

C. Vision and Thermal

Nandhakumar and Aggarwal [101]–[104] have presented a technique for the classification of objects in outdoor scenes using thermal and visual sensors. A thermal camera is used to acquire an infrared image of a scene and a vision camera is used to acquire an intensity image. Both cameras are adjusted so that their images are in spatial correspondence. Three features are used in a decision tree to label the objects in a scene. The first, the conductive heat flux of each region in the scene, is determined by integrating complementary information provided by both sensors. Although this feature, characterizing the intrinsic thermal behavior of the imaged object, provides the greatest amount of discriminatory information, two additional features are used to identify the objects unambiguously: the surface reflectance of a region as determined from the visual image and the average region temperature as derived from the thermal image. Production rules are used to implement the decision tree. An object label is assigned in the consequent of each rule and logical combinations of heuristically determined intervals for the value of each feature are used in the antecedent.

D. Range and Tactile

Grimson and Lozano-Pérez [105] describe a technique that uses tactile and range sensors to provide measurements of position and surface normals that can be used to identify and locate objects from among a group of known objects. The objects are modeled as polyhedrons, and constraints are used to keep the number of hypotheses as to an object's identity small. The only assumptions made about the sensors required are that they be able to provide information concerning an object's surface points and surface normals; as a result, the technique should be applicable to a wide variety of different types of sensors besides range and tactile.

E. Laser Radar and Forward Looking Infrared

Roggemann *et al.* [106] and Tong *et al.* [107] have developed a method of fusing the complementary information provided by laser radar and forward looking infrared sensors to segment and enhance features of man-made objects such as tanks and trucks in a cluttered background. Once separated from the background, these features can be used by other methods to identify the object. The laser radar image provides range information for areas within a field of view, and the infrared image can show special features of the area. The gradients of both images are used for both object enhancement and segmentation. Boundaries enclosing areas of small range gradient can be used to identify possible objects because natural backgrounds such as mud, grass, and trees exhibit large range differences from one pixel to another that can be modeled as random noise. The gradient of the infrared image will sharpen the infrared signature and temperature characteristics of objects that are infrared sources and can also serve to identify cold objects in a hot background because only temperature differences are noted. Initially, a binary mask of the laser radar image is created that indicates the location, area, and boundary of possible objects. Assuming that objects exhibit small gradients relative to the background and occupy a small percentage of the overall image, the first derivative of the histogram of the gradient magnitudes in the image is used to estimate a threshold for the mask. The threshold alone is not discriminatory enough to separate every pixel in the image. To distinguish object from background pixels further, both the segmented infrared and segmented range gradient images are first "anded," pixel to pixel, with the mask. The resulting images are then multiplied to produce a final image that shows where the range and infrared gradients match, emphasizing object features and deemphasizing the background.

V. Multisensor-based Mobile Robots

The mobility of robots and other vehicles is required in a variety of applications. In simple well-structured environments, automatic control technology is sufficient to coordinate the use of the required sensor systems (e.g., automatic guided vehicles, missiles, etc.) [108]. When a vehicle must operate in an uncertain or unknown dynamic environment—usually in close to real time—it becomes necessary to consider integrating or fusing the data from a variety of different sensors so that an adequate amount of information from the environment can be quickly per-

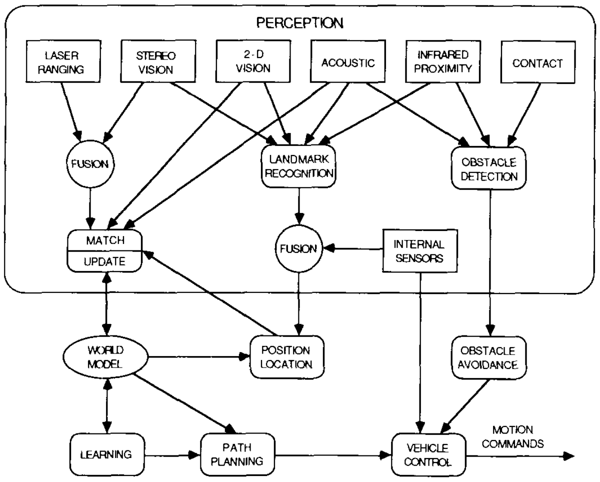

Fig. 11. Role of perception function in hypothetical architecture for mobile robot.

ceived. Because of these factors, mobile robot research has proved to be a major stimulus to the development of concrete approaches to multisensor integration and fusion. Luo and Kay [109] have reviewed the role of multisensor integration and fusion in the operation of mobile robots, and Levi [110] has discussed some of the multisensor fusion techniques appropriate for mobile robot navigation and has reviewed their use in a number of applications. Multiple sensors have been used in mobile robots to enable them to operate in environments ranging from roadways [111], [112] to unstructured indoor environments [113] to unknown natural terrain [114], [115] and to be used for applications including assembly [116] and nuclear power station maintenance [117].

A. The Role of Perception in a Hypothetical Architecture

Fig. 11 illustrates the role of the perception function in a hypothetical architecture for a mobile robot. Perception, together with vehicle control, obstacle avoidance, position location, path planning, and learning, are generic functions necessary for intelligent autonomous mobility [108]. Six different external sensor types are shown in the figure as part of the perception function. Subsets of these sensors are used to perform three tasks that usually comprise the perception function: the matching of sensory data to a world model (or map) representing the environment and then updating the model to reflect the matching results, the recognition of landmarks in the environment for use in determining the location of the robot, and the detection of obstacles so that they can be avoided. The degree of integration and fusion of the sensory data required for each of these tasks can differ. In a simple system, each sensor used for obstacle detection might operate independently of the other sensors if the detection of nearby obstacles is required; in a more complex system, some of the sensor data might be also fused so as to extend the range and accuracy of possible detection [114]. One of the simplest techniques used for position location is trajectory integration, where the location is calculated from the accumulated rotational and translational motion of the vehicle as determined by internal sensors like an odometer. Due to the inherent inaccuracies in any sensor, locational error continues to accumulate as the robot moves. To deduce this cumulative error most mobile robots periodically determine the location of some external landmark. As shown in the figure, the results from landmark recognition are sometimes fused with the location determined by internal sensors after each has been transformed to common coordinates. Any of the techniques of multisensor object recognition presented in this paper may be used first to recognize the landmark. The world model matching and updating task requires that the sensor information and any associated measure of its uncertainty correspond to the representation used in the world model so that integration can take place. Depending on the representational format used in the world model, the information will in most cases have to be made commensurate by applying appropriate space and time transformations. Information from the different sensors might be fused or otherwise transformed before reaching the matching task to reduce the communication bandwidth required or the complexity of the matching process. In the figure, information from the laser ranging and stereo vision sensors are fused before being matched to the world model.

While the hypothetical architecture described represents in broad outline most current approaches to the design of mobile robots, the MIT mobile robot project has adopted a radically different layered approach for autonomous control that they term a "subsumption" architecture [118], [119]. Each layer in this architecture consists of a complete control system similar to that in Fig. 11 for a simple task achieving behavior like avoiding obstacles or wandering. Starting with low-level tasks, new task-achieving behavior can be added incrementally because the layers operate asynchronously communicating over low-bandwidth lines without a central locus of control, central data structure, or global plan.

B. High-Level Representations

Considerable research has been directed at the development of either a single representation or multiple hierarchical levels of representations suitable for use by a mobile robot to perform the reasoning required for its control, path planning, and learning functions. The representations are usually at a high enough level so as not to be sensor specific. As mentioned before, information from different sensors is usually transformed to the common high-level representation and then added to a world model. The function of the world model and the particular form of the high-level representation depend both on the control architecture used in the robot and the complexity of the required reasoning—extremes range from road-following vehicles where sensor information is dynamically processed using feedback loops to produce control commands without ever using a world model, to pure production rule-based representations that assume a static and perfect model of the world that is difficult to modify [120]. In practice, the representations used for robots operating in unknown or unstructured environments allow for their world models to be dynamically modified and updated with uncertain sensor information. Except for explicit learning procedures used in many production rule-based representations, learning takes place implicitly as the world model is updated with new information as the robot traverses the environment.

Included below are some examples of different high-level representations. Many of the papers referenced are distinguished by a discussion of the multisensor integration and fusion issues relevant to their proposed representation. Other representations are discussed as part of the descriptions of different mobile robots found in Section V-D of this section.

1) Spherical Octree: Chen [121]–[124] has proposed a "spherical octree" representation for use in mobile robot navigation. A spherical octree is an 8-ary tree structure that at its first level separates a spherical perspective view of the environment into eight octants corresponding to the children of the root node (the entire spherical environment perceived by the robot). Objects in the environment can be represented in the octree by recursively subdividing octants containing part of the object into eight more octants at the next lowest level and merging octants that are completely contained by the object into one octant at the next highest level, repeating this process to represent the object at increasingly finer resolution. The use of a spherical perspective view eliminates some of the limitations of the typical orthographic and planar perspectives used in optical sensing. It is also appropriate for sensors providing range information because range values can be represented as the radial distances from the sensor to the object. Information from each different sensor is used to reconstruct three-dimensional surfaces using a knowledge base of typical patterns for the sensor. The reconstructed surfaces are then fused together as part of the process of being represented in the octree structure.

2) Occupancy Grids and Neural Nets: Elfes [12]–[127] originally developed a cellular world model representation called the "occupancy grid" for use with a sonar equipped mobile robot. This representation has been extended to allow for the integration of information from many different types of sensors [128], [129]. Bayesian estimation is used to fuse together each sensor's probabilistic estimate as to whether a cell in the grid is occupied by an object. The resulting grid can then be used to determine paths through unoccupied areas of the grid. Jorgensen [130] has proposed dividing the environment into equal-size volumetric cells and associating each with a neuron. In a manner similar to the occupancy grid, the magnitude of each neuron's activation corresponds to the probability that the cell it represents is occupied. The neurons are trained using sensory information from different perspectives. Associative recall (Section III-A-2) can then be used to recognize objects in the environment and simulated annealing can be used to find optimal global paths for navigation.

3) Graphs: Graph structures have been used to represent the local and topological features of the environment to avoid having to define a global metric relation between nonadjacent nodes (or points) in the graph. When landmarks or beacons are not used to correct cumulative position error, a global metric would contain too much uncertainty to be useful. Graph structures allow the topological features to be represented and reasoned with in an efficient manner. Kak *et al.* [131] and Andress and Kak [132] used an attributed graph and Shafer–Dempster evidential reasoning (Section III-E-6) to integrate sensory information for hierarchical spatial reasoning. Brooks [133] proposed the use of a graph to represent regions of potential collision-free motion termed "freeways" and "meadows." Each point in the graph, represented as an "uncertainty manifold," corresponds to the location of a robot in configuration space at which sensory information was acquired; each arc is labeled with a local measurement of the distance traveled between endpoints. The cumulative uncertainty of the robot as it moves from point to point in the graph is taken into account through the cascading of successive uncertainty manifolds.

4) Labeled Regions: Sensory information can be used to segment the environment into regions with properties that are useful for spatial reasoning. The known characteristics of different types of sensory information can be used to label some useful property of each region so that symbolic reasoning can be performed at higher levels in the control structure. Asada [134] proposed a method for building world models that uses the range images from sensors to create height maps of the local environment. Grey levels represent the height of points in the map with respect to an assumed ground plane. The map is then segmented into regions and labeled. Sequences of maps, created as a robot moves, are integrated into a global map by overlaying pairs of height maps and then replacing the labels of corresponding regions in the height maps with a label determined according to a precedence procedure (e.g., if a region is labeled as unexplored in one height map and as an obstacle in another, the global map might label the region as an obstacle). The global map is then used for obstacle detection and path planning. Miller [135] developed a spatial representation that divided a map of an indoor environment into labeled regions. Each label is one of four possible types—each type referring to the number of degrees of freedom in the region (i.e., the two planar dimensions and the robot's orientation) that information from a sensor could be used to eliminate (e.g., the empty space in the center of the room is of type zero, a region near a corner is of type three, etc.). "Voronoi diagrams" are then used to group the regions into areas that correspond to a specific edge (or wall) in the environment.

5) Production Rules: The use of production rules in a control structure allows for a wide range of well-known artificial intelligence methods to be used for path planning and learning purposes. Lawton *et al.* [136] have used

TABLE II
SELECTED EXAMPLES OF MULTISENSOR-BASED MOBILE ROBOTS

Mobile Robot	References	External Sensors	Operating Environment	World Model Representation	Fusion Method
HILARE (1979)	[113], [149]–[154]	vision acoustic laser range finder	unknown man-made	polygon objects in graph of locations	weighted average
Crowley's mobile robot (1984)	[155]–[157]	rotating ultrasonic tactile	known man-made	connected sequences of line segments in two dimensions	best correspondence using integer valued confidence factors
Ground surveillance robot (1984)	[114], [158]	high-resolution grey level vision Low-resolution color vision acoustic laser range finder	unknown natural terrain	triangular segments in blackboard	variety possible (data put in spatial and temporal correspondence)
Stanford mobile robot (1987)	[159]	stereo vision tactile ultrasonic	unknown man-made	hierarchical sensor measurements and symbols	Kalman filter
CMU's ALV's NAVLAB and Terregator (1986)	[111], [160]	color vision sonar laser range finder	unknown roadway	polygon tokens with attribute-value pairs in whiteboard	variety possible (data put in spatial and temporal correspondence)
DARPA ALV (1985)	[112], [115], [161]–[163]	color vision sonar laser range finder	unknown natural terrain	Cartesian elevation maps (CEM's)	average elevation change over small CEM areas

production rules to create schemas to represent both objects in terrain models and certain generic object types. Network hierarchies are created from the schemas that allow inference and matching procedures to take place at multiple levels of abstraction, with each level using an appropriate combination of the available sensors. Isik and Meystel [137] have used fuzzy-valued linguistic variables to represent the attributes of objects as part of a fuzzy logic-based (Section III-E-7) production system for mobile robot control.

C. Sensor Combinations

Due to the advantages and limitations of each type of sensor, most mobile robots use some combination of different sensor types to perform each task of the perception function. Some sensors cannot be used in a particular environment due to their inherent limitations (e.g., acoustic sensors in space), while others are limited due to either technical or economic factors. Obstacle detection with contact sensors necessarily limits the speed of a robot because contact must be made before detection can take place [138]; laser sensors require an intense energy source, and they have a short range and slow scan rate (their use can also cause eye problems [139]); and vision sensors are critically dependent on ambient lighting conditions, and their scene analysis and registration procedures can be complex and time-consuming [138]. Shaky, one of the first autonomous vehicles, used vision together with tactile sensors for obstacle detection [140]. JASON combined acoustic and infrared proximity sensors for obstacle detection and also used these sensors for path planning [141]. The Stanford University Cart used acoustic and infrared sensors together with stereo vision for navigating over a flat terrain while avoiding obstacles [142]. Bixler and Miller [143] used simple low-resolution vision in their autonomous mobile robot to locate the direction of an obstacle, and then used an ultrasonic range finder to determine its depth and shape. Other combinations of sensors used in mobile robot systems have included contact, infrared, and stereo vision [144]; sonar and infrared [145]; contact and acoustic [146]; acoustic and stereo vision [147]; and stereo vision and laser range finding [148].

D. Selected Examples

Short descriptions of a number of different mobile robots are provided next. Emphasis is given to the role of multisensor integration and fusion in their navigation and control. Table II summarizes for comparison the relevant multisensor integration and fusion features of each mobile robot. Included under the name of each mobile robot in the table is the year the initial publication appeared. In cases where there have been major modifications to the mobile robot, the features listed in the table correspond to the most recent published research.

1) Hilare: The mobile robot Hilare combines contact, acoustic, two-dimensional vision, and laser range finding sensors so that it can operate in unknown environments [113], [149]–[154]. Hilare was the first mobile robot to create a world model of an unknown environment using information from multiple sensors [108]. Acoustic and vision sensors are used to create a graph partitioned into a hierarchy of locations. Vision and laser range-finding sensors are then used to develop an approximate three-dimensional representation of different regions in the environment—constraints being used to eliminate extraneous features of the representation. The laser range finder is then used to obtain more accurate range information for each region. To provide a robust and accurate estimation of the robot's position, three independent methods are used: absolute position referencing by use of a beacon, trajectory integration without external reference, and relative position referencing with respect to landmarks in the environment [154]. Each of these methods is used in a complementary fashion to correct or reduce errors and uncertainties in the other methods. The information from a variety of sensors is integrated to provide the position of known objects and places relative to the robot. The shape of each object is represented as a polygon. Depending on the features of an object and its distance from the robot, an appropriate group of redundant sensors is selected to measure the object. The uncertainty of each sensor is modeled as a Gaussian distribution. If the standard deviations of all the sensor's measurements have the same magnitude, a weighted average (Section III-E-1) of their values is used as the fused estimate of a vertex of the object; otherwise, the measurement from the sensor with the smallest standard deviation is used. The estimated vertices of the object can then be matched to known regions of the world model by finding an object in the model that minimizes the weighted sum of the distance between corresponding vertices.

2) Crowley's Mobile Robot: Crowley [155]–[157] describes a mobile robot with a rotating ultrasonic range sensor and a touch sensor capable of autonomous navigation in a known domain. Information from a prelearned global world model is integrated with information from both sensors to maintain dynamically a composite model of the local environment. Obstacles and surfaces are represented as connected sequences of line segments in two dimensions. Confidence as to the actual existence of each line segment is represented by an integer ranging from one (transient) to five (stable and connected). The uncertainty as to the position, orientation, and length of each segment is accounted for by allowing tolerances in the value of these attributes. Information from the global model and the sensors is matched to the composite local model by determining which line segment in the local model has the best correspondence with a given line segment from either the global model or sensors. The best correspondence is found by performing a sequence of tests of increasing computational cost based on the position, orientation, and length of the line segments relative to each other. The results of the matching process are then used to update the composite local model by either adding newly perceived line segments to the model or adjusting the confidence value of the existing segments. The local model can then be used for obstacle detection, local path planning, path execution, and learning.

3) Ground Surveillance Robot: The ground surveillance robot described by Harmon [114], [158] is an autonomous vehicle designed to transit from one known location to another over unknown natural terrain. Vision and acoustic ranging sensors are used for obstacle detection. A laser range finder together with a high-resolution gray-level camera and a low-resolution color camera are used for distant terrain and landmark recognition. The information from the obstacle detection sensors is fused into a single estimate of the position of nearby obstacles by superpositioning distributions that represent each sensor's *a priori* probability of detection. A distributed blackboard is used both to control the various subsystems of the vehicle and as a mechanism through which to integrate and fuse various types of sensor data. As part of the blackboard, a world model is used to organize the data into a class tree with inheritance properties. Each element in the world model has a list of properties to which values are assigned, some values being determined by sensory data. Terrain data are represented as triangular segments with the properties of absolute position, orientation, adjacency, and type of ground cover. To allow for a variety of fusion methods to be used, each element in the world model includes a time stamp and measures of its accuracy and confidence. When two or more sensor values are functionally dependent, changes in one value will propagate throughout the blackboard so that its dependent values reflect the change. When sensor values refer to the same property of an element, either a decision is made as to which of the competing values is to be used (e.g., the most recent value of the most accurate sensor) or the values are fused.

4) Stanford Mobile Robot: The Stanford mobile robot [159] uses tactile, stereo vision, and ultrasonic sensors for navigation in unstructured man-made environments. A hierarchical representation is used in its two-dimensional world model, with features close to the actual sensor measurements at the lowest level and more abstract or symbolic features at the higher levels. The uncertainty as to the location of the robot and the features in the environment is modeled with a Gaussian distribution. A Kalman filter (Section III-E-2) is used to fuse the measurements from a sensor as the robot moves. An example of the application of this method is shown in Fig. 12. Fig. 12(a) shows two points $\boldsymbol{a}$ and $\boldsymbol{b}$ as measured by the stereo vision sensor—first at location $\boldsymbol{p}_1$ and then at location $\boldsymbol{p}_2$. The uncertainty ellipses around the point measurements are elongated toward the sensor because the uncertainty due to distance is much greater than the angular uncertainty when calculated through stereo vision. The uncertainty due to distance is also greater for points further from their location of measurement (i.e., $\boldsymbol{b}$ is more uncertain than $\boldsymbol{a}$). The two measurements of each point (e.g., $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$) are not coincident because of the inherent error in the internal odometer sensor of the robot. In Fig. 12(b), the uncertainty of the points and location $\boldsymbol{p}_2$ are shown with respect to $\boldsymbol{p}_1$. The uncertainty ellipse around $\boldsymbol{p}_2$ is elongated perpendicular to the direction of motion because the angular error of the odometer sensor is greater than its error in determining distance. The uncertainty with respect to $\boldsymbol{p}_1$ of measurements $\boldsymbol{a}_2$ and $\boldsymbol{b}_2$ has increased because their uncertainty with respect to $\boldsymbol{p}_2$ has been compounded with $\boldsymbol{p}_2$ locational uncertainty. In Fig. 12(c), the Kalman filter is used to determine new fused estimates for both points ($\boldsymbol{a}^*$ and $\boldsymbol{b}^*$) that have a reduced uncertainty with

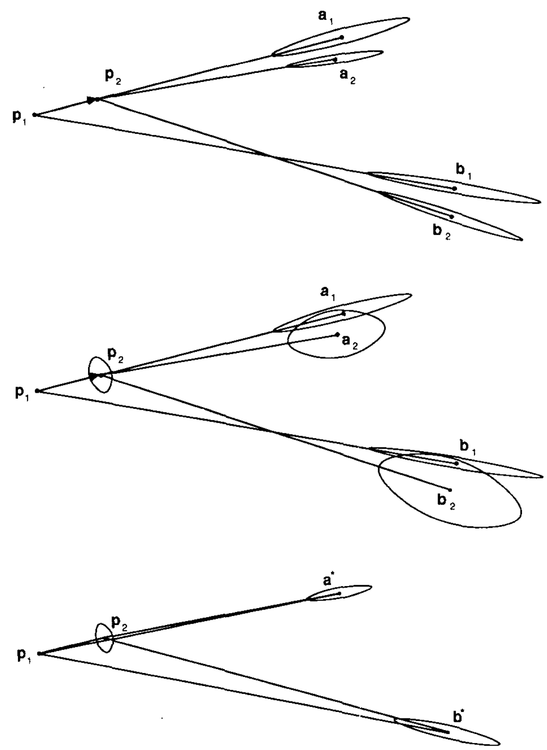

Fig. 12. Reduction in uncertainty as to location of robot ($\boldsymbol{p}$) and two points ($\boldsymbol{a}$ and $\boldsymbol{b}$) in environment through use of Kalman filter to fuse measurements as robot moves from $\boldsymbol{p}_1$ to $\boldsymbol{p}_2$. (a) Uncertainty before fusion of points $\boldsymbol{a}$ and $\boldsymbol{b}$ as measured from $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. (b) Uncertainty of $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{p}_2$ with respect to $\boldsymbol{p}_1$. (c) Fused estimates for $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{p}_2$ with respect to $\boldsymbol{p}_1$. (Adapted from [159, fig. 10].)

respect to both locations $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. The uncertainty of $\boldsymbol{p}_2$ with respect to $\boldsymbol{p}_1$ has also been reduced.

5) CMU's Autonomous Land Vehicles: The NAVLAB and Terregator are two vehicles developed at Carnegie–Mellon University's Robotics Institute as part of their research on autonomous land vehicles [111], [160]. Each vehicle is equipped with a color TV camera, laser range finder, and sonar sensors. The sonar sensors are used to detect nearby obstacles. The design of an architecture able to support parallel processing and the development of multisensor integration and fusion techniques have been major goals of the research. The current system consists of several independently running modules that are tied together in what is termed a "whiteboard" control structure, which differs from a blackboard in that each module continues to run while synchronization and data retrieval requests are made. Data in a local world model are represented as tokens with attribute–value pairs. Tokens representing physical objects and geometric locations consist of a two-dimensional polygonal shape, a reference coordinate frame that can be used to transform the location to other frames, and time stamps that record when the token was created and the time at which sensor data were received that led to its creation. When range data, measured by the camera and laser range finder at different times and locations on the vehicle, are to be fused, the coordinate frames of the tokens created by each sensor for these data are first transformed to a common vehicle frame and then transformed forward to the same point in time. The data are now fused, resulting in the creation of a new token representing the fused data.

6) The DARPA Autonomous Land Vehicle: The Autonomous Land Vehicle (ALV) [112], [115], [161]–[163] built by Martin Marietta is part of the Defense Advanced Research Projects Agency's (DARPA) Strategic Computing Program. The ALV is intended to be a testbed designed for demonstrating the state of the art in autonomous vehicle research [163]. A number of companies and universities are currently working on different research aspects of the project. In the initial stages of the project, the ALV was used in road-following applications [112], [161], [164]; in more recent stages, obstacle avoidance [163] and autonomous cross-country navigation [163], [165] capabilities have been demonstrated. Future research is aimed at enhancing the operational speed and robustness of the ALV, and adding capabilities like landmark recognition [163].

In road-following applications [112], [163], the ALV uses sonar to determine its height, tilt, and roll with respect to road surfaces directly beneath it. Complementary information from a laser range scanner and two color video cameras is used for obstacle detection. Color video information is used to locate roads because the laser range information can easily be confused if there is very little difference in depth between the road and surrounding areas. Laser range information is used to obtain accurate descriptions of the geometrical features of obstacles on the road because, unlike the video information, it is not sensitive to poor lighting conditions and shadows. After being transformed to a common world coordinate system, the video information is used both to determine the boundaries of the road for path planning and, after being integrated with similarly transformed laser range information, for obstacle recognition.

In autonomous cross-country navigation applications [115] a hierarchical control system is used to provide the ALV with the flexibility needed for operation over natural terrain. At the lowest level in the hierarchy, "virtual sensors" and "reflexive behaviors" are used as real-time operating primitives for the rest of the control system [165]. Functioning in a similar transparent manner as the logical sensors described in Section III-A-3, virtual sensors combine information from physical sensors with appropriate processing algorithms to provide specific information to associated reflexive behaviors. Combinations of behavior and virtual sensors are used to handle specific subproblems that are part of the overall navigation task. In these applications a laser range-scanner, together with orientation sensors to determine the pitch, roll, and x and y position of the vehicle, were used to provide information needed to create overhead map view representations of the terrain called Cartesian elevation maps (CEM's); other range sensors such as stereo vision could also be used to create CEM's. Smoothing procedures are applied to the CEM's to fill in detail not provided by the sparse laser range information. As the ALV travels over the terrain, CEM's are fused together to provide a means for selecting traversable trajectories for the vehicle.

VI. Applications

This section discusses a variety of intelligent systems in different areas of application to illustrate the role of multisensor integration and fusion in their overall operation. A description of the use of multisensor-based mobile robots in different areas of application can be found in Section V.

A. Industrial

The addition of sensory capabilities to the control of industrial robots can increase their flexibility and allow for their use in the production of products with a low volume or short design life [166]. In many industrial applications, the use of multiple sensors is required to provide the robot with sufficient sensory capabilities. Most of the multisensor integration and fusion techniques discussed in this paper are suitable for industrial applications because the industrial environment is usually well-structured, and descriptions of many of the objects in the environment are available from the data bases of computer-aided design systems.

Nitzan *et al.* [167] divide industrial robot applications into four general areas: material handling, part fabrication (e.g., spot and arc welding, forging, etc.), inspection, and assembly. Material handling is usually the simplest area of application and assembly the most complex.

1) Material Handling: Industrial robots can be used for in-process workpiece handling and the loading and unloading of industrial trucks (e.g., automatic guided vehicles, tow tractors, etc.) and conveyors—the two major material handling equipment types. Much of the research on mobile robots (Section V) can readily be applied to existing industrial trucks to increase their capabilities in areas such as route planning and obstacle avoidance. Choudry *et al.* [168] have developed a simulation system to test designs for the sensory control of an autonomous material handling vehicle. Their sensory control designs take advantage of the relatively well-structured shop floor environment to avoid having to use the more cumbersome hierarchical control structures used in most mobile robots. The hierarchical phase-template paradigm (Section III-A-1) summarizes the integration and fusion issues resulting from the use of a multisensor robot for handling workpieces. Sensory capabilities can enable a robot to grasp workpieces that are randomly oriented in a bin or on a conveyor. Hitachi Ltd. [169] has developed a robot which uses three-dimensional vision and force sensors to pick up randomly positioned connectors and mount them on a printed circuit board. One of the projects of the ESPRIT program (the European Strategic Programme for Research and Development into Information Technology) is developing a system that combines vision and tactile sensors for real-time applications in material handling [170]. Miller [45], [46] has applied adaptive learning (Section III-B-3) to an experiment involving having a robot use sensory feedback to track and intercept an object on a moving conveyor. Results of the experiment showed that, by the tenth attempt to teach the robot to follow the object to the end of the conveyor, the gripper was able to approach the object to within a small (1–4 cm) degree of accuracy.

2) Part Fabrication: As of 1985, almost half of the robots in U.S. industry were being used for welding [171] —the majority being used in spot welding applications because arc welding robots without sensory capabilities cannot track a seam with randomly variable gaps. Kremers *et al.* [172] developed a robot that used both a vision sensor and wrist-mounted laser scanner range sensors to guide the arm of the robot during the one pass arc welding of workpiece joints that had random gaps along their seams. Howarth and Guyote [173] describe work being done at Oxford that uses eddy current and ultrasonic sensors for robot arc welding. Nitzan *et al.* [167] provide a plan for a sensor guided arc welding system as a specific example to illustrate the use of generic robot functions (e.g., "recognize," "place," "grasp," etc.) and their associated high-level properties (e.g., "identity," "location," etc.) determined from sensory information.

3) Inspection: Inspection can be divided into two different types [167]: explicit and implicit. Explicit inspection verifies the integrity of workpieces, as a separate operation, either during or after the manufacturing process. Depending on the nature of the work pieces involved, any of the multisensor integration and fusion techniques described in this paper could be of potential use for explicit inspection. Implicit inspection verifies the integrity of work pieces while handling them during the manufacturing process. Many of the object-recognition approaches that combine vision with tactile sensors (Section IV-B) would be especially appropriate if applied to implicit inspection (as well as assembly) operations because the manipulator would already be in position to grasp and inspect the workpieces. Manufacturing research at Georgia Tech [174] is focused on integrating vision and tactile information for the adaptive control of a robot manipulator that would be useful in just such applications.

4) Assembly: Assembly is the most complex area of industrial robot application because, in addition to operations like insertion that are unique to assembly, different aspects of the other three application areas can be part of the overall assembly process. Smith and Nitzan [175] describe an assembly station consisting of two robots with wrist-mounted vision and force sensors, overhead cameras, and a general-purpose parts feeder. Printer carriages were assembled by first locating components on the feeder using the overhead cameras and then transporting them to the carriage and snapping them in place using the robots. The force and wrist-mounted vision sensors were used to verify that the components were correctly in place. One of the projects of the ESPRIT program is to develop vision and tactile sensors that can be used in an integrated manner for assembly operations [170].

Groen *et al.* [166], [176] describe a multisensor robotic assembly station equipped with vision, ultrasonic, tactile, and force/torque sensors. The assembly process is represented as a sequence of stages that are entered when certain sensor determined conditions are satisfied. A hierarchical control structure, modeled after the NBS control hierarchy (Section III-B-1), is used to enable the entire process to be executed using a set of modular low-level peripheral processes that perform dedicated tasks like sensory processing, robot control, and data communication. The system has been applied to the assembly of three different types of hydraulic lift assemblies for gas water heaters. In operation, vision sensors are used to recognize different parts of the assemblies as they arrive in varying order and at undefined positions. Feedback information from the force/torque sensors and the passive compliance of the robot's gripper are used for bolt insertion operations and to transport and place, with great precision, assembly housings on work spots so that the housing can be used as a reference for the remaining assembly operations. Final inspection is performed with the vision sensors.

Ruokangas *et al.* [177] describe an experimental hierarchically controlled multisensor robot work station developed using Rockwell's Automation Sciences Testbed. Vision, acoustic, and force/torque sensors, all mounted on the end effector of a robot, were used both separately and in different combinations to demonstrate the limitations of the information each sensor was able to provide and the advantages to be gained when two or more sensors are integrated. In one demonstration, distance information from the acoustic sensor was used to position the end effector so that the camera was at the correct focal distance for the visual inspection of workpieces. In another demonstration, the force/torque sensors were used to provide modifications during task execution to the measurement of hole positions that were initially determined by the integrated range information provided by stereo cameras and an acoustic sensor—the acoustic information being used to provide redundancy to the distance information provided by the cameras.

B. Military

As the complexity, speed, and scope of warfare has increased, the military has increasingly turned to automated systems to support many activities traditionally performed manually [178]. The use of large numbers of diverse sensors as part of these systems has resulted in the issues of multisensor integration and fusion assuming critical importance. As an example of the need for highly automated systems, the typical average information requirements for the command and control of tactical air warfare have been estimated to be 25–50 decisions/min based on 50 000–100 000 reports from 156 separate sensor platforms concerning as many as 1000 hostile targets being tracked up to an altitude of 20 km over an area of 800 km^2 [69]. As part of a comprehensive survey of the possible military applications of artificial intelligence technologies, Franklin *et al.* [179] listed multisensor integration and fusion as being of major importance in the areas of general operations, intelligence analysis and situation assessment, force command and control, autonomous vehicles, avionics, and electronic warfare; together with areas of minor importance, multisensor integration and fusion were listed in 20 of the 25 areas of application. Given this wide scope, multisensor integration and fusion will be one of the technologies necessary for the development of the Autonomous Land Vehicle (Section V-D-6), an intelligent Pilot's Associate, and a command and control system for naval battle management—three of the initial projects supported by DARPA's Strategic Computing Program aimed at exploring the potential of AI-based solutions to important military problems [180].

The use of multisensor integration and fusion for object recognition in military applications requires that consideration be given to some additional factors that are not present in nonmilitary applications. Object recognition can be considered as a two-person zero-sum statistical game played against nature by the system performing the recognition [181]. In the terms of game theory, at each move in the game both players select a strategy. Associated with each pair of possible strategies is a payoff. In a zero-sum game, the gain of one player is equal to the loss of the other. The system's strategies are the possible decisions as to the identity of an object; nature's strategies are the *a priori* probabilities corresponding to the occurrence of each of the possible classes to which an object can belong. In most nonmilitary applications, nature can be considered to be indifferent (i.e., the strategies it selects are based upon the *a priori* probability of each object, and they remain constant even though they may not be optimal in the game theoretic sense); in military applications, by contrast, it is possible that the moves of the other player (e.g., the design or operation of the hostile aircraft being recognized) are being selected with the goal of maximizing the potential payoff.

The possibility of a game against a "malevolent nature" (i.e., the enemy) highlights the necessity of being able to make decisions through the use of a variety of different sensors. Any system relying on the information provided by just a single sensor type could be more vulnerable to what might be called "meta-sensor" error. Although the sensor might be functionally providing correct information, it could be spoofed as to the true identity of an object being sensed (e.g., a hostile aircraft using a friendly radar signature). Through the use of redundant information provided by sensors of different types together with the ability to integrate this information intelligently, the system can minimize the effect of the spoofed sensor in the overall determination of the object's identity. Garvey and Fischler [182] discuss the use of multiple sensors for inductive and interpretative perceptual reasoning in hostile environments, Comparato [183] describes the potential role of multisensor fusion in the next generation of tactical com-

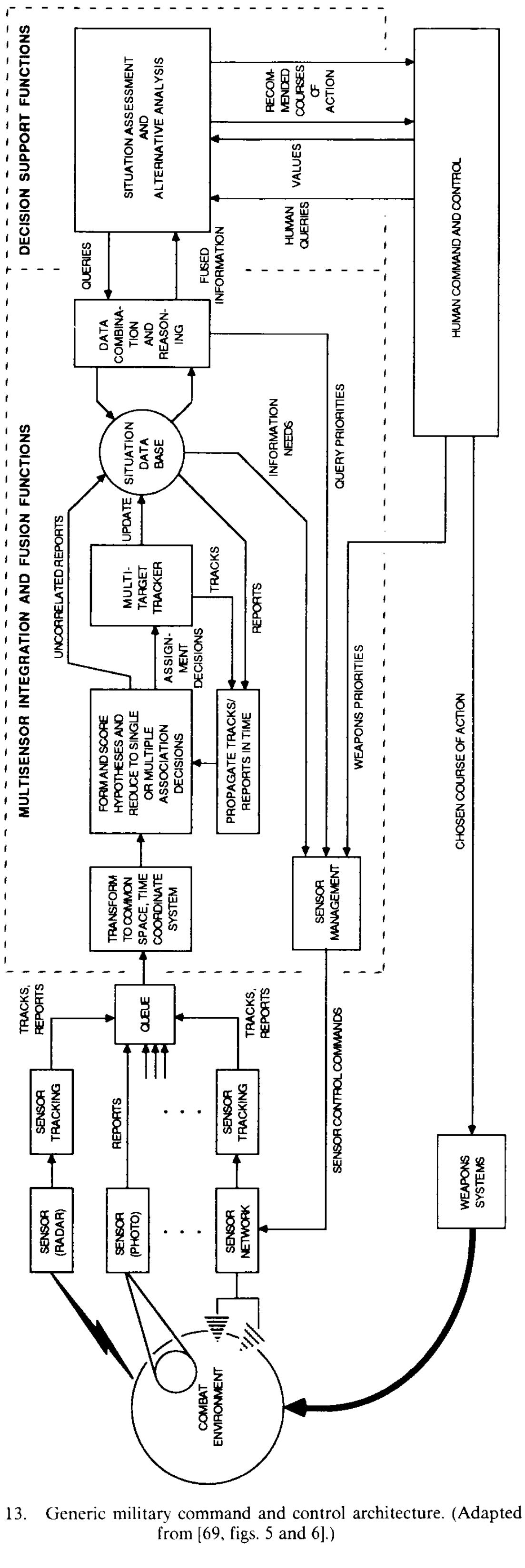

Fig. 13. Generic military command and control architecture. (Adapted from [69, figs. 5 and 6].)

bat platforms, and Mayersak [184] reviews the multisensor fusion aspects of munition (e.g., guided missile) design.

1) A Generic Command and Control Architecture: Waltz and Buede [69] proposed an architecture for a generic command and control system that includes multisensor integration and fusion functions as one of its two major subsystems (see Fig. 13). The operation of the system can be divided into four main steps of a feedback loop. First, a variety of different sensors collect data from the combat environment and then transmit these data on to the multisensor integration and fusion subsystem. The functions comprising this subsystem integrate and fuse the data so that any targets or events can be located and identified. The fused information, representing the possible targets or events that comprise the current situation, is then sent to the decision support subsection where it is used to create, analyze, and rank alternative courses of action. A human commander completes the feedback loop by selecting a course of action which possibly changes the environment. The system initiates operation with a query from the human commander to the decision support subsystem for recommended courses of action. Using certain key parameter values supplied by the human commander and the system's assessment of current situation, the decision support subsystem analyzes alternative courses of action, possibly querying the multisensor integration and fusion subsystem for additional information, to select those actions to recommend to the human commander. The human commander can then either select one of the current recommended actions or query the system for additional information. Although the human commander can always make the final decision, certain routine or time-critical actions can automatically be determined by the system.

The process of data fusion starts with each sensor or network of sensors in the combat environment sending detection reports or target tracks to a queue from which they are made commensurate by being transformed to a common space and time coordinate system. A local target track is maintained by a sensor if the measurement time of the sensor is small relative to the ability of the system to process the data. (In Fig. 13, both the radar and the network of sensors (e.g., ground sensors measuring seismic or acoustic events) may generate a track containing dozens of measurements in the time it takes the photographic sensor to generate a single image.) Because the system is generic, the common spatial reference (e.g., *XYZ*; latitude/longitude, range/azimuth/elevation, etc.) depends on the specific environment, and the methods of transforming the sensed data depend on the differing locations, resolutions, fields of view, and measurement times of the particular sensors used. After the reports or tracks are transformed to a common coordinate system, hypothetical pairwise assignments are made to each existing report from the situation data base and track from the multitarget dynamic tracking filter—each of which has been propagated in time to coincide with the current data. The assignments are then scored using a metric defined over some feature of the data (e.g., the spatial distance between the data and existing reports or tracks). The resulting scores are then used in selecting either a single hypothesis or a small set of hypotheses that are sent to the multitarget tracker. Probabilistic methods are sometimes needed to make the selection because of uncertainties in the sensor measurements or the large number of hypotheses generated when there are many sensors and possible targets. Reports that remain uncorrelated to any existing reports are sent to the situation data base for possible use in the future. When an assignment of a new report is made, the multitarget tracker updates the estimate of the associated target in the data base. The situation data base contains the reports and tracks that correspond to the most likely grouping of the data so far collected. The process of actually combining the collected data to enable the attributes of a target (e.g., its identity, intent, future behavior, etc.) to be inferred is the core function in the multisensor integration and fusion subsystem. While all of the methods of fusion mentioned in this paper can be used in implementing this function, the use of production rules to infer higher-level information (e.g., intent and future behavior) allows for a convenient interface to the decision support subsystem. The arrival of new sensor data causes a forward-chaining process that may result in the antecedent of a production rule being satisfied and the situation data base being updated with the rule's consequent. Reversing the direction of inference, queries from the decision support subsystem as to the support for a hypothetical situation can cause a backward-chaining process that searches to determine if any of the antecedents in the data base support the consequents implied by the query. If the required antecedents to not exist or have too large of an associated uncertainty, the data base can direct the sensor management function to redirect the sensors to search for data to support the required antecedents; for example, if the current possible enemy threat to a specific location X is queried, sensors near to X could be redirected to focus on that location. Priorities can be sent to the sensor management function when multiple queries exist.

2) Analyst: Analyst [185], [186] is a prototype expert system developed for the U.S. Army that generates real-time battlefield tactical situation descriptions using the information provided by multiple sensor sources. (This summary follows that given in [174].) Analyst determines as output a situation map showing the suspected location of enemy units through the use of intelligence input from battlefield sensors, photographs, and intercepted enemy communications and radar transmissions. Each intelligence report input is represented as a frame. The use of frames allows default values to be attached to the slots of the frame in the case where the intelligence report was incomplete. Each frame is then applied simultaneously to production rules contained in the first two of six different knowledge bases. The initial fusion of the sensor information is performed by these rules in the process of inferring a possible battlefield entity (also represented as a frame). Associated with each entity frame is a slot containing a likelihood factor to indicate the strength of the evidence used to infer the presence of the entity. The third knowledge base eliminates possible multiple frames corresponding to the same entity that may have been created from the reports of different sensors. The fourth and fifth knowledge bases refine and reinforce the entity frames through the use of tactical and terrain data together with the presence of other possible related entities in the knowledge bases. The sixth knowledge base serves to remove from consideration those entities that have not been reinforced with additional information for a sufficient length of time. Each knowledge base is applied sequentially to each piece of sensor information as it becomes available, enabling Analyst continuously to provide the most current estimate of the battlefield situation.

C. Space

NASA's permanently manned space station will be the United States' major space program in coming years [187]. Previous NASA programs, including the space shuttle, have used a high degree of participation by both the crew and ground-based personnel to perform the sensing and perception functions required for many tasks. Both to increase productivity and allow tasks beyond the capability of the crew to be performed, the space station will make increasing use of autonomous systems for the servicing, maintenance, and repair of satellites and the assembly of large structures for use as production facilities and commercial laboratories. Probably the most important factor promoting the use of autonomous systems for these applications is the cost of having a human in space. In addition to these economic factors, aspects of the space environment can make the use of multiple sensors an especially important part of these systems. The sensing of objects in space using just optical sensors is difficult because the lack of atmosphere invalidates some of the common assumptions concerning surface reflectance used in many visual recognition methods; also, images in space frequently have deep shadows, missing edges, and intense specular reflections [188].

Shaw *et al.* [188] have presented a system that uses TV images to guide a microwave radar unit in the determination of the shape of objects in space. Microwave radar information serves to complement optical information for a number of important features of objects typically found in space. Scattering cross-sectional data from radar can be used to determine the shape of the metallic surfaces typically found on satellites; optical sensors and even laser range-scanners have difficulty with metallic surfaces because they reflect light in a specular direction. Optical sensors would be more useful in determining the shape of the slightly matte surfaces found on the space shuttle because these surfaces reflect light more equally in all directions. As compared to optical wavelengths, the longer wavelength of radar can be used to penetrate the solar blankets on space objects that sometimes cover important surface details needed by a robot for grasping and may sometimes diffract around objects that are occluded along the optical sensor's line of sight. The optical image is used to provide an initial estimate of the shape of a surface that serves to reduce the ambiguity inherent when interpreting narrow-band radar-scattering cross-sectional data. In operation, the system uses equations defining the orthogonal directions of the polarized radar-scattering cross section to determine surface shapes. Initially, occluding contours are derived from the optical image of the surface by thresholding, and a partial surface shape description is derived from shape-from-shading or stereo. The surface is then either matched to some simple geometric shape or, if no match is found, a grid is constructed over the surface. If a simple match is found, a closed-form expression can be used for the scattering equations; if no match is found, an iterative nonlinear least squares technique is used to approximate the equations.

D. Target Tracking

A variety of different filtering techniques, together with radar, optical, and sonar sensors, have been used for tracking targets (e.g., missiles, planes, submarines, etc.) in the air, space, and under water. Bar-Shalom and Fortmann [189], [190] have surveyed a number of tracking methods that can be used when there are multiple targets in the environment. Recently, researchers have been developing methods of multitarget tracking that integrate and/or fuse the information (or measurements) provided by a number of identical sensors. The key problem in multitarget tracking in general, and in multisensor multitarget tracking in particular, is "data association"—the association of sensor information to a single target [191] (in general multisensor integration and fusion, this problem is termed the registration problem (Section VII-A-1)). For the purposes of this discussion, the different multisensor tracking methods can be distinguished based on the complementary versus redundant nature of the information provided by the sensors (Section II-C).

1) Complementary Tracking Information: The position and velocity of targets can be derived through the use of the complementary information provided by the time of arrival and frequency of signals sent by a small group of sensors. The difference between both the time and frequency of the signals arriving at each sensor can be used to determine the tracks of targets. Arnold *et al.* [192] and Mucci *et al.* [193] have used the Fisher information matrix to evaluate the performance of this approach and have found that a single pair of omnidirectional sensors are sufficient to determine the location of moving targets when they are in the vicinity ("near field") of the sensors; three sensors are required when they are distant or stationary.

2) Redundant Tracking Information: Redundant tracking information can be provided by a network of sensors distributed over a large geographic area to increase the overall reliability and survivability of the tracking system [191]. Two different processing architectures have been developed for the association of the information provided by each sensor in the network. In the first, the measurements of all the sensors are transmitted to a centralized site for processing; in the second, the processing is distributed to local nodes in the network. As compared to the first, the second architecture provides the additional benefits of an increased survivability and the possibility of using a smaller bandwidth for communication within the network.

Chang *et al.* [191] and Chong *et al.* [194] have presented a distributed processing version of the Joint Probabilistic Data Association algorithm, first applied in a centralized processing architecture [195], that reduces the network's bandwidth requirements by fusing, at each local node in the network, measurements from a small number of sensors. The compressed higher level fused information is then propagated to the other local nodes in the network. Each local node can use the fused information it receives from the other nodes to arrive at a solution to the global tracking problem. The algorithm used for fusion was developed by Speyer [196]. It is based on the use of a Kalman filter (Section III-E-2) at each local node and requires the propagation of an additional data-dependent vector beyond the usual filter equations.

In a series of papers Thomopoulos *et al.* [197]–[200] have developed algorithms for fusing target detection information from a distributed network of sensors. In the first paper [197], optimal and suboptimal algorithms are developed for the situation in which a group of parallel sensors transmit their detection decision to a fusion center where, assuming the sensors are conditionally independent, a Neyman–Pearson test maximizes the fused probability of detection for a fixed probability of false alarm. A likelihood-ratio test is used for detection at each sensor. In the second paper [198], a pair of sensors is considered in which one sensor is the primary sensor responsible for the final detection decision and the other is a consulting sensor. Based on an estimated cost associated with any communication between the sensors and the quality of its own raw data, the primary sensor makes a decision as to whether to consult with the other sensor as part of its overall detection decision. In the third paper [199] delays in the network and channel errors are considered, and in the final paper [200] the time origin of the information from each sensor is assumed to be uncertain. An extension to the Kalman filter is developed to account for this uncertainty.

E. Inertial Navigation

The inertial system used in spacecraft and many advanced aircraft for navigation consists of a gyroscope mounted on a platform suspended in a gimbal structure that allows the vehicle to change its angular orientation while maintaining the platform fixed with respect to a reference coordinate frame [66], [201]. The position and velocity of the vehicle can be determined by integrating the signals from accelerometers mounted on the platform. Due to inherent gyro characteristics that cause errors in posi-

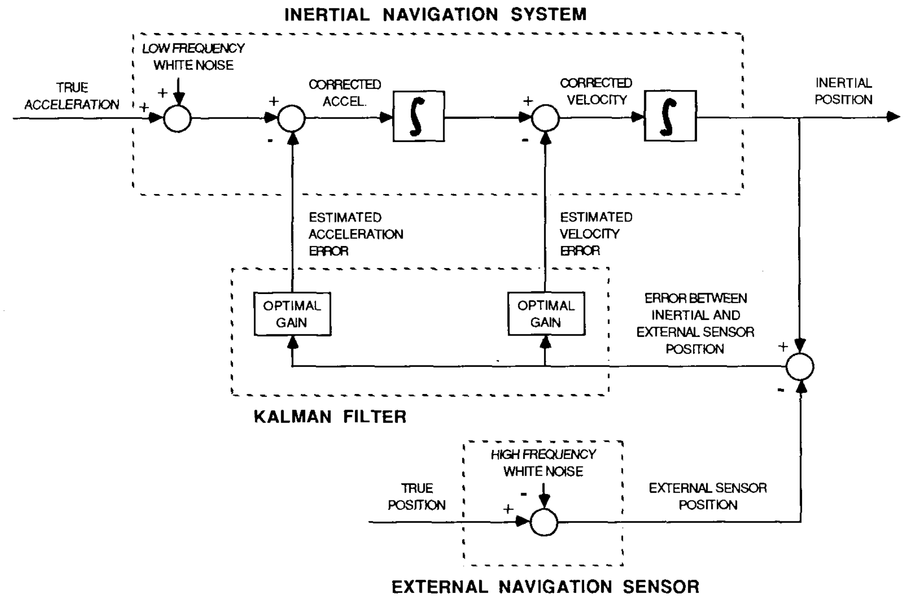

Fig. 14. External sensor-aided indirect feedback inertial navigation system configuration. (Adapted from [66, fig. 6.9(a)].)

tion and velocity to grow unbounded slowly over time (i.e., low-frequency noise), the inertial navigation system requires the aid of other external navigation sensors to bound or dampen these errors. Typical external sensors used to aid the inertial system include on-board or ground-based radar, radio, navigation, satellites, landmarks or star sightings, laser ranging, and altimeters. As opposed to the inertial system which accurately follows the high-frequency motions of the vehicle, each of these external sensors provides information which is good on average (i.e., its error does not increase over time) but is subject to considerable high-frequency noise (i.e., each measurement has considerable error). A Kalman filter (Section III-E-2) can be used to determine the fusion of inertial system and external sensor information that will statistically minimize the error in the estimate of the vehicle's position.

Fig. 14 shows a typical external sensor-aided inertial navigation system configuration. In the figure, the Kalman filter is used in an indirect (error state space) feedback configuration to generate estimates of the errors in the inertial system and then feed these estimates back into the inertial system to correct it. The indirect feedback filter configuration is used in most inertial navigation systems because 1) as opposed to the direct formulation where the vehicle's actual position is estimated, the indirect formulation increases the overall reliability of the navigation system because the inertial system can still operate if the filter should fail due to a temporary computer or external navigation sensor failure; 2) the feedback, as opposed to feedforward, mechanization is used so that inertial errors do not grow unchecked; and 3) the sample rate for this filter is low because only low frequency linear dynamics are modeled.

F. Remote Sensing

In aerial photo mapping over land, known ground points can be used to establish the orientation of photographs; in aerial mapping over water, the orientation must be determined by accurately knowing the position and attitude of the camera at the time the photograph is taken because known ground points are not generally available. Gesing and Reid [202] describe a system that fuses information from multiple navigation sensors to estimate an aircraft's position and attitude accurately enough for use in the mapping of shallow coastal waters. An inertial navigation system is mounted to the top of the aerial survey camera in the aircraft used for mapping. During flight, information is recorded from a number of auxiliary navigation sensors including: a laser bathymeter to measure water depth, a microwave ranging system, barometric and radar altimeters, a radio navigation system, and a Doppler radar. A U-D covariance factorization filter (Section III-E-2) and a modified Bryson–Frazier smoother are then used for post-flight processing to produce time-correlated sensor error estimates that can be subtracted from the recorded inertial system data to yield highly accurate position and attitude information that serves to orient the photographs taken during the flight.

VII. Conclusion

The issues of and approaches to the problem of multisensor integration and fusion presented above demonstrate the wide scope of present research efforts in this area. To conclude this paper, a discussion of the possible problems and future research directions in the area of multisensor integration and fusion is provided.

A. 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 1) the integration and fusion process, 2) the sensory information, and 3) the operation of the overall system including the sensors. For the potential advantages in integrating multiple sensors (Section II-C) 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 [77] and data association [191] problem in stereo vision and multitarget tracking research, respectively.) Barniv and Casasent [203] 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. Lee and Van Vleet [204] and Holm [205] have studied the registration errors between radar and infrared sensors. Lee and Van Vleet have presented an approach that is able both to estimate and minimize the registration error, and Holm has developed a method that is able to compensate autonomously 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 are 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 (Section VI-E-1)) it is still sometimes possible to retain the whiteness assumption through the use of a shaping filter [66]. 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 in Sections III-E-4 and -5) 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 (cf. paragraph 3, below).

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 [66]. 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. [23], [206]).

B. 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 [207]. Many integration and fusion techniques will likely be implemented on recently developed highly parallel computer architectures (e.g., the Connection Machine [208], etc.) 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 [9]. 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 [209], [210]. As current progress in VLSI technology [211] continues, it is likely that so-called "smart sensors" [212] 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" [213] 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" [214] robots will necessarily entail the advancement of the state of the art in multisensor integration and fusion.

REFERENCES

[1] I. P. Howard, *Human Visual Orientation*. Chichester, UK: Wiley, 1982, ch. 11.

[2] E. A. Newman and P. H. Hartline, "The infrared 'vision' of snakes," *Sci. Amer.*, vol. 246, no. 3, pp. 116–127, Mar. 1982.

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

[4] A. Mitiche and J. K. Aggarwal, "Multiple sensor integration/fusion through image processing: a review," *Opt. Eng.*, vol. 25, no. 3, pp. 380–386, 1986.

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

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

[7] R. C. Luo and M. G. Kay, "Multisensor integration and fusion: issues and approaches," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 42–49.

[8] S. S. Blackman, "Theoretical approaches to data association and fusion," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 50–55.

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

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

[11] M. A. Fischler and O. Firschein, *Intelligence: The Eye, the Brain and the Computer*. Reading, MA: Addison-Wesley, 1987, pp. 241–242.

[12] J. M. Richardson and K. A. Marsh, "Fusion of multisensor data," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 78–96, 1988.

[13] R. C. Luo, M. Lin, and R. S. Scherp, "The issues and approaches of a robot multisensor integration," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 1941–1946.

[14] ____, "Dynamic multi-sensor data fusion system for intelligent robots," *IEEE J. Robot. Automat.*, vol. RA-4, no. 4, pp. 386–396, 1988.

[15] ____, "Multi-sensor based intelligent robot system," in *Proc. Int. Conf. Indust. Elect. Contr. and Instrumentation*, Milwaukee, WI, Sept. 1986, pp. 238–243.

[16] R. C. Luo and M. Lin, "Multi-sensor integrated intelligent robot for automated assembly," in *Proc. Workshop Spatial Reasoning and Multi-Sensor Fusion*, St. Charles, IL, Oct. 1987, pp. 351–360.

[17] ____, "Robot multi-sensor fusion and integration: optimum estimation of fused sensor data," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 1076–1081.

[18] G. E. Hinton and J. A. Anderson, Eds., *Parallel Models of Associative Memory*. Hillsdale, NJ: Erlbaum, 1981.

[19] D. E. Rumelhart and J. L. McClelland, Eds., *Parallel Distributed Processing: Explorations into the Microstructure of Cognition*. Cambridge, MA: MIT Press, 1986.

[20] T. Kohonen, *Self-Organization and Associative Memory*. Berlin, Germany: Springer-Verlag, 1984.

[21] A. J. Pellionisz, "Sensorimotor operation: A ground for the co-evolution of brain-theory with neurobotics and neurocomputers," in *Proc. IEEE 1st Int. Conf. Neural Networks*, San Diego, CA, June 1987, pp. IV-593–IV-600.

[22] J. C. Pearson, J. J. Gelfand, W. E. Sullivan, R. M. Peterson, and C. D. Spence, "Neural network approach to sensory fusion," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 103–108.

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

[24] W. B. Dress, "Frequency-coded artificial neural networks: An approach to self-organizing systems," in *Proc. IEEE 1st Int. Conf. Neural Networks*, San Diego, CA, June 1987, pp. II-47–II-54.

[25] T. Henderson and E. Shilcrat, "Logical sensor systems," *J. Robot. Syst.*, vol. 1, no. 2, pp. 169–193, 1984.

[26] T. Henderson, E. Shilcrat, and C. Hansen, "A fault tolerant sensor scheme," in *Proc. 7th Int. Conf. Pattern Recognition*, Montreal, PQ, Canada, July 1984, pp. 663–665.

[27] T. Henderson, C. Hansen, and B. Bhanu, "A framework for distributed sensing and control," in *Proc. 9th Int. Joint Conf. Artificial Intell.*, Los Angeles, CA, Aug. 1985, pp. 1106–1109.

[28] ____, "The synthesis of logical sensor specifications," in *Proc. SPIE*, vol. 579, *Intell. Robots and Comput. Vision*, Cambridge, MA, Sept. 1985, pp. 442–445.

[29] ____, "The specification of distributed sensing and control," *J. Robot. Syst.*, vol. 2, no. 4, pp. 387–396, 1985.

[30] T. C. Henderson and E. Weitz, "Multisensor integration in a multiprocessor environment," in *Proc. ASME Int. Comput. in Engr. Conf. and Exhibition*, R. Raghavan and T. J. Cokonis, Eds., New York, NY, Aug. 1987, pp. 311–316.

[31] T. Henderson, E. Weitz, C. Hanson, and A. Mitiche, "Multisensor knowledge systems: interpreting 3-D structure," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 114–137, 1988.

[32] T. C. Henderson and C. Hanson, "Multisensor knowledge systems," in *Real-Time Object Measurement and Classification*, A. K. Jain, Ed. Berlin, Germany: Springer-Verlag, 1988, pp. 375–390.

[33] R. C. Luo and T. Henderson, "A servo-controlled robot gripper with multiple sensors and its logical specification," *J. Robot. Syst.*, vol. 3, no. 4, pp. 409–420, 1986.

[34] J. C. Rodger and R. A. Browse, "An object-based representation for multisensory robotic perception," in *Proc. Workshop Spatial Reasoning and Multi-Sensor Fusion*, St. Charles, IL, Oct. 1987, pp. 13–20.

[35] P. K. Allen, "A framework for implementing multi-sensor robotic tasks," in *Proc. ASME Int. Computers in Engr. Conf. and Exhibition*, R. Raghavan and T. J. Cokonis, Eds., New York, NY, Aug. 1987, pp. 303–309.

[36] H. G. McCain, "A hierarchically controlled, sensory interactive robot in the automated manufacturing research facility," in *Proc. IEEE Int. Conf. Robotics and Automat.*, St. Louis, MO, Mar. 1985, pp. 931–939.

[37] E. W. Kent and J. S. Albus, "Servoed world models as interfaces between robot control systems and sensory data," *Robotica*, vol. 2, pp. 17–25, 1984.

[38] J. S. Albus, *Brains, Behavior, & Robotics*. Peterborough, NH: Byte Books, 1981.

[39] M. Shneier *et al.*, "Robot sensing for a hierarchical control system," in *Proc. 13th Int. Symp. Industrial Robots and Robots 7*, Chicago, IL, Apr. 1983, pp. 14.50–14.64.

[40] A. J. Barbera, M. L. Fitzgerald, J. S. Albus, and L. S. Haynes, "RCS: The NBS real-time control system," in *Robots 8 Conf. Proc.*, Detroit, MI, June 1984, pp. 19.1–19.33.

[41] J. S. Albus, "Data storage in the Cerebellar Model Articulation Controller (CMAC)," *J. Dynamic Syst., Meas. Contr.*, vol. 97, no. 3, pp. 228–233, 1975.

[42] ____, "A new approach to manipulator control: The cerebellar model articulation controller (CMAC)," *J. Dynamic Syst., Meas. Contr.*, vol. 97, no. 3, pp. 220–227, 1975.

[43] S. Y. Harmon, G. L. Bianchini, and B. E. Pinz, "Sensor data fusion through a distributed blackboard," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1449–1454.

[44] W. T. Miller, III, F. H. Glanz, and L. G. Kraft, III, "Application of a general learning algorithm to the control of robotic manipulators," *Int. J. Robot. Res.*, vol. 6, no. 2, pp. 84–98, 1987.

[45] W. T. Miller, III, "Sensor-based control of robotic manipulators using a general learning algorithm," *IEEE J. Robot. Automat.*, vol. RA-3, no. 2, pp. 157–165, 1987.

[46] ____, "A nonlinear learning controller for robotic manipulators," in *Proc. SPIE*, vol. 579, *Intell. Robots and Comput. Vision*, Cambridge, MA, Sept. 1986, pp. 416–423.

[47] G. Beni, S. Hackwood, L. A. Hornak, and J. L. Jackel, "Dynamic sensing for robots: an analysis and implementation," *Int. J. Robot. Res.*, vol. 2, no. 2, pp. 51–61, 1983.

[48] S. A. Hutchinson, R. L. Cromwell, and A. C. Kak, "Planning sensing strategies in a robot work cell with multisensor capabilities," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1068–1075.

[49] W. E. L. Grimson, "Disambiguating sensory interpretations using minimal sets of sensory data," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 286–292.

[50] G. E. Taylor and P. M. Taylor, "Dynamic error probability vectors: a framework for sensory decisionmaking," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 1096–1100.

[51] T. C. Henderson, W. S. Fai, and C. Hansen, "MKS: A multisensor kernel system," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-14, no. 5, pp. 784–791, 1984.

[52] T. Henderson, B. Bhanu, and C. Hansen, "Distributed control in the multi-sensor kernel system," in *Proc. SPIE*, vol. 521, *Intell. Robots and Comput. Vision*, Cambridge, MA, Nov. 1984, pp. 253–255.

[53] M. O. Shneier, R. Lumia, and E. W. Kent, "Model-based strategies for high-level robot vision," *Comp. Vision Graph. Image Process.*, vol. 33, pp. 293–306, 1986.

[54] E. W. Kent, M. O. Shneier, and T. H. Hong, "Building representations from fusions of multiple views," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1634–1639.

[55] M. O. Shneier, E. W. Kent, and P. Mansbach, "Representing workspace and model knowledge for a robot with mobile sensors," in *Proc. 7th Int. Conf. Pattern Recognition*, Montreal, PQ, Canada, July 1984, pp. 199–202.

[56] L. F. Pau, "Fusion of multisensor data in pattern recognition," in *Pattern Recognition Theory and Applications*, J. Kittler, K. S. Fu, and L. F. Pau, Eds. Reidel, 1982, pp. 189–201.

[57] ____, "Knowledge representation for three-dimensional sensor fusion with context truth maintenance," in *Real-Time Object Measurement and Classification*, A. K. Jain, Ed. Berlin, Germany: Springer-Verlag, 1988, pp. 391–404.

[58] H. F. Durrant-Whyte, "Consistent integration and propagation of disparate sensor observations," *Int. J. Robot. Res.*, vol. 6, no. 3, pp. 3–24, 1987.

[59] ____, "Sensor models and multisensor integration," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 97–113, 1988.

[60] ____, "Consistent integration and propagation of disparate sensor observations," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1464–1469.

[61] ____, "Sensor models and multi-sensor integration," in *Proc. Workshop Spatial Reasoning and Multi-Sensor Fusion*, St. Charles, IL, Oct. 1987, pp. 303–312.

[62] R. McKendall and M. Mintz, "Robust fusion of location information," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 1239–1244.

[63] M. Zeytinoglu and M. Mintz, "Robust fixed size confidence procedures for a restricted parameter space," *Ann. Statist.*, vol. 16, no. 3, pp. 1241–1253, 1988.

[64] ____, "Optimal fixed size confidence procedures for a restricted parameter space," *Ann. Statist.*, vol. 12, no. 3, pp. 945–957, 1984.

[65] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[66] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, vols. 1 and 2. New York: Academic, 1979 and 1982.

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

[68] P. L. Bogler, "Shafer-Dempster reasoning with applications to multisensor target identification systems," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-17, no. 6, pp. 968–977, 1987.

[69] E. L. Waltz and D. M. Buede, "Data fusion and decision support for command and control," *IEEE Trans. Syst. Man Cybern.*, vol. 16, no. 6, pp. 865–879, 1986.

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

[71] T. L. Huntsberger and S. N. Jayaramamurthy, "A framework for multi-sensor fusion in the presence of uncertainty," in *Proc. Workshop Spatial Reasoning and Multi-Sensor Fusion*, St. Charles, IL, Oct. 1987, pp. 345–350.

[72] L. A. Zadeh, "Fuzzy sets," *Inform. Contr.*, vol. 8, pp. 338–353, 1965.

[73] S. J. Kamat, "Value function structure for multiple sensor integration," in *Proc. SPIE*, vol. 579, *Intell. Robots and Computer Vision*, Cambridge, MA, Sept. 1985, pp. 432–435.

[74] R. Belknap, E. Riseman, and A. Hanson, "The information fusion problem and rule-based hypotheses applied to complex aggregations of image events," in *Proc. IEEE Conf. Computer Vision and Pattern Recog.*, Miami Beach, FL, June 1986, pp. 227–234.

[75] A. R. Hanson, E. M. Riseman, and T. D. Williams, "Sensor and information fusion from knowledge-based constraints," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 186–196.

[76] D. Ballard and C. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[77] B. K. P. Horn, *Robot Vision*. Cambridge, MA: MIT Press, 1986.

[78] M. J. Magee and J. K. Aggarwal, "Using multisensory images to derive the structure of three-dimensional objects—A review," *Comp. Vision Graph. Image Process.*, vol. 32, pp. 145–157, 1985.

[79] M. J. Magee, B. A. Boyter, C. Chien, and J. K. Aggarwal, "Experiments in intensity guided range sensing recognition of three-dimensional objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 6, pp. 629–636, 1985.

[80] J. K. Aggarwal and M. J. Magee, "Determining motion parameters using intensity guided range sensing," *Pattern Recognition*, vol. 19, no. 2, pp. 169–180, 1986.

[81] E. Hildreth, "Optical flow," in *Encyclopedia of Artificial Intelligence*. New York: Wiley, 1986, pp. 684–687.

[82] Y. P. Hung, D. B. Cooper, and B. Cernuschi-Frias, "Bayesian estimation of 3-D surfaces from a sequence of images," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 906–911.

[83] G. M. Flachs, J. B. Jordan, and J. J. Carlson, "Information fusion methodology," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 56–63.

[84] N. Ayache and O. D. Faugeras, "Building, registrating, and fusing noisy visual maps," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 45–65, 1988.

[85] O. D. Faugeras, N. Ayache, and B. Faverjon, "Building visual maps by combining noisy stereo measurements," in *Proc. IEEE Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1433–1438.

[86] S. T. Barnard and M. A. Fischler, "Stereo vision," in *Encyclopedia of Artificial Intelligence*. New York: Wiley, 1986, pp. 1083–1090.

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

[88] J. Porrill, "Optimal combination and constraints for geometrical sensor data," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 66–77, 1988.

[89] R. C. Luo and M. Lin, "3-D object recognition using combined overhead and robot eye-in-hand vision system," in *Proc. SPIE*, vol. 856, *IECON: Indust. Applic. of Robots and Machine Vision*, Cambridge, MA, Nov. 1987, pp. 682–689.

[90] P. K. Allen, "Integrating vision and touch for object recognition tasks," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 15–33, 1988.

[91] ____, *Robotic Object Recognition using Vision and Touch*. Boston, MA: Kluwer, 1987.

[92] ____, "Sensing and describing 3-D structure," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 126–131.

[93] ____, "Surface descriptions from vision and touch," in *Proc. IEEE Int. Conf. Robotics*, 1984, pp. 394–397.

[94] P. Allen and R. Bajcsy, "Integrating sensory data for object recognition tasks," in *Proc. SPIE*, vol. 595, *Computer Vision for Robotics*, Cannes, France, Dec. 1985, pp. 225–232.

[95] ____, "Object recognition using vision and touch," in *Proc. 9th Int. Joint Conf. Artificial Intell.*, Los Angeles, CA, Aug. 1981, pp. 1131–1137.

[96] R. C. Luo and W. Tsai, "Object recognition using tactile array sensors," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1248–1253.

[97] S. A. Stansfield, "A robotic perceptual system utilizing passive vision and active touch," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 138–161, 1988.

[98] ____, "Visually-aided tactile exploration," in *Proc. IEEE Int. Conf. Robotics and Automation*, Raleigh, NC, Mar. 1987, pp. 1487–1492.

[99] ____, "Primitives, features, and exploratory procedures building a robot tactile perception system," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1274–1279.

[100] ____, "Representing generic objects for exploration and recognition," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 1090–1095.

[101] N. Nandhakumar and J. K. Aggarwal, "Integrated analysis of thermal and visual images for scene interpretation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-10, no. 4, pp. 469–481, 1988.

[102] ____, "Thermal and visual information fusion for outdoor scene perception," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 1306–1308.

[103] ____, "Thermal and visual sensor fusion for outdoor robot vision," in *Proc. ASME Int. Comput. in Engr. Conf. Exhibition*, R. Raghavan and T. J. Cokonis, Eds., New York, NY, Aug. 1987, pp. 301–302.

[104] ____, "Synergetic integration of thermal and visual images for computer vision," in *Proc. SPIE*, vol. 782, *Infrared Sensors and Sensor Fusion*, R. G. Buser and F. B. Warren, Eds., Orlando, FL, May 1987, pp. 28–36.

[105] W. E. L. Grimson and T. Lozano-Pérez, "Model-based recognition and localization from sparse range or tactile data," *Int. J. Robot. Res.*, vol. 3, no. 3, pp. 3–35, 1984.

[106] M. C. Roggemann, J. P. Mills, S. K. Rogers, and M. Kabrisky, "Multisensor information fusion for target detection and classification," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 8–13.

[107] C. W. Tong, S. K. Rogers, J. P. Mills, and M. K. Kabrisky, "Multisensor data fusion of laser radar and forward looking infrared (FLIR) for target segmentation and enhancement," in *Proc. SPIE*, vol. 782, *Infrared Sensors and Sensor Fusion*, R. G. Buser and F. B. Warren, Eds., Orlando, FL, May 1987, pp. 10–19.

[108] S. Y. Harmon, "Autonomous vehicles," in *Encyclopedia of Artificial Intelligence*. New York: Wiley, 1986, pp. 39–45.

[109] R. C. Luo and M. G. Kay, "The role of multisensor integration and fusion in the operation of mobile robots," presented at the IEEE Int. Workshop Intelligent Robots, Tokyo, Japan, Oct. 1988.

[110] P. Levi, "Principles of planning and control concepts for autonomous mobile robots," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 874–881.

[111] Y. Goto and A. Stentz, "The CMU system for mobile robot navigation," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 99–105.

[112] M. A. Turk, D. G. Morgenthaler, K. D. Gremban, and M. Marra, "Video road-following for the Autonomous Land Vehicle," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 273–279.

[113] G. Giralt, R. Chatila, and M. Vaisset, "An integrated navigation and motion control system for autonomous multisensory mobile robots," in *Robotics Research: The First International Symposium*, M. Brady and R. Paul, Eds. Cambridge, MA: MIT Press, 1984, pp. 191–214.

[114] S. Y. Harmon, "The Ground Surveillance Robot (GSR): an autonomous vehicle designed to transit unknown terrain," *IEEE J. Robot. Automat.*, vol. RA-3, no. 3, pp. 266–279, 1987.

[115] M. Daily *et al.*, "Autonomous cross-country navigation with the ALV," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 718–726.

[116] U. Rembold, "The Karlsruhe autonomous mobile assembly robot," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 598–603.

[117] R. C. Mann, W. R. Hamel, and C. R. Weisbin, "The development of an intelligent nuclear maintenance robot," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 621–623.

[118] R. A. Brooks, "A hardware retargetable distributed layered architecture for mobile robot control," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 106–110.

[119] ____, "A robust layered control system for a mobile robot," *IEEE J. Robot. Automat.*, vol. 2, no. 1, pp. 14–23, 1986.

[120] J. L. Crowley, "Path planning and obstacle avoidance," in *Encyclopedia of Artificial Intelligence*. New York: Wiley, 1986, pp. 708–715.

[121] S. Chen, "Multisensor fusion and navigation of mobile robots," *Int. J. Intelligent Syst.*, vol. 2, no. 2, pp. 227–251, 1987.

[122] ____, "A geometric approach to multisensor fusion," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 22–25.

[123] ____, "A geometric approach to multisensor fusion and spatial reasoning," in *Proc. Workshop on Spatial Reasoning and Multi-Sensor Fusion*, St. Charles, IL, Oct. 1987, pp. 201–210.

[124] ____, "Adaptive control of multisensor systems," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 98–102.

[125] A. Elfes, "Sonar-based real-world mapping and navigation," *IEEE J. Robot. Automat.*, vol. RA-3, no. 3, pp. 249–265, 1987.

[126] ____, "A sonar-based mapping and navigation system," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1151–1156.

[127] H. P. Moravec, and A. Elfes, "High resolution maps from wide angle sonar," in *Proc. IEEE Int. Conf. Robotics and Automat.*, St. Louis, MO, Mar. 1985, pp. 116–121.

[128] L. Matthies and A. Elfes, "Integration of sonar range data using a grid-based representation," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 727–733.

[129] H. P. Moravec, "Sensor fusion in certainty grids for mobile robots," *AI Mag.*, vol. 9, no. 2, pp. 61–74, 1988.

[130] C. C. Jorgensen, "Neural network representation of sensor graphs in autonomous robot path planning," in *Proc. IEEE 1st Int. Conf. Neural Nets*, San Diego, CA, June 1987, pp. IV-507–IV-515.

[131] A. C. Kak, B. A. Roberts, K. M. Andress, and R. L. Cromwell, "Experiments in the integration of world knowledge with sensory information for mobile robots," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 734–740.

[132] K. M. Andress and A. C. Kak, "Evidence accumulation & flow of control in a hierarchical spatial reasoning system," *AI Mag.*, vol. 9, no. 2, pp. 75–94, 1988.

[133] R. A. Brooks, "Visual map making for a mobile robot," in *Proc. IEEE Int. Conf. Robotics and Automat.*, St. Louis, MO, Mar. 1985, pp. 824–829.

[134] M. Asada, "Building a 3-D world model for a mobile robot from sensory data," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 918–923.

[135] D. Miller, "A spatial representation system for mobile robots," in *Proc. IEEE Int. Conf. Robotics and Automat.*, St. Louis, MO, Mar. 1985, pp. 122–127.

[136] D. T. Lawton, T. S. Levitt, C. McConnell, and J. Glicksman, "Terrain models for an autonomous land vehicle," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 2043–2051.

[137] C. Isik and A. Meystel, "Decision making at a level of a hierarchical control for unmanned robots," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1772–1778.

[138] S. Harmon, "Robots, mobile," in *Encyclopedia of Artificial Intelligence*. New York: Wiley, 1986, pp. 957–963.

[139] ____, "A report on the NATO workshop on mobile robot implementation," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 604–610.

[140] N. J. Nilsson, "A mobile automation: an application of artificial intelligence techniques," in *Proc. 1st Int. Joint Conf. Artificial Intell.*, Washington, DC, May 1969, pp. 509–520.

[141] M. H. Smith *et al.*, "The design of JASON, a computer controlled mobile robot," in *Proc. Int. Conf. Cybern. Society*, New York, NY, Sept. 1975, pp. 72–75.

[142] H. Moravec, "The Stanford Cart and the CMU Rover," *Proc. IEEE*, vol. 71, no. 7, pp. 872–884, 1983.

[143] J. P. Bixler and D. P. Miller, "A sensory input system for autonomous mobile robots," in *Proc. Workshop Spatial Reasoning and Multi-Sensor Fusion*, St. Charles, IL, Oct. 1987, pp. 211–219.

[144] H. R. Everett, "A second-generation autonomous sentry robot," *Robotics Age*, pp. 29–32, Apr. 1985.

[145] A. M. Flynn, "Combining sonar and infrared sensors for mobile robot navigation," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 5–14, 1988.

[146] R. Berry, K. Loebbaka, and E. Hall, "Sensors for mobile robots," in *Proc. 3rd Conf. Robot Vision and Sensory Control*, Cambridge, MA, Nov. 1983, pp. 584–588.

[147] R. Wallace *et al.*, "First results in robot road-following," in *Proc. 9th Int. Joint Conf. Artificial Intell.*, Los Angeles, CA, Aug. 1985, pp. 1089–1095.

[148] J. Miller, "A discrete adaptive guidance system for a roving vehicle," in *Proc. IEEE Conf. Decision and Contr.*, New Orleans, LA, Dec. 1977, pp. 566–575.

[149] G. Giralt, R. Sobek, and R. Chatila, "A multi-level planning and navigation system for a mobile robot: A first approach to HILARE," in *Proc. 6th Int. Joint Conf. Artificial Intell.*, Tokyo, Japan, Aug. 1979, pp. 335–337.

[150] M. Ferrer, M. Briot, and J. C. Talou, "Study of a video image treatment system for the mobile robot HILARE," in *Proc. 1st Conf. Robot Vision and Sensory Contrl*, Stratford-upon-Avon, UK, Apr. 1981.

[151] A. R. de Saint Vincent, "A 3D perception system for the mobile robot HILARE," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1105–1111.

[152] M. Briot, J. C. Talou, and G. Bauzil, "The multi-sensors which help a mobile robot find its place," *Sensor Rev.*, vol. 1, no. 1, pp. 15–19, 1981.

[153] G. Bauzil, M. Briot, and P. Ribes, "A navigation sub-system using ultrasonic sensors for the mobile robot HILARE," in *Proc. 1st Conf. Robot Vision and Sensory Contr.*, Stratford-upon-Avon, UK, Apr. 1981.

[154] R. Chatila and J. Laumond, "Position referencing and consistent world modeling for mobile robots," in *Proc. IEEE Int. Conf. Robotics and Automat.*, St. Louis, MO, Mar. 1985, pp. 138–145.

[155] J. L. Crowley, "Navigation for an intelligent mobile robot," *IEEE J. Robot. Automat.*, vol. RA-1, no. 1, pp. 31–41, 1985.

[156] ——, "Dynamic world modeling for an intelligent mobile robot," in *Proc. 7th Int. Conf. Pattern Recognition*, Montreal, PQ, Canada, July 1984, pp. 207–210.

[157] ——, "Dynamic world modeling for an intelligent mobile robot using a rotating ultrasonic ranging device," in *Proc. IEEE Int. Conf. Robotics and Automat.*, St. Louis, MO, Mar. 1985, pp. 128–135.

[158] S. Harmon, "USMC ground surveillance robot: A testbed for autonomous vehicle research," in *Proc. 4th Univ. Alabama Robotics Conf.*, Huntsville, AL, Apr. 1984.

[159] D. J. Kriegman, E. Triendl, and T. O. Binford, "A mobile robot: Sensing, planning and locomotion," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 402–408.

[160] S. A. Shafer, A. Stentz, and C. E. Thorpe, "An architecture for sensory fusion in a mobile robot," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 2002–2011.

[161] J. M. Lowrie, M. Thomas, K. Gremban, and M. Turk, "The autonomous land vehicle (ALV) preliminary road-following demonstrations," in *Proc. SPIE*, vol. 579, *Intelligent Robots and Computer Vision*, D. P. Casasent, Ed., Sept. 1985, pp. 336–350.

[162] T. A. Linden, J. P. Marsh, and D. L. Dove, "Architecture and early experience with planning for the ALV," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 2035–2042.

[163] R. T. Dunlay, "Obstacle avoidance perception processing for the autonomous land vehicle," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 912–917.

[164] A. M. Waxman *et al.*, "A visual navigation system for autonomous land vehicles," *IEEE J. Robot. Automat.*, vol. RA-3, no. 2, pp. 124–141, 1987.

[165] D. W. Payton, "An architecture for reflexive control," in *Proc. IEEE Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1838–1845.

[166] F. C. A. Groen, E. R. Komen, M. A. C. Vreeburg, and T. P. H. Warmerdam, "Multisensor robot assembly station," *Robotics*, vol. 2, pp. 205–214, 1986.

[167] D. Nitzan, C. Barrouil, P. Cheeseman, and R. Smith, "Use of sensors in robot systems" in *Proc. Int. Conf. Adv. Robotics*, Tokyo, Japan, Sept. 1983, pp. 123–132.

[168] A. Choudry, W. Duinker, L. O. Hertzberger, and F. Tuijnman, "Simulating the sensory-control of an autonomous vehicle," in *Proc. 5th Int. Conf. Robot Vision and Sensory Controls*, Amsterdam, The Netherlands, Oct. 1985, pp. 147–154.

[169] J. Mochizuki, M. Takahashi, and S. Hata, "Unpositioned workpieces handling robot with visual and force sensors," in *Proc. IEEE Int. Conf. Indust. Elect. Contr. Instrumentation*, San Francisco, CA, Nov. 1985, pp. 299–302.

[170] A. Ikonomopoulos, "Trends in computer vision research for advanced industrial automation," in *Advances in Image Processing and Pattern Recognition*, V. Cappellini and R. Marconi, Eds. New York: Elsevier, 1986.

[171] A. L. Porter and F. A. Rossini, "Robotics in the year 2000," *Robotics Today*, pp. 27–28, June 1987.

[172] J. H. Kremers *et al.*, "Development of a machine-vision based robotic arc-welding system," in *Proc. 13th Int. Symp. Industrial Robots and Robots*, Chicago, IL, Apr. 1983, pp. 14.19–14.33.

[173] M. P. Howarth and M. F. Guyote, "Eddy current and ultrasonic sensors for robot arc welding," *Sensor Rev.*, vol. 3, no. 2, pp. 90–93, 1983.

[174] R. J. Didocha, D. W. Lyons, and J. C. Thompson, "Integration of tactile sensors and machine vision for control of robotic manipulators," in *Robots 9 Conf. Proc.*, Detroit, MI, June 1987, pp. 37–71.

[175] R. C. Smith and D. Nitzan, "A modular programmable assembly station," in *Proc. 13th Int. Symp. Industrial Robots and Robots*, Chicago, IL, Apr. 1983, pp. 5.53–5.75.

[176] F. C. A. Groen, E. M. Petriu, M. A. C. Vreeburg, and T. P. H. Warmerdam, "Multi-sensor robot assembly station," in *Proc. 5th Int. Conf. Robot Vision and Sensory Contr.*, Amsterdam, The Netherlands, Oct. 1985, pp. 439–448.

[177] C. C. Ruokangas, M. S. Black, J. F. Martin, and J. S. Schoenwald, "Integration of multiple sensors to provide flexible control strategies," in *Proc. IEEE Int. Conf. Robotics and Automat.*, San Francisco, CA, Apr. 1986, pp. 1947–1953.

[178] C. A. Fowler and R. F. Nesbit, "Tactical C^3, counter C^3 and C^5," *J. Electron. Defense*, Nov. 1980.

[179] J. Franklin, L. Davis, R. Shumaker, and P. Morawski, "Military applications," in *Encyclopedia of Artificial Intelligence*. New York: Wiley, 1986, pp. 604–614.

[180] P. J. Klass, "DARPA envisions new generation of machine intelligence technology," *Aviation Wk. Space Tech.*, vol. 122, no. 16, pp. 46–84, Apr. 22, 1985.

[181] J. T. Tou and R. C. Gonzales, *Pattern Recognition Principles*. Reading, MA: Addison-Wesley, 1974, pp. 111–112.

[182] T. D. Garvey and M. A. Fischler, "Perceptual reasoning in a hostile environment," in *Proc. 1st Annu. Nat. Conf. Artificial Intell.*, Stanford Univ., CA, Aug. 1980, pp. 253–255.

[183] V. G. Comparato, "Fusion—The key to tactical mission success," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 2–7.

[184] J. R. Mayersak, "An alternate view of munition sensor fusion," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 64–73.

[185] R. P. Bonasso, Jr., "ANALYST: An expert system for processing sensor returns," The MITRE Corp., McLean, VA, Tech. Rep. MTP-83W 00002, 1984.

[186] ——, "ANALYST II: A knowledge-based intelligence support system," The MITRE Corp., McLean, VA, Tech. Rep. MTP-84W 00220, 1985.

[187] K. Krishen, R. J. P. de Figueiredo, and O. Graham, "Robotic vision/sensing for space applications," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Raleigh, NC, Mar. 1987, pp. 138–150.

[188] S. W. Shaw, R. J. P. de Figueiredo, and K. Krishen, "Fusion of radar and optical sensors for robotic vision," in *Proc. IEEE Int. Conf. Robotics and Automat.*, Philadelphia, PA, Apr. 1988, pp. 1842–1846.

[189] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Boston, MA: Academic, 1988.

[190] Y. Bar-Shalom, "Tracking methods in a multitarget environment," *IEEE Trans. Automat. Contr.*, vol. AC-23, no. 4, pp. 618–626, 1978.

[191] K. C. Chang, C. Y. Chong, and Y. Bar-Shalom, "Joint probabilistic data association in distributed sensors networks," *IEEE Trans. Automat. Contr.*, vol. AC-31, no. 10, pp. 889–897, 1986.

[192] J. F. Arnold, Y. Bar-Shalom, R. Estrada, and R. A. Mucci, "Target parameter estimation using measurements acquired with a small number of sensors," *IEEE J. Oceanic Eng.*, vol. OE-8, no. 3, pp. 163–171, 1983.

[193] R. Mucci, J. Arnold, and Y. Bar-Shalom, "Track segment association with a distributed field of sensors," *J. Acoust. Soc. Amer.*, vol. 78, no. 4, pp. 1317–1324, 1985.

[194] C. Y. Chong, S. Mori, and K. C. Chang, "Information fusion in distributed sensor networks," in *Proc. American Control Conf.*, Boston, MA, June 1985, pp. 830–835.

[195] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE J. Oceanic Eng.*, vol. OE-8, no. 3, pp. 173–184, 1983.

[196] J. L. Speyer, "Computation and transmission requirements for a decentralized Linear-Quadratic-Gaussian control problem," *IEEE Trans. Automat. Contr.*, vol. AC-24, no. 2, pp. 266–269, 1979.

[197] S. C. A. Thomopoulos, D. K. Bougoulias, and L. Zhang, "Optimal and suboptimal distributed decision fusion," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 26–30.

[198] S. C. A. Thomopoulos and N. N. Okello, "Distributed detection with consulting sensors and communication cost," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 31–40.

[199] S. C. A. Thomopoulos and L. Zhang, "Networking delay and channel errors in distributed decision fusion," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 154–160.

[200] ____, "Distributed filtering with random sampling and delay," in *Proc. SPIE*, vol. 931, *Sensor Fusion*, C. W. Weaver, Ed., Orlando, FL, Apr. 1988, pp. 161–167.

[201] K. R. Britting, *Inertial Navigation System Analysis*. New York: Wiley, 1971.

[202] W. S. Gesing and D. B. Reid, "An integrated multisensor aircraft track recovery system for remote sensing," *IEEE Trans. Automat. Contr.*, vol. AC-28, no. 3, pp. 356–363, 1983.

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

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

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

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

[207] J. A. Adam, "Star Wars in transition," *IEEE Spectrum*, vol. 26, no. 3, pp. 32–38, Mar. 1989.

[208] W. D. Hillis, *The Connection Machine*. Cambridge, MA: MIT Press, 1985.

[209] K. D. Wise, "Intelligent sensors for semiconductor process automation," in *Proc. Int. Conf. Indust. Elect. Contr. Instrumentation*, Milwaukee, WI, Sept. 1986, pp. 213–217.

[210] R. N. Stauffer, "Integrated sensors extend robotic system intelligence," *Robotics Today*, Aug. 1987.

[211] R. J. Offen, *VLSI Image Processing*. New York: McGraw-Hill, 1985.

[212] S. Middelhoek and A. C. Hoogerwerf, "Smart sensors: When and where?" *Sensors and Actuators*, vol. 8, pp. 39–48, 1985.

[213] NATO Advanced Research Workshop on Highly Redundant Sensing in Robotic Systems, Il Ciocco, Italy, May 1988.

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

Ren C. Luo (M'83–SM'88) received the B.S. degree in mechanical engineering and the M.S. degree in electrical engineering both from Feng Chia University, Taiwan, Republic of China, in 1973 and 1975, and the M.S. and Ph.D. degrees in electrical engineering both from Technische Universitat Berlin, Berlin, Germany, in 1980 and 1982, respectively.

From 1974 to 1977 he was employed in industry and served as a Chief Engineer and Research Engineer in Victor, Inc., Taiwan, and Waldrich Siegen GmbH, Seigen, West Germany. From 1977 to 1983 he served on the scientific research staff in the Fraunhofer Institute fur Produktionstechnik and Automatisierung, Division of Industrial Robots and in the Institute fur Mess- und Regelungstechnik in the area of intelligent contact and no-contact sensors in Berlin, West Germany. From 1982 to 1983 he was a Research Project Leader in sensor-based robotics at the NSF-sponsored Robotics Research Center of the University of Rhode Island, Kingston. From 1983 to 1984 he was an Assistant Professor in the Department of Electrical Engineering and Computer Science at the University of Illinois, Chicago. Since September 1984 he has been with the Department of Electrical and Computer Engineering at North Carolina State University, Raleigh, where he is currently an Associate Professor. He is the author and coauthor of a number of technical publications in refereed journals, book chapters, national, and international conference proceedings. These articles deal with a variety of sensor design, sensor data processing, computer vision based 3-D object recognition, and multisensor fusion and integration applied to intelligent robotics systems.

Dr. Luo is a member of Verein Deutsche Ingineure (VDI), West Germany, and SPE and a senior member of RI/SME. He is also the Technical Committee Chairman for transducers and actuators at the IEEE Industrial Electronics Society and Associate Editor of IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS. He is listed in *Who's Who in the World* and *Who's Who in Frontiers of Science and Technology*.

Michael G. Kay received the B.A. degree in economics and the M.S. degree in industrial and systems engineering both from the University of Florida, Gainesville, in 1981 and 1984, respectively. He is currently pursuing the Ph.D. degree in industrial engineering at North Carolina State University, Raleigh.

He has worked as a Consultant to municipal utilities for rate making and planning purposes.