Sensor Fusion: From Cognitive Theory to Robotic Architecture

Robin R. Murphy
School of Mathematical and Computer Sciences
Colorado School of Mines
Golden, CO 80401
rmurphy@mines.colorado.edu

Introduction

Our research has focused on *intelligent sensor fusion* for autonomous mobile robots, which we consider to be a process with three abilities: the ability to combine the perceptual information extracted from multiple sensors, the ability to adapt the sensor fusion process to major environmental changes and sensor malfunctions, and the ability to develop the sensing plans and objectives for a particular sensor fusion process.

Our approach to sensor fusion has been to construct a broad working theory specifying the types of knowledge and the operations involved in fusion [7]. This particular approach is unique in that it has relied on insights from cognitive and behavioral psychology. However, there are other alternative approaches to autonomous sensor fusion; the reader is directed to [2,3,9,11].

This research provides a case study of some of the issues involved in transferring cognitive theories to robotics. There are several advantages to considering cognitive and behavioral models as a basis for a robotic theory. First, they offer existence proofs that a particular activity does occur. Further, these models abstract the issues from implementation details, allowing a roboticist to work on the "big picture". And finally, they may lead to a non-intuitive solution or approach to a problem.

However, there are many problems incumbent in trying to exploit the cognitive plausible models offered by psychology. The difference in implementations (e.g. neurons vs. silicon) is frequently cited. A more subtle problem is that the strategies employed by biological systems, being directed at survival of the species rather than the individual, are not necessarily robust enough for robotics. A robotic theory, on the other hand, does not have to be constrained by the same concerns, i.e., it does not have to be cognitive plausible. Therefore it can address any shortcomings in cognitive models by identifying the needs of robots in a particular area and then adapting the cognitive theory. It should be noted that having a robotic theory does not guarantee that the implementation on a robot will be straightforward. Our robotic theory of sensor fusion could not be mapped directly to a robotic architecture without compromises because the technology was not available to support the types of planning and reasoning activities outlined in the theory.

This paper summarizes the insights from the cognitive and behavior literature in sensor fusion and how they formed the basis for a robotic theory. Experiments previously reported in [6,8,7], and briefly described here, confirm the utility of *fusion states*, which were the central contribution from psychology. The rest of the paper presents the work in progress in implementing the theory as the Sensor Fusion Effects (SFX) architecture.

This work was supported in part by a Rockwell Doctoral Fellowship, the Georgia Tech CIMS Program, and by the Georgia Tech Material Handling Research Center.

Cognitive Theory to Robotic Theory

The venue of the cognitive and behavioral sciences include sensor fusion in animals (termed multi-sensory integration). Studies from these disciplines do not specify the particular mechanisms for sensor fusion, but do offer insights into the control of the activities involved in sensor fusion. These insights suggest that biological sensor fusion proceeds along a well structured control scheme consisting of a few mechanisms intended for classes delineated by the interaction between sensors, the performance of these sensors in the environment, and the task.

Insights from Cognitive and Behavioral Sciences

The cognitive and behavioral literature suggests four useful insights for the design of robotic sensor fusion systems. First, Marks demonstrates a commonality among the senses and sensory integration [5], the biological equivalent of sensors and sensor fusion. His work argues that sensor fusion is implemented as a small set of general mechanisms operating independently of the contributing sensor modalities. Second, the basic orienting behavior postulated by Lee [4] separates the configuration and execution activities of a perceptual process from its actual performance of perception. Third, the choice of fusion mechanism is based on how the perceptual input is being used to accomplish a task, as illustrated by the work of Pick and Saltzman with perceptual modes [10].

The most important insight into sensor fusion is Bower's speculation [1] that the perceptual mode is defined by the types of discordances expected to occur between the sensors. His taxonomy of levels of sensor fusion can be mapped onto three states of robotic sensor configurations:

$State_1$: *complete sensor fusion*
All features are included in the fusion process, because no discordance is expected.

$State_2$: *fusion with the possibility of discordance and resultant recalibration of dependent perceptual sources*
In this state, certain features are known to be subject to systematic errors, such as slippage in shaft encoders. Rather than tolerate an discordant observation, the perceptual process uses feedback to recalibrate or retune parameters.

$State_3$: *fusion with possibility of discordance and resultant suppression of discordant perceptual sources*
The state, like $State_2$, permits a set of previously identified features to be temporarily removed from the fusion step. This state handles sensors, such as ultrasonics, which exhibit spurious readings over time. Feedback is used to discount or suppress these readings.

A fusion state has two parts: a *feedback rule*, which responds to an expected discordance in one of the methods above, and *a set of state failure conditions*, which determine when a discordance cannot be accommodated by feedback and triggers exception handling. The concept of fusion state differs from typical uncertainty management systems for fusion because it enables the fusion process to categorize and explicitly accommodate the types of discordances that can occur between features. The states monitor for potential discordances; remove or mitigate their impact on the fused belief; and provide feedback to deficient sensors or algorithms. States have the potential for making the process computationally tractable, improving the final belief in the fusion output, and for improving sensing quality.

Shortcomings of Cognitive Theory

While offering valuable, non-intuitive insights for sensor fusion, the cognitive theory falls short for what is needed for a robotic theory. In particular, the cognitive theory does not promote robustness for an autonomous mobile robot. What happens when the assumptions which led to a sensor fusion state for a particular task are invalidated? The cognitive theory suggests that the fusion continues on, erroneously, in the same state. This may be reasonable for an animal where the assumptions leading to the choice of state are sufficient for the survival of the species at the cost of the individual; but it is not a desirable response for a one-of-a-kind Mars Rover. Therefore a robotic theory must include some mechanism for monitoring for sensor malfunctions or unanticipated changes in the sensing environment (exception detection) and provide some type of response (exception handling).

Resulting Robotic Theory

The robotic theory is based on a perceptual process which is responsible for providing the perception needed to accomplish a task. The heart of the perceptual process is a state-based control scheme guiding the sequencing and execution of all perceptual activities.

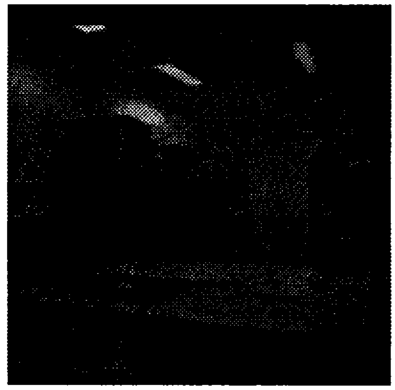
According to this theory, the perceptual process is an autonomous agent. It begins in an initial *investigatory phase* where it configures the sensing plan, using the current status of the sensor suite, and *a priori* knowledge about what are the relevant, action-oriented descriptions of the percept. It also determines the appropriate fusion state.

After configuration is complete, the perceptual process enters the *performatory phase* where it collects observations according to the sensing plan, preprocesses them, and fuses them into a value for the percept with an associated measure of uncertainty. The perceptual process applies feedback according to the state-specific mechanism. This feedback is represented as a *feedback rule*. When a *belief criterion* is satisfied, the perceptual process performs the corresponding *feedback activity*. The perceptual process remains in the performatory phase repeating the cycle of collection, preprocessing, and fusion until either the motor behavior terminates eliminating the need for the process, or it detects a state failure. If a state failure occurs, the perceptual process interrupts the performatory phase and reinvokes the investigatory phase in order to reconfigure the sensing plan.

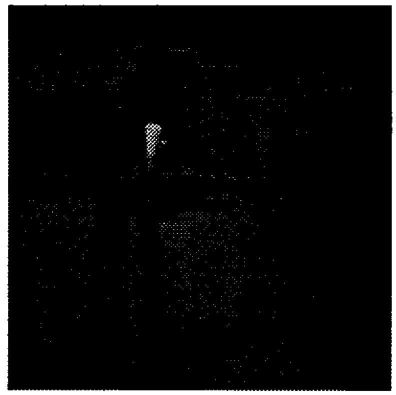
Experiments with Robotic Theory

Experiments demonstrating the theory have been conducted in the security guard domain, where the Georgia Tech mobile robot determined if three different areas in a cluttered tool room remained unchanged after each visit. The robot positioned itself at roughly the same location and viewing angle each scene during every visit to the tool room. These experiments are detailed in [6,8]. Sensor data for each scene was collected from a Sony Hi8 color camcorder, a Pulnix black and white camer, an Inframetrics true infrared camera, and Polaroid ultrasonic transducers, and Hamamatsu Ultraviolet light camera. Each scene was modeled manually beforehand using three representative sensor data sets of each scene.

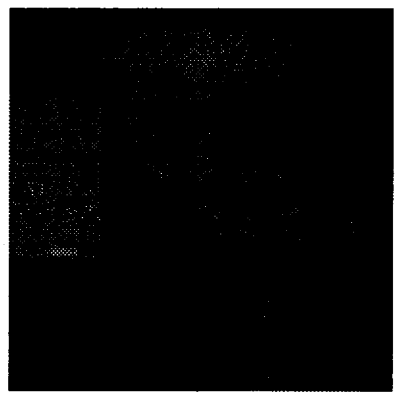
These experiments were primarily concerned with demonstrating the operation of the three fusion states and detection of state failures. The C code combined observations from each sensor using a Dempster-Shafer evidential system, detected state failures which were defined as thresholds on the evidence, and performed feedback according to the feedback rule which implemented both the belief criteria and feedback activity as functions. Qualitative and quantitative features were used as input.



a.



b.



c.

Figure 1: Three scenes used in experiments.
a. The **student desks scene.**
b. The **drill press scene.**
c. The **vcr and monitor scene.**

$State_1$

The **student desk scene** shown in Fig.1a. was used to illustrate the operation of $State_1$, complete sensor fusion. The perceptual process correctly confirmed that the scene was either changed or unchanged in each trial. The model for this scene consisted of four descriptions, one for each sensor. The description of the scene used by the Hi8 camcorder was based on two features: a prominent blue equipment cabinet and the color histogram of the image. The description of the scene used by the Pulnix camera clustered the 30 most significant interest points, identified by an interest operator, into three groups using a clustering algorithm. The centers of each of the groups formed a *constellation* consisting of the topological relation and the distance between pairs of the centers. The relative distance to the nearest surface was the description used by the ultrasonic transducers, and the thermal description of the scene was ambient temperature. $State_1$ was chosen for the fusion state because there was no indication of any significant relationships between the descriptions for this scene.

$State_2$

Figure 1b. shows the **drill press scene**, which was used to demonstrate $State_2$, recalibration of discordant sensing. The scene was described for the Pulnix black and white camera in terms of the relationship between three regions. The centroid of pixels in each of these three regions formed a constellation. The thermal image was similarly modeled as the relationship between two thermal regions. The distance to the nearest surface was used for the ultrasonic description.

The perceptual process for this configuration was assigned to $State_2$ because it was known that the Inframetrics thermal camera drifts. The camera uses liquid nitrogen as a reference temperature and after 1 to 2 hours, the nitrogen completely evaporates. When that happens the values of the image are shifted and and the histogram compressed. Fortunately, the information can be recovered with a simple recalibration routine (i.e., the feedback activity). The question becomes under what conditions can it be recalibrated (i.e., the belief criteria). If the thermal image was continuously recalibrated without considering the observations of the other sensors, it could recalibrate around an intruder's heat signature! Therefore the belief criteria was to recalibrate when the other sensors showed a strong consensus that the scene was unchanged.

Experiments with the **drill press scene** showed that it correctly determined when the belief criteria was met and applied the recalibration feedback. If the perceptual process had been "state-less", it would have posted an incorrect interpretation of the evidence based on a lack of consensus; whereas after recalibration, the thermal observation was able to agree with the others and correctly confirm that the scene had not changed.

$State_3$

Figure 1c. shows the **vcr and monitor scene**, which was used to demonstrate how $State_3$, suppression of discordant sensing, can be used to improve the total belief in the fused percept. The description of the scene used by the Pulnix camera extracted the upper right hand corners of both the vcr and the monitor and noted that the topological relation between the corners (above and to the left). The description used by the infrared camera was based on the topological relationship between two segmented thermal regions. The nearest surface description was again used by the ultrasonic sensors.

During the collection of the input data for constructing the model, it was noted that the ultrasonic sensors gave erratic readings. If the sensors were configured as $State_1$, complete sensor fusion, the occasional erratic reading would cause a significant conflict between the sensors, triggering a failure even though the scene had not changed. However, because it was known that the ultrasonics are behaving erratically, the fusion process was configured as $State_3$, suppression of discordant sensing, with the selection of an appropriate feedback rule.

The state feedback rule chosen for this configuration suppresses the belief contributed by the ultrasonics when it disagrees with the other sensors. The feedback activity replaced the belief contributed by the ultrasonics with complete uncertainty. This configuration was able to correctly determine when the scene was unchanged, while a "state-less" fusion was not be able to generate a belief about the scene due to the high conflict caused by the erroneous ultrasonic observation.

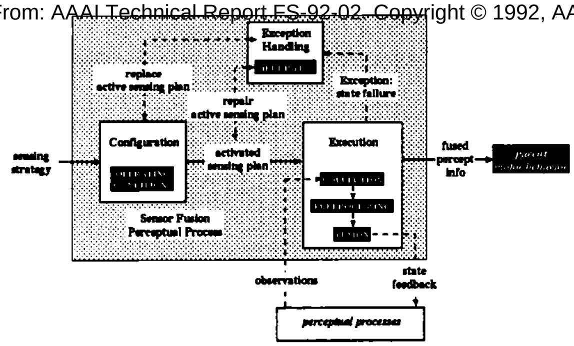


Figure 2: Software layout of the SFX architecture.

Detection of State Failures

State failures may result from high conflict between sensor observations or by missing or highly uncertain evidence. Experiments showed that the state-base control scheme could detect when the sensor observations disagreed due to sensor malfunctions or changes in the scene that only was detected by a single sensor, and disagreements due to environmental failures (e.g. lights turned off).

Work in Progress: Robotic Theory to Robotic Architecture

The above experiments serve as a proof of concept for the theory, but do not necessarily aid a roboticist in applying the theory to a particular domain. Therefore, there needs to be an sensor fusion architecture (a commitment to a set of implementation details) and a methodology for the designer to follow in applying that implementation to the problem. It should be noted that an architecture is one way of interpreting the theory, but not the only way. Also, because an architecture is concerned with pragmatics such as reliability, computational speed, and software maintainability, it may have to compromise the theory.

Our current work is concentrating on developing and testing the Sensor Fusion Effects (SFX) architecture. One goal of SFX is that it can be readily customized for an application. The software layout for a perceptual process under SFX is shown in Figure 2. The uncertainty management and exception handling modules are generic; they operate according to the sensing strategy which is supplied by the designer.

One challenge of developing an architecture is that it must try to be consistent with theory despite gaps in available technology. Two particularly acute gaps are *planning*, which is used for the configuration of the sensing strategy and by exception handling to respond to a state failure, and *evidential reasoning* for computing and propagating the uncertainty in the sensor observations.

To cope with the lack of a suitable planner, the SFX architecture relies on the designer to provide the perceptual process with a set of sensing plans in advance. The perceptual process uses a finite automaton to select the appropriate sensing plan for the conditions at runtime. The set of plans and transition relations, taken together form the sensing strategy.

The sensing strategy can also be used by the exception handling module to respond to a state failure by substituting one of the other plans. If there is no suitable substitute, the exception handling uses general heuristics to attempt to add or eliminate sensors from the sensing plan, at the cost of reducing certainty.

To cope with gaps in evidential reasoning techniques for sensor fusion, SFX uses Dempster-Shafer theory which allows the designer to embed domain knowledge into a set of rules for transferring evidence of features into evidence for a percept.

Conclusions

Based on our experience with sensor fusion, we believe that there are significant advantages to considering cognitive and behavioral models in formulating robotic theories. However, psychological models may not map directly to robotics, so these models will likely have to be recast in order to incorporate the explicit needs of a robot in a particular application.

References

[1] Bower,T.G.R. , "The Evolution of Sensory Systems," *Perception: Essays in Honor of James J. Gibson*, ed. by R.B. MacLeod and H.L. Pick, Jr., Cornell University Press, Ithaca, NY, 1974, pp. 141-153.

[2] Durrant-Whyte, H.F., *Integration, Coordination, and Control of Multi-Sensor Robot Systems*, Kluwer Academic Publishers, Boston, 1988.

[3] Hager, G.D., *Task-Directed Sensor Fusion and Planning*, Kluwer Academic Publishers, Norwell, MA 1990.

[4] Lee, D., "The Functions of Vision", *Modes of Perceiving and Processing Information*, ed. by H.L. Pick, Jr., and E. Saltzman, John Wiley and Sons, NY, 1978, pp. 159-170.

[5] Marks, L. E., *The Unity of the Senses: Interrelations among the Modalities*, Academic Press, New York, 1978.

[6] Murphy, R.R., and Arkin, R.C., "SFX: An Architecture for Action-Oriented Sensor Fusion", to appear in proceedings *IROS 1992*.

[7] Murphy, R.R., "An Architecture for Intelligent Robotic Sensor Fusion", *PhD thesis*, College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280.

[8] Murphy, R.R., "State Based Sensor Fusion for Surveillance", *Proceedings SPIE Sensor Fusion IV*, Boston, MA, Nov. 1991.

[9] Pau, L.F., "Behavioral Knowledge in Sensor/Data Fusion Systems", *Journal of Robotic Systems*, vol. 7, no. 3, 1990, pp. 295-372.

[10] Pick, H.L, and Saltzman,E., "Modes of Perceiving and Processing Information," *Modes of Perceiving and Processing Information*, ed. by H.L. Pick, Jr., and E. Saltzman, John Wiley and Sons, NY, 1978, pp. 1-20.

[11] Shafer, S., Stentz, A., and Thorpe, C.E., "An Architecture for Sensor Fusion in a Mobile Robot", proceedings *IEEE International Conference on Robotics and Automation*, San Francisco, 1986, pp. 2002-2010.