Dempster-Shafer Theory for Sensor Fusion in Autonomous Mobile Robots

Robin R. Murphy

Department of Mathematical and Computer Sciences
Colorado School of Mines
Golden, CO 80404-1887
rmurphy@mines.colorado.edu

Abstract

This article presents the uncertainty management system used for the execution activity of the Sensor Fusion Effects (SFX) architecture. The SFX architecture is a generic sensor fusion system for autonomous mobile robots, suitable for a wide variety of sensors and environments. The execution activity uses the belief generated for a percept to either proceed with a task safely (e.g., navigate to a specific location), terminate the task (e.g., can't recognize the location), or investigate the situation further in the hopes of obtaining sufficient belief (e.g., what has changed?). Dempster-Shafer (DS) theory serves as the foundation for uncertainty management. The SFX implementation of DS theory incorporates evidence from sensor observations and domain knowledge into three levels of perceptual abstraction. It also makes use of the DS weight of conflict metric to prevent the robot from acting on faulty observations. Experiments with four types of sensor data collected by a mobile robot are presented and discussed.

1 Introduction

Autonomous mobile robots have accomplished basic tasks in navigation, such as moving to a goal, avoiding obstacles, and docking at a workstation, using perception from one sensor. However, these single sensor perceptual systems have not been entirely successful for more demanding tasks in navigation [36], target or goal recognition [9,17,24,35,53], and general scene interpretation [10,30]. This has limited the potential benefits of mobile robots for applications in space, defense, and manufacturing.

Perceptual systems based on a single sensor have an inherent weakness: they generally cannot reduce uncertainty. Sensor uncertainty [51], as distinguished from imprecision, largely depends on what is being observed rather than the camera. Uncertainty arises when features are missing (e.g., occlusions), when the sensor cannot measure all relevant attributes of the percept (e.g., a video camera cannot measure thermal radiation), and when the observation is ambiguous (e.g., an edge detected by a camera may be a part of a desk or the artifact of a shadow). Active perception techniques [3,4,5,6,54], where the robot tries to get a "better look," can sometimes compensate for missing observations. But a different view may not make up for observations which are inherently incomplete or ambiguous.

To overcome these problems, some researchers have proposed turning to perceptual systems which rely on multiple general purpose sensors. These systems would combine the observations from each sensor and produce a single *percept*, or coherent perception of an object, scene, or event, through a process commonly referred to as *sensor fusion* at the symbolic level [1]. In addition to reducing uncertainty, a sensor fusion system can be expected to provide less costly perception because of the potential for distributing the demands across processors dedicated to each sensor [34].

While the potential benefits of sensor fusion have motivated much research, no general purpose method for accomplishing sensor fusion has emerged. One reason is because the sensors' output may have very little in common. For example, they may offer different resolutions of data and have little or no correspondence, all of which is exacerbated by sensor noise. Consider the output shown in Figures 1 and 2 from four different sensors, a Sony Hi8 color video camcorder, a Pulnix b&w video camera, an Inframetrics infra-red camera (thermal), and Polaroid ultrasonic transducers (range). First, note that the output from the color and b&w cameras each have different fields of view and resolutions. These two cameras measure the same modality (visible light), yet the graduate student figures prominently in one image but not in the other. More importantly, the desk in the foreground of the b&w image doesn't appear in the color and thermal images, the only item that appears in all three is the student. The range reading in Fig. 2b. is the distance to the desk which is seen only by the b&w camera. There is no feature common to each of these observations, yet the sensors are observing the same region. It should also be noted that the image from the thermal camera (Fig.2a.) exhibits bands of digitization noise, which may further complicate the fusion process.

As a result, many sensor fusion systems, for both object recognition and autonomous mobile robots, treat sensor observations as *evidence* and use evidential reasoning techniques to infer the percept. The majority of these systems represent the sensor evidence *probabilistically* and use Bayes Rule to infer the percept. A significant portion are *possibilistic*; they consider sensor evidence to be *belief* and rely on a Dempster-Shafer Theory framework [48]. Other approaches include fuzzy logic and other logics. However, each of these implementations are difficult to adapt to new sensing configurations, and/or was unable to detect that one or more the sensors was providing suspect or unreasonable observations (e.g., sensor has malfunctioned).

This article describes how the Sensor Fusion Effects (SFX) architecture overcomes these shortcomings within a Dempster-Shafer (DS) framework. DS theory is suitably expressive; it is able to explicitly represent ignorance, permitting the robot to differentiate between ambiguious sensing results and not have sensed at all. The *weight of conflict* metric inherent in DS theory is used to measure the amount of consensus between different sensors. A lack of consensus leads the robot to either compensate within certain limits or investigate the problem further, adding robustness to the robot's operation. The modular decomposition of evidence into three frames of discernment (feature, description, percept) allows the dynamic substitution of sensors and algorithms in response to sensor failures. This decomposition

a.

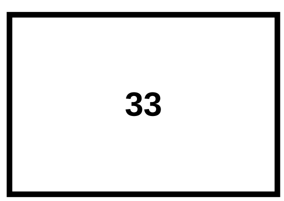
b.

Figure 1: A student desk area as seen by

a.) a Sony Hi8 color camcorder.

b.) a Pulnix b&w video camera.

a.



b.

Figure 2: The same student desk area as observed by

a.) an Inframetric true infra-red camera.

b.) a Polaroid ultrasonic range transducer (range is in inches).

offers the advantages of perceptual abstraction given by *logical sensors* [22] and *equivalence classes* [14]. This decomposition also permits expert knowledge about the domain to be embedded in the frames of discernment, simplifying its application and maintenance.

The paper concentrates on the principles of evidential reasoning used in SFX. The reader is directed to [37,39] for details about how the SFX system works. It is laid out as follows. First, the various evidential methods for sensor fusion are presented, emphasizing the differences between Bayesian and DS theories. Given the relative strengths and weaknesses of each and the demands of autonomous mobile robots, DS theory emerges as the most appropriate choice for this application. Next, the salient aspects of DS theory are reviewed, including ignorance, the Con weight of conflict metric, and the concept of enlargement of frames of discernment. The review is followed by a formal development of the evidential representation used by SFX. Six different demonstrations with sensor data collected from a mobile robot illustrate the flow of evidence from observations of feature into a belief in a percept. The experiments also illustrates how SFX uses evidence to control or suspend the fusion process, as well as the importance of domain knowledge in deriving the correct interpretation of evidence.

2 Evidential Methods for Sensor Fusion

Research in evidential representations for sensor fusion has been primarily concerned with developing a particular system for evidential reasoning first, then adapting the reasoning system for the requirements of sensor fusion activities and the types of evidence offered by actual sensors. Bayesian and DS theory are discussed in detail because they are established theories and have a history of being successfully used for sensor fusion.

Probability-based evidential reasoning systems represent the certainty in an attribute-value as a point probability. Bayesian theory [41] provides the foundation for most systems in the probability category where the certainty in a feature is represented as a probability function. There are a few notable exceptions which use non-Bayesian specifically for sensor fusion: Nakamura and Xu [40], Luo, *et al* [33], and Beckerman *et al* [7,8] have each used point probabilities for sensor fusion with alternate application-dependent decision rules. Bayesian theory has also been used for reasoning about the "uncertainty" in sensor observations resulting from imprecision, most notably [50].

Possibilistic, or fuzzy [55], systems have been discussed for sensor fusion [10]. Qualitative logics [21], and epistemic logics [43,44], which aim to be the superordinate formalism for all uncertainty management techniques, (both under development) have not yet been applied to sensor fusion. Boolean logic has also been implemented in sensor fusion systems [7].

DS theory is another popular technique for reasoning about sensor observations, which is sometimes classified as a probabilistic technique. It represents evidence as a Shafer belief function over $\langle 0.0, 1.0 \rangle$ for convenience, which makes the belief function appear to be a point probability. There has been a great deal of debate in the literature about the conceptual differences between Bayesian and DS theories, if the belief function is considered to be a probability function; the reader is directed to [9,25,41,46,48] for further details. However, belief functions have been shown not to be probabilities about sample spaces, despite the fact that the theory refers to them as probabilities [27]. Instead, DS theory serves as a *model for transferring belief* [49], where belief functions capture an interpretation of the evidence afforded by some observed event. Subsequent references to "probabilities" will be reserved for true statistical probabilities (i.e., Bayesian) while "belief function" will be used for the DS probabilities.

2.1 Bayesian Methods

Bayesian theory is based on the work of Thomas Bayes during the 1700's. It has been used for perceptual activities such as target recognition [35], scene interpretation [10,42], and determining the free space available for the navigation of a mobile robot [7,36]. Hager [18,19] and Durrant-Whyte [16] both use a Bayesian framework for their work in sensor fusion.

The primary advantage of Bayesian theory for sensor fusion is that it is a well developed formalism, where evidence is treated as a probability density function. A percept, p, is inferred from bodies of evidence, e_i, according to Bayes' Rule:

$$P(p|e_1, e_2, \ldots, e_n) = \frac{P(e_1, e_2, \ldots, e_n|p)P(p)}{P(e_1, e_2, \ldots, e_n)} \tag{1}$$

where $P(p|e_1, e_2, \ldots, e_n)$ is the *a posteriori* probability that p is actually present given the observed evidence, e; $P(e_1, e_2, \ldots, e_n|p)$ is the conditional probability that the evidence is observable given the presence of p; and $P(p)$ and $P(e_1, e_2, \ldots, e_n)$ are the probabilities of p and the evidence occurring (*a priori* probabilities).

The conditioning of the bodies of evidence is frequently used in conjunction with a directed, acyclic graph model of the percept [2,28,29,42]. Nodes show PART-OF relationships between the component features of a percept, and observations of these features contribute the evidence used by the inference process. These graphical representations are generally referred to as Bayesian belief nets and/or influence diagrams.

While Bayesian theory is clearly the oldest established formalism, it has many disadvantages for sensor fusion in autonomous mobile robots. First, there is no explicit representation of ignorance. The probability function cannot explicitly indicate whether an observation is missing or ambiguous.

Second, as seen by Eqn. 1, it requires that *a priori* knowledge about the occurrence of each feature. As will be seen later, this information is deemed more semantically suitable for the planning activity in SFX, which predicts what the robot is expected to perceive and what sensors are best suited for that task, rather than for execution. Rimey and Brown [42] in particular have developed a system which exploits Bayesian belief networks for predicting what information should be extracted from a scene.

The reader should also be made aware that, unlike medical domains, priors may not be available at all. Consider a rescue robot looking for a trapped miner. It has no prior probability that a miner will be in a particular area being sensed. One way to overcome this is to assume the "non-informative prior" [20] ($P(miner) = P(\neg miner) = 0.5$). However, this leads to difficulties in using evidence to control the robot's actions. Does $P(miner) = 0.5$ mean that the region is unsensed or that it has been sensed with ambiguous results and a more detailed examination with different sensors is called for?

Third, the structure of the percept model may not correspond to probabilisitic dependencies. For example, a model of a COFFEE CUP may consist of two parts, a cylinder and a handle. Finding the cylinder will increase the probability of finding the handle, so despite the two parts being structurally distinct, evidentially they are conditionally dependent. One common solution to this non-conformance of the evidence to the structural model is to assume conditional independence and tolerate the resulting inaccuracy.

A fourth disadvantage is that Bayes' rule acts as a smoothing function on multiple sources of evidence. Mechanisms for quantifying whether posterior probability reflects a consensus of the bodies of evidence or that one source of evidence for p dominated evidence to the contrary are either computationally intractable or heuristic [26]. This could cause the robot to unknowingly enter a dangerous area rather than cautiously explore it.

Contextual based interpretation of the evidence is a fifth problem. For example, in recognizing a red ball in a bin of blue balls, "red" carries more weight than "blue" since it is distinctive. However, if the red ball is in a bin of red blocks, the evidential information contributed by red is much less. This reflects a change to the conditional probabilities (or link matrices) which is not currently supported by updating rules for belief nets such as the one given by Pearl [41].

Finally, knowledge which is not probabilistic in nature is difficult to include. For example, it may be known that a particular sensor is susceptible to drift, but not in a predictable way or time. This knowledge has to be available to the uncertainty management system if it is to explain discordant sensor observations. Team decision theoretic implementations of Bayesian theory such as [16] attempt to incorporate this type of knowledge, but at the cost of restricting the reasoning to geometric features. This restriction eliminates non-geometric features such as "blue" from contributing evidence.

2.2 Analysis of Dempster-Shafer Methods

The theory of evidence proposed by Glenn Shafer [48] is an extension of the ideas developed by Dempster for use with Bayesian probabilities during the 1960's [12,13] to subjective evidence. This work is commonly referred to as either Dempster-Shafer or Shafer-Dempster Theory[1]; the term *Dempster-Shafer theory* will be used in this paper.

DS theory has been considered for a variety of perceptual activities including sensor fusion [9,10,17], scene interpretation [30], object/target recognition [23,24], and object verification [45]. The systems of [23,24,45] are particularly relevant because they view DS in terms of sensor characteristics.

The most frequently cited advantages of DS theory are that it can represent ignorance, and it does not need *a priori* probabilities. These two advantages are related. Evidence is represented as a Shafer belief function which allows any portion of the belief mass to be explicitly assigned to ignorance. Prior to the acquisition of evidence, the total belief about the world is ignorance. Therefore when evidence is acquired about the world, that evidence replaces the ignorance, and so Dempster-Shafer theory eliminates the need for *a priori* belief functions. Alternatively, Dempster-Shafer theory can be viewed as requiring a *a priori* belief function of ignorance (which Shafer calls the *vacuous* belief function [48]). The vacuous belief function acts as an identity function during the combination of belief, that is, the combination of a belief function and the vacuous belief function is the first function. Regardless of the viewpoint, DS theory does not mandate any explicit *a priori* knowledge about the world. However, as will be seen in Section 4, substantial implicit knowledge is encoded in the choice of the structure for the frames of discernment.

DS theory has other advantages for this particular application. Using the approach in Section 4, it can be used to modularize the incorporation of the influence of domain knowledge on the transfer of evidence from a feature to a percept. For example, two sensors may be able to sense the same percept, however, the believability of the results may depend on the current condition of the environment, what is known about each sensor's interaction with the

[1]Shafer merely refers to his contribution as a reinterpretation of Dempster's rule [47,48] for evidence.

other sensors aboard the robot, etc. Domain knowledge is not easily expressed as frequency-based probability density functions; however, it can usually be expressed subjectively by a knowledge engineer or learned [31]. Next, each time evidence is combined with other evidence, the amount of disagreement is measured by the Con metric alerting the robot that a sensing exception has occurred.

Some disadvantages of DS theory cited in the literature are:

1. It is computationally intractable because a belief function must distribute belief to the power set of all the features in the world [30,45,47]. However, DS theory is not alone in this regard; Bayesian formulations suffer from a similar intractability.

2. The conditioning rule, which governs the normalization of the combined evidence, leads to counterintuitive averaging of conflicting evidence [55].

These two disadvantages are not significant for SFX as will be seen in the next section.

3 Review of Salient Aspects of Dempster-Shafer Theory

A brief review of the salient aspects of DS theory may be needed at this point to aid the reader in following the development of the evidential system for SFX. A more complete exposition appears in [48]. DS theory represents the relevant characteristics of the world as a finite set of mutually exclusive propositions and assumptions called the *frame of discernment* (FOD). Traditionally, the notation of a capital Greek letter (e.g., Θ) is used for both the frame of discernment and the set of propositions within any FOD. Since this article discusses multiple frames of discernment with different assumptions, FOD will be used when discussing a frame of reference and the use of capital Greek letters will be reserved for the contents of a FOD, unless otherwise noted.

A belief function, Bel, distributes a quantum of belief among the 2^{Θ} subsets of a FOD. It is generally written as $Bel|2^{\Theta}$ when the frame of discernment is not obvious (the $|$ does not mean "conditional" in the sense of a conditional probability, rather that the belief is over the elements in the set 2^{Θ}).

The actual distribution of belief mass among the subsets is by a function $m : 2^{\Theta} \rightarrow [0,1]$. $m(A \subset \Theta)$ is called A's *basic belief number* or the mass of A. The total belief contributed by Bel must be equal to 1, therefore: $\sum_i m(A_i \subset \Theta) = 1$. The total belief committed to a particular subset A is the sum of the belief committed to all proper subsets B of A: $Bel(A) = \sum_{B \subset A} m(B)$. Any subset A of Θ with $m(A) > 0.0$ for a particular belief function is called a *focal element* of that function.

3.1 Dempster's Rule of Combination

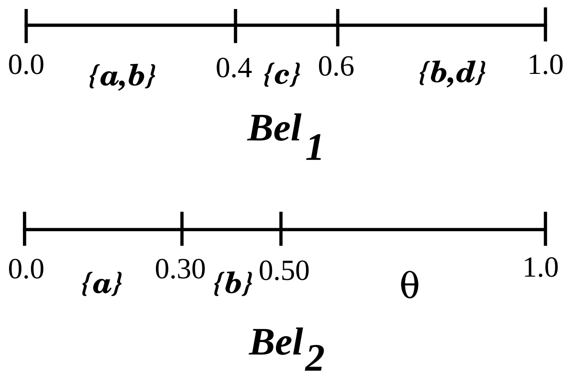
Any two belief functions A and B over the same FOD with at least one focal element in common may be combined into a new belief function over that FOD using Dempster's rule of combination. Dempster's rule specifies the combined belief mass assigned to each C_k, where C is the set of all subsets produced by $A \cap B$. The rule is:

$$m(C_k) = \frac{\sum_{A_i \cap B_j = C_k; C_k \neq \emptyset} m(A_i)m(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m(A_i)m(B_j)} \tag{2}$$

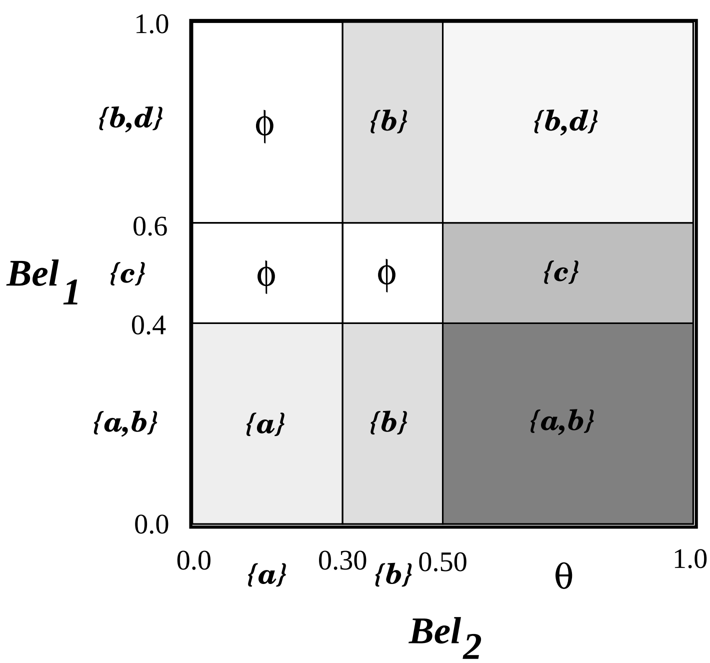
where the focal elements of $Bel_1{=}A = \{A_1, \ldots, A_i\}$ and $Bel_2{=}B = \{B_1, \ldots, B_j\}$. The combination of two belief functions is also known as taking the *orthogonal sum*, $\oplus$, and is written as: $Bel_3 = Bel_1 \oplus Bel_2 = (m(C_1), \ldots, m(C_k))$.

Conceptually, this can be represented by a square. The total belief of two belief functions must occupy an area of 1^2. Bel_1 contributes one dimension, Bel_2 the other. Since $\sum m(A_i) = 1$, a belief function can be represented as a line of length 1, with subsegments representing the belief of each subset as shown in Figure 3a. The basic belief assignment "lines" of both functions serve to partition the square. Each partition of size $m_1(A_i)m_2(B_j)$ represents the amount of belief (as a belief mass) associated with $A_i \cap B_j$.

Consider the combination of two belief functions shown in Figure 3b. Bel_1 is focused on ($\{a, b\}$, $\{c\}$, and $\{b, d\}$), with belief assignments of m_1= $(0.40, 0.20, 0.40)$ respectively. Bel_2



a.



b.

Figure 3: Example of two belief functions

a.) represented as lines.

b.) combined, represented as a square.

is focused on $(\{a\}, \{b\}, \Theta)$, with m_2=$(0.30, 0.20, 0.50)$ respectively. Note that one partition contains mass for $\{a\}$, and is of area 0.12 ($m_1(\{a,b\}) = 0.40 \times m_2(\{a\}) = 0.30$). Three partitions resulted from Bel_1 specifying where some of the uncommitted belief (Θ) from Bel_2 should be allocated; these are $\{a,b\}$, $\{c\}$, and $\{b,d\}$. Two partitions allocated belief mass to $\{b\}$. According to the numerator of Dempster's rule, these two areas would be summed together when computing $m(\{b\})$:

$$\sum_{A_i \cap B_j = b} m_1(A_i) m_2(B_j) \tag{3}$$

which can be generalized by $\{b\} = C_k$ to form the numerator of Dempster's rule. Note that three partitions are $\emptyset$. But, by definition, belief cannot be assigned to $\emptyset$. Therefore the square does not accurately represent the distribution of the belief mass. Conceptually, the mass associated with the area of those $\emptyset$ partitions has been evenly moved to the non-null partitions. The sum of Equation (3) must be *renormalized* by: $1 - \sum_{A_i \cap B_j = \emptyset} m(A_i) m(B_j)$.

which forms the denominator of Dempster's rule. The renormalization allows the belief "area" of the two functions to be projected back onto a new belief "line". In terms of sensing, the need to renormalize may arise from noise in the belief functions (e.g., sensors aren't perfect) and from sensor limitations (e.g., all sensors don't observe the same things).

As Zadeh [15,55] has pointed out, renormalization in extreme cases may lead to counter-intuitive results. Consider the case shown in Figure 4 where a feature must be either a, b, or c. Sensor 1 provides an observation which gives strong evidence for feature a and some evidence of feature c: If Bel_1 ($m(\{a\}) = 0.80$,$m(\{c\}) = 0.20$) is combined with contradictory evidence, Bel_2 ($m(\{b\}) = 0.80$,$m(\{c\}) = 0.20$), the result is total support for c. Smets [49] argues that this is actually reasonable: if the feature must be either a, b or c, then Bel_1 and Bel_2 eliminate a and b by cancelling each other, leaving the only possible conclusion that the feature must be c.

Renormalization of contradictory evidence may produce a justifiable measure of evidence, but a robot needs to be aware of such discordances. Instead, the renormalization term can be viewed as a measure of the conflict between the pair of belief functions. The larger the

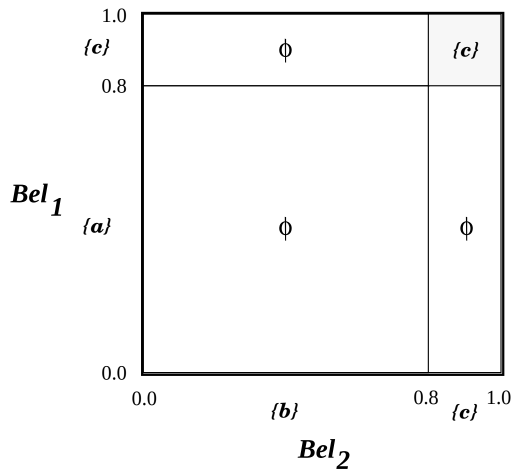


Figure 4: Example of the combination of contradictory belief functions.

area assigned to $\emptyset$, the more disagreement between the two beliefs about the FOD. Shafer defines such a measure in the *weight of conflict* metric, Con [48]:

$$Con(Bel_1, Bel_2) = \log(\frac{1}{1-\kappa});\ where\ \kappa = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) \tag{4}$$

Con takes a value between 0 and ∞; as $\kappa \to 0.0$, $Con \to 0.0$, and as $\kappa \to 1.0$, $Con \to \infty$. It is additive.

3.2 Multiple Frames of Discernment

As noted by [32,17,37], belief about the evidence may form a different belief "space" from that of the hypothesis or percept. The belief in a hypothesis is based on the belief in the observed evidence, but also influenced by knowledge about the reliability of the sensors, their behavior for the current, etc., which is difficult to represent probabilistically. DS theory permits belief to be transferred across multiple frames of discernment, for example evidence for a component to be transformed into evidence for an object. Such a mapping is called an enlargement [48]. Other more restrictive mappings are possible.

4 The SFX Evidential Representation

The Sensor Fusion Effects (SFX) architecture is a reusable generic control scheme for intelligent sensor fusion, targeted for use by autonomous mobile robots operating in unknown or partially known environments. Sensor fusion in SFX consists of three distinct activities: *planning*, *execution*, and *exception handling*. The planning activity is concerned with using the task goals of a robot to generate expectations of percepts and to predict what features of the percept will be observable to which sensors and the significance of their contribution. The execution activity collects observations and computes the total belief in the percept. If the belief for the percept is high, the robot proceeds with a behavior to accomplish a goal, if ambiguous, it persists by exploring further or adding more sensing resources. If there is high belief against the percept, the robot will reconsider its goals. The execution activity also monitors for patterns of evidence which indicate a sensing anomaly, such as a sensor malfunction. If an anomaly is detected, fusion is suspended and exception handling is invoked. For brevity, this discussion of SFX will be restricted to its management of uncertainty during the execution activity. The reader is directed to [37,39] for more specifics of the architecture.

4.1 Assumptions

The uncertainty management component of SFX is based on four assumptions about the characteristics of sensor fusion for autonomous mobile robots. The first assumption is that *in order for a robot to function robustly in an open world, it must be capable of being surprised.* For the purposes of this paper, surprise is the determination that an unexpected event with the potential for a negative impact on performance has occurred. This is different from reacting to unmodeled, but reasonable, variations in the world by a reactive motor control architecture. There are three general categories of unexpected events: plan failures, sensor malfunctions, and environmental changes. Plan failures are due to a discrepancy between the expectated state of the world and robot's actions and the observed state. An uncertainty management system which does not differentiate between expectations and observations (e.g., Bayesian) is subject to hallucinations, and could cause the robot to erroneously continue a dangerous activity. Another ramification of the need for surprise is that the uncertainty management system should support the detection of sensor malfunctions and environmental changes. SFX supports surprise by keeping predictions about what to perceive separate from observations and measures of conflict within those observations. If the belief in the expected percept is not high or the DS conflict metric exceeds a threshold, the exception handling component is alerted.

The second assumption is that *a percept can be observed in a variety of ways, each contributing a different amount of evidence depending on the circumstances.* SFX views a percept p as being composed of a family of *descriptions* where each description, d, is a collection of *features*, d_i. A description is a model of the percept for recognition by a particular sensor modality. Returning to the STUDENT DESK SCENE in Figure 1, consider that the scene could have a passive vision-based description and a thermal description. A description is abstract; different types of cameras may be able to extract the features in the vision-based description. Likewise, different descriptions can be derived from the same modality. Consider that the percept STUDENT DESK SCENE can be modeled with 2D features (a blue region, etc.) or 3D features (range profiles, generalized cylinders, etc.). A family of descriptions for a percept can be represented as: $p = \{d^1, d^2, \ldots, d^n\}$.

Each description is composed of *features* with a specified relationship to each other. A feature is the fundamental unit of symbolic perceptual information. It is defined as an attribute-value pair and is written as $\langle attribute : value \rangle$. When features are discussed in terms of their membership in a particular description, d^i, then each feature in the description is arbitrarily ordered from 1 to m as: $d^i = \{d^i_1, d^i_2, \ldots, d^i_m\}$.

The ramification of this assumption for an uncertainty management system is two-fold. The system must be able to account and propagate evidence across these different frames of references. And the propagation is most likely to involve domain knowledge or heuristics.

The third assumption is that *perception of an object (or event or scene) is sufficiently represented by belief for the object or belief against the object.* This assumption is based on the principles of action-oriented sensor fusion given in [37], which note that biological systems use expectations to filter irrelevant perception and avoid computational overloading. The

SFX system employs a perceptual process to gather evidence to support or deny the expected object, event, or scene. However, the perceptual process does not attempt to label the source of evidence if it against the expected percept. If the robot expects to see a desk where an equipment cabinet is, the evidence will be against the desk. The robot can then engage another perceptual process(es) to determine the cause of the evidential discrepancy, or run several separate recognition processes in parallel (*is it a desk? is it a cabinet? is it chair?*). This drastically reduces the computational complexity involved in evidential management and inference for recognition.

The final assumption in SFX is that *in order to reason about the use of sensors to reduce uncertainty and to respond dynamically to surprises, the contributions of uncertainty by algorithms and sensors must be modular.* As argued in [14,22], many sensors and/or algorithms can acheive the same perception. However, these logical sensors will vary in the amount of evidence they can contribute. Therefore, a representation is needed which acknowledges the perceptual equivalence yet expresses their evidential differences. SFX provides modularity via a three level taxonomy of evidence, from which the system is able to infer the percept from observations of features. The evidential flow of this taxonomy is formalized as an evidential representation consisting of *domain knowledge* and an evidential *operator*. The taxonomy is presented first, followed by the formal representation.

4.2 Accrual of Evidence

The accrual of evidence, ε, in SFX follows a three-level hierarchy. Evidence at any level of the hierarchy is represented as a Shafer belief function. Accrual begins with the collection of *evidence for a feature*. The quality of the evidence of a feature depends both on the feature and the sensing process observing it, therefore it is written as $\varepsilon^{s^k}_{d^i_j}$ to reflect the impact of the sensor s^k. Since each feature is extracted independently of other features, it has a unique FOD, $\Omega_{d^i_j}$. The belief function for the evidence contributed by the kth sensor's observation of feature d^i_j, $\varepsilon^{s^k}_{d^i_j}$, is represented as: $Bel^{s^k}_{d^i_j}|2^{\Omega} : m(d^i_j),\ m(\overline{d}^i_j),\ m(\Omega)$.

The evidence for a feature is based on how well the observed value of the feature matches the expected value. By using descriptions constructed from example sensor data, matching is able to take into account the uncertainty and variation arising from the particular sensors and feature-extraction algorithms used in the observation. Many feature observation algorithms in the literature are for model-based perception and come with goodness-of-fit functions that can be readily adapted to assign belief.

Not all feature observations necessarily contribute evidence for the percept. Some observations of features may be treated as *structural*, where the observation is only used as a component of another observation. For example, an observation of an edge may contribute evidence for the model of a DESK. But an equally valid model of the same desk can be based on lines formed from the edges. In this case the edges themselves are integral structurally to the recognition of the desk but are not directly involved in contributing evidence.

The second level is the *evidence for a description of a percept.* The evidence from each feature in a particular description of a percept leads to a *body of evidence* for the percept. The body of evidence is a single value which is written as $\varepsilon_p^{d^i s^k}$, where the superscripts indicate the source of the evidence and the subscript indicates the evidence is for the percept. Each description is over a FOD Φ_{d^i}. The belief function for each description d^i is given by: $Bel_{d^i}^{s^k}|2^\Phi : m(d^i),\ m(\overline{d}^i),\ m(\Phi)$.

An enlargement function called the **description_interpretation**, r_{d^i}, is responsible for propagating the evidence for the features into a body of evidence: $\varepsilon_p^{d^i s^k} = r_{d^i}\langle \varepsilon_{d_1^i}^{s^k}, \varepsilon_{d_2^i}^{s^k}, \ldots, \varepsilon_{d_m^i}^{s^k} \rangle$.

The description interpretation plays an important part in generating the body of evidence because it may weigh or bias the features' evidence differently. For example, in recognizing a place to sit given imperfect sensing, evidence of the functional feature "thermal profile" is more important than the evidence of "blue region". Essentially, it serves to express the conditioning of evidence for a feature based on the evidence for other features while respecting the dichotomy between structural and evidential features.

The third level concerns the *evidence for an observation of the entire percept.* Each body of evidence for the descriptions in the expected percept is used to generate the total measure of evidence: $\varepsilon_p = \bigoplus_{i,i=1\ldots n} \varepsilon_p^{d^i s^k}$ where $\bigoplus$ is used the represent the uncertainty management function which will combine the evidence, in this case via the orthogonal sum given by Dempster's rule. The belief function for a percept is given by: $Bel_p^{d^i s^k}|2^\Theta : m(p),\ m(\overline{p}),\ m(\Theta)$.

The uncertainty management function must take into account that each description may not contribute equally, regardless of the certainty in its observation. For example, what is the contribution of very high uncertainty in an observation of the visual description of the STUDENT DESK SCENE versus a very low uncertainty in an observation of the thermal description? A second enlargement function, the percept_interpretation, r_p represents the contributions of each description of the percept that is being observed and how that influences the combination mechanism.

4.3 Evidential Mappings

The interpretation rules permit the propagation of evidence between the three levels of the hierarchy. Other applications of DS theory have encountered this problem of transferring belief across disparate FODs. Lowrance, Garvey, and Strat [17] addressed this in 1986 by introducing a new knowledge structure called a *compatibility relation*, which they used to propagate belief in such cases. No theoretical justification was given for the compatibility relations, although they are conceptually similar to the enlargement of FODs through formal assumptions about the differences between FODs. Keller and Hobson [24] in 1989 implemented *interpretation rules* for translating the belief in one FOD to another, again with no explicit theoretical justification. At about the same time, Hutchinson, Cromwell, and Kak [23] noted that for their application some FODs could be considered *refinements* of another and transferred belief according to [48], permitting an implementation that would remain fully consistent with DS theory. Liu, Hughes, and McTear formalized the representation of heuristic knowledge as an evidential mapping between frames of discernment in [32].

The enlargement functions in SFX encode the evidential mappings between each frame of discernment. Evidential mappings in SFX are implemented as weight vectors which are selected according to rules. This implementation is similar to the evidential mapping matrices

Figure 5: George, the Georgia Tech Denning mobile robot.

proposed by Lui, Hughes, and McTear [32], which form an implicit set of rules about interpreting evidence. SFX takes advantage of the sparseness of the matrix for scene recognition and simplifies the matrix into vectors and explicit rules.

5 Experiments

Experiments with the SFX uncertainty management system were carried out with the Georgia Tech mobile robot (Figure 5) acting as a security guard. The robot would enter the room, move to within a foot of a known location, and then scan the room at $20°$ increments, comparing the current views for changes. Each scene was considered a percept, and evidentially the task was for the robot to determine the belief in the percept. If the belief in a scene is low and the sensors are operating properly, the room was defined as having experienced a significant change and security notified.

In Section 5.1, the collection and propagation of evidence for the DRILL PRESS SCENE is detailed as an exemplar of uncertainty management using DS theory in SFX. It shows how the evidential component operates under nominal conditions. Section 5.2 describes the five additional experiments used to demonstrate the propagation of evidence under various scenarios with the STUDENT DESK SCENE.

Figure 6: The DRILL PRESS SCENE of a tool room.

5.1 Detailed Example of Inference in SFX

Figure 6 shows a view of a tool room, the DRILL PRESS SCENE. In this experiment the robot visited the room when the scene was unchanged. The robot observed the scene and generated a concensus from the four sensors that it was unchanged with a belief of 0.99.

5.1.1 Feature Evidence

The sensing plan for the DRILL PRESS SCENE is based on four descriptions: *color* (d^1), *b&w* (d^2), *thermal* (d^3), and *range* (d^4). The color description consists of a single feature, the color histogram [52] for the scene. As noted earlier, many feature extraction algorithms come with a matching function that can be used to assign belief. In this case, the color histogram (feature d^1_1), computes the "intersection" of a color histogram from the observed image, I, with the modeled histogram, M. The intersection is the number of pixels in I which have the same number of pixels for each bucket in M, and can be converted to a fraction by dividing the intersection by the total number of pixels in M. This intersection can be used as a matching function where the fraction of pixels which match are the belief mass *for* the color histogram. The fraction which do not match constitute the belief mass *against* the color histogram. The entire matching function is given by:

$$\begin{aligned} m(d_1^1) &= \frac{\sum_{j=1}^{n} min(I_j, M_j)}{\sum_{j=1}^{n} M_j} \\ m(\overline{d}_1^1) &= 1.0 - m(d_1^1) \\ m(\Omega) &= 0.0 \end{aligned}$$

where the FOD is $\Omega = \{d_1^1, \overline{d}_1^1\}$

It should be noted that a matching function may assign ignorance; $m(\Omega)$ does not have to be 0.0.

The b&w description consists of 4 features based on the spatial relationships between three peaks in the image intensity histogram. The centroid (in image coordinates) of the pixels in each of the three peaks, d_2^2, d_3^2, d_4^2 form a constellation d_1^2, a specific topological and distance relationship. The difference in pixels between the modeled value relationship and the observed value is the belief against the feature. Features d_2^2, d_3^2, d_4^2 are structural, while d_1^2 is evidential.

The thermal description also relied on the topological d_1^3 and distance relationship d_2^3 between the centroids of two prominent thermal regions (d_3^3, d_4^3). The matching function was computed as above for the b&w constellation. d_1^3, d_2^3 are evidential.

The range description was a single feature: the distance to the nearest surface in the scene. Evidence for the feature was inversely proportional to the difference between the observed distance and the modeled difference.

5.1.2 Evidence for a Description

The description_interpretation enlargement function is implemented in SFX as r_{d^i}, a vector of scalars (0.0–1.0), reflecting the weight of each feature contributing evidence to the description. The vector has one element per evidence-contributing feature in the description;

the jth element is the weight for the belief function associated with the jth feature. The translation of belief from a belief function $Bel^{s^k}_{d^i_j}$ focused on Ω to a belief function $Bel^{s^k}_{d^i}$ focused on Φ is given by:

$$
\begin{aligned}
r_{d^i}: &\quad Bel^{s^k}_{d^i_j}|2^{\Omega} \rightarrow Bel^{s^k}_{d^i}|2^{\Phi} \\
Bel^{s^k}_{d^i_j}|2^{\Omega}: &\quad m(d^i_j), m(\overline{d}^i_j), m(\Omega) \\
Bel^{s^k}_{d^i}|2^{\Phi}: &\quad m(d^i), m(\overline{d}^i), m(\Phi)
\end{aligned}
$$

where the **description_interpretation** vector $r^{d^i} = \langle r^{d^i_1}, r^{d^i_2}, \ldots, r^{d^i_n} \rangle$ is applied:

$$
\begin{aligned}
m(d^i) &= r_{d^i_j}(m(d^i_j)) \\
m(\overline{d}^i) &= r_{d^i_j}(m(\overline{d}^i_j)) \\
m(\Phi) &= 1 - m(d^i) - m(\overline{d}^i) \qquad (5)
\end{aligned}
$$

It should be noted that SFX filters at this point in process for two types of sensing anomalies. It compares the belief mass assigned to ignorance in the body of evidence ($m(\Phi)$); if it is 1.0 the system flags a missing evidence error. Likewise, the ignorance is compared to the high uncertainty threshold; if the ignorance is high, the entire belief is uncertain, and a sensing exception is signaled.

5.1.3 Evidence for the Percept

SFX combines the evidence by first transferring the belief focused on individual description FODs, Φ, to the percept FOD, Θ, then combining the resulting belief functions. The belief transfer is done according to the **percept_interpretation** enlargement function r_p. The new belief functions are combined using Dempster's rule of combination.

The **percept_interpretation** is implemented as a collection of weighting vectors and a set of *if-then* rules regarding when to use a particular vector. The fusion step selects the *if-then* rule which is satisfied by the current conditions, and the corresponding vector, denoted by r_p. These conditions allow the designer to permit different weightings to be applied to the transfer of evidence based on the current circumstances. Consider again the DRILL PRESS SCENE percept, where the bodies of evidence from a thermal camera, color camera, black and white camera, and ultrasonics were used to determine whether the scene was the same (evidence *for* the percept) or had changed (*against* the percept). There are at least three different ways that a change to the scene is manifested: first, an intruder could be present, in which case all description/sensors should report a consensus *against* the scene being the same; second, an intruder could have disturbed the scene and then left (e.g., dropped the stolen goods and fled), in which case the visual sensors would report evidence *against* the scene, while the thermal sensor would not; and third, the drill could be overheating or the wall could be hot indicating an incipient fire, in which case the thermal sensor would report evidence *against* the scene, but without consensual evidence from the vision sensors.

These three cases arise because evidence for the expected thermal signature doesn't reveal much about the drill press scene under normal circumstances; only when there is *not* a match does the thermal sensor appear to be useful. One conclusion that may be drawn is that when the thermal evidence is certain and *against* the percept, it should be heavily weighted, while when it is *for* the percept its contribution should be minimal.

So it can be seen that within a single sensing plan there are circumstances when one source of evidence is more important than the others, and that these circumstances lead to different interpretations of the evidence. Specifying interpretations potentially requires a great deal of domain knowledge which may not always be available to the designer. If the designer does have access to this type of domain knowledge, the question then becomes how to implement a structure to support the selection of the appropriate weighting of the evidence?

The **percept_interpretation** for the **drill press scene** example (shown below) encodes the circumstances which change the weighting of the thermal evidence. This rule is supplied by the knowledge engineer. The transference of belief from each description/sensor belief function $Bel_{d^i}^{s^k}|2^\Phi$ to a new belief function $Bel_p^{d^i s^k}|2^\Theta : m(p), m(\overline{p}), m(\Theta)$ is given by the *percept vector* $r_p = \langle r_p^{d^1}, r_p^{d^2}, \ldots, r_p^{d^m} \rangle$. If the contribution of a body of evidence is discounted, the belief "discount" must be moved to ignorance ($m(\Phi)$) in order to keep the total belief mass equal to one. The percept vector is applied as in Equation 5.

if	$m(\overline{d}^1) \geq 0.30$
	$m(\Phi) \leq 0.20$
then	$r_p = \langle 1.00, 0.50, 0.50, 0.25 \rangle$
else	$r_p = \langle 0.50, 1.00, 1.00, 0.25 \rangle$

After the bodies of evidence have been transferred to the percept FOD, the belief functions are then combined using Dempster's rule of combination, and Con, the weight of conflict metric, is computed. Checks are made to see if a fusion step failure condition has been met; SFX currently checks for Con exceeding a high conflict threshold and the total evidence *for* the percept ($m(p)$) reaching the minimum certainty threshold.

In this experiment the bodies of evidence are:

$$\begin{aligned}
Bel_{d^1}^{s^1}|2^\Phi &: m(d^1) = 0.75,\ m(\overline{d}^1) = 0.25,\ m(\Phi) = 0.00 \\
Bel_{d^2}^{s^2}|2^\Phi &: m(d^2) = 0.79,\ m(\overline{d}^2) = 0.11,\ m(\Phi) = 0.10 \\
Bel_{d^3}^{s^3}|2^\Phi &: m(d^3) = 0.97,\ m(\overline{d}^3) = 0.03,\ m(\Phi) = 0.00 \\
Bel_{d^4}^{s^4}|2^\Phi &: m(d^4) = 0.93,\ m(\overline{d}^4) = 0.07,\ m(\Phi) = 0.00 \qquad (6)
\end{aligned}$$

where d^3s^3 is the thermal description, d^1s^1 is the color video description, d^2s^2 is the black & white video description, and d^4s^4 is range profile from the ultrasonics.

Belief transfer begins by selecting the appropriate percept rule vector from the *percept rule*. The antecedent is not satisfied so $r_p = \langle 0.50,\ 1.00,\ 1.00,\ 0.25 \rangle$ is selected. It is applied to the belief functions in Eqn.(6) yielding four new belief functions:

$$Bel_p^{d^1 s^1}|2^\Theta : m(p) = 0.38,\ m(\overline{p}) = 0.12,\ m(\Theta) = 0.50$$

$$
\begin{aligned}
Bel_p^{d^2 s^2}|2^\Theta &: \; m(p) = 0.79, \; m(\overline{p}) = 0.11, \; m(\Theta) = 0.10 \\
Bel_p^{d^3 s^3}|2^\Theta &: \; m(p) = 0.97, \; m(\overline{p}) = 0.03, \; m(\Theta) = 0.00 \\
Bel_p^{d^4 s^4}|2^\Theta &: \; m(p) = 0.23, \; m(\overline{p}) = 0.02, \; m(\Theta) = 0.75
\end{aligned} \tag{7}
$$

The new belief functions in Eqn.(7) reflect that neither the thermal nor the ultrasonics observations contribute much evidence either for or against the percept under these conditions. The final belief functions are combined according to Eqn.(2):

$$
\begin{aligned}
Bel_p &= ((((Bel_p^{d^1 s^1} \bigoplus Bel_p^{d^2 s^2}) \bigoplus Bel_p^{d^3 s^3}) \bigoplus Bel_p^{d^4 s^4}) \\
&: \; m(p) = 0.99, \; m(\overline{p}) = 0.01, \; m(\Theta) = 0.0 \\
& Con(Bel_p^{d^1 s^1}, Bel_p^{d^2 s^2}, Bel_p^{d^3 s^3}, Bel_p^{d^4 s^4}) = 0.29
\end{aligned} \tag{8}
$$

SFX then checks the fusion step failure conditions and sees that the Con value is below the predetermined high conflict threshold, and $m(p) = 0.99$ exceeds the minimum certainty.

5.2 Other Experiments

Seven additional experiments were conducted with the STUDENT DESK SCENE (Figure 1). Two of the experiments used observations from three sensors, the Pulnix b&w camera, the Inframetrics thermal camera, and the Polaroid ultrasonics; the remaining five experiments included observations from a Hi8 color camera to serve as a competing sensor [16] for the b&w camera. The features and descriptions were similar to those for the DRILL PRESS SCENE.

5.2.1 Detection of Sensing Anomalies

Three experiments tested the ability of the execution mechanism to detect sensing anomalies and suspend execution pending further investigation by the exception handling mechanism. If undetected, the execution mechanism will fuse the percept based on potentially erroneous observations.

Experiments 2, 3, and 4 presented the robot respectively with a different problem: a wrong scene replaced the expected scene, a sensor malfunction changed the robot's sensing capabilities, and turning off the lights violated the assumption that the environment would remain unchanged. The execution mechanism of SFX is expected to post a failure in response to each invalidation. SFX has three categories of failures: missing evidence, a high conflict between bodies of evidence, and the total belief below a prescribed minimum. The high conflict threshold was set at 0.69 for these experiments, which represents a null set of 50%. The minimum belief in p was set at 0.75 based on an analysis of what the average value of belief was for unchanged scenes.

As seen in the tables below all three experiments successfully detected such a failure, where p is belief *for*, $\sim p$ is *against*, and ? is ignorance. Furthermore, experiment 3 tested the additional hypothesis that configurations with competing sensors (sensors which extract observations from the same sensing modality such as visible light) are more robust than configurations without competition. The experiments showed that this was not necessarily true, as the color camera led only to a high conflict, not an acceptable belief that the scene was unchanged.

Experiment 2: wrong scene

description/sensor	$Bel\|2^\Theta$	p	$\sim p$	?
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.00	1.00	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.50	0.00	0.50
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.24	0.01	0.75
total	Bel_p	*failure: high conflict* Con=0.97$\geq$0.69		

Experiment 3: b&w camera malfunctioning, no corroboration

description/sensor	$Bel\|2^\Theta$	p	$\sim p$	?
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.44	0.56	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.50	0.00	0.50
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.25	0.00	0.75
total	Bel_p	*failure: minimum belief* $m(p)$=0.68 $<$ 0.75		

Experiment 3: b&w camera malfunctioning, with corroboration

description/sensor	$Bel\|2^{\Theta}$	p	$\sim p$	?
color/Hi8	$Bel_p^{d^1 s^1}$	0.68	0.32	0.00
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.44	0.56	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.50	0.00	0.50
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.25	0.00	0.75
total	Bel_p	*failure: conflict* Con=1.00≥0.69		

Experiment 4: lights turned off

description/sensor	$Bel\|2^{\Theta}$	p	$\sim p$	?
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.00	0.00	1.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.50	0.00	0.50
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.24	0.01	0.75
total	Bel_p	*failure: missing evidence from b&w/Pulnix*		

5.2.2 Utility of Interpretation Rules

The purpose of the remaining two experiments was to test how different observations lead to different weighting vectors being selected from the *percept if-then rules*, and how these weighted bodies of evidence would outperform a single fixed weighting. The *percept if-then rule* for the sensing plan for the **student desk scene** has two weight vectors. One vector discounts the body of evidence from the thermal camera under normal circumstances, while the other applies a full weight if heat is detected. Experiment 5 observed the scene with a person sitting near a desk, which was expected to satisfy the condition in the *percept if-then rules* for the STUDENT DESK SCENE shown below that would weight the thermal body of evidence with a value of 1.0, or no discount:

		$r_p^{d^1}$	$r_p^{d^2}$	$r_p^{d^3}$	$r_p^{d^4}$
if $m(\overline{d}^3) \geq 0.30$	then	1.00	1.00	1.00	0.25
	else	1.00	1.00	0.50	0.25

The experimental results below showed that this vector from the percept rule was employed and that the thermal evidence was weighted accordingly. The descriptions of the percept, however, were not sensitive to the change, resulting in bodies of evidence which where ambiguous. This triggered a state failure due to high conflict between all the bodies of evidence. The experiment also compared this result to what would have happened if the other weighting vector in the *percept if-then rules* had been used. In this case, the execution mechanism would have produced a high belief ($m(p) = 0.97$) that the scene was unchanged, which is incorrect, which is consistent with the experimental predictions.

Experiment 5: intruder, with appropriate rule

description/sensor	$Bel\|2^{\Theta}$	p	$\sim p$	?
color/Hi8	$Bel_p^{d^1 s^1}$	0.77	0.23	0.00
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.87	0.13	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.58	0.42	0.00
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.24	0.01	0.75
total	Bel_p	*failure: conflict* $Con = 0.92 \geq 0.69$		

Experiment 5: intruder, without appropriate rule

description/sensor	$Bel\|2^{\Theta}$	p	$\sim p$	?
color/Hi8	$Bel_p^{d^1 s^1}$	0.77	0.23	0.00
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.87	0.13	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.29	0.21	0.50
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.24	0.01	0.75
total	Bel_p	0.97	0.03	0.00
		$Con = 0.62 \leq 0.69$		

In Experiment 6, a 6 ft. Godzilla pool toy was introduced into a less cluttered area in the scene; this should cause a change in the **student desk scene** not discernible by the thermal sensor, and result in the *percept if-then rules* applying a low weight to the thermal evidence. The experiment did show that the appropriate weighting vector was chosen and the thermal evidence was weighted as expected. However, only one description produced an observation that the scene had changed, due to a lack of sensitivity in the descriptions. This triggered a state failure due to the conflict between the sensors ($Con = 2.44 \geq 0.69$).

Experiment 6: Godzilla in scene, with appropriate rule

description/sensor	$Bel\|2^{\Theta}$	p	$\sim p$	?
color/Hi8	$Bel_p^{d^1 s^1}$	0.77	0.23	0.00
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.00	1.00	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	0.50	0.00	0.50
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.24	0.01	0.75
total	Bel_p	*failure: conflict* Con=2.44$\geq$ 0.69		

Experiment 6: Godzilla in scene, without appropriate rule

description/sensor	$Bel\|2^{\Theta}$	p	$\sim p$	?
color/Hi8	$Bel_p^{d^1 s^1}$	0.77	0.23	0.00
b&w/Pulnix	$Bel_p^{d^2 s^2}$	0.00	1.00	0.00
thermal/Inframetrics	$Bel_p^{d^3 s^3}$	1.00	0.00	0.00
range/Ultrasonics	$Bel_p^{d^4 s^4}$	0.24	0.01	0.75
total	Bel_p	*failure: conflict* $Con=\infty \geq 0.69$		

These experiments were successful in that they demonstrated the role of the *percept if-then rules* in adapting the interpretation of the body of evidence to the context. They showed that no single rule worked for both Experiments 5 and 6 since applying a discount to the thermal body of evidence gave better results for the circumstances in Experiment 6 but was not appropriate for the conditions in Experiment 5. The experiments were unsuccessful in that they reported state failures rather than strong belief *against* the percept (i.e., that it had changed) due to the poor model of the percept. On the other hand, it should be noted that in Experiment 5 the failures prevented SFX from erroneously confirming that the scene was unchanged.

5.2.3 Overall Results

In general, the experiments confirmed the expectations. The main exception was Experiment 3 which showed that competing sensors did not necessarily improve robustness. Some portions of the experiments were inconclusive due to the lack of sensitivity of the descriptions to some of the changes. In addition to the specific goals of each experiment, these experiments showed:

State failures from high conflict can prevent incorrect perception. Updating evidence with either Bayes' Rule or Dempster's Rule will normalize out the disagreements between bodies of evidence, producing a high belief *for* the percept when in fact the belief *for* is just slightly more that the belief *against*. One of the advantages of explicitly monitoring for *high conflicts* between bodies of evidence was to prevent this from happening, where the robot would be acting on what was reported with high belief but was in fact an artifact of the rule of combination.

Adaptive weighting in the percept if-then rules *makes the transfer of belief more robust.* The use of the appropriate weighting vector in Experiment 5 prevented the execution mechanism from posting a high belief in the scene being unchanged when in fact an intruder was present, which would have been incorrect.

6 Discussion

The previous section shows the success of the SFX in propagating and combining belief, detecting a lack of consensus in the accumulated belief, and the incorporation of heuristic domain knowledge. The uncertainty management mechanism for SFX also overcomes two disadvantages of Dempster-Shafer theory cited in Section 2.2. First, the potential computational intractability of a Dempster-Shafer formalization is never realized in the SFX implementation. As a consequence of the selection of models by the planning activity, and the dichotomy between structural and evidental features, evidence is not collected for every feature in the world, only for components of the descriptions of the percept which are relevant for the task. Therefore the frame of discernment for features and descriptions is a small subset of what *could* be observed. So while the computational complexity associated with assigning belief mass and combining belief functions is 2^{FOD}, the size of FOD is generally small. The second problem, counterintuitive averaging of conflicting evidence cited by [55], is not a concern for this application because the execution mechanism monitors for conflicts.

The biggest advantage of Dempster-Shafer theory for SFX is also its biggest disadvantage: it allows the designer to quantify the domain knowledge into matching, description, and percept interpretations. As was seen in Section 5.1.3, domain knowledge is necessary to weight the contributions of descriptions and arrive at a correct interpretation. However, the quantification of domain knowledge relies on the competence of the knowledge engineer; while this is no different than other implementations, either DS or Bayesian, it certainly calls for additional research. Our preliminary efforts [31] indicate that such knowledge can be learned.

The overall operation of the uncertainty management mechanism illustrates several practical contributions of the three level evidential taxonomy for the application of any evidential reasoning system for sensor fusion. Evidence in the taxonomy starts with a measured difference between an observed and an expected value, then weights that difference according to the description and percept interpretations. This separation of measurement and interpretation preserves the ability of a sensor fusion system to share sensing information between perceptual processes; two processes can simultaneously share the same features but interpret them differently. It also allows the interpretation to change with context without changing the way evidence at the feature level is measured. Also it suggests the possibility that the description and percept interpretation can be adaptively modified or even learned.

Besides avoiding the computational tractability problem for DS theory, the model of the expected percept as a small collection of descriptions avoids another issue. Applications of evidential representations all share a subtle concern: while the evidential system delimits the representation and combination of evidence, it does not offer any way of measuring when there is *enough* evidence. In some sense, these systems are prone to an evidential "horizon effect" where if two sources of evidence are good, three, four, or more might be better. Configuring the sensing process around descriptions of the expected percept which are expected to be sufficient by the planning activity and the conditions under which these expectations are violated (e.g., the state failure conditions) eliminates this issue.

7 Summary

This article presents the evidential representation used by the Sensor Fusion Effects (SFX) architecture for autonomous mobile robots. Evidence in SFX is used not only for recognition of a percept, but also to control whether the robot proceeds with an action, investigates the situation further, or terminates the task. The article compares Bayesian and DS methods, and concludes that a DS framework is more appropriate for reasoning about the interpretation of sensor observations. This comparison is expected to guide others in developing an evidential representation for their application.

SFX exploits two less well known aspects of DS theory. It uses the Con weight of conflict metric, which quantifies the discordance in a set of evidence, to indicate that one (or more) of the sensors has malfunctioned, the environment has changed, or the percept model is incorrect. SFX also incorporates domain knowledge into the inference process in a natural way by enlarging belief. Another advantage is that it is practical; belief in the form of a goodness-of-fit measure is readily available in machine perception while probability density functions are not. The major disadvantage of Dempster-Shafer theory for this application is that it relies on the knowledge engineer to quantify the domain knowledge into interpretation functions. This is problematic for Bayesian implementations as well.

The evidential representation is based on a three level taxonomy of evidence that is independent of sensor and feature extraction algorithms. The major advantage of this taxonomy is that it structures the transfer of belief in a such a way as to allow the observations to be interpreted differently by different perceptual processes or contexts, yet allows the algorithms themselves to be treated as logical sensors.

Six experiments with four types of sensor data collected from a mobile robot illustrate the accrual of evidence in SFX and the detection of sensing anomalies. One demonstrated the transfer of belief under nominal conditions, three experiments showed how the fusion process was suspended when SFX detected a conflict between sensors, and two the importance of the interpretation rules in incorporating the influence of contextual knowledge.

Other extensions of this research on uncertain reasoning for intelligent sensor fusion is concentrating on reasoning about sensing exceptions, the quantification of the domain knowledge through learning, and reasoning about sensor observations collected over time.

Acknowledgements

Portions of this work were supported by a Rockwell Doctoral Fellowship, the Georgia Tech CIMS Program, and by the Georgia Tech Material Handling Research Center. The author would like to thank Dr. John Lemmer and Dr. Ron Arkin for their helpful discussions, and Jim Hoffman, Tye Hooley, Sean Kerstiens, Dr. John Lindner, Elizabeth Nitz, and Tye Hooley for their comments and editoral suggestions.

References

[1] Abidi, M.A., and Gonazalez, R.C., *Data Fusion in Robotics and Machine Intelligence*, Academic Press, Inc., San Diego, 1992.

[2] Agosta, J.M., "Probabilistic Recognition Newtworks, An Application of Influence Diagrams to Visual Recognition", PhD thesis, Stanford University, 1991.

[3] Allen, P., "Integrating Vision and Touch for Object Recognition Tasks," *International Journal of Robotics Research*, vol. 7, no. 6, Dec 1988, pp. 15-33.

[4] Aloimonos, J., "Visual Shape Computation", *Proceedings of IEEE*, vol. 76, no. 8, Aug 1988, pp. 899-916.

[5] Bajcsy, R., "Active Perception", *Proceedings of IEEE*, vol. 76, no. 8, Aug. 1988, pp. 886-1005.

[6] Ballard, D., "Eye Fixation and Early Vision: Kinetic Depth", *Second International Conference on Computer Vision*, Tampa FL, 1988, pp. 524-531.

[7] Beckerman, M., and Oblow, E.M., "Treatment of Systematic Errors in the Processing of Wide-Angle Sonar Sensor Data for Robotics Navigation" *IEEE Transactions on Robotics and Automation*, vol. 6., no. 2, April 1990, pp. 137-145.

[8] Beckerman, M., Farkas, L.A., and Johnston, S.E., "Treatment of Systematic Errors II: Fusion of Ultrasound and Visual Sensor Data", *Technical Memo ORNL/TM-11349/CESAR-89/38*, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1989.

[9] Bolger, P.L., "Shafer-Dempster Reasoning with Applications to Multisensor Target Identification Systems", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-17, no. 6, Nov/Dec 1987, pp. 968-977.

[10] Chatterjee, P.S., and Huntsberger, T.L., "Comparison of Techniques for Sensor Fusion under Uncertain Conditions", *Proceedings SPIE Sensor Fusion: Spatial Reasoning and Scene Interpretation*, vol. 1003, 1988, pp. 194-199.

[11] Chavez, G.T., and Murphy, R.R., "Exception Handling for Sensor Fusion", *SPIE Sensor Fusion VI*, Boston, MA, Sept. 7-10, 1993, pp. 142–153.

[12] Dempster, A.P., "A Generalization of Bayesian Inference", *Journal of the Royal Statistical Society: Series B*, vol. 30, 1968, pp. 205-247.

[13] Dempster, A.P., "Upper and Lower Probabilities Induced by a Multivalued Mapping", *Annals of Mathematical Statistics*, vol. 38, no. 2, April, 1967, pp. 325-339.

[14] Donald, B., and Jennings, J., "Perceptual, Limits, Perceptual Equivalence Classes, and a Robot's Sensori-Computational Capabilities", *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, Osaka, Japan, 1991.

[15] Dubois, D. and Prade, H., "An Introduction to Possibilistic and Fuzzy Logics", *Readings in Uncertain Reasoning*, edited by G. Shafer and J. Pearl, Morgan Kaufmann Publishers, San Mateo, CA, 1990.

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

[17] Garvey, T.D., Lowrance, J.D., and Fischler, M.A., "An Inference Technique for Integrating Knowledge from Disparate Sources", *Proceedings IJCAI-81*, August 24-28, 1981, Vancouver, B.C.,Canada, vol. 1, pp. 319-325.

[18] Hager, G., "Active Reduction of Uncertainty in Multi-sensor Systems", Ph.D. thesis, University of Pennsylvania, Philadelphia, PA, *MS-CIS-88-47*, 1988.

[19] Hager, G., and Mintz, M., "Task-Directed Multi-Sensor Fusion", *Proceedings 1989 IEEE International Conference on Robotics and Automation*, vol. 1, 1989, pp. 662-667.

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

[21] Halpern, J.Y., and McAllester, D.A., "Likelihood, probability, and knowledge", *Research Report RJ 4313 (47141)*, MIT Artificial Intelligence Laboratory, Cambridge, MA, June 6, 1984.

[22] Henderson, T. and Shilcrat, E., "Logical Sensor Systems", *Journal of Robotic Systems*, vol. 1, no. 2, 1984, pp. 169-193.

[23] Hutchinson, S.A., Cromwell, R.L., and Kak, A.C., "Applying Uncertainty Reasoning to Model Based Object Reasoning", proceedings *CVPR '89*, San Diego, CA, June 4-8, 1989, pp. 541-548.

[24] Keller, J.M., and Hobson, G., "Uncertainty Mangement in a Rule-Based Automatic Target Recognizer", *Proceedings SPIE Applications of Artificial Intelligence VII*, vol. 1095, 1985, Orlando, FA, March 28-30, 1989, pp. 126-137.

[25] Kyburg, H.E., "Bayesian and Non-Bayesian Evidential Updating", *Artificial Intelligence*, vol. 31, no. 3, March, 1987, pp. 271-293.

[26] Laskey, K.B., "Conflict and Surprise: Heuristics for Model Revision", *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, UCLA, July 13-15, 1991, pp. 197-204.

[27] Lemmer, J. F., "Confidence Faction, Empiricism and the Dempster-Shafer Theory of Evidence", *Uncertainty in Artificial Intelligence*, L.N. Kanal and J.F. Lemmer, Elsevier Science Publishers, B.V., 1986, pp. 117-125.

[28] Levitt, T.S., Agosta, J.M., and Binford, T.O., "Model-Based Influence Diagrams for Machine Vision", *Uncertainty in Artificial Intelligence 5*, M. Henrion, R.D. Shacter, L.N. Kanal, and J.F. Lemmer, ed., Elsevier Science Publishers, B.V., 1990, pp. 371-388.

[29] Levitt, T.S., Binford, T.O., and Ettinger, G.J., "Utility-Based Control for Computer Vision", *Uncertainty in Artificial Intelligence 4*, R.D. Shacter, L.N. Kanal, T.S. Levitt, and J.F. Lemmer, ed., Elsevier Science Publishers, B.V., 1990, pp. 407-422.

[30] Lehrer, N., Reynolds, G., and Griffith, J., "A Method for Initial Hypothesis Formation in Image Understanding", *COINS Technical Report 87-04*, University of Massachusetts, Amherst, 1987.

[31] Lindner, J., Murphy, R., and Nitz, R., "Learning the Expected Utility of Sensors and Algorithms", to appear in the proceedings of *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Las Vegas, NV, Oct. 2-5, 1994.

[32] Liu, W., Hughes, J.G., and McTear, M.F., "Representing Heuristic Knowledge in D-S Theory", *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, Stanford University, July 17-19, 1992, pp. 182-190.

[33] Luo, R.C., Lin, M., Scherp, R.S., "The Issues and Approaches of a Robot Multi-sensor Integration", *Proceedings IEEE Robotics and Automation Conference*, Raleigh, NC, March 30-April 3, 1987, pp. 1941-1946.

[34] Luo, R.C., and Kay, M.G., "Multisensor Integration and Fusion in Intelligent Systems", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-19, no. 5, Sept/Oct 1989, pp. 901-931.

[35] Miltonberger, T., Morgan, D., and Orr, G., "Multisensor object recognition from 3D Models", *Proceedings SPIE: Sensor Fusion I*, 1988, pp. 161-169.

[36] Moravec, H.P., "Sensor Fusion in Certainty Grids for Mobile Robots", *AI Magazine*, vol. 9, no. 2, Summer, 1988, pp. 61-74.

[37] Murphy, R.R., "An Architecture for Intelligent Robotic Sensor Fusion", PhD thesis, Technical Report #GIT-ICS-92/42, College of Computing, Georgia Institute of Technology, Atlanta, GA, 30332-0280.

[38] Murphy, R.R., "Learning to Eliminate Background Effects in Object Recognition", *AAAI 1993 Fall Symposium on Machine Learning in Computer Vsion: What, Why and How?*, Raleigh, NC, Oct. 22-24, 1993.

[39] Murphy, R.R, and Arkin, R.C., "SFX: An Architecture for Action-Oriented Sensor Fusion", *1992 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Raleigh, NC, July 7-10, 1992.

[40] Nakamura, Y., and Xu, Y., "Geometrical Fusion Method for Multi-sensor Robotic Systems", *Proceedings 1989 IEEE International Conference on Robotics and Automation*, vol. 1, 1989, pp. 668-673.

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

[42] Rimey, R.D., and Brown, C. M., "Control of Selective Perception Using Bayes Nets and Decision Theory", to appear *International Journal of Computer Vision*, special issue on Active Vision.

[43] Ruspini, E.H., "Epistemic logics, Probability, and the Calculus of Evidence", *Proceedings IJCAI87*, 1987.

[44] Ruspini, E.H., "The Logical Foundations of Evidential Reasoning", *SRI Technical Note 408*, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, Dec. 20, 1986.

[45] Safranek, R.J., Gottschlich, S., and Kak, A.C., "Evidence Accumulation Using Binary Frames of Discernment for Verification Vision", *IEEE Transactions on Robotics and Automation*, vol. 4, no. 4, August, 1990, pp. 405-417.

[46] Schocken, S., and Kleindorfer, P.R., "Artificial Intelligence Dialects of the Bayesian Belief Revision Language", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-19, no. 5, Sept/Oct 1989, pp. 1106-1121.

[47] Shafer, G., and Logan, R., "Implementing Dempster's Rule for Hierarchical Evidence", *Artificial Intelligence*, vol. 33, no. 3, Nov., 1987, pp. 271-298.

[48] Shafer, G., *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[49] Smets, P., "The Combination of Evidence in the Transferable Belief Model", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, May 1990, pp. 447-458.

[50] Smith, R. and Cheeseman, P., "On the Representation and Estimations of Spatial Uncertainty", *SRI International Robotics Laboratory Techincal Paper*, September, 1985.

[51] Stephanou, H.E., and Sage, A.P., "Perspectives on Imperfect Information Processing", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-17, no. 5, Sept/Oct 1987, pp. 780-798.

[52] Swain, M., and Ballard, D.H., "Color Indexing", *International Journal of Computer Vision*, vol. 7, no. 1, Nov., 1991.

[53] Waltz, E.L., and Buede, D.M., "Data Fusion and Decision Support for Command and Control", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-16, no. 6, Nov/Dec 1986, pp. 865-879.

[54] Wang, Y.F., and Aggarwal, J.K., "Integration of Active and Passive Sensing Techniques for Representing 3D Objects", *IEEE Transactions on Robotics and Automation*, vol. 5, no. 4, Aug 1989, pp. 460-471.

[55] Zadeh, L.A., "A Simple View of the Dempster-Shafer Theory of Evidence and its Implication for the Rule of Combination", *AI Magazine*, Summer, 1986, pp. 85-90.