An Investigation of the Applicability of Fuzzy Logic Techniques to Real-Time Temporal Classification for Threat Warning Systems

by

Timothy James Boyd

B.S., Mathematics with Computer Science Option
University of Colorado, Boulder, Colorado, 1982

Submitted to the
Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

at the

Massachusetts Institute of Technology

May, 1994

Signature of Author ______________________________
Department of Electrical Engineering and Computer Science
May, 1994

Certified by ______________________________
J. Francis Reintjes, Professor of Electrical Engineering
Academic Thesis Supervisor

Certified by ______________________________
Kurt Nordhaus, Loral Infrared and Imaging Systems
Company Thesis Supervisor

Accepted by ______________________________
Frederic R. Morgenthaler, Chairman, Committee on Graduate Students

An Investigation of the Applicability of Fuzzy Logic Techniques to Real-Time Temporal Classification for Threat Warning Systems

by

Timothy James Boyd

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of Master of Science
in Electrical Engineering and Computer Science

ABSTRACT

An investigation is performed to determine the applicability of fuzzy logic techniques to temporal classification for real-time threat warning systems. Existing discrimination algorithms apply a series of classifiers designed to detect threatening targets at particular ranges in signals generated from multiple electro-optic sensors. Recent growth in the temporal complexity of this system has restricted its performance due in part to two inherent problems in the discrimination approach. Fuzzy logic is applied to a key subset of the classifiers to exploit complex temporal characteristics of signals from evolving targets at various ranges.

Fuzzy logic is a discipline that utilizes approximate reasoning and multi-valued logic in dealing with complex problems that must deal with imprecise information in some tractable manner. Real-world knowledge is represented by a continuum of logical truth values ranging between false and true. These truth values simultaneously stimulate a set of fuzzy rules to varying degrees, the consequences of which are interpolated into a single crisp result. Fuzzy rules are based on both common sense laws and expert knowledge about the system being designed.

The fuzzy system designed for this investigation converts a set of input signal measurements to corresponding fuzzy sets and applies a Mamdani-based fuzzy inference engine to generate output fuzzy alarm levels. The fuzzified inputs simultaneously stimulate a set of if-then fuzzy rules that are designed to capture the decision relevant expert knowledge about the discrimination process.

This investigation demonstrates that by applying fuzzy logic to a key portion of the temporal discrimination problem, moderate performance improvements can be achieved. These improvements are primarily due to the ability of fuzzy logic to represent and exploit complex temporal knowledge fundamental to the threat warning process in a robust and meaningful manner. Further improvements are anticipated by converting the entire system to a fuzzy logic paradigm.

Acknowledgement

First and foremost I wish to thank my wife, Carrie, to whom I was wed during this undertaking. Only with her constant support, patience, and understanding was I able to complete this thesis. Although sacrifices were made, it was her warmhearted encouragement that made this research possible.

I wish to thank Kurt Nordhaus for his outstanding role as my company thesis supervisor. His seemingly unlimited expertise in a wide range of technical disciplines was invaluable. As the champion of the algorithm development effort for threat warning systems at Loral, Kurt was the driving force behind this work. His inspirational guidance and support throughout this project will always be greatly appreciated.

I wish to thank Professor J. Francis Reintjes who generously offered his time and expertise as my academic thesis supervisor to make the successful completion of this work possible. He was solely responsible for providing a cooperative and stimulating environment at MIT and for establishing a fruitful relationship between MIT and Loral. His rigorous reviews of the content of this thesis in addition to his technical insight substantially improved the final result.

I wish to thank all my co-workers in the Signal and Image Processing Lab here at Loral. Their valuable assistance and friendship was very important in making this project successful. In particular, I wish to thank Robin Williams, Jim Cornman, Jennifer Roy, and Mike Bulpett for their much appreciated support and technical expertise in temporal classification systems. Special thanks go to Neeli Sharma for her undying enthusiasm and valued friendship that was invigorating and often relied upon throughout this undertaking.

I am also grateful to the people at Loral Infrared and Imaging Systems who helped sponsor this thesis work and made the Masters in Industry Program available to someone who would not otherwise have had the opportunity to pursue such a goal. In particular, I am grateful to Tom Mascoli, Dick Manning, and Jill Wittels for their support.

Table of Contents

1 Introduction

1.1 Statement of the Problem

Discrimination algorithms have been developed for real-time threat warning systems that detect airborne targets by analyzing the signals gathered from multiple electro-optic sensors. Each sensor is a wide field of view staring receiver that collects inband radiation and converts it to a digital signal. The signal from each sensor is processed independently. The algorithms are intended to discriminate between signals generated from threatening targets and signals generated from other sources that are non-threatening. Threatening targets are known to emit or reflect varying patterns of energy over the course of the engagement. The discrimination algorithms are designed to detect these patterns by continuously applying a series of temporal[1] classifiers to the signals. Each classifier is designed to measure a particular set of temporal characteristics and test for the presence of a threatening target. If any classifier detects a target then an alarm is reported and the appropriate countermeasure action is taken. The goal of this system is to provide a high probability of target detection with adequate early warning capability while maintaining a sufficiently low false alarm rate. This system is currently being used in a variety of real-time situations.

These discrimination algorithms are subject to several inherent problems due in part to the growth in the complexity of the system requirements as more threat data became available. The algorithms were originally designed to classify a limited number of well understood threats in a real-time system when existing hardware provided relatively limited processing capability. Under these circumstances, a straight forward yet highly successful approach was taken that applied a small number of tests for each classifier using a series of one-dimensional thresholds established for the appropriate temporal characteristics. A relatively few number of classifiers were needed. However, as additional tactical scenarios and target designations became available, a higher level of complexity was demanded from the discrimination algorithms. This simple approach became inadequate in terms of performance and growth capability. The problem is to develop an approach that achieves the required performance and growth capability while preserving, as much as possible, simplicity in the system

1.2 Investigation of Fuzzy Logic

The goal of this thesis is to present an investigation into the applicability of fuzzy logic techniques to the temporal discrimination problem for threat warning systems. Fuzzy logic is a discipline that utilizes approximate reasoning and multi-valued logic in dealing with complex problems that must deal with imprecise information and vagueness in

[1] The term "temporal" is used throughout this report in reference to processes or events that are dependent on time and whose characteristics evolve as a function of time. The term temporal classification refers to a classifier which takes advantage of the temporal nature of the process or event of interest both for the purposes of discrimination and for computational simplification.

some tractable manner (Zadeh, [1965])[1]. It allows us to represent and reason with forms of real-world knowledge where observations and decisions are often based on uncertainty by allowing a continuum of logical values representing degrees of truth ranging between false (0) and true (1). These "fuzzy truth values" simultaneously stimulate a set of "fuzzy rules" to varying degrees, the consequences of which are interpolated into a single "crisp" result. Fuzzy rules are based on both common sense laws and expert knowledge about the system being designed. Fuzzy logicians maintain that fuzzy logic closely mimics the ability of the human mind to summarize ambiguous data and focus on decision-relevant information.[2]

This investigation of the applicability of fuzzy logic to the temporal discrimination problem is executed by converting a key portion of the existing threat warning system to fuzzy logic and analyzing the effects of fuzzy logic on detection performance. Specifically, three primary tasks are performed. First, a fuzzy logic subsystem is designed to replace a subset of the sequential classifiers in the current system that are temporally related. This includes the design of a fuzzy rule base that captures the expert knowledge about the system and exploits the temporal information available to each classifier. Next, the impact of the fuzzy logic subsystem on classification performance and early warning capability is evaluated. And finally, conclusions are drawn as to the overall applicability of fuzzy logic to the entire threat warning system. These three tasks are presented in greater detail in later sections.

Research indicates that this is a novel application of fuzzy logic techniques to real-time temporal classification for threat warning systems. This investigation should advance our understanding of the applicability of these techniques in terms of probability of target detection, probability of false alarm, and earlier warning capability.

1.3 Summary of the Results

The application of the fuzzy logic subsystem to the discrimination problem has shown a moderate improvement in detection performance and early warning capability. The procedure used to characterize this performance is summarized below, followed by some specifics of the performance improvement obtained.

Test data are available for a variety of engagements with threatening targets and other sources that are known to be non-threatening. Each engagement (either a threat or a non-threat) is processed independently. Detection performance is measured as the percentage of threats correctly classified as threats (probability of detection) and the percentage of non-threats incorrectly classified as threats (probability of false alarm). These statistics are gathered only for the threat and non-threat signals that are actually processed by the subset of existing classifiers selected for this investigation. Due to the temporal nature of the processing structure, some signals are never processed by these classifiers, and thus are not considered during performance characterization. In addition, these performance statistics characterize only the "local" performance of the se-

[1] Citations such as this are in reference to the bibliography at the end of this report.

[2] Attributed to Lotfi A. Zadeh

lected classifiers and the fuzzy logic subsystem. They do not reflect the overall "global" performance of the entire system.

The detection performance for the selected subset of classifiers is known for the existing algorithms. Specifically, the probability of detection (P_D) is 0.82 for the test data sets. This means that 82 percent of the threats that were processed by the selected classifiers were successfully detected. The probability of false alarm (P_{FA}) is 0.40 for the same data sets, indicating that 40 percent of the non-threat were incorrectly classified as threats.

The detection performance for the fuzzy subsystem shows moderate improvement in these statistics. With the same probability of detection, the fuzzy subsystem can reduce the probability of false alarm to 0.22, cutting the number of false alarms almost in half. On the other hand, by allowing the same number of false alarms, the fuzzy subsystem can increase the probability of detection to 0.91. This improvement in performance for the fuzzy subsystem is presented in more detail in section 5 in terms of receiver operator (ROC) curves and other performance metrics for the fuzzy logic subsystem.

In addition, an improvement in early warning capability was demonstrated by the fuzzy subsystem. Warning time is a measure of the time between the detection of a threat and the estimated time of impact (or closest point of approach). An average warning time improvement of about 6.7 percent was observed for the successfully detected threats. Even small increases in warning time may prove to be significant for certain threat warning applications. Also, it is recognized that warning times are inherently limited by the temporal characteristics of the evolving threat signals. These and other performance issues will be discussed in more detail in sections 4 and 5.

This investigation has demonstrated that by applying fuzzy logic to a key portion of the temporal discrimination problem, moderate performance improvements can be achieved. These improvements are primarily due to the ability of fuzzy logic to represent and exploit complex temporal knowledge fundamental to the threat warning system in a meaningful and efficient manner. Further improvements are anticipated by converting the entire system to a fuzzy logic paradigm.

2 Background

2.1 Details of the Problem

2.1.1 Temporal Aspects of the Threat Warning System

The threat warning system uses multiple electro-optic sensors to detect threatening targets at the longest possible range to provide the earliest warning time so that the appropriate countermeasure action can be taken. Each sensor is a wide field of view staring electro-optic receiver that collects radiation in a specific waveband and converts it to digital form. The result from each sensor is a one-dimensional time series signal. Figure 1 shows a signal generated from a simulated threat for a single sensor.

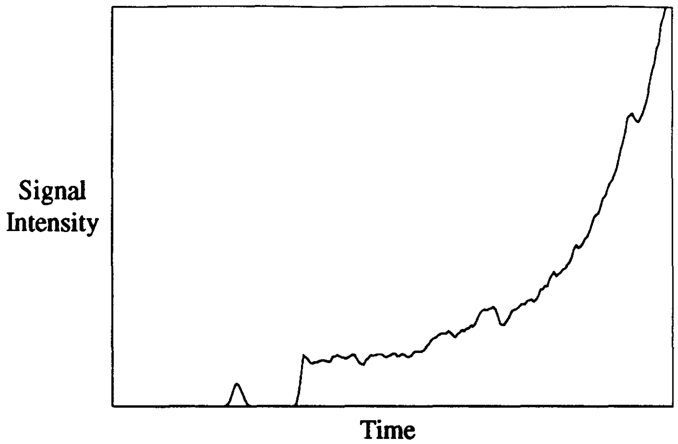

Figure 1. Simulated threat signal.

The signal from each sensor is processed independently by the discrimination algorithms to determine whether the signal is generated from a threatening target or if the signal is the result of another source that is known to be non-threatening.

Threatening targets are known to emit or reflect varying patterns of energy during the course of their flight. These patterns of energy can be detected by the system's optical detectors and converted to digital signals with characteristic temporal profiles. The level of visible detail and the strength of these threat signals are highly dependent on range to the target. For instance, temporal patterns characteristic of targets early in their flight are only visible for targets initiated at close range. These preliminary patterns are not visible for targets initiated at long range. On the other hand, signals from long range targets exhibit temporal patterns characteristic of targets later in their flight. These temporal patterns tend to increase in strength as the target approaches.

The temporal characteristic of the threat signals is the basis for the design of the threat warning system. The goal of the discrimination algorithms is to accurately detect these temporal characteristics and declare an alarm with sufficient warning time to be able to avoid the threat.

2.1.2 Classification Modes

The existing discrimination algorithms consist of a series of tests that are continuously applied to the signals from each sensor. These tests are designed to detect temporal characteristics of signals from threatening targets at various ranges. Each test is referred to as a classification "mode", or simply a "mode". Modes that are designed to detect close range threats look for temporal patterns that are characteristic of targets early in their flight. Since these targets pose the most immediate threat, close range modes are applied at every sample using full resolution data. Conversely, modes that detect longer range threats look for longer range characteristics that are only visible well into their flight and are applied at a reduced rate using lower resolution data. The temporal nature of this multiple mode discrimination process is designed to detect threats at the earliest possible time given the signal processing resources available.

A total of eight temporal modes are currently used in the threat warning system. If any mode detects a threat, then an alarm is declared and the appropriate countermeasure action is taken. More than one mode may declare an alarm at the same time. If no modes declare an alarm for the current sample, then the system proceeds to the next sample and the appropriate classification modes are re-applied. The decision made by one mode does not effect the decision made by any subsequent mode.

Each classification mode measures a set of predefined signal characteristics, or *features*, that represent the unique temporal patterns for that particular mode. Up to fifteen features may be measured for any one classification mode. Each feature measurement for a particular mode is compared to either a single or a pair of one-dimensional thresholds to determine if that feature measurement falls within the expected limits for the corresponding temporal characteristic established for threat signals. All feature measurements for a particular mode must pass their respective threshold tests in order for an alarm to be declared by that mode. If any single feature measurement fails its threshold test then that sample is rejected as a possible threat for that classification mode.

This series of classification mode tests can be thought of as a cascade of binary decisions where each classification mode thresholds a set of features measurements to determine whether the signal represents a threat or a non-threat. The binary decision made by one mode does not effect the decision made by any subsequent mode. However, some signal characteristics measured for close range modes are used as prerequisites for applying tests for longer range modes.

2.1.3 Data Base Training

The thresholds applied to each feature are determined *a priori* through years of training and analysis by an experienced engineering team. For these purposes, an extensive database of actual field data containing both threats and non-threats plus numerous simulated threat signals is maintained and available for this investigation.

Field data are gathered from digital recorders connected to the threat warning system during actual flight tests. These flight tests consist of a coordinated set of threat engagements with a variety of target types, approach vectors, and atmospheric conditions. In addition, data are collected for a number of non-threatening sources that are known to produce signals that are similar to those produced by threatening targets. These non-threat signals are intended to stress the discrimination capability of the threat warning system.

Numerous simulated threat signals are available from a sophisticated set of target, atmospheric, and sensor modeling utilities. Target models are based on spatial, spectral, and temporal characteristics measured for a variety of known targets. Flight profiles for threatening targets are controlled by kinematic models that simulate a variety of realistic approach paths. Atmospheric transmission models are designed to simulate various absorption, scattering, and radiance levels of the target models. The threat warning sensors are modeled to gather target radiation from a variety of moving platforms. Noise is added to the simulated data by modeling the digital signal as a Poisson process. The combination of these modeling utilities provides a matrix of simulated threat signals for various target, atmospheric, and sensor modeling parameters.

Each set of data contains an engagement of either a single threat or one or more non-threat sources. The duration of each engagement varies. Some data sets may contain less than five hundred samples, while data sets for longer flights may contain more than five thousand samples for each sensor. Up to four sensors may be active during any particular engagement. The signals from each sensor are gathered concurrently. Ancillary information about each data set, including source designations, flight dynamics, number of sensors, and threat range and closing time, is available for each engagement as *ground truth* data.

Table 1 summarizes the data sets available for this investigation.

CATEGORY	*# THREATS*	*# NON-THREATS*
Field Data	61	286
Simulated Data	1936	-
TOTAL	1997	286

Table 1. Available test data sets.

2.2 Need for an Alternative Method

2.2.1 Growing Complexity of System Requirements

Performance demands on the threat warning system have become increasingly complex as more target designations and non-threatening sources are now available and longer detection ranges are required for earlier warning times. Fortunately, advances in hardware technology are also available to provide higher processing throughput. However, it has been necessary to incorporate additional classification modes and feature measurements into the discrimination algorithms in an attempt to handle the added complexity of the system requirements.

Original algorithm designs employed only three classifiers that measured a small number of characteristic features to detect a relatively few target types. The current system applies up to eight classifiers to each signal with anywhere from six to fifteen features measured for each mode. Several of these classifiers attempt to capture some of the temporal characteristics of the signal by retaining features measured during previous time intervals. These temporal features are typically time tags of when a particular temporal event was previously detected. The result is a complex discrimination process that attempts to detect temporally evolving signals in a high dimensional decision space.

The current approach to this complex system is difficult to model and audit in a purely numeric framework. Not only does this complexity limit performance, but periodic revisions to the existing algorithms in response to new data or additional system requirements are increasingly difficult to manage. Minor changes in feature characteristics or threshold levels can have an unpredictable impact on the overall response of the system. Any change to the temporal nature of the discrimination process would require a major restructuring of the algorithms. A new approach is sought.

Performance limitations are primarily due to the fact that the current discrimination algorithms have two inherent problems. These two problems, coined the "binary decision" problem and the "brick wall" problem, are discussed in the next two sections, followed by a description of the governing design philosophy for the threat warning system that is partially responsible for its past success in real-world applications.

2.2.2 The "Binary Decision" Problem

The classification result for any particular mode is a binary decision based on a series of one-dimensional tests applied to the set of features measured for that mode. Each feature measurement is independently compared to either a single or a pair of thresholds to determine if the feature falls within the expected limits established for a threat signal at the appropriate range. If any single feature measurement fails its threshold test then the classifier rejects the current sample as coming from a threat signal for that mode, regardless of how threat-like the other feature measurements appear. Only if all feature measurements pass their respective threshold tests is a threat declared for that mode.

This results in an overall decision boundary that can be described as a "hyper-rectangle" aligned with the coordinate axes of the feature space. In many cases this may be highly suboptimal since this decision function does not exploit possible relationships between the features measurements within each mode. For instance, two features that exhibit a stronger linear correlation for threat data than for the non-threat data may offer more discrimination power when using a quadratic decision boundary that better reflects this relationship[1]. A rectangular boundary may exhibit substantial "holes" in the decision surface. Figure 2 shows a scatter plot of two features for a particular classification mode measured for the threat data with a three-sigma Gaussian ellipse superimposed. Clearly a rectangular decision surface is suboptimal in this case.

[1] A more classic approach to this problem was addressed by the author in a 1993 spring term project for *Pattern Analysis and Pattern Recognition* at MIT in which signal patterns were classified by applying multivariate two-class discriminant functions for Gaussian densities at each mode.

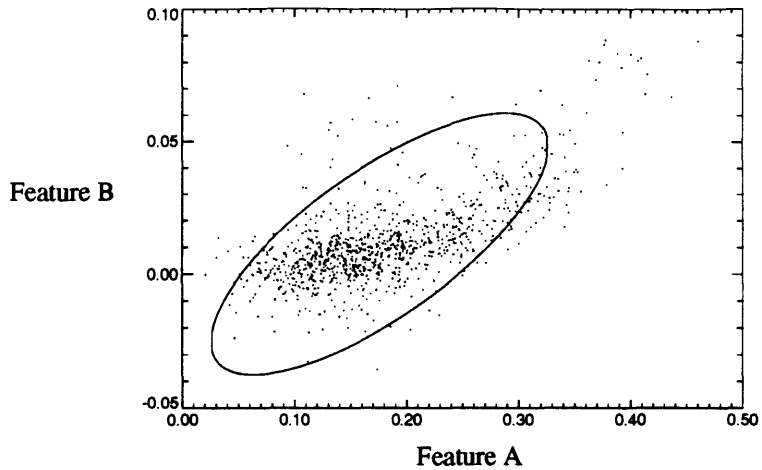

Figure 2. Distribution of two correlated features for a particular classification mode with a three-sigma Gaussian ellipse superimposed.

2.2.3 The "Brick Wall" Problem

The second inherent problem is that classification modes do not exploit temporal information obtained during previous classification modes. Mode decisions are strictly binary so that no partial or "shaded" information from previously applied modes is available to help detect temporal trends and mode transitions for the current mode. Some classifiers do measure the time at which a particular temporal feature was previously detected. But again, this is based on a binary decision as to whether or not the particular event occurred, not the degree to which it may have been detected.

This temporal limitation is referred to as the "brick wall" problem, which clearly limits the robustness of the discrimination algorithm. For instance, suppose one particular classification mode is close to declaring an alarm in that only a few features do not completely satisfy the thresholds for that mode. And suppose that the subsequent classification mode is even closer to declaring an alarm. It may be that this combination is indicative of a temporal trend that is indeed worthy of an early alarm declaration, yet the current algorithms are not capable of declaring an alarm in this situation.

2.2.4 Design Philosophy

The design philosophy of this system has been and will continue to be based on maintaining a thorough understanding of the physical meaning behind the discrimination process. That is to say, any decision criteria should be based on measurements and thresholds that can be related to underlying physical phenomena. The idea here is to design the system using our knowledge base rather than relying solely on the available database so as to avoid training on artifacts that may be present within the data and to provide the system with manageable growth capabilities when presented with new target designations, longer detection ranges, and earlier warning requirements. Although our database is quite extensive, data for certain classification modes is fairly sparse and can not be considered truly representative.

The continuation of this design philosophy is one reason why techniques in fuzzy logic are being pursued over other more conventional methods such as those based on probability theory and those using neural networks. These methods are subject to data dependencies and lack the physical traceability that this system requires.

2.3 *Benefits of Fuzzy Logic*

Fuzzy logic is most useful for complex systems that are not easily described by an exact numerical model and that must respond to imprecise or ambiguous real-world information in a controlled and meaningful manner. Fuzzy logic provides a simple but robust method of representing and implementing expert knowledge about a particular system. It logically combines structured knowledge that can be expressed in linguistic terms with a powerful interpolative mechanism that simultaneously evaluates multiple rules to various degrees of satisfaction. These rules are designed to capture expert knowledge about the system and are represented in simple linguistic terms. In this way, fuzzy systems are easily managed and understood so as to provide an efficient and robust classification system.

The benefit of applying fuzzy logic to the temporal discrimination problem for threat warning is that the complex temporal relationships for this system can be easily represented and manipulated. Signal characteristics relating to real-world physical concepts that are measured by this system are inherently imprecise. The current system treats these characteristics as numeric quantities and applies strict thresholds to make binary decisions. These decisions are clearly lacking in robustness. Fuzzy logic, on the other hand, provides the ability to represent and manipulate these physical characteristics as fuzzy concepts that capture the meaning of the characteristics but tolerate a certain degree of imprecision. Furthermore, these concepts can be logically related using a set of fuzzy rules that combine common sense and expert knowledge about the threat warning system. These fuzzy rules are designed to capture the complex temporal relationships that are characteristic of signals generated from real target threats.

In this way, fuzzy logic provides a powerful technique for temporal discrimination that is based on the underlying physics behind the signals being processed. This approach also provides future growth potential when the system is faced with additional system requirements. New threat data with possibly more complex temporal relationships may be handled in the fuzzy logic framework by modifying the fuzzy rule base and, if necessary, adding new fuzzy concepts to the fuzzy data base. This growth potential is critical for threat warning systems that must provide a long term growth capability.

3 Principles of Fuzzy Logic

Following a brief history of fuzzy logic, a discussion of the basic concepts of fuzzy set theory and fuzzy logic that are applicable to the temporal classification problem is presented. Specifically, the concepts of *fuzzy sets* and *linguistic variables*, *fuzzy relations* and *fuzzy implications*, and *fuzzy rules* are discussed. A more detailed discussion of fuzzy logic may be found throughout the literature. (See Kandel, [1982] and Dubois, Prade and Yager, [1993] for voluminous bibliographies in fuzzy sets and fuzzy logic applications.)

3.1 Brief History

The birth of fuzzy logic can be attributed to Lotfi A. Zadeh in 1965 when he published his seminal paper on fuzzy sets, the first to generalize the concept of a characteristic function of a set to allow for continuous degrees of membership. Zadeh's work is responsible for stimulating extensive interest in fuzzy logic over the last two decades with research and applications spanning many disciplines.

The application of fuzzy set theory to pattern recognition problems was originally discussed in the 1966 paper by Bellman, Kalaba, and Zadeh which dealt with interpolation of fuzzy set membership functions for decision making processes. The concept of a fuzzy partition for pattern analysis was introduced by Ruspini in 1969, and was expanded to cluster analysis by Bezdek and Dunn in their landmark work on fuzzy ISODATA and the fuzzy *c*-means algorithms in the early seventies. In 1979, S. K. Pal applied fuzzy sets to problems in speech recognition. Subsequently, Bezdek and Pal have been major contributors to the theory and application of fuzzy models in image processing.

Tremendous interest in applying fuzzy logic to marketable control systems has grown throughout Europe and Japan in the last decade, pioneered by work in approximate reasoning by Mamdani and his colleagues, and is more recently becoming more accepted in the United States. Their work and the work of Rutherford, Ostergaard, and Tanaka have demonstrated the utility of fuzzy control systems in home appliances, automotive electronics, manufacturing robotics, and automatic train operation, to name only a few. Stability of fuzzy control systems has been analyzed by Gupta *et al* in the mid-eighties. Nonlinear fuzzy control was advanced by the work of Sugeno and Murakami who applied systematic fuzzy control to the parking of a model car. Recent advances have been made by Kosko in adaptive fuzzy systems that modify the fuzzy rule base by learning from examples.

Fuzzy systems have recently become more practical with the development of the fuzzy chip by Togai and Watanabe in 1985, followed by Yamakawa's complete fuzzy controller hardware system in 1986. Voluminous bibliographies in fuzzy sets and fuzzy applications have been complied in many sources including those by Dubois, Prade and Yager [1993] and Kandel [1982]. Zadeh himself has published more than a hundred papers on the subject since 1965.

3.2 Fuzzy Logic Concepts

The following discussion on fuzzy concepts is broken down into sections on fuzzy sets and linguistics variables, fuzzy relations and fuzzy implications, and fuzzy rules. Each section is an overview of broad areas in the discipline of fuzzy logic, but should be sufficient for establishing the necessary framework for this thesis project.

3.2.1 Fuzzy Sets and Linguistic Variables

Let U be a collection of objects with some common characteristic that defines the *universe of discourse* (or *domain*) for the objects with a generic element u so that $U = \{u\}$. The elements of U may be discrete (and finite) or continuous (and infinite). Often, as in our case, the universe of discourse will be the real numbers. A fuzzy set A defined in a universe of discourse U is characterized by its *membership function* μ_A which maps each point in U to a real number in the interval [0, 1] such that the value of $\mu_A(u)$ represents the degree of membership of u in the set A. A fuzzy set may be viewed as a generalization of the ordinary (classical or *crisp*) set whose membership function only takes on the two values $\{0, 1\}$. A fuzzy set A is completely determined by the set of tuples

$$A = \{(u, \mu_A(u)), \forall\, u \in U\}$$

and, depending on whether U is continuous or discrete, can be written concisely as follows:

U continuous: $$A = \int_U \mu_A(u) \,/\, u$$

U discrete: $$A = \sum_{i=1}^{n} \mu_A(u_i) \,/\, u_i$$

where the symbol / denotes a tuple. For example, a fuzzy set A representing the set of real numbers approximately equal to 5 may have a membership function defined as

$$\mu_A(u) = \frac{1}{1 + (u - 5)^2}$$

where $U = \Re$, the set of real numbers. More typically, membership functions are chosen from a standard set of parametric functions that are either increasing, decreasing, or approximating (see Figure 3).

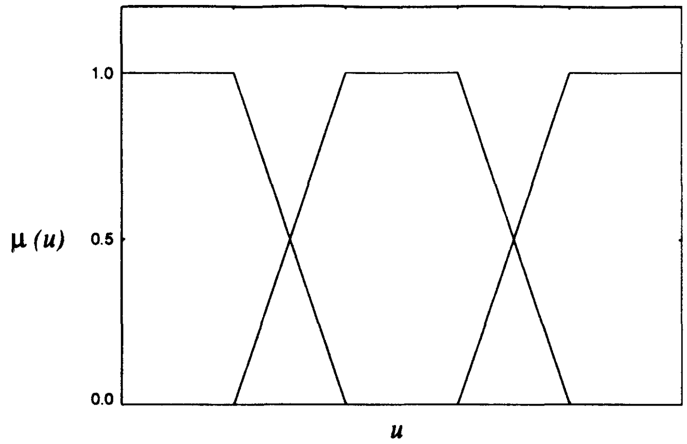

Figure 3. Three piecewise linear membership functions. Left: *Decreasing.* Middle: *Approximating.* Right: *Increasing.*

For this investigation, the following two parametric forms for piecewise linear (generalized trapezoid) and second order membership functions are considered:

piecewise linear:

$$\mu_\rho(u;a,b,c,d) = \begin{cases} 0 & \text{for } u \le a, \\ (u-a)/(b-a) & \text{for } a < u \le b, \\ 1 & \text{for } b < u < c, \\ (u-d)/(c-d) & \text{for } c \le u < d, \\ 0 & \text{for } d \le u \end{cases}$$

second order:

$$\mu_\sigma(u;a,b,c,d) = \begin{cases} 0 & \text{for } u \le a, \\ 2[(u-a)/(b-a)]^2 & \text{for } a < u < (a+b)/2, \\ 1-2[(u-b)/(b-a)]^2 & \text{for } (a+b)/2 \le u < b, \\ 1 & \text{for } b \le u \le c, \\ 2[(u-d)/(c-d)]^2 & \text{for } c < u \le (c+d)/2, \\ 1-2[(u-c)/(c-d)]^2 & \text{for } (c+d)/2 < u < d, \\ 0 & \text{for } d \le u, \end{cases}$$

where the parameters (a, b, c, d) are nondecreasing (*i.e.*, $a \le b \le c \le d$). Figure 4 shows examples of fuzzy membership functions defined using these parametric forms.

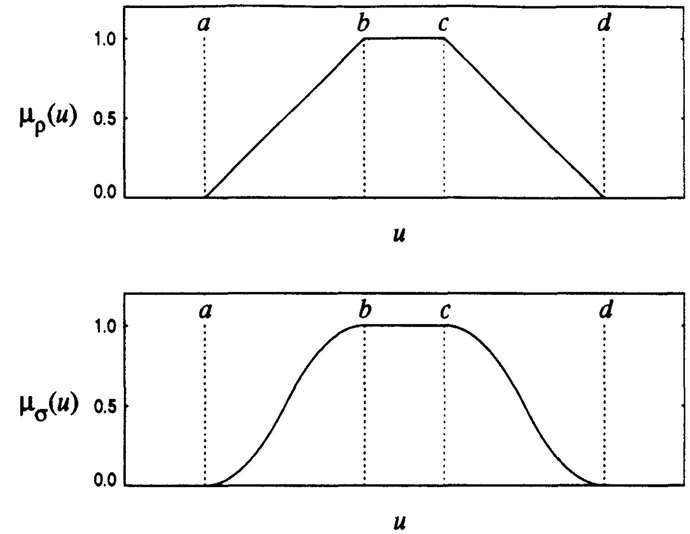

Figure 4. Two membership functions defined in parametric form.
Top: *Piecewise Linear*. Bottom: *Second Order*.

Often fuzzy sets are used to represent imprecise linguistic notions like "tall", "thin", and "young" on the domains *Height*, *Weight*, and *Age* respectively. These domains represent the physical measurement space for what is known as a *linguistic variable*, the fundamental unit of knowledge representation in fuzzy logic.

The linguistic variable is a variable whose values are words or phrases in a natural language and can be characterized by a tuple (x, T_x, U_x, M_x) where x is the name of the variable, T_x is the *term set*, or set of linguistic values that x can take on, U_x is the physical domain over which x is defined, and M_x is a semantic function which gives "meaning" to the variable by mapping each term to its corresponding fuzzy set representation via a membership function. For example, the linguistic variable *temperature* t could be characterized by a term set T_t = {cold, cool, comfortable, warm, hot} where each term in T_t is represented by a fuzzy set defined over the domain U_t = [−10°C, 40°C] and the meaning of the function M_t is represented by membership functions as shown in Figure 5.

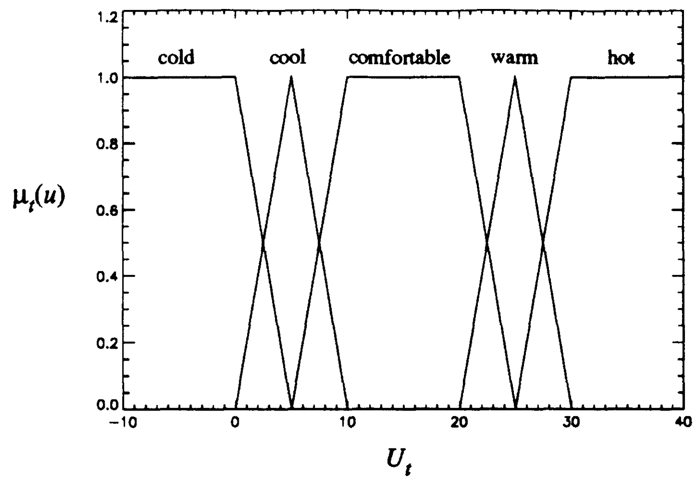

Figure 5. Possible term set for linguistic variable representing temperature.

Basic operations on fuzzy sets, such as *union*, *intersection*, and *complement*, can be defined in terms of their membership functions in a variety of ways. Let A and B be two fuzzy sets with membership functions μ_A and μ_B, respectively, over a universe of discourse U. Then, Zadeh proposes a simple extension of the classical set operations as follows:

union:	$\mu_{A \cup B}(u) = \min(\mu_A(u), \mu_B(u))$,	$\forall\, u \in U$,
intersection:	$\mu_{A \cap B}(u) = \max(\mu_A(u), \mu_B(u))$,	$\forall\, u \in U$,
complement:	$\mu_{A'}(u) = 1 - \mu_A(u)$,	$\forall\, u \in U$.

Other extensions are possible as well and can be generalized into a family of *triangular norms* and *co-norms* that satisfy certain properties for each operation (Weber, [1983]).

With these set operations, it is easy to show that fuzzy sets follow most of the same basic properties as do classical sets, such as *commutativity*, *associativity*, *distributivity*, *DeMorgan's laws*, etc. (Zadeh, [1965]). Because of this, classical sets can be considered a special case of fuzzy sets. (One property that is not shared is the *excluded middle law*, since fuzzy sets, and their complements, are allowed to overlap.)

3.2.2 Fuzzy Relations and Fuzzy Implications

An important concept in fuzzy set theory is the *fuzzy relation*, another extension from classical set theory. An n-ary fuzzy relation is a fuzzy set that relates n fuzzy variables in the product space $(U_1 \times \cdots \times U_n)$ of the variables. For example, let the binary fuzzy relation R define the fuzzy notion "approximately equal to" over the space $U \times U$ with U representing the continuous interval [0, 100]. Then a possible membership function for R may be defined as

$$\mu_R(u,v) = \begin{cases} 1 - |u-v|/10, & \text{for } |u-v| \le 10, \\ 0, & \text{for } |u-v| > 10, \end{cases}$$

which can be expressed parametrically as $\mu_R(u,v) = \mu_p(|u-v|;-10,0,0,10)$, a piecewise linear triangle. Thus, the truth value of the relationship u is "approximately equal to" v is 1 when $u = v$, linearly decreases towards 0 as their absolute difference approaches 10, and is 0 when their absolute difference is more than 10.

Fuzzy relations are important because they can describe interactions between variables and in particular can represent the meaning behind fuzzy *inference rules*, which will be used in our fuzzy discrimination system. Two inference rules of particular importance in approximate reasoning are the *generalized modus ponens* and the *compositional rule of inference.* The generalize modus ponens has an if-then inference structure represented symbolically as

$$\begin{array}{l} x_1 \text{ is } P_1 \\ \text{if } x_1 \text{ is } Q_1 \text{ then } x_2 \text{ is } Q_2, \\ \therefore\ x_2 \text{ is } P_2, \end{array}$$

where x_1 and x_2 are objects and P_1, P_2, Q_1 and Q_2 are fuzzy object properties.

The compositional rule of inference, on the other hand, uses a fuzzy relation to represent the if-then rule (and thus is considered a superset of the generalized modus ponens rule of inference) and can be described symbolically as

$$\begin{array}{l} x_1 \text{ is } P_1, \\ x_1\, R\, x_2 \\ \therefore\ x_2 \text{ is } P_2, \end{array}$$

where the second line reads "x_1 is in relation R to x_2" and R is characterized by its membership function μ_R. This combination of fuzzy sets and fuzzy relations is called *composition*, denoted by the symbol o, and is performed with the aid of the fuzzy operations *conjunction* (or *cylindrical extension)* and *projection*, which basically map the domains of fuzzy sets onto higher and lower dimensional spaces so that inferences can be made (Zadeh, [1977], Dubois and Prade, [1980], Kandel, [1982]).

The composition operation has two forms. If A is a fuzzy set defined on U and R is a fuzzy relation defined on $U \times V$, then the composition of A and R resulting in a fuzzy set B defined on V is $B = A \circ R$. This is done by taking the intersection of R and the cylindrical extension of A onto $U \times V$, and projecting the result onto V. With the projection operation typically taken as the supremum function (SUP) for continuous domains (or the maximum function for discrete domains), composition is referred to as the *sup-star* composition where "star" denotes a fuzzy intersection operation (Lee, [1990a]). If the minimum operator is used for intersection , then we have *sup-min composition* with

$$\mu_B(v) = \underset{u \in U}{\mathrm{SUP}}\ (\mathrm{MIN}\ (\mu_A(u), \mu_R(u,v))),$$

and if the product operator is used, then we have *sup-product composition* with

$$\mu_B(v) = \underset{u \in U}{\mathrm{SUP}}\ (\mu_A(u) \cdot \mu_R(u,v))).$$

Other interpretations of the compositional rule of inference are found in the literature but the sup-min (attributed to Zadeh) and the sup-product are the most frequently used due to their computational advantages (Lee, [1990b]).

The inference mechanism considered for this fuzzy system is based on a set of rules that take the form of the fuzzy implication "if p then q" or "$p \rightarrow q$", where p is the rule *antecedent*, which is a (possibly compound) fuzzy proposition, and q is the rule *consequent*, also a fuzzy proposition. These fuzzy implications describe the causal relationship between the system input measurement variable(s) and the output decision variable, and can be described by a suitable fuzzy relation. There are a variety of ways of interpreting the meaning of these fuzzy implications (Driankov, *et al*, [1993]). Four of the most common methods used in fuzzy decision systems will be described in terms of their representative fuzzy relations.

Consider the inference rule "if X is A then Y is B" where the fuzzy sets A and B are characterized by their membership functions μ_A and μ_B respectively. Two fuzzy implications are based on the classical equivalence $p \rightarrow q \equiv \neg p \vee q$. The *Kleene-Dienes implication* interprets the "or" operation as the maximum of p and q, and is described by the fuzzy relation R_D which is characterized by its membership function

$$\mu_D(x,y) = \text{MAX}\ (1 - \mu_A(x), \mu_B(y)).$$

Similarly, the *Lukasiewicz implication* interprets the "or" operation as the bounded sum so that the relation R_L is characterized by

$$\mu_L(x,y) = \text{MIN}\ (1, 1 - \mu_A(x) + \mu_B(y)).$$

The *Zadeh implication* uses the equivalence $p \rightarrow q \equiv (p \wedge q) \vee \neg p$ to define the relation R_Z characterized by

$$\mu_Z(x,y) = \text{MAX}\ (\text{MIN}\ (\mu_A(x), \mu_B(y)), 1 - \mu_A(x)).$$

Two of the better known implications in fuzzy logic literature are the *Mamdani implication* and the *Gödel implication.* The Mamdani implication, which uses the equivalence $p \rightarrow q \equiv p \wedge q$ to define the relation R_M, is characterized by

$$\mu_M(x,y) = \text{MIN}\ (\mu_A(x), \mu_B(y)),$$

and is popular due to its computational simplicity. The Gödel implication defines the relation R_G characterized by the membership function

$$\mu_G(x,y) = \begin{array}{ll} 1 & \text{for } \mu_A(x) \leq \mu_B(y)), \\ \mu_B(y) & \text{otherwise.} \end{array}$$

Dubois and Prade [1980] establish an ordering for fuzzy implications in terms of their strength, which gives us $R_L \supseteq R_D \supseteq R_G \supseteq R_M$ and $R_L \supseteq R_Z \supseteq R_M$, so that R_M (the Mamdani implication) is the strongest of these implications. Choice of these implications is dependent on how the inference mechanism is used to represent the set of fuzzy rules in the system's rule base.

3.2.3 Fuzzy Rules

As mentioned previously, the inference mechanism for this system is based on a set of rules that describe the relationship between the input measurements and the output decision variable. These rules are derived from expert knowledge about the system and its environment. Multiple-input single-output (MISO) rules take the form

$$\text{if } x_1 \text{ is } A_{k,1} \text{ and } x_2 \text{ is } A_{k,2} \text{ and ... and } x_m \text{ is } A_{k,m} \text{ then } y \text{ is } B_k, \quad k = 1,\ldots,n,$$

for n rules with m input fuzzy variables x_i having fuzzy values $A_{k,i}$ and one output fuzzy variable y having fuzzy value B_k. The rule antecedent (x_1 is $A_{k,1}$ and x_2 is $A_{k,2}$ and ... and x_m is $A_{k,m}$) forms a fuzzy set in the product space $U_{k,1}\times\cdots\times U_{k,m}$ for the $A_{k,i}$. The rule consequence (y is B_k) forms a fuzzy set in V_k for B_k. These n rules are represented as fuzzy relations using a suitable implication strategy (*e.g.*, the Mamdani implication) and can be combined to produce a single inference mechanism in a variety of ways. For instance, *composition based inference* first combines all rules into a single overall relation R_m and then "fires" the relation with the fuzzy input variables via a sup-star composition operator. Another way is *individual rule based inference* which first "fires" each rule individually and then combines the resulting fuzzy sets into a single output. It has been shown that for the Mamdani type of inference these two methods are equivalent (Lee, [1990b], Driankov, *et al*, [1993]).

The development of the fuzzy rule base is a critical step in the design of a robust fuzzy system. Although there is no formal development methodology for a fuzzy rule base, several properties of the set of rules are considered. These properties include *completeness*, *consistency*, *continuity*, and *interaction* (Driankov, *et al*, [1993]), which are discussed in subsequent sections.

4 Research

4.1 Investigation Procedure

The current system uses a set of eight temporal classification modes to perform threat versus non-threat target discrimination. For this investigation, a subset of these classification modes is converted to a fuzzy logic subsystem and integrated with the other modes of the current system. This new system is referred to as the *fuzzy system*. Based on the relative performance of this fuzzy system and the original system, the applicability of fuzzy logic is assessed. The following sections discuss the procedures for this investigation in more detail.

4.1.1 Classification Mode Selection

Two sequential classification modes are selected for this investigation. Specifically, *classification mode 4* and *classification mode 5* are used for several clear reasons. First, mode 4 and mode 5 are two of the key modes that are applied during the temporal evolution of the medium to long range target signals. They measure strong signal characteristics that are evident at most detectable ranges. These signal characteristics, however, are similar for many non-threatening sources, which can cause a false alarm problem for these modes. Application of fuzzy logic to these modes will attempt to reduce these false alarms.

Secondly, the classification structure for modes 4 and 5 are very similar. In fact, the same features are measured for both modes. These features measure signal strength and various shape characteristics over a specific (finite) time interval. The difference between the two modes is that they apply a different set of thresholds as part of the classification tests. Mode 4 is designed to detect closer range threats than mode 5. As such, the thresholds for mode 5 tend to be more restrictive than those for mode 4, since the signals from longer range threats must match more closely the target model to keep the false alarm rate down. This similarity in classification structure easily lends itself to a combined fuzzy logic subsystem.

Thirdly, the design of these two modes is based on the detected occurrence of a previous temporal event. In fact, the application of mode 4 and mode 5 is mutually exclusive. If the occurrence of this temporal event is detected within some prior time interval, then classification mode 4 is applied. If, on the other hand, the temporal event is detected prior to this time interval or not detected at all, then classification mode 5 is applied. The detection of this temporal event is binary. Features that measure the strength and shape of the desired temporal event, as well as the error in these measurements, are tested against nominal thresholds. If these tests are satisfied then the temporal event is declared and its time is retained for the subsequent choice between the application of mode 4 or mode 5 classification tests. The features used to measure the occurrence of this temporal event can be incorporated into the fuzzy subsystem. By allowing for a continuous degree of occurrence of the temporal event, the fuzzy subsystem provides a continuous transition between mode 4 and mode 5 classification criteria.

For these reasons, classification mode 4 and classification mode 5 are selected for this investigation into the applicability of fuzzy logic. The next section discusses the selection of features for fuzzy logic processing. This is followed by a section on the implementation of the fuzzy logic subsystem and how it is integrated with the rest of the system. Finally, a description of the evaluation strategy used for this thesis project is presented.

4.1.2 Feature Selection

As discussed in above, classification mode 4 and classification mode 5 measure the same set of features. Specifically, eight features are measured by the original classifiers. Four of these features are incorporated into the fuzzy logic subsystem as fuzzy variables. These four features, referred to as $\{x_1, x_2, x_3, x_4\}$, measure specific signal characteristics related to signal shape and feature measurement quality.

The other four features are not incorporated into the fuzzy logic process and remain subjected to the same strict threshold tests as in the original system. This is because two of these features measure very general characteristics that must be evident in the signal for a threatening target to be present. Loose thresholds are applied to these features to eliminate those signals which do not exhibit these general characteristics. A third feature measurement is tested simply to ensure that the computational integrity of the measurements of the other features is intact. A fourth feature is not incorporated into the fuzzy logic process because it represents the discrete time at which the contingent temporal event (described in the previous section) was previously detected. This feature is replaced in the fuzzy subsystem by two new features, $\{x_5, x_6\}$, that measure the size and shape of the temporal event in a continuous manner.

Thus, a total of six features are incorporated into the design of the fuzzy logic subsystem. These features define the input vector $\underline{x} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ used for this investigation.

4.1.3 Fuzzy Logic Subsystem Implementation and Integration

A software simulation of the original system has been developed in the FORTRAN programming language. All test data sets are digitally stored in binary data files that are directly accessible by the FORTRAN simulation. Due to the graphical interface and feedback conducive to the development of fuzzy systems, the fuzzy logic subsystem for this investigation is developed in a visualization and analysis environment implemented in the C programming language.

The fuzzy logic subsystem directly replaces the mode 4 and mode 5 classification modules. The other six classification modules are integrated with the fuzzy logic subsystem to form a fuzzy simulation used for this investigation. In addition, this simulation is integrated with all relevant post-processing and input/output modules from the original system. The fuzzy simulation is designed such that with the appropriate "crisp" definition of fuzzy set membership functions and the proper fuzzy rule base, identical results can be obtained as from the original system.

Various graphical interface utilities are embedded in the fuzzy simulation as well. These utilities are described in more detail in section 4.5.1

4.1.4 Evaluation Strategy

Evaluation of the applicability of fuzzy logic techniques to temporal discrimination is performed by comparing the classification results from the original mode 4 and mode 5 classifiers with the classification results of the fuzzy logic subsystem. These results are considered "local" results in that they do not measure the overall system performance. An evaluation of these local results allow conclusions to be drawn as to the "global" applicability of fuzzy logic techniques to the entire threat warning system.

The test data sets are split into two classes, namely, the *threat* class and the *non-threat* class. Each class is processed separately by both the original system simulation and the fuzzy system simulation. The simulations have the ability to capture local result statistics. Specifically, the number of correctly classified threats and the number of incorrectly classified non-threats are measured. In addition, detection times are recorded for early warning time assessment. Performance metrics are discussed in more detail in section 4.5.2.

Each data set consists of up to four signals, one for each active sensor, for a particular engagement. Each data set in the threat class contains the signal (or signals) for a single threat engagement. Each data set in the non-threat class contains the signal (or signals) for at least one non-threatening source. No single data set contains both a threat and a non-threat source. As shown in table 1, 1997 threat data sets and 286 non-threat data sets are available. The classification systems are applied to the samples in the data sets at the appropriate sample rate.

The threshold tests applied to the four features not incorporated into the fuzzy logic subsystem, as described in section 4.1.2, are used to define the data sets used for accumulating local performance statistics. Since these tests are common to both the original mode 4 and mode 5 classification tests and the fuzzy classification tests, data sets that do not pass these cursory tests are locally rejected by both systems. Of the 1997 threat data sets, 1651 contain signals that satisfy this criteria for at least one sample. Of the 286 non-threat data sets, 86 contain signals that also satisfy this criteria. Performance measures are only applied to these restricted data sets. In this way, the performance measures are further localized to the common processes that are relevant to this investigation.

Local performance statistics count the number of data sets for which a threat was declared. If multiple threats are declared for a particular data set, then that data set is counted as being a threat declaration exactly once. If a threat is detected in a particular data set, the earliest time at which it was detected is recorded for warning time assessment. For the original system, local performance statistics count the number of data sets for which threats are detected by either classification mode 4 or classification mode 5. For the fuzzy system, these statistics count the number of data sets for which threats are detected by the fuzzy subsystem.

4.2 System Design

The principal components of a fuzzy logic subsystem are *fuzzification*, the *inference engine*, and *defuzzification* as shown in the block diagram in Figure 6. These components interact with a *knowledge base* which contains a *data base* of membership functions and normalization factors, and a *rule base* of fuzzy decision rules. The input to the fuzzy subsystem is a feature measurement vector of crisp values. The output of the subsystem is a single crisp value indicating the degree of alarm to be reported. Design procedures for these components shall be summarized in the next few sections.

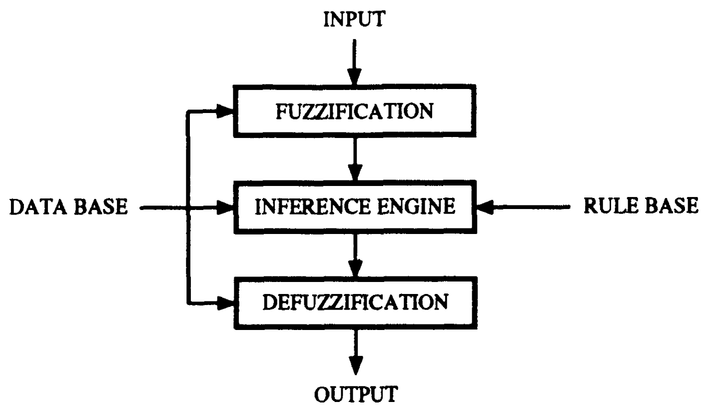

Figure 6. Fuzzy logic subsystem block diagram.

4.2.1 Fuzzification

The purpose of the fuzzification module is twofold. First, the input vector is optionally transformed from its physical domain into an arbitrary normalized domain. Input normalization is usually, although not necessarily, a linear scaling. This normalization can simplify the definition of fuzzy membership functions at the possible expense of removing the design parameters from a more familiar physical domain. For this investigation, input normalization is not performed since the expert knowledge about the system has historically been expressed in terms of the physical domain of the input features.

Secondly, each crisp value is converted to suitable fuzzy set representation of the rule base antecedents. This fuzzification of the inputs will be done in one of two ways, depending on the type of inference mechanism selected. If composition based inference is used, then the crisp value must be converted into a fuzzy set so that the fuzzy relation representing the meaning of the overall set of rules can be stimulated by the fuzzified input using a composition operator. In this case, a crisp input value $*x$ is converted to a singleton fuzzy set S represented by a membership function $\mu_S(x)$ equal to 1 at the point $*x$ and 0 everywhere else. (Note that crisp values shall be preceded by asterisk superscripts to distinguish them from fuzzy values.)

If individual rule based inference is used, each rule is fired to the degree determined by how well the crisp input values match the rule's antecedent. This is accomplished by evaluating the membership functions for each of the fuzzy sets in the antecedent at the crisp input values (*i.e.*, $\mu_A(*x)$ for fuzzy set A). How these two methods utilize the fuzzified inputs and combine their results is discussed in the next section.

4.2.2 Inference Engine

There are two basic approaches used in the design of the inference engine in a fuzzy discrimination system, namely, composition based inference and individual rule based inference. The fuzzy inference engine applies the fuzzified inputs to the antecedents of the rules in the rule base using the selected inference mechanism. The rule consequences are then inferred and combined into a single fuzzy set that represents the response of the system to the fuzzy inputs.

Let the relation R_k represent the following compound implication that describes the k-th rule in the fuzzy rule base:

$$\text{if } x_1 \text{ is } P_{k,x,1} \text{ and } x_2 \text{ is } P_{k,x,2} \text{ and ... and } x_m \text{ is } P_{k,x,m} \text{ then } y \text{ is } P_{k,y},$$

where the x_i's and y are fuzzy variables and the $P_{k,x,i}$'s and $P_{k,y}$ are respective fuzzy properties of the fuzzy variables.

Using the Mamdani interpretation of this implication, for example, the membership function characterizing relation R_k is

$$\mu_k(\underline{x},y) = \text{MIN}\,(\mu_{k,x,1}(x_1),\, \mu_{k,x,2}(x_2),\, ...,\, \mu_{k,x,m}(x_m),\, \mu_{k,y}(y))),$$

where $\mu_{k,x,i}(x_i)$ is the membership function describing the i-th proposition in the antecedent of the k-th rule and $\mu_k(\underline{x},y)$ is shorthand for $\mu_k(x_1,x_2,...,x_m,y)$. If a composition base inference engine is used, then all the rules are combined into a single relation representing the meaning of the whole set of rules and this relation is fired by the fuzzified input vector via the composition operator. Under the Mamdani interpretation, the relations describing the individual rules are combined using a union operator to form the overall relation

$$R_{CB} = \bigcup_{k=1}^{n} R_k.$$

Using sup-min composition, the resulting membership function characterizing R_{CB} can then be expressed as

$$\mu_{CB}(y) = \underset{k}{\text{SUP}}\,(\text{MIN}\,(\mu_{k,x,1}(^*x_1),\, \mu_{k,x,2}(^*x_2),\, ...,\, \mu_{k,x,m}(^*x_m),\, \mu_{k,y}(y))),$$

where the $\mu_{k,x,i}(^*x_i)$'s are the fuzzified inputs. Other continuous (or discrete) interpretations can be used as well.

If individual rule based inference is used, each rule is fired individually by computing the degree of match between the crisp input values $\{^*x_i\}$ and the fuzzy sets $\{\mu_{k,x,i}(x_i)\}$ representing the rule's antecedent. Using the Mamdani interpretation, the result is a fuzzy set representing the rule's consequence that has been clipped to the degree to which the rule's antecedent has been matched by the crisp input. The resulting sets for all the rules are then combined using a suitable union operator. If the union is performed with the supremum operator, then the relation representing the individual rule based inference is characterized by the same membership function

$$\mu_{IRB}(y) = \underset{k}{\text{SUP}}\,(\text{MIN}\,(\mu_{k,x,1}(^*x_1),\, \mu_{k,x,2}(^*x_2),\, ...,\, \mu_{k,x,m}(^*x_m),\, \mu_{k,y}(y))),$$

indicating that under the Mamdani interpretation, composition based inference and individual rule based inference are equivalent. (For a proof of this equivalence, see Lee [1990b].)

Given this equivalence, which eliminates the possibility of rule interaction (Hellendoorn [1992]), and considering its computational simplicity, the Mamdani implication, as described above, is selected as the basis for the fuzzy inference engine used for this investigation.

4.2.3 Defuzzification

The defuzzification module is responsible for two functions. First, the output of the inference engine must be *defuzzified*, or converted from a fuzzy set representation to a single crisp output value. For the fuzzy subsystem designed for this investigation, each fuzzy rule applied by the fuzzy inference engine produces a fuzzy consequence set that is clipped to the level at which the rule's antecedent is satisfied. These clipped fuzzy sets are combined to form an overall output fuzzy set representing the degree of alarm indicated for the given input vector. The defuzzification process converts this output fuzzy set to a single crisp output value that indicates whether or not a threat detection should be declared. A graphical representation of this defuzzification process is shown in figure 7. Here, three rules have been partially satisfied, resulting in three clipped consequence sets, and the crisp value (*y) is computed using a particular defuzzification method.

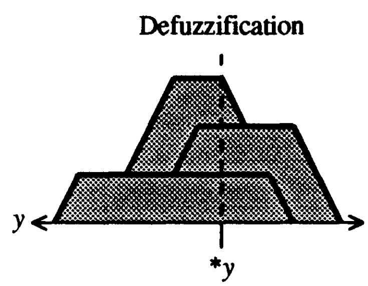

Figure 7. Illustration of the defuzzification Proc-

A variety of defuzzification methods are discussed in the literature, including *center-of-area*, *center-of-sums*, *center-of-largest-area*, *first-*, *middle-*, *last-*, and *mean-of-maxima*, and ***height*** defuzzification. Although choice of which defuzzification method to use is highly dependent on the individual application, it is suggested that several important defuzzification criteria be considered (Driankov, *et al*, [1993]). For instance, the defuzzification procedure should be *continuous*, so that small changes in the input should not produce large changes in the output. Defuzzification should also be *unambiguous* so that a unique crisp value can be determined from any fuzzy set. *Computational complexity* is an obvious consideration for some systems. Another important criteria is *weight counting*. Weight counting is the property that dictates how overlapping consequence sets are aggregated into a single crisp output (Driankov, *et al*, [1993]). The issue here is whether or not overlapping areas are counted more than once during defuzzification. For instance, the center-of-area method computes the centroid of the clipped consequence fuzzy sets by counting overlapping areas only once (no weight counting). The center-of-sums method, on the other hand, computes a weighted average of all the overlapping consequence sets regardless of whether or not they overlap (weight counting). Satisfaction of the weight counting property is dependent on how the if-then rules are defined in the fuzzy rule base (see section 4.4).

For this investigation, the center-of-sums defuzzification method is used. The center-of-sums method satisfies the weight counting property. This is desirable since the the particular form of the fuzzy rules design for this fuzzy subsystem expect the consequence of each rule to contribute to the overall results individually. Thus, the overlapping areas must be counted separately.

Three other desirable properties are satisfied by the center-of-sums defuzzification method. For instance, this method is continuous, so that small changes in the position and height of the clipped consequence sets produce small changes in the crisp value of the defuzzified output variable. This provides a certain degree of consistency in alarm declaration for the temporal discrimination process. The center-of-sums method is also unambiguous, so that the crisp output value is always unique. And finally, the center-of-sums method is computationally more efficient than, for example, the center-of-area method.

For piecewise linear membership functions defined in the parametric form $\mu_\rho(u;a,b,c,d)$, described above, the consequence fuzzy sets can be represented by a generalized trapezoidal function. Computing the center-of-sums for these trapezoids involves first measuring the center-of-gravity for each trapezoid as

$$COG_i = \frac{1}{3} \times \frac{a^2 + b^2 + ab - c^2 - d^2 - cd}{a + b - c - d}.$$

Then to compute the overall center-of-sums, the individual centers-of-gravity are weighted by their clipped heights (h_i) and accumulated as

$$COS_i = \frac{1}{3} \times \frac{\Sigma\, h_i (a^2 + b^2 + ab - c^2 - d^2 - cd)}{\Sigma\, h_i (a + b - c - d)},$$

where, the summations are computed over all the consequence sets.

The second function of the defuzzification module is to optionally transform the output value from a normalized domain to its actual physical domain. For this system, the domain of the output variable is alarm level, which can be defined on an arbitrary scale. Specifically, alarm level is defined on a continuous scale of [0, 1], where a value near 0 indicates a very low alarm level and a value near 1 indicates a very high alarm level (see section 4.3). Thus, denormalization of the defuzzified output is not performed for this investigation.

4.2.4 Knowledge Base

The *knowledge base* of the fuzzy logic subsystem consists of a *data base* and a *rule base*. The purpose of the knowledge base is to provide the necessary information for the proper functioning of the fuzzy system. The design of both the data base and the rule base is subjective in nature and requires experienced engineering knowledge of the how the classification system is to function. Application of this knowledge should adhere to the threat warning design philosophy presented in section 2.2.4 by exploiting the underlying phenomenology behind the discrimination process. Several important aspects of this design are considered during development of the knowledge base, as indicated below. The specific design of the knowledge base for this investigation is presented in sections 4.3 and 4.4.

A) Data Base

The data base contains the information necessary to characterize the imprecise knowledge used in the fuzzy system. In particular, the data base defines the fuzzy sets representing the meaning of the linguistic input and output variables in terms of their membership functions. In this investigation, these membership functions are defined for each fuzzy set in the term set of a linguistic variable using the parametric forms described previously in section 3.2.1. For each fuzzy set, the parameters $\{a,b,c,d\}$ control the shape and position of the membership function for that fuzzy set along the (normalized) domain of the linguistic variable. The number of terms in the term set is chosen to represent the desired granularity, or resolution, for a particular variable and is usually determined heuristically. The total number of terms in the input space and the output space defines the maximum number of rules that can be constructed in the rule base. For example, for a system with three input variables with two, five and six terms respectively, and one output variable, then the maximum number of rules would be $2\times5\times6 =$ 60 rules. However, in practice it is not necessary to completely "cover" the rule space since some antecedent term set combinations may either never occur or are considered unimportant (Kosko, [1992]). The *support* (or width) of each fuzzy set can be related to the measurement noise associated with that variable. Other membership function design parameters include *peak value*, *degree of overlap* (defined in terms of *cross-point level* and *cross-point ratio*), and *symmetry*. These issues are considered during the data base design described in section 4.3 below.

B) Rule Base

The rule base contains the set of rules that describe the relationship between the input variables and the output variable. These multiple-input-single-output (MISO) if-then rules take the form

$$\text{if } x_1 \text{ is } P_{k,x,1} \text{ and } x_2 \text{ is } P_{k,x,2} \text{ and ... and } x_m \text{ is } P_{k,x,m} \text{ then } y \text{ is } P_{k,y},$$

for the k-th rule, as describe above. Rules are typically stated in linguistic (rather than numerical) terms selected from the term sets of the linguistic variables defined in the fuzzy data base. There are a variety of ways that the rule base can be developed. Most of these are based on common sense and engineering judgment derived from expert knowledge about the system. Self-organizing systems have been demonstrated that exhibit a human-like learning ability to create and modify rules based on system performance (Sugeno, [1985]). The design of this discrimination system capitalizes on the extensive experience gathered over a period of twelve years by a crew of engineers who have an intimate understanding of how the system responds to the available test data. Rule base design is described in section 4.4 below.

4.2.5 The Fuzzy Logic Subsystem

A graphical representation of the principal components of the fuzzy logic subsystem are shown in Figure 8 on the following page. This simplified system uses a fuzzy inference engine based on the Mamdani implication method. Three multiple-input-single-output if-then fuzzy rules are fired simultaneously. Each fuzzy rule contains two antecedent propositions and a single output proposition, or consequence. The fuzzy rules describe the casual relationships between two input fuzzy variables and an output fuzzy variable. Two crisp inputs are fuzzified according to the membership functions representing the meaning of the input fuzzy variables. The degree to which each rule antecedent is satisfied is computed using the minimum operator for the Mamdani-based inference engine. The degree of satisfaction of the entire rule antecedent determines the degree to which the corresponding consequence is satisfied. This output degree of satisfaction is represented graphically as a clipped version of the membership function representing the meaning of the output fuzzy variable. All three rules, in this case, have varying degrees of output satisfaction. A single crisp output value is obtained from a defuzzification process that combines the clipped fuzzy membership functions using some centroid technique. This crisp output represents the response of the fuzzy system to the two crisp inputs.

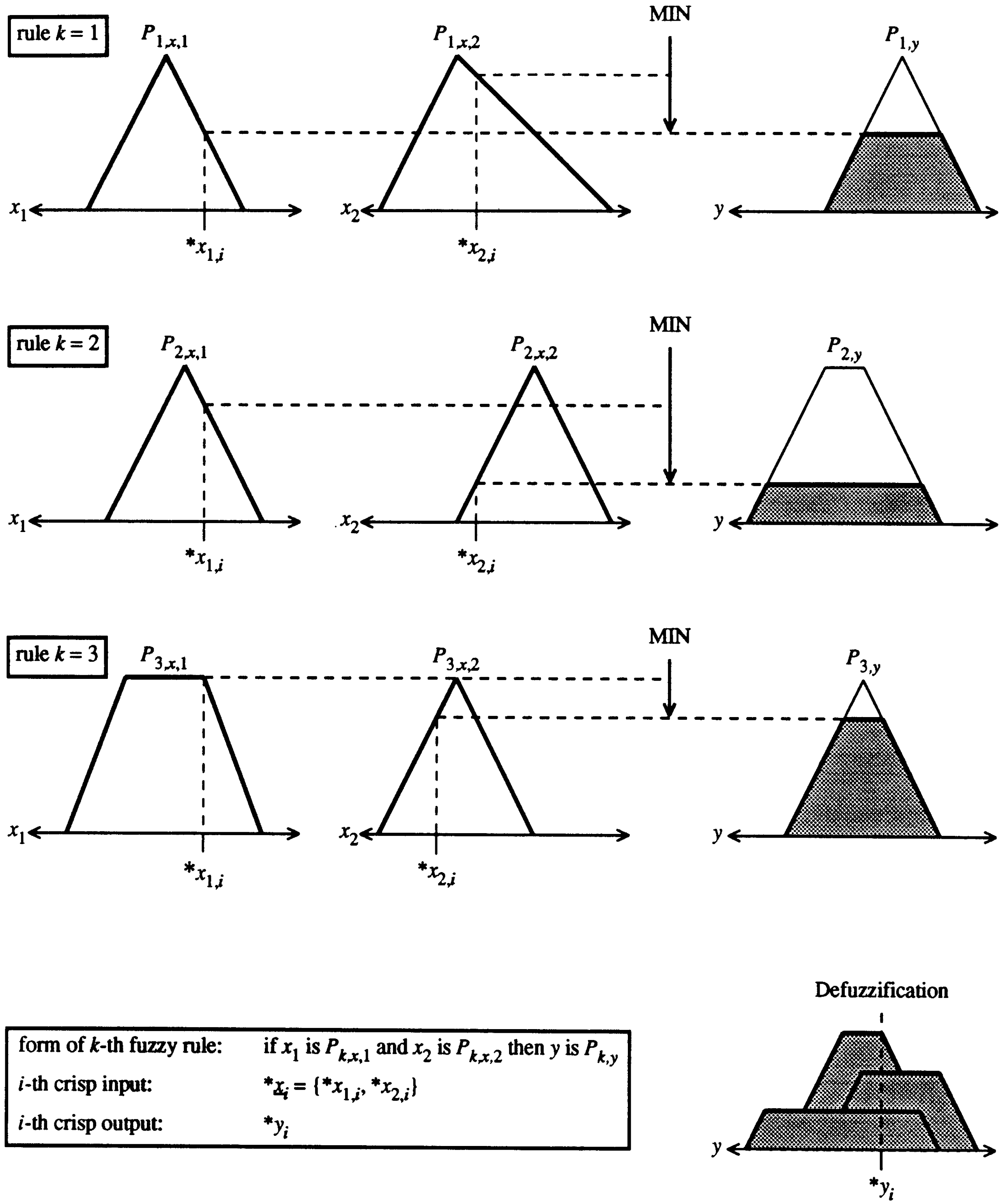

Figure 8. Graphical representation of Mamdani inference mechanism showing three MISO rules each with two fuzzy antecedent propositions and one clipped fuzzy consequence.

4.3 Data Base Design

The data base design strategy is based on expert knowledge developed during years of analysis on the discrimination problem for threat warning systems. This experience has led to the definition of the features used in the existing discrimination algorithms and the establishment of the thresholds designed to test these features for threat-like characteristics. The fuzzy data base design is based on these features and feature thresholds.

A linguistic variable is defined for each input and output feature used by the fuzzy subsystem. In particular, six input linguistic variables (corresponding to the input features $\underline{x} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$) and one output linguistic variable (corresponding to the output alarm level) are defined. For each linguistic variable, a set of fuzzy sets are defined to partition the variable's domain into distinguishable fuzzy concepts. These fuzzy sets are represented in terms of their respective membership functions.

The parameters used to define the fuzzy data base include the number of fuzzy sets in the term set of each linguistic variable, the position and general shape of each membership function, and its parametric form. The design of each of these parameters is based on the established thresholds for each feature. Up to four thresholds are defined for each feature to be used by the fuzzy subsystem. This is because each feature may have a minimum and a maximum bound for each classification mode, and each feature is common to both classification mode 4 and classification mode 5 (*i.e.*, they have common feature domains). For a particular such feature, x_i, these thresholds typically partition the feature domain as shown in figure 9,

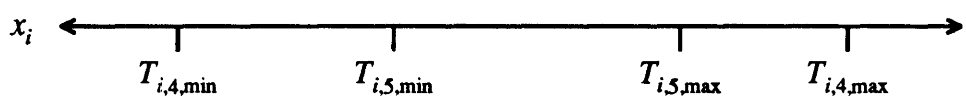

Figure 9. Typical threshold partition for feature x_i for classification modes 4 and 5.

where $\{T_{i,4,\min}, T_{i,4,\max}\}$ are the mode 4 bounds for x_i and $\{T_{i,5,\min}, T_{i,5,\max}\}$ are the mode 5 bounds for x_i. Typically, the mode 5 thresholds are more strict (and thus fall between) the mode 4 thresholds since they represent criteria for which there is no prior (in the temporal sense) collaborating evidence.

A direct conversion strategy is used to define the linguistic variable representing this feature. At most five fuzzy sets are defined in the term set for this linguistic variable, one for each of the possible intervals defined by the threshold partitioning of the feature domain. Each fuzzy set represents a fuzzy concept, or property, that can be associated with this feature. If a mode 4 and mode 5 minimum or maximum threshold coincide, then the corresponding interval does not exists and no fuzzy set is defined for that threshold pair. The resulting number of fuzzy sets provides the necessary and sufficient descriptive power to capture the meaning behind the expert knowledge developed for this system. Additional terms may be added to the fuzzy subsystem, but these four fuzzy sets for each feature represent the level and complexity to which the current system has been designed. This direct conversion strategy provides a strong basis for comparison and evaluation suitable for this investigation.

Another advantage to this direct approach is that by defining the fuzzy sets with "crisp" boundaries, as represented in figure 10(a), identical results can be achieved from the original and the fuzzy system. This is useful for testing purposes during development and for studying the effects of incremental fuzzification of the fuzzy data base for design analysis.

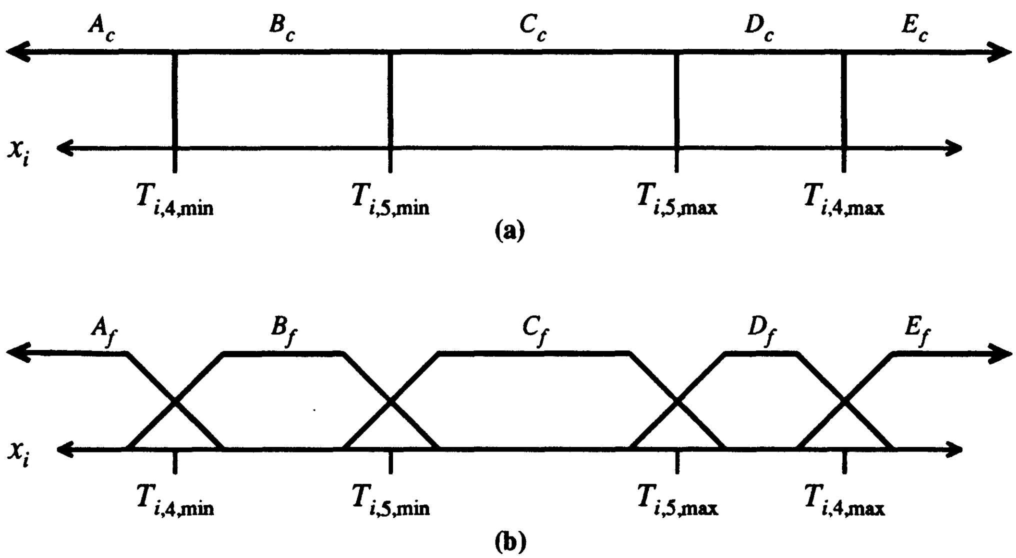

Figure 10. (a) Five "crisp" fuzzy sets (A_c, B_c, C_c, D_c, E_c) corresponding to the threshold partition for feature x_i. (b) Five piecewise linear "fuzzified" fuzzy sets (A_f, B_f, C_f, D_f, E_f) corresponding to the same threshold partition.

Fuzzy sets are "fuzzified" by allowing elements to have variable degrees of membership, as indicated in figure 10(b). Fuzzy sets are allowed to overlap so that a single crisp value may be a member of more than one fuzzy set. The degree of overlap is determined by the position and shape of the fuzzy set membership functions. For this fuzzy subsystem, membership functions for adjacent fuzzy sets intersect at the established threshold values that partition the two sets. This conforms to the direct conversion strategy that partitions each fuzzy domain into distinguishable concepts as determined by the available expert knowledge about the system.

The membership value at this point of intersection is set to 0.5 for all adjacent fuzzy sets so that at least one fuzzy set has a dominant membership value for any one input. (In the extreme case, two adjacent membership functions can both have membership values of exactly 0.5 for a single element.) This also provides a smooth transition in the response of the fuzzy subsystem as inputs transition from one fuzzy concept to another and allows multiple rules to be fired for any given input vector.

The support, or width, of the membership functions controls how "fuzzy" the fuzzy sets are. This can be determined by a combination of a thorough understanding of the behavior of the corresponding feature measurements, an analysis of the noise in the feature measurement (as suggested by Devi and Sarma, [1985]), and experimental analysis. It has been found in this investigation that the support of the fuzzy set membership functions has only a small effect on the performance of the fuzzy subsystem (see section 5.2).

Several fuzzy data bases are defined for this investigation. The i-th fuzzy data base is referred to as Ω_i. Figure 11 shows the membership functions for the input and output fuzzy sets defined for fuzzy data base Ω_1. The intersection points of the membership functions for the six input variables correspond to the thresholds established for the original system. Fewer than five fuzzy sets are defined for the first four variables (x_1, x_2, x_3, and x_4) since some of the mode 4 and mode 5 thresholds coincide. Each fuzzy set has a linguistic name that described the fuzzy concepts referred to by the if-then rules in the fuzzy rule base. The membership functions for the two extreme fuzzy sets "big" and "small" extend towards plus and minus infinity, respectively. The term sets for the two input variables (x_5 and x_6) that characterize the previous temporal event, each contain only two fuzzy sets. These two sets are highly fuzzified to provide a smooth transition between classification mode 4 related tests and classification mode 5 related tests.

The output fuzzy variable for Ω_1 is partitioned into a sufficient number of fuzzy sets to adequately describe the alarm level for each fuzzy rule. These output fuzzy sets are defined to be symmetric about their peak values since the center-of-sums method is used for output defuzzification. If this were not the case, then the defuzzified output value, when a single rule is fired, may not correspond to the peak membership value. This would suggest a lack of *plausibility* in the fuzzy rule base (Driankov, *et al*, [1993]). In addition, the overlap and support of these output fuzzy sets, defined as shown, provides a smooth response to continuously varying inputs. The same output fuzzy sets are defined for all fuzzy data bases Ω_i used in this investigation.

Figure 12 shows an alternate set of membership functions for the same input and output variables that define fuzzy data base Ω_2. The input variables in this case have fuzzy sets with "fuzzier" membership functions. That is, some fuzzy sets have wider support so there is more overlap between adjacent fuzzy sets. No significant difference in detection performance is noticed when using Ω_1 or Ω_2 (see section 5.2).

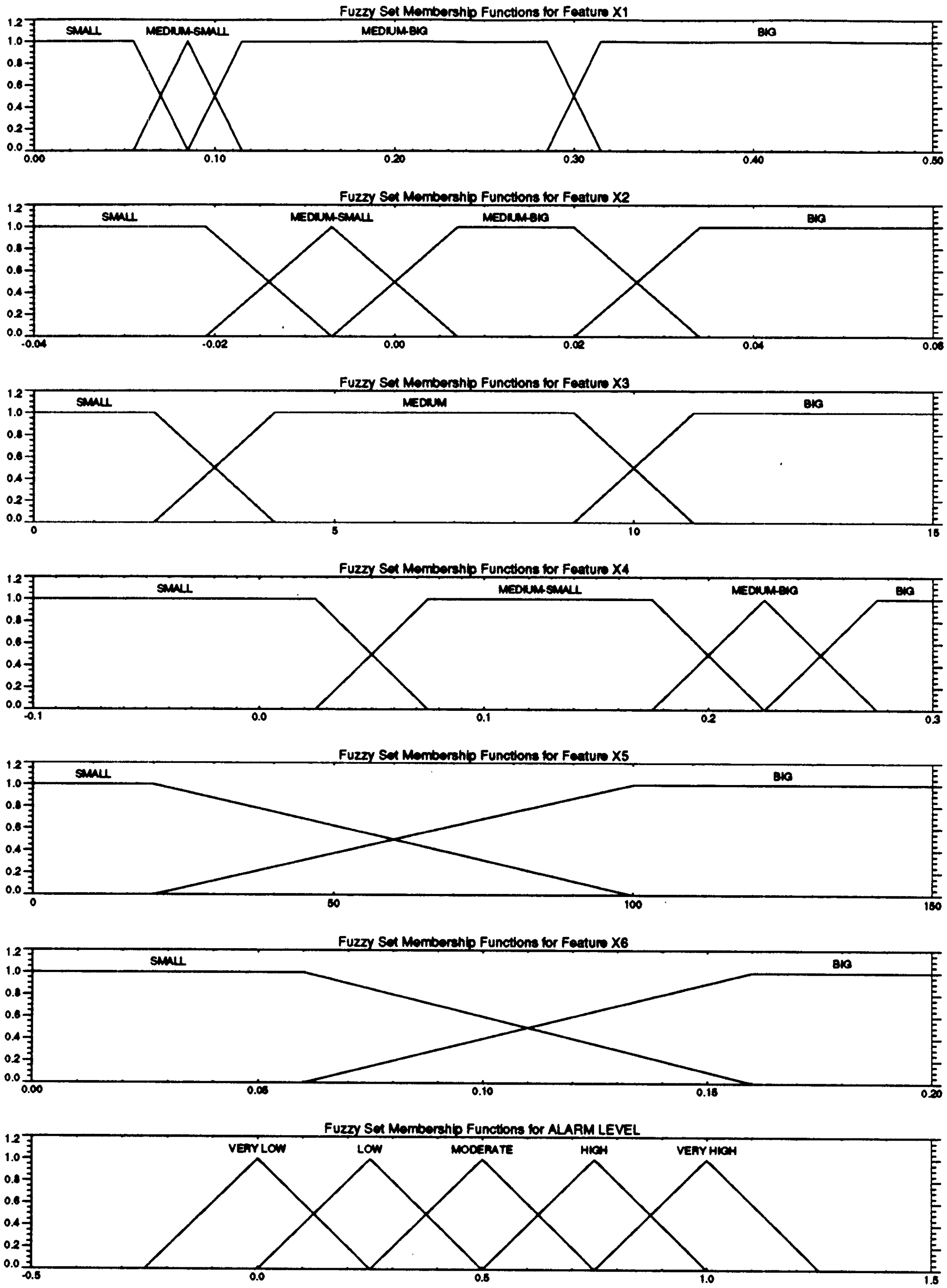

Figure 11. Membership functions for fuzzy data base Ω_1.

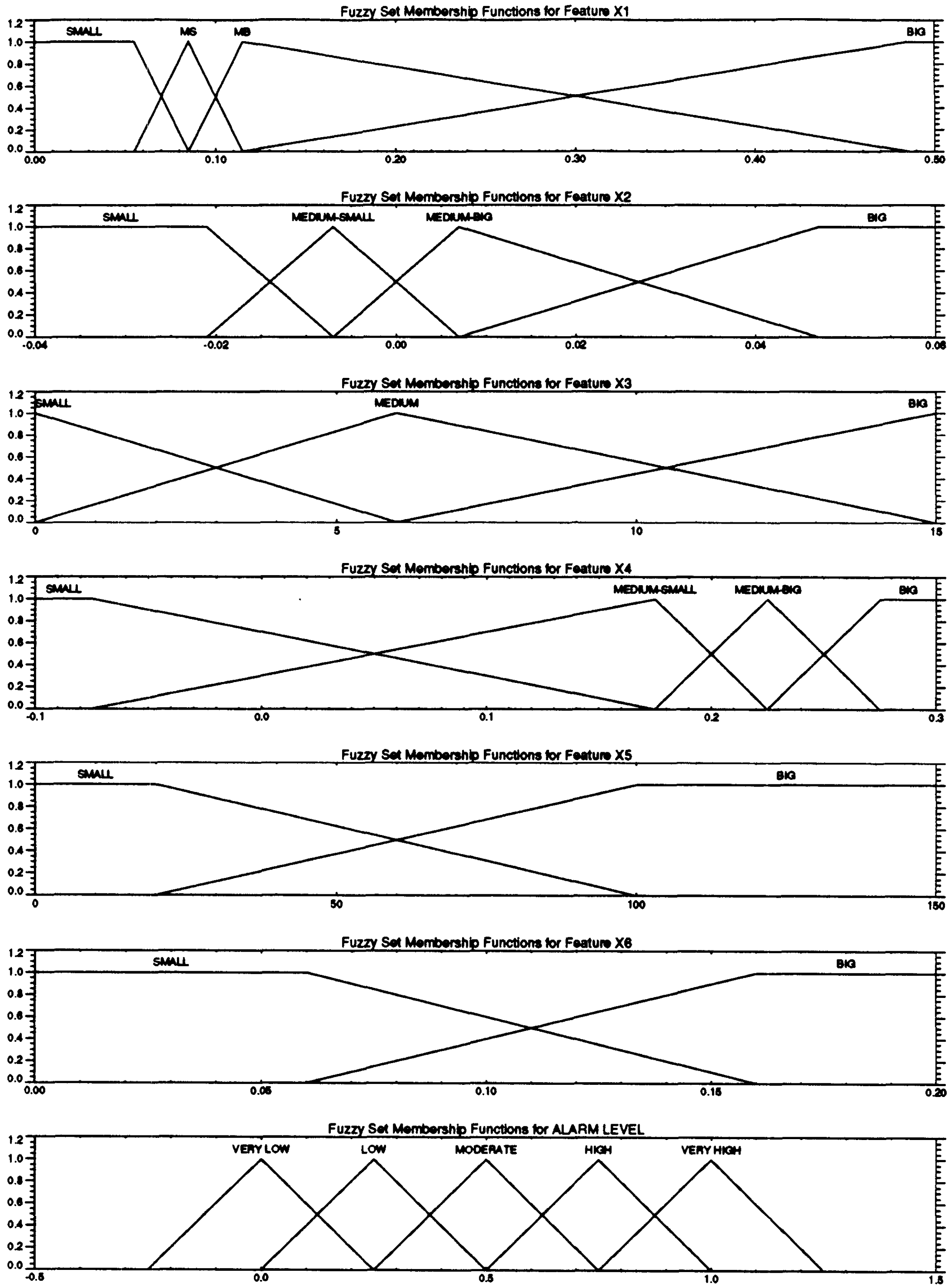

Figure 12. Membership functions for fuzzy data base Ω_2.

Two additional fuzzy data bases, Ω_3 and Ω_4, are defined for performance analysis. The fuzzy set membership functions for these fuzzy data bases are shown in Appendix A. The design of the Ω_3 and Ω_4 data bases is discussed in section 5.3 below.

4.4 Rule Base Design

The design of fuzzy if-then rules for this fuzzy subsystem is based on common sense and engineering judgment derived from expert knowledge about the system. This is a heuristic process with no formal development methodology available[1]. However, several design issues and desirable properties are considered during the design of the rule base. The following sections discuss the strategy used for defining the if-then rules in the fuzzy data base designed for this investigation, the visualization of decision surfaces generated from this rule base, and some basic properties considered during this rule base design.

4.4.1 Rule Definition Strategy

The maximum number of possible rules in the rule base is determined by the number of input variables and the number of fuzzy sets in the term set of each corresponding linguistic variable. For the fuzzy data bases used for this investigation, a total of $4\times4\times3\times4\times2\times2 = 768$ rules are possible. However, a considerable reduction in the total number of rules is achieved by understanding and exploiting the relationships between the input variables and, in particular, their relative importance.

For this fuzzy subsystem, a hierarchical strategy based on expert knowledge about the system is used for developing the fuzzy rule base. It is recognized that two of the input features, specifically x_1 and x_2, are the key features that measure the *shape* of the signal to determine how characteristic it is of signals generated from threatening targets at the appropriate ranges. A set of if-then rules is defined that captures the well understood relationships between these shape features for target ranges considered by each classification mode. For instance, it is known that shape characteristics for threat signals at relatively close range (for classification mode 4 detection) produce "medium" levels for x_1 and x_2 feature measurements. If either feature measurement is "low" or "high", then the shape characteristics indicating the presence of a close range threat is not evident in the signal. These relationships for classification mode 4 criteria are translated into fuzzy rules of the form

> if x_1 is $\text{SMALL}_{x,1}$ and x_2 is $\text{SMALL}_{x,2}$... then y is LOW
> if x_1 is $\text{MEDIUM-SMALL}_{x,1}$ and x_2 is $\text{MEDIUM-SMALL}_{x,2}$... then y is VERY HIGH
> if x_1 is $\text{MEDIUM-BIG}_{x,1}$ and x_2 is $\text{MEDIUM-BIG}_{x,2}$... then y is VERY HIGH
> if x_1 is $\text{BIG}_{x,1}$ and x_2 is $\text{BIG}_{x,2}$... then y is LOW
> ... *etc.*,

where the output variable y represents alarm level and the fuzzy sets ($\text{SMALL}_{x,i}$, $\text{MEDIUM-SMALL}_{x,i}$, $\text{MEDIUM-BIG}_{x,i}$, $\text{BIG}_{x,i}$, LOW, and VERY HIGH) are defined in the fuzzy data bases described above. These relationships are completely described by a set of $4\times4 = 16$ rules for classification mode 4 detection ranges (see table A1). A different set of rules describe the relationship between these two features for classification mode 5 detection ranges.

[1] Several attempts at developing formal rule definition methodologies for fuzzy control systems are discussed in Lee, [1990a].

Additional rules are defined to capture the effects that other features have on the relationships between these two key features. Specifically, feature x_3 measures the *error* in the measurements of the two shape features. For example, a "high" level of x_3 indicates that the x_1 and x_2 values are unreliable. Accordingly, the possibility of a threat signal is less and the resulting alarm level should be reduced. A set of three fuzzy rules are defined to capture this relationship (see table A1). Since the tolerance level in measurement error is lower for longer range targets, only one rule is required for mode 4 detection ranges, while two rules are required for mode 5 detection ranges.

Feature x_4 is used to further refine the decision space for the shape features. In fact, feature x_4 measures a linear relationship between x_1 and x_2 that when satisfied also reduces the probability of a threat being present. Four additional fuzzy rules are defined to capture this relationship (see table A1), one for mode 4 detection ranges and three for mode 5 detection ranges.

Finally, the relationship between classification mode 4 criteria and classification mode 5 criteria is defined in terms of fuzzy sets. Features x_5 and x_6 measure the size and shape of the temporal event that controls the degree to which the mode 4 and mode 5 tests are applied. For relatively close range threats, x_5 tends to be "big" and x_6 tends to be "small," in which case mode 4 tests are emphasized. If x_5 is "small" or if x_5 is "small" and x_6 is also "small," then the system emphasizes classification mode 5 tests. This knowledge is incorporated into the fuzzy rule base by defining one set of fuzzy rules for the relatively close range target conditions and two sets of fuzzy rules for longer range target conditions. Each of these sets of rules captures the knowledge behind the relationships described above between the first four features (x_1, x_2, x_3, x_4). The linguistic variables corresponding to features x_5 and x_6 are highly fuzzified to provide a smooth transition between classification mode 4 detection ranges and classification mode 5 detection ranges.

This hierarchical approach to designing the rule base for this investigation results in $(16+1+1) = 18$ rules for classification mode 4 detection ranges and $2\times(16+2+3) = 42$ rules for classification mode 5 detection ranges for a total of 60 rules used for this investigation. The rules in this fuzzy rule base are articulated in Appendix B.

4.4.2 Rule Base Decision Surfaces

The fuzzy rule base defines a complex decision surface in what is referred to as "fuzzy hyper-space." That is, for each crisp input vector, the fuzzy subsystem responds with a crisp output value indicating the detected degree of alarm. This output value is located at a point in fuzzy hyper-space. The domain of each linguistic variable represents one degree of freedom in this space, for which a continuous decision function is defined. Obviously, this fuzzy hyper-space is impossible to visualize. However, due to the hierarchical strategy used to define the rule base, a meaningful representation of the response of the fuzzy subsystem can be obtained by examining the two-dimensional fuzzy decision surface representing the relationships between the shape features, x_1 and x_2, at discrete values of the other input features. For example, figure 13 shows this decision surface for the shape features for mode 4 detection ranges with low measurement error. The membership functions for the fuzzy sets defined for the x_1 and x_2 linguistics variables are shown along the edges of plot. Each fuzzy set is labeled with a number (to simplify the presentation format) that represents the linguistic name of the fuzzy set in sequence (*e.g.*, 0 represents SMALL, 1 represents MEDIUM-SMALL, *etc.*). The horizontal and vertical lines overlaying the decision surface correspond to the discrete thresholds established for each input feature domain.

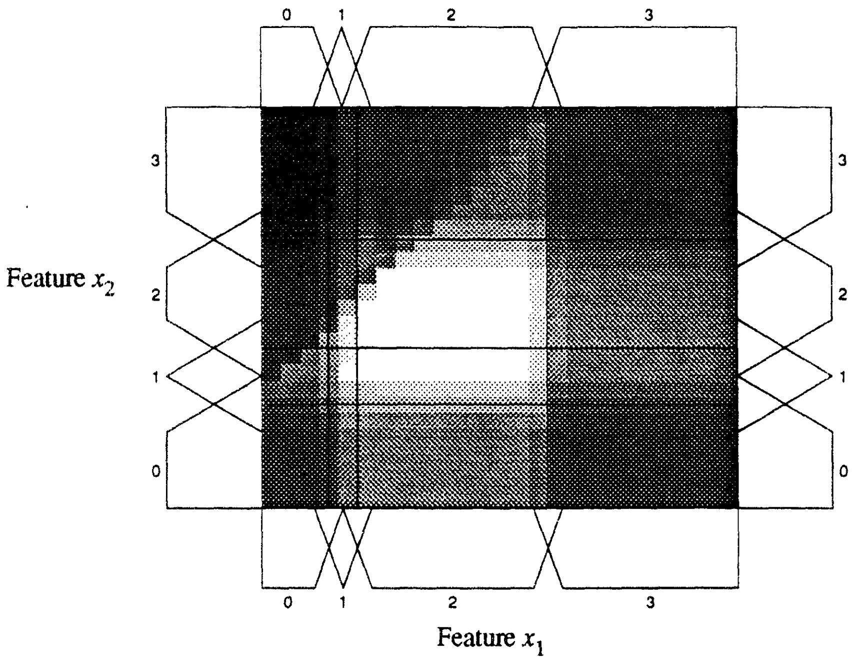

Figure 13. Decision surface for shape features at mode 4 ranges with low measurement error

Figures 14 through 16 show similar decision surfaces for other discrete values of the other input features. As the values of the other features continuously change, the x_1 versus x_2 decision surface smoothly varies in height and shape. These decision surfaces provide an understanding of the system response function in a meaningful format. In addition they illustrate that the rule base designed for this investigation exhibits the properties of *completeness*, *consistency*, and *continuity*.

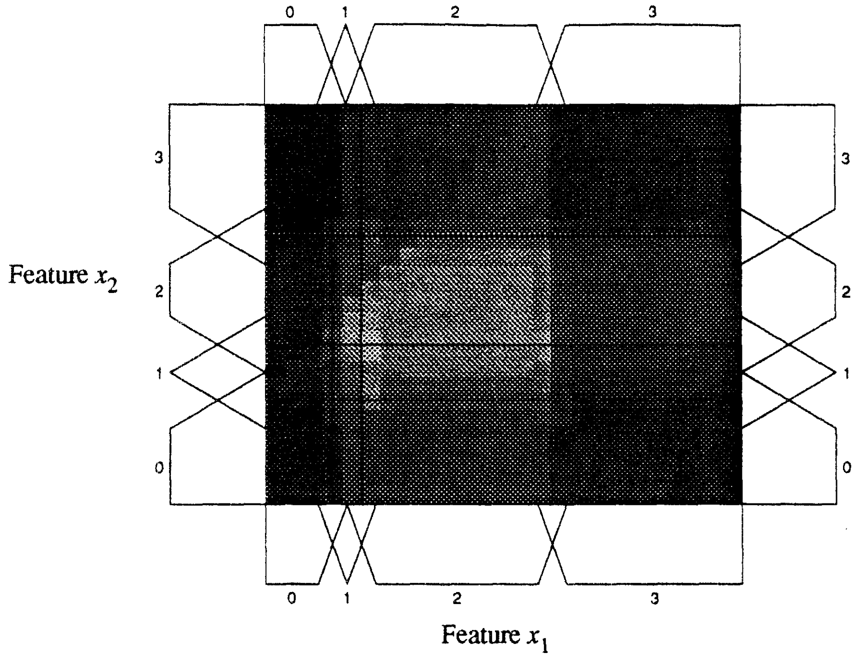

Figure 14. Decision surface for shape features at mode 4 ranges with high measurement error

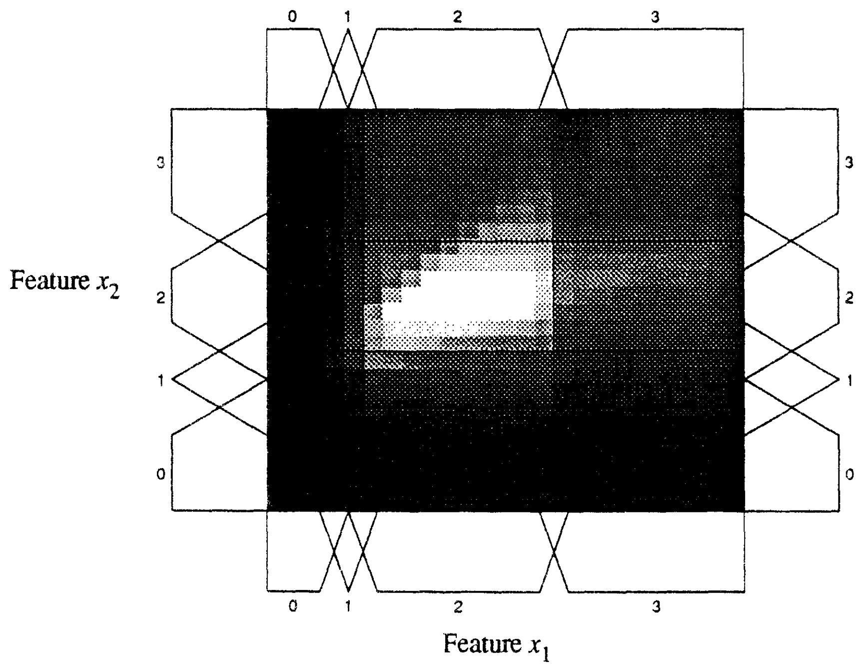

Figure 15. Decision surface for shape features at mode 5 ranges with low measurement error

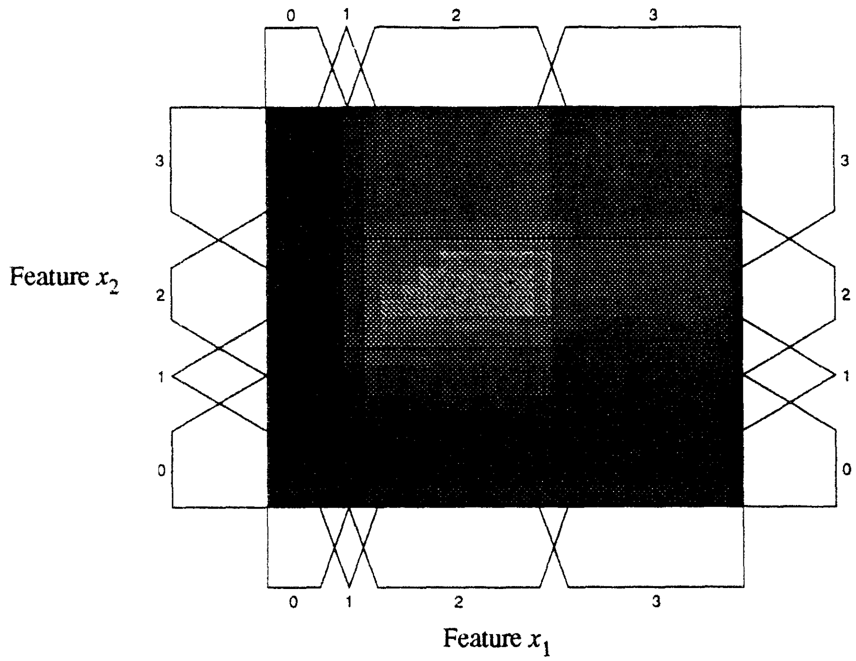

Figure 16. Decision surface for shape features at mode 5 ranges with high measurement error

4.4.3 Rule Base Properties

Completeness ensures that the system fires at least one rule to a dominant degree for any given input so that a decisive response is always produced. That is, the antecedent of at least one rule is satisfied to a degree of at least 0.5 so that the rule consequence is applied to the corresponding degree. This property is ensured by the fact that the fuzzy sets for the x_1 and x_2 input variables have membership functions that intersect at a level of 0.5 so that these two input domains are fully "covered" in the corresponding decision space[1].

The rule base is said to be inconsistent if a contradiction between rules exists. This can only happen when two rules share the same antecedent but have difference consequences. This is clearly not the case for the rule base designed for this fuzzy subsystem. The hierarchical design approach ensures that no two fuzzy rules in the rule base have the same antecedent.

The continuity of the rule base ensures that smoothly varying inputs elicit a continuous response from the fuzzy subsystem. Satisfaction of this property is indicated (although not insured) by the smoothness in the decision surfaces (figures 13 through 16). Note

[1] It is not necessary for the rule base for a fuzzy system to be complete. Due to the fuzziness of the input variables, many fuzzy control systems show no significant reduction in performance using only a subset of control rules.

that the apparent blockiness in the presentation format of these surfaces is simply due to the resolution at which they are displayed.

4.5 Analysis and Evaluation Tools

The following sections discuss the analysis and evaluation tools used for this investigation. These tools include a useful graphical feedback utility that is embedded into the simulation of the fuzzy logic subsystem and several utilities to evaluate the performance of the fuzzy logic subsystem in comparison with the existing algorithms.

4.5.1 Graphical Feedback Utility

Analysis of any fuzzy process is a difficult task. The fuzzy logic subsystem designed for this investigation converts a crisp input vector to a set of fuzzy variables, applies an inference engine that simultaneously fires multiple fuzzy rules, and interpolates the fuzzy result to get a crisp output value. This process is performed for every appropriate sample (*i.e.*, every sample that passes the cursory tests described in section 4.1.4) in each signal. For each of these samples, it is useful to know which rules are fired and to what degree they are satisfied by the inputs. In addition, it is useful to monitor this behavior from sample to sample so as to gain insight into the temporal behavior of the system.

For these purposes, a dynamic graphical feedback utility has been developed for this investigation that is embedded into the simulation of the fuzzy logic subsystem. This utility provides four graphical displays that are updated during the processing of each sample. Figure 17 shows a snapshot of this display capability for a particular sample taken during the processing of a threat signal.

The top display shows the degree to which each rule is satisfied by the current input vector. The x-axis represents the rule number (see table A1) and the y-axis is the degree of membership in the fuzzy relation defined by that particular rule. For this sample, four rules are satisfied to various degrees. The rule numbers indicate that only rules defined for mode 4 detection ranges are satisfied. The information in this display relates directly to the second display which shows the clipped membership functions for the fuzzy sets that represent the consequences of the applied rules. The level to which these membership functions are clipped is equal to the level to which the corresponding rules are satisfied. The vertical dashed line on this display indicates the defuzzified output value for the fuzzy subsystem. This crisp output value represents the current level of alarm indicated by the fuzzy subsystem. The third display shows a time history of these crisp output values over a finite set of samples. This display is useful for monitoring the temporal trend of the alarm level as each sample is processed. The horizontal line represents an arbitrary threshold that could be applied to the alarm level for threat declaration. The bottom display shows a normalized histogram of the rules that are satisfied during the processing of a particular signal. The height of each vertical bar represents the number of times the corresponding rule has been satisfied, normalized to a scale between 0 and 1. The histogram shown in figure 17 indicates that six rules for mode 4 detection ranges and the two feature x_3 and x_4 rules are satisfied at the indicated frequencies. No rules for mode 5 detection ranges are satisfied for this signal.

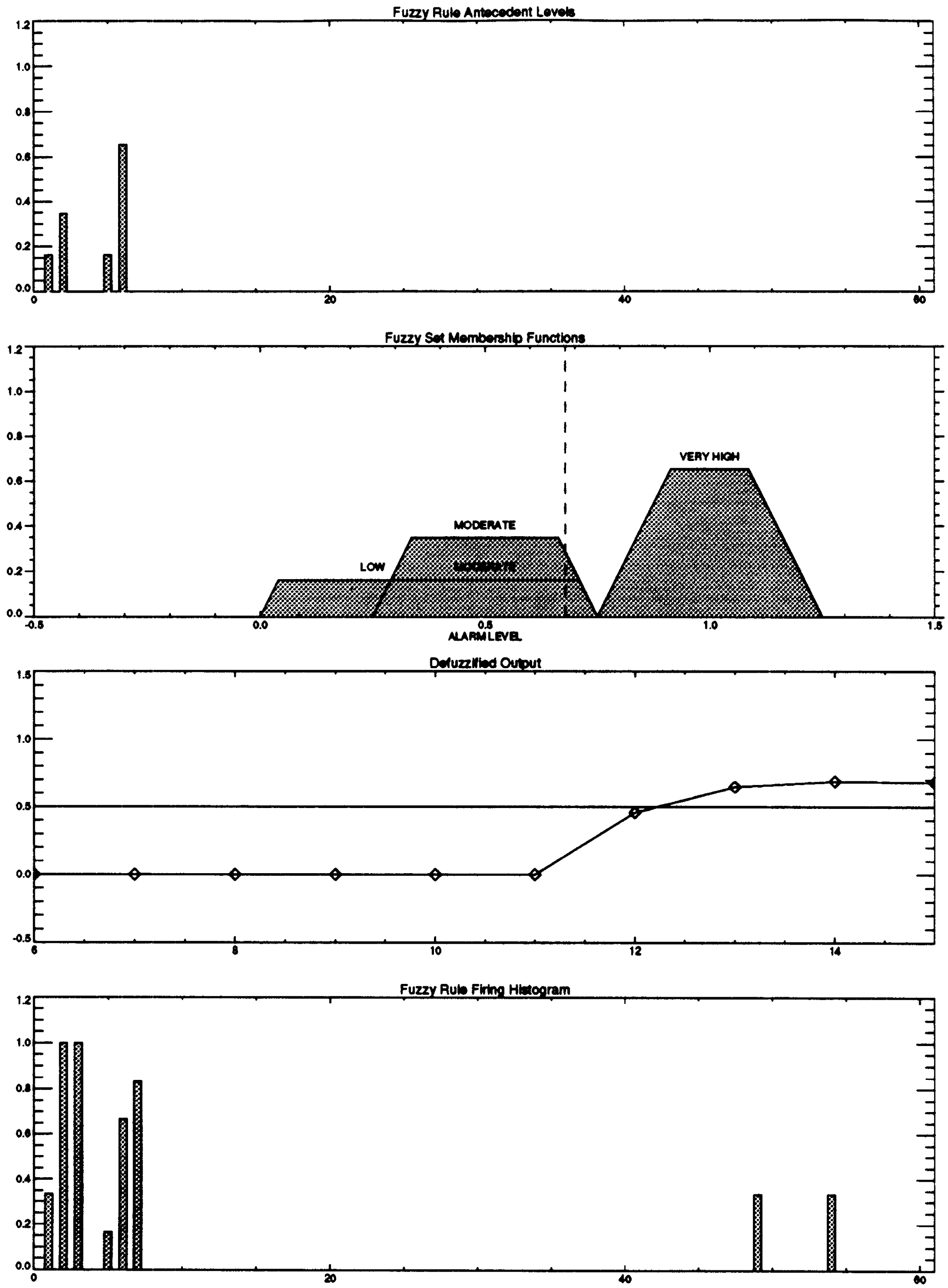

Figure 17. Snapshot of dynamic graphical display utility for one sample from a threat signal.

4.5.2 Performance Evaluation Utilities

Several evaluation utilities were developed for this investigation to compare the performance of the fuzzy logic subsystem with the performance of the existing algorithms.

Performance is measured both in terms of detection probabilities and early warning capability. Detection probabilities are measured as the percentage of threats correctly classified as threats (probability of detection) versus the percentage of non-threats incorrectly classified as threats (probability of false alarm). Note that these probability measures are not based on known density functions, but are generated from estimated density functions based on histograms of the results obtained from processing the test data sets. These detection statistics are gathered from the threat and non-threat signals that are processed by both the fuzzy subsystem and the existing mode 4 and mode 5 classifiers. As indicated in section 4.1.4, these performance statistics characterize only the "local" performance of these classifiers in that they do not reflect the overall "global" performance of the entire integrated system.

The detection performance for the existing mode 4 and mode 5 classifiers is known. Specifically, the estimated probability of detection (P_D) is 0.82 and the estimated probability of false alarm (P_{FA}) is 0.40 for the processed data sets. The output of the fuzzy logic subsystem is the defuzzified alarm level for each test sample. To determine the detection performance of this system, a threshold is applied to the output alarm level. If the alarm level for any sample for a signal in a particular data set exceeds that threshold, then an alarm is declared for that data set and the earliest time at which this occurs is recorded.

To compare the false alarm performance of the fuzzy logic subsystem with the existing algorithms, the detection threshold can be set at a level that yields the same probability of detection as obtained by the existing algorithms for the fuzzy subsystem. Then, the probability of false alarm is measured by counting the data sets with alarm level that exceed this threshold across the non-threat data sets. Conversely, the threshold can be set such that the same probability of false alarm is achieved, so the probability of detection can be compared. These performance numbers for the test data sets are presented in the next section.

A more continuous method of comparing detection performance statistics is to compute what is known as receiver operator (ROC) curves[1] for the fuzzy logic subsystem. The probability of detection and the probability of false alarm measured for a particular setting of the detection threshold can be plotted on a graph. By allowing this threshold to continuously vary, the relationship between these two parameters forms a curve. Figure 18 shows the form of a typical ROC curve. As the threshold increases, the corresponding position along the curve moves from the point (1,1) to the point (0,0). Superior performance is indicated when the ROC curve falls higher in the upper left-hand corner of the plot. In such a case, higher probabilities of detection are achieved with lower probabilities of false alarm. ROC curves for the fuzzy logic subsystem are presented in the next section.

[1] ROC curves have historically been used for two-class pattern problems for radar systems.

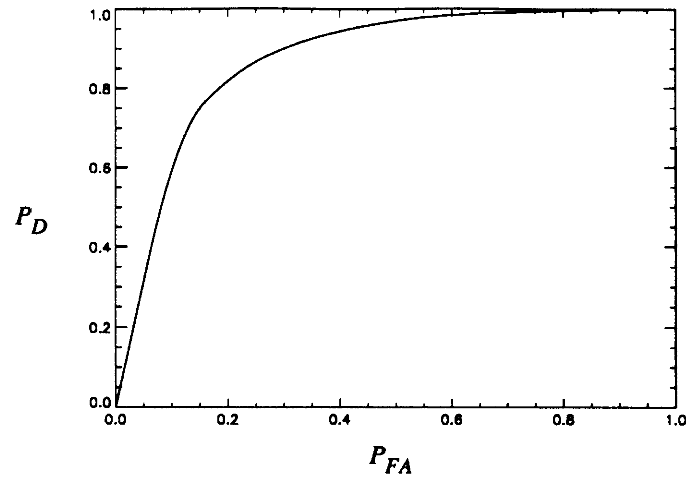

Figure 18. Representative ROC curve for threat versus non-threat discrimination problem.

An assessment of the warning time for each system is made as well. The warning time for a threat engagement in a particular data set is a measure of the amount of time between the earliest time a threat is declared by the classifiers considered for this investigation and the time remaining to target intercept. Clearly, the earliest warning time possible is desired for such a system so that sufficient time is available for appropriate countermeasure action. Due to the temporal nature of the classifiers considered in this investigation, there are inherent limits as to how early a threat can be detected. For instance, these classifiers are dependent on the measurement of a previously occurring temporal event that determines the range characteristics of the potential threat signal. This measurement operates on a certain number of previously processed data samples. Since detection can not occur before these samples are processed, the warning times for these temporal classifiers are limited. Warning time performance for each system is measured by simply averaging the warning times across all the successfully classified simulated threat data sets. Warning times are presented in terms of processing sample units. That is, the number of samples to be processed after the sample at which the detection occurred but before the sample at which intercept occurs, is presented. Warning time results for the existing classifiers and the fuzzy subsystem are presented in the next section.

5 Results

The next few sections describe the results obtained during this investigation. First, the known results for the existing mode 4 and mode 5 classification algorithms are presented. This is followed by a presentation of various experiments using the fuzzy subsystem. These results are analyzed and compared using the performance evaluation utilities developed for this investigation.

5.1 Performance of the Existing Classifiers

As described in section 4.1.4, test sets containing 1651 threat data sets and 86 non-threat data sets are used for performance evaluation and system comparison for this investigation. These data sets contain the signals that are actually processed by both the fuzzy subsystem and either of the existing mode 4 or mode 5 classifiers. "Local" performance measures are computed for these two processes using the evaluation utilities described in the previous section. Performance statistics for the existing mode 4 and mode 5 classifiers are shown in table 2, which is often referred to as a *confusion matrix*.

EXISTING SYSTEM	Classified as Threats	Classified as Non-Threats
Actual Threats	1350 (82%)	301 (18%)
Actual Non-Threats	34 (40%)	52 (60%)

Table 2. Confusion matrix for existing mode 4 and mode 5 classifiers.

These performance statistics represent a single point on the ROC plots for the fuzzy subsystem since the existing classifiers can only makes binary decisions (*i.e.*, there is no decision threshold to vary).

The warning times for the mode 4 and mode 5 classifiers are also known for the existing algorithms. In particular, the average warning time (in processing sample units) for these combined modes is measured at 7.16 samples. That is, on average, either the mode 4 or the mode 5 classifier for the existing algorithms detected a threat a little more than seven samples prior to the sample at which target intercept is expected to occur.

5.2 Performance of the Fuzzy Subsystem

Various experiments were performed for this investigation to assess the performance of the fuzzy subsystem. A set of four fuzzy data bases $\{\Omega_1, \Omega_2, \Omega_3, \Omega_4\}$ were developed and used in conjunction with the fuzzy rule base designed for the fuzzy subsystem (as described in section 4.4). For the fuzzy data base Ω_1 (see figure 11), the ROC curve shown in figure 19 was generated. A single point, designated by the asterisk, is the probability of detection and probability of false alarm for the existing mode 4 and mode 5 classifiers.

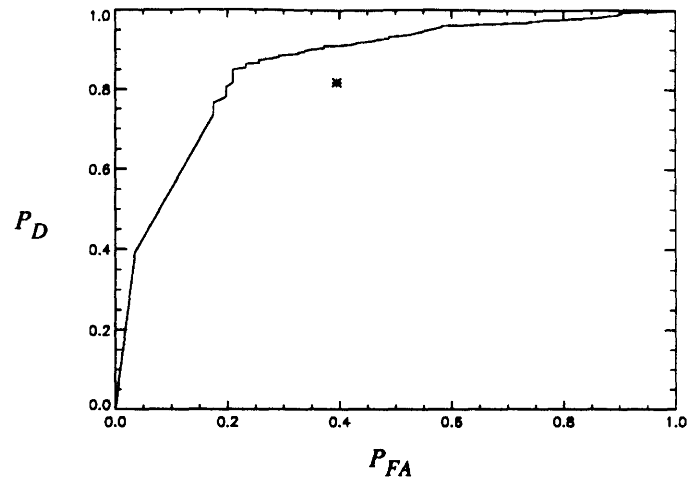

Figure 19. ROC curve for fuzzy subsystem using fuzzy data base Ω_1. The asterisk indicates the P_D and the P_{FA} for the existing mode 4 and mode 5 classifiers.

This plot indicates a moderate level of performance improvement over the existing algorithms is achieved when using the Ω_1 data base for the fuzzy logic subsystem. With the same probability of detection, this fuzzy subsystem has a lower probability of false alarm at about 0.22, which is equivalent to reducing the number of false alarms by almost one half. In addition, by allowing the same number of false alarms, the fuzzy logic subsystem has an increased the probability of detection from 0.82 to about 0.91.

Note that for this investigation, a reduction in the number of false alarms is judged to be more critical than an increase in the number of detections. This is because other classification modes (modes 1-3 and modes 6-8) are available to possibly declare some of the missed detections as threats. However, once an alarm occurs, whether it be induced from an actual threat signal or a non-threat signal, it can only be retracted by post-processing suppression algorithms. In fact, it may be desirable to select an operating point with a probability of detection lower than the present value of 0.82 in order to preserve an even lower number of false alarms. This trade-off must be made in the overall context of total system performance and is beyond the scope of this investigation.

The average warning time measured for the fuzzy subsystem using the Ω_1 data base is equivalent to the time of 7.64 samples. This corresponds to an average warning time improvement over the existing classifiers of about 6.7 percent. Even such small increases in warning time may prove to be significant in certain threat warning situations when swift countermeasure action is essential.

A similar experiment was performed using the fuzzy data base Ω_2 (see figure 12). The corresponding performance statistics are summarized by the ROC curve shown in figure 20.

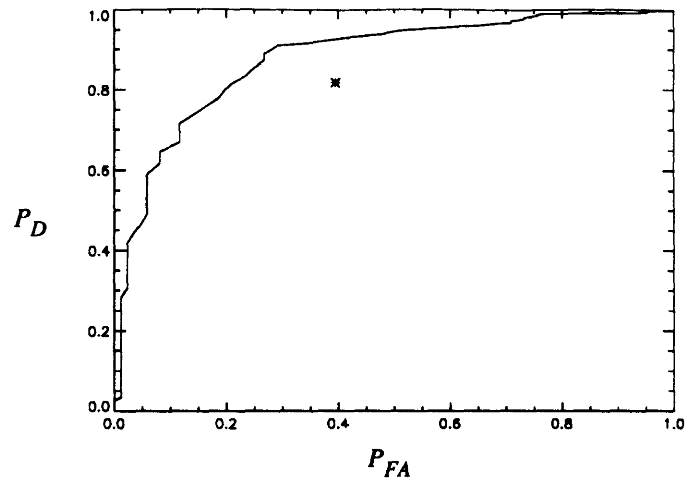

Figure 20. ROC curve for fuzzy subsystem using fuzzy data base Ω_2.

This plot indicates a similar level of performance improvement over the existing algorithms as observed using the Ω_1 fuzzy data base. In particular, the same probability of false alarm (about 0.22) is indicated using both Ω_1 and Ω_2 when the probability of detection is fixed at the value obtained from the exiting classifiers. When the probability of false alarm is fixed at the value allowed by the existing classifiers, the fuzzy logic subsystem using the Ω_2 data base has a probability of detection of about 0.93. This statistic is slightly higher than that obtained using the Ω_1 fuzzy data base. It may be to appropriate to select the point on the operating curve just prior to the point where the slope of the curve begins to flatten. Low slope regions of the ROC curve indicate small gains in probability of detection for relatively large numbers of false alarms.

The average warning time measured for the fuzzy subsystem using the Ω_2 data base is 8.34 samples which is a moderate improvement over the average warning time of the existing algorithms by about 16.4 percent. This average warning time is even higher than the average warning time for the Ω_2 data base. This is expected since the member functions defined for the Ω_2 data base have a higher degree of fuzziness.

The similarity in the performance of the fuzzy subsystem using these two fuzzy data bases suggests a certain degree of robustness in the application of fuzzy logic to this discrimination problem. The fuzzy data bases Ω_1 and Ω_2 consist of fuzzy sets that are designed with substantially different levels of fuzziness as indicated by their membership functions (see figures 11 and 12). This suggests that by fuzzifying the input boundaries to any degree and by applying a fuzzy combination of the mode 4 and mode 5 classification criteria, information about the relationships between the features and the temporal knowledge inherent in this process is more effectively utilized.

5.3 Fuzzy Subsystem Variations

Two alternate fuzzy data bases, Ω_3 and Ω_4, were also designed. The fuzzy sets contained in these data bases are represented in figures A1 and figures A2, respectively, located in Appendix A. The purpose of these designs is to determine the source of the performance improvement demonstrated by the fuzzy logic subsystem. In the Ω_3 fuzzy data set, the linguistic variables for the first four features $\{x_1, x_2, x_3, x_4\}$ contain identical fuzzy sets as the Ω_1 fuzzy data base. These four features measure signal characteristics related to signal shape and feature measurement quality. The other two features $\{x_1, x_2\}$ that measure occurrence of the previous temporal event are designed as "crisp" fuzzy sets so that the distinction between mode 4 and mode 5 classification tests reverts back to a binary decision. This way, the contribution of fuzzy logic applied to the shape and measurement features for each current sample can be isolated. This corresponds to the fuzzification of the "binary decision" problem (see section 2.2.2). The performance of the fuzzy logic subsystem using this fuzzy data base is presented in terms of the ROC curve shown in figure 21.

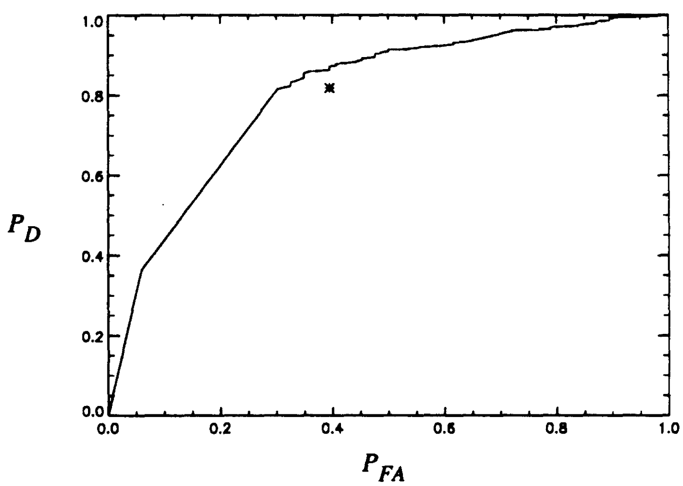

Figure 21. ROC curve for fuzzy subsystem using fuzzy data base Ω_3.

Although improved performance over the existing classifiers is indicated, the amount of improvement is about half as much as was observed using the Ω_1 and Ω_2 fuzzy data base designs.

Conversely, the Ω_4 fuzzy data base was designed to isolate the contribution of fuzzy logic applied only to the previous temporal event features. This corresponds to the fuzzification of the "brick wall" problem (see section 2.2.3). The performance of the fuzzy logic subsystem using the Ω_4 data base is also presented in terms of an ROC curve, as shown in figure 22.

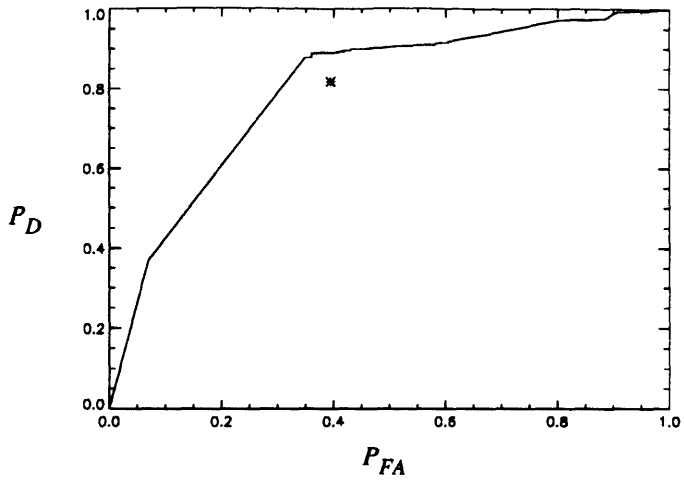

Figure 22. ROC curve for fuzzy subsystem using fuzzy data base Ω_4.

Again, some improved performance over the existing classifiers is indicated. The amount of improvement here is about the same as that indicated using the Ω_3 fuzzy data base design.

These results indicate that the improved performance gained by applying the fuzzy logic subsystem is derived from both the fuzzification of the signal shape features and the fuzzification for the previous temporal event features. The improved performance using the Ω_3 fuzzy data base indicates that by fuzzifying the decision boundaries for classification mode 4 and classification mode 5 features, the binary decision problems induced by strict thresholds can be relieved. In addition, the improved performance using the Ω_4 fuzzy data base indicates that by allowing a continuous transition between classification mode 4 and classification mode 5 detection criteria, temporal information can be exploited to reduce the exclusive nature of the existing classification structure. These results also suggest that these performance gains are complementary in that the degree of improvement in performance statistics appears to be additive.

6 Conclusions and Recommendations

6.1 Investigation Summary

This investigation has demonstrated that by applying fuzzy logic techniques to a key portion of the temporal discrimination problem, moderate performance improvements can be achieved. These improvements can be attributed to two reasons. The first reason is due to the ability of fuzzy logic to represent the input features for each individual classification mode as fuzzy linguistic variables. As discussed in section 3.2.1, fuzzy linguistic variables partition the domain of the variable into a number of overlapping fuzzy sets. Each fuzzy set is defined in terms of a continuous membership function representing truth values between 0 and 1. In this way, it is possible for each input feature to belong to multiple fuzzy sets with variable degrees of membership.

In contrast, the existing algorithms, against which the fuzzy logic approach is being compared, apply each individual classifier as a binary decision based on a series of one-dimensional threshold tests. This results in a strictly binary decision surface that does not allow for "shaded" decisions and does not capture the relationships between the input features. Instead, each individual classification mode rejects the current sample if any one of the input feature measurements fails its corresponding threshold test.

The application of fuzzy logic to each individual classification mode provides continuous and overlapping boundaries between the fuzzy concepts defined along the domain of the input feature. A set of fuzzy rules defined in terms of these fuzzy concepts are simultaneously applied to the input features. Some of these fuzzy rules capture decision relevant information that relates to all of the features for each classification mode. Thus, a more continuous decision can be made at each sample, which provides an improvement in the performance statistics as discussed in section 5.2 above.

The second reason for the improved performance is due to the ability of fuzzy logic to better exploit the temporal characteristics inherent in this discrimination process than do the existing algorithms. The existing algorithms suffer from what is called the "brick wall" problem which limits the temporal information that is shared between individual classification modes. In particular, a binary decision is made as to whether to apply classification mode 4 or classification mode 5 at each sample based on the values of features that measure the size and quality of a previously occurring temporal event. If these feature measurements pass a set of strict threshold tests then classification mode 4 is applied. Otherwise, classification mode 5 is applied.

The fuzzification of this process allows a continuous transition between classification mode 4 and classification mode 5 decision criteria. With fuzzy logic, both of these modes can be applied simultaneously but with different levels of confidence. The level to which each classification mode is applied is based on the degree to which the temporal event features satisfy the temporal criteria for that mode. This fuzzification in the temporal domain provides an improvement in the performance statistics as discussed in section 5.3 above.

These two fuzzification processes are combined in the fuzzy logic subsystem designed for this investigation. The improved performance demonstrated by this subsystem indicates that fuzzy logic is able to benefit from both of these processes in a complementary fashion. With the same number of detections obtained from the existing classifiers, a moderate reduction in the number of false alarms (almost by a factor of 2) is achieved. In addition, an increase in the average warning time (ranging from 6 to 16 percent) for simulated threat data is demonstrated. These results are presented in detail in section 5.

The fuzzy logic subsystem developed for this investigation is based on expert knowledge about threat warning systems gathered from years of experience. A direct conversion of this knowledge was used to define the linguistic variables that represent the features defined in the original system. Well established thresholds were used to partition these linguistic variables into fuzzy sets. Several fuzzy data bases (*i.e.*, Ω_1, Ω_2, Ω_3, and

Ω_4) were tested with the fuzzy subsystem. The fuzzy data bases Ω_1 and Ω_2 were designed such that every input feature is fuzzified to various degrees of fuzziness. Similar performance improvement was achieved from both of these experiments, which suggests a certain degree of robustness in the design of the fuzzy subsystem. The fuzzy data base Ω_3 was designed to determine the effects of fuzzy logic applied only to the individual classifiers. As such, only the individual classification features (x_1, x_2, x_3, and x_4) are fuzzified. About half as much performance improvement over the Ω_1 and Ω_2 tests was observed using the Ω_3 fuzzy data base. Conversely, data base Ω_4 was designed to determine the effects of fuzzy logic applied only in the temporal domain. In this case, only the temporal event features (x_5 and x_6) are fuzzified. Again, about half as much performance improvement was observed using the Ω_4 fuzzy data base. Complete fuzzification of this fuzzy subsystem, using either fuzzy data base Ω_1 or Ω_2, is the preferred method of applying fuzzy logic to the temporal classification problem.

A set of fuzzy rules was designed to capture the expert knowledge that was used to develop the original system. These rules were defined using a hierarchical strategy that provides a straight forward method of translating this high dimensional discrimination problem into a manageable number of fuzzy rules. Desirable rule base properties were shown to be satisfied by these fuzzy rules.

6.2 Applicability to the Entire System

The positive results from this investigation suggest that fuzzy logic techniques are indeed applicable to this temporal classification problem. The key portion of the existing algorithms that have been investigated here are judged to be representative of the temporal structure of the entire system. The same techniques that were used to apply fuzzy logic to the classification mode 4 and mode 5 algorithms could be used to convert the complete set of classifiers into a single fuzzy logic system.

It is anticipated that similar benefits will be achieved from the fuzzification of each additional classification mode. Each existing classifier measures a different set of features that are designed to capture distinctive temporal characteristics of threats at specific ranges and applies a set of binary thresholds to these features. Fuzzification of each classifier could potentially improve detection performance as seen in this investigation.

In addition, the unification of all these individual classifiers into a single fuzzy system that applies fuzzy logic across the entire temporal domain has the potential for significant performance improvement. The fuzzy combination of classification modes 4 and 5 is based on the fuzzification of a single temporal event. In actuality, many temporal events occur in a threat signal over the course of the engagement. Each classifier in the original system was designed to detect these temporal events using relatively limited binary classification tests. By applying fuzzy logic to the entire system, the relationships between all these temporal events can be exploited in a unified and more continuous process.

The additional complexity involved in integrating this entire temporal classification system into a single process is well suited to fuzzy logic techniques. New fuzzy linguistic variables can be defined for each additional feature in the decision space. Fuzzy sets can be defined along the domain of each feature to effectively partition the feature domain into an appropriate number of fuzzy concepts that represent feature values in linguistic terms. Additional fuzzy rules can then be defined using these linguistic terms and added to the fuzzy rule base. These fuzzy rules should capture the appropriate relationships between the linguistic variables and exploit the additional temporal information available in the threat signal. Similar hierarchical design techniques can be used to develop a meaningful set of rules and to control the overall size of the fuzzy rule base.

6.3 Recommendations

It is recommended that the fuzzification of the entire threat warning system be pursued in an attempt to further improve the performance of the discrimination process and to provide the manageable growth capability that is desirable for an evolving system such as this. This would include combining all classification modes into a single fuzzy process using the fuzzy logic techniques considered during this investigation. In addition, information from concurrent signals from other sensors could also be incorporated into the fuzzy process with the addition of the appropriate fuzzy rules to the fuzzy rule base. Continued analysis of variations in the fuzzy knowledge data should be performed in an attempt to optimize performance and to further establish the fuzzy relationships between the fuzzy variables.

It is also recommended that the issue of processing time for the fuzzy subsystem should be considered before advanced fuzzy logic techniques are pursued. The simulation for this fuzzy subsystem was designed with a high degree of flexibility both in fuzzy design parameters and graphical display options. As such, no effort at optimizing this processing has been performed. This simulation takes about five times longer than the simulation of the original algorithms to process the entire set of test data. However, the underlying operations required to perform the fuzzy logic operations are quite straight forward. It is anticipated that by assuming that certain design parameters (*e.g.*, inference method, membership function form, defuzzification method, *etc.*,) will be fixed, and by considering recent developments in fuzzy logic hardware (Lee, [1990b]), real-time implementation of a fuzzy logic system for threat warning systems is a possibility.

Various fuzzy hybrid techniques that incorporate classical pattern recognition techniques and adaptive learning techniques (including neural network hybrids) should also be considered in future research.

For this threat warning system, however, it is essential that any approach conform to the prescribed design philosophy that is based on maintaining a thorough understanding of the physical meaning behind the discrimination process and to be able to relate the design parameters back to underlying physical phenomena. The fuzzy logic techniques applied during this investigation conform to this design philosophy.

Appendix A

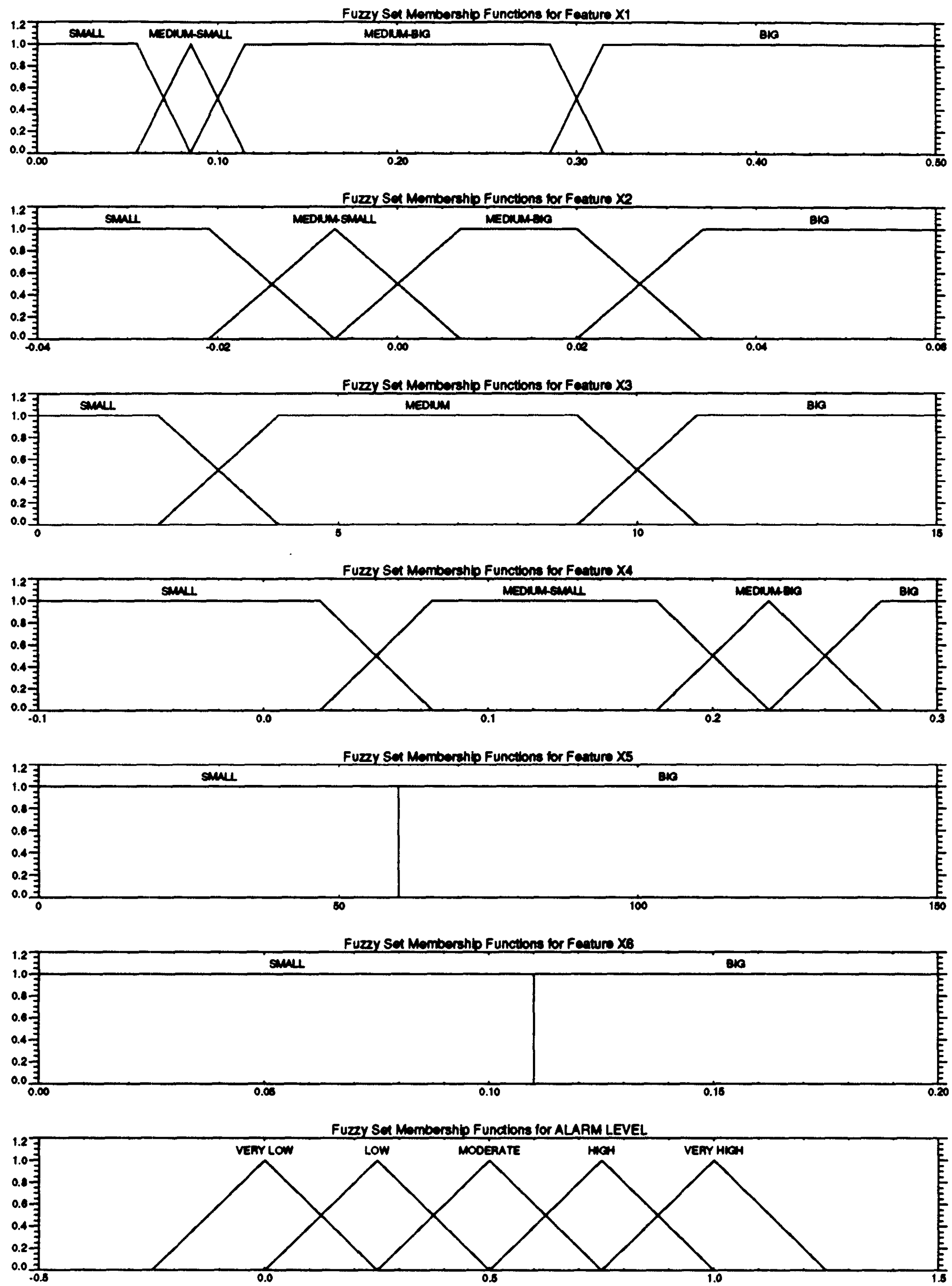

Figure A1. Membership functions for fuzzy data base Ω_3.

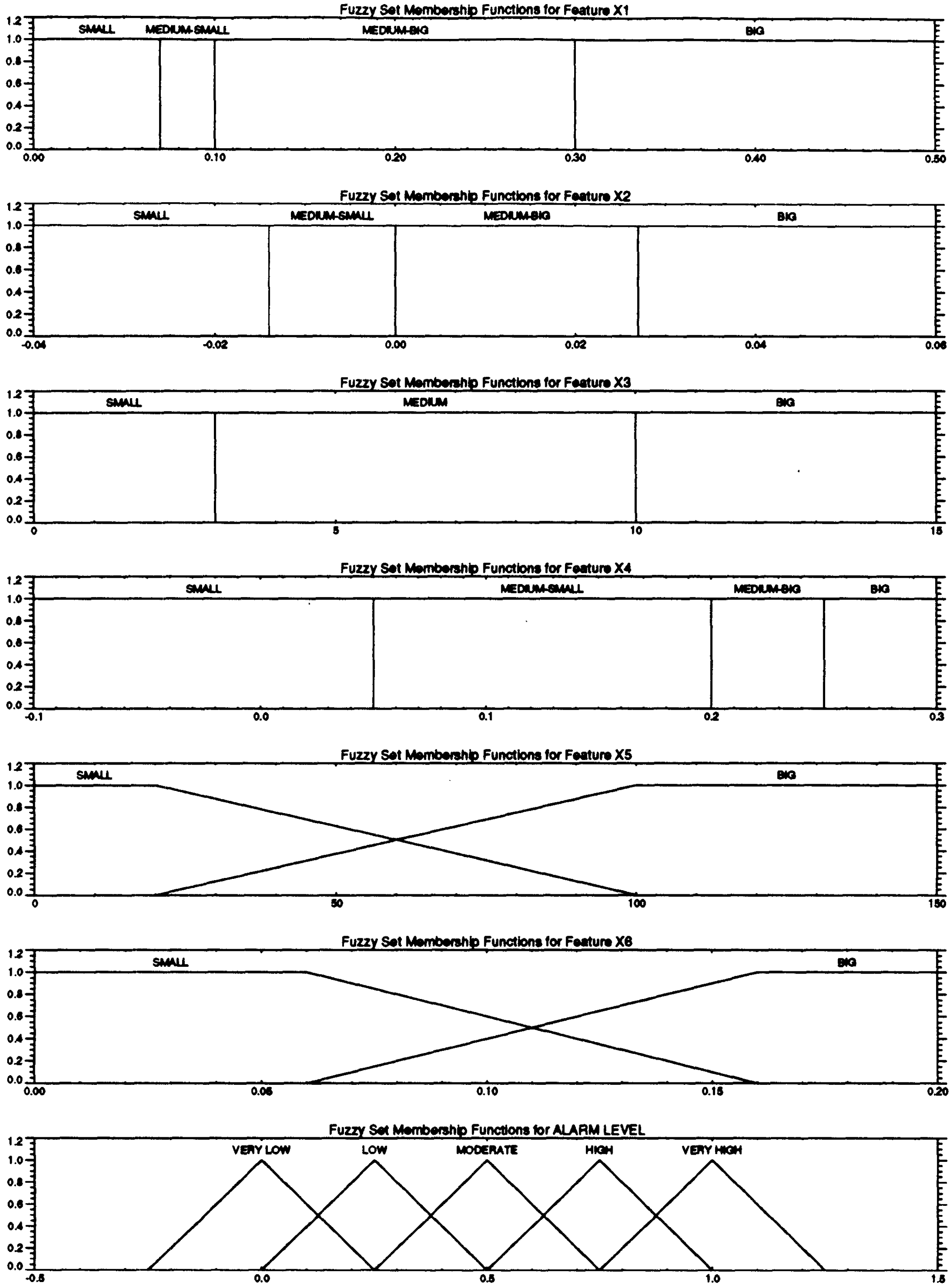

Figure A2. Membership functions for fuzzy data base Ω_4.

Appendix B

Classification mode 4 detection rules:

1 if x_1 is SMALL and x_2 is SMALL and x_5 is BIG and x_6 is SMALL then y is LOW
2 if x_1 is SMALL and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is SMALL then y is MODERATE
3 if x_1 is SMALL and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is SMALL then y is MODERATE
4 if x_1 is SMALL and x_2 is BIG and x_5 is BIG and x_6 is SMALL then y is LOW

5 if x_1 is MEDIUM-SMALL and x_2 is SMALL and x_5 is BIG and x_6 is SMALL then y is MODERATE
6 if x_1 is MEDIUM-SMALL and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is SMALL then y is VERY HIGH
7 if x_1 is MEDIUM-SMALL and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is SMALL then y is VERY HIGH
8 if x_1 is MEDIUM-SMALL and x_2 is BIG and x_5 is BIG and x_6 is SMALL then y is MODERATE

9 if x_1 is MEDIUM-BIG and x_2 is SMALL and x_5 is BIG and x_6 is SMALL then y is MODERATE
10 if x_1 is MEDIUM-BIG and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is SMALL then y is VERY HIGH
11 if x_1 is MEDIUM-BIG and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is SMALL then y is VERY HIGH
12 if x_1 is MEDIUM-BIG and x_2 is BIG and x_5 is BIG and x_6 is SMALL then y is MODERATE

13 if x_1 is BIG and x_2 is SMALL and x_5 is BIG and x_6 is SMALL then y is LOW
14 if x_1 is BIG and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is SMALL then y is MODERATE
15 if x_1 is BIG and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is SMALL then y is MODERATE
16 if x_1 is BIG and x_2 is BIG and x_5 is BIG and x_6 is SMALL then y is LOW

Classification mode 5 detection rules:

17 if x_1 is SMALL and x_2 is SMALL and x_5 is SMALL then y is VERY LOW
18 if x_1 is SMALL and x_2 is MEDIUM-SMALL and x_5 is SMALL then y is VERY LOW
19 if x_1 is SMALL and x_2 is MEDIUM-BIG and x_5 is SMALL then y is VERY LOW
20 if x_1 is SMALL and x_2 is BIG and x_5 is SMALL then y is VERY LOW

21 if x_1 is MEDIUM-SMALL and x_2 is SMALL and x_5 is SMALL then y is VERY LOW
22 if x_1 is MEDIUM-SMALL and x_2 is MEDIUM-SMALL and x_5 is SMALL then y is LOW
23 if x_1 is MEDIUM-SMALL and x_2 is MEDIUM-BIG and x_5 is SMALL then y is MODERATE
24 if x_1 is MEDIUM-SMALL and x_2 is BIG and x_5 is SMALL then y is LOW

25 if x_1 is MEDIUM-BIG and x_2 is SMALL and x_5 is SMALL then y is VERY LOW
26 if x_1 is MEDIUM-BIG and x_2 is MEDIUM-SMALL and x_5 is SMALL then y is MODERATE
27 if x_1 is MEDIUM-BIG and x_2 is MEDIUM-BIG and x_5 is SMALL then y is VERY HIGH
28 if x_1 is MEDIUM-BIG and x_2 is BIG and x_5 is SMALL then y is MODERATE

29 if x_1 is BIG and x_2 is SMALL and x_5 is SMALL then y is VERY LOW
30 if x_1 is BIG and x_2 is MEDIUM-SMALL and x_5 is SMALL then y is LOW
31 if x_1 is BIG and x_2 is MEDIUM-BIG and x_5 is SMALL then y is MODERATE
32 if x_1 is BIG and x_2 is BIG and x_5 is SMALL then y is LOW

33 if x_1 is SMALL and x_2 is SMALL and x_5 is BIG and x_6 is BIG then y is VERY LOW
34 if x_1 is SMALL and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is BIG then y is VERY LOW
35 if x_1 is SMALL and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is BIG then y is VERY LOW
36 if x_1 is SMALL and x_2 is BIG and x_5 is BIG and x_6 is BIG then y is VERY LOW

37 if x_1 is MEDIUM-SMALL and x_2 is SMALL and x_5 is BIG and x_6 is BIG then y is VERY LOW
38 if x_1 is MEDIUM-SMALL and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is BIG then y is LOW
39 if x_1 is MEDIUM-SMALL and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is BIG then y is MODERATE
40 if x_1 is MEDIUM-SMALL and x_2 is BIG and x_5 is BIG and x_6 is BIG then y is LOW

41 if x_1 is MEDIUM-BIG and x_2 is SMALL and x_5 is BIG and x_6 is BIG then y is VERY LOW
42 if x_1 is MEDIUM-BIG and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is BIG then y is MODERATE
43 if x_1 is MEDIUM-BIG and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is BIG then y is VERY HIGH
44 if x_1 is MEDIUM-BIG and x_2 is BIG and x_5 is BIG and x_6 is BIG then y is MODERATE

45 if x_1 is BIG and x_2 is SMALL and x_5 is BIG and x_6 is BIG then y is VERY LOW
46 if x_1 is BIG and x_2 is MEDIUM-SMALL and x_5 is BIG and x_6 is BIG then y is LOW
47 if x_1 is BIG and x_2 is MEDIUM-BIG and x_5 is BIG and x_6 is BIG then y is MODERATE
48 if x_1 is BIG and x_2 is BIG and x_5 is BIG and x_6 is BIG then y is LOW

Measurement error rules:

49 if x_3 is BIG and x_5 is BIG and x_6 is SMALL then y is VERY LOW

50 if x_3 is M and x_5 is SMALL then y is VERY LOW
51 if x_3 is BIG and x_5 is SMALL then y is VERY LOW
52 if x_3 is M and x_5 is BIG and x_6 is BIG then y is VERY LOW
53 if x_3 is BIG and x_5 is BIG and x_6 is BIG then y is VERY LOW

Linear shape feature rules:

54 if x_4 is BIG and x_5 is BIG and x_6 is SMALL then y is VERY LOW

55 if x_4 is SMALL and x_5 is SMALL then y is VERY LOW
56 if x_4 is MEDIUM-BIG and x_5 is SMALL then y is VERY LOW
57 if x_4 is BIG and x_5 is SMALL then y is VERY LOW
58 if x_4 is SMALL and x_5 is BIG and x_6 is BIG then y is VERY LOW
59 if x_4 is MEDIUM-BIG and x_5 is BIG and x_6 is BIG then y is VERY LOW
60 if x_4 is BIG and x_5 is BIG and x_6 is BIG then y is VERY LOW

Bibliography

L. J. Bain and M. Engelhardt, *Introduction to Probability and Mathematical Statistics.* Boston, MA: PWS-Kent, 1992.

J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York: Plenum Press, 1981.

J. C. Bezdek, *Analysis of Fuzzy Information, Vol. II and Vol. III.* Boca Raton, FL: CRC Press, 1987

J. C. Bezdek and S. K. Pal, eds., *Fuzzy Models for Pattern Recognition.* Piscataway, NJ: IEEE Press, 1992

B. B. Devi and V. V. S. Sarma, "Estimation of Fuzzy Memberships From Histograms," *Information Sciences,* vol. 35, pp. 43-59, 1985.

B. B. Devi and V. V. S. Sarma, "A Fuzzy Approximation Scheme for Sequential Learning in Pattern Recognition," *IEEE Trans. Syst. Man Cybern.*, vol. 16, no. 5, pp. 668-679, Sept./Oct. 1986.

D. Driankov, H. Ellendoorn, and M. Reinfrank, *An Introduction to Fuzzy Control.* Berlin: Springer-Verlag, 1993.

D. Dubois and H. Prade, *Fuzzy Sets and Systems.* New York: Academic Press, 1980.

D. Dubois and H. Prade, "Processing Fuzzy Temporal Knowledge," *IEEE Trans. Syst. Man Cybern.*, vol. 19, no. 4, pp. 729-744, July/Aug. 1989.

D. Dubois, H. Prade, and R. R. Yager, eds., *Fuzzy Sets for Intelligent Systems.* San Mateo, CA: Morgan Kaumann, 1993.

R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis.* New York: Wiley-Interscience, 1973.

M. M. Gupta, *et al*, eds., *Approximate Reasoning in Expert Systems.* Amsterdam: North-Holland, 1985.

M. M. Gupta and T. Yamakawa, eds. *Fuzzy Logic in Knowledge-Based Systems, Decision and Control.* Amsterdam: North-Holland, 1988.

H. Hellendoorn, "The Genralized Modus Ponens Considered as a Fuzzy Relation," *Fuzzy Sets and Systems,* vol. 46, pp. 29-48, 1992.

M. Jamshidi, N. Vadiee, and T. J. Ross, eds., *Fuzzy Logic and Control.* Englewood Cliffs, NJ: Prentice Hall, 1993.

A. Kandel, *Fuzzy Techniques in Pattern Recognition.* New York: Wiley-Interscience, 1982.

A. Kandel, *Fuzzy Mathematical Techniques with Applications.* Reading, MA: Addison-Wesley, 1986.

G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information.* Englewood Cliffs, NJ: Prentice Hall,
1988.

B. Kosko, *Neural Networks and Fuzzy Systems.* Englewood Cliffs, NJ: Prentice Hall, 1992.

C. C. Lee, "Fuzzy Logic in Control Systems: Fuzzy Logic Controller - Part I," *IEEE Trans. Syst. Man Cybern.*, vol. 20, no. 2, pp. 404-418, March/April 1990.

C. C. Lee, "Fuzzy Logic in Control Systems: Fuzzy Logic Controller - Part II," *IEEE Trans. Syst. Man Cybern.*, vol. 20, no. 2, pp. 419-435, March/April 1990.

S. K. Pal and D. K. D. Majumder, *Fuzzy Mathematical Approach to Pattern Recognition.* New York: Halsted Press, 1986.

M. Sugeno, "An Introductory Survey of Fuzzy Control," *Information Sciences,* vol. 36, pp. 59-83, 1985.

C. W. Therrien, *Decision Estimation and Classification.* New York: Wiley and Sons, 1989.

J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles.* Reading, MA: Addison-Wesley, 1974.

S. Weber, "A Genral Concept of Fuzzy Connectives, Negations and Implications Based on *t*-norms and *t*-co-norms," *Fuzzy Sets and Systems,* vol. 11, pp. 115-134, 1983.

L. A. Zadeh, "Fuzzy Sets," *Information and Control,* vol. 8, pp. 338-353, 1965.

L. A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Trans. Syst. Man Cybern.*, vol. 3, no. 1, pp. 28-44, Jan. 1973.

L. A. Zadeh, "Fuzzy Logic," Report No. CSLI-88-116, Center for the Study of Language and Information, Stanford University, Stanford, CA, 1988.