Dirty Secrets in Multisensor Data Fusion

David L. Hall[1]
and
Alan Steinberg[2]

[1]The Pennsylvania State University
Applied Research Laboratory
P. O. Box 30
State College, PA 16801-0030
dlh28@psu.edu

[2]Veridian ERIM International
14150 Newbrook Drive
Suite 300
Chantilly, VA 20151

Abstract

Multisensor data fusion systems seek to combine information from multiple sources and sensors in order to achieve inferences that cannot be achieved with a single sensor or source. Applications of data fusion for Department of Defense (DoD) applications include automatic target recognition (ATR), identification-friend-foe-neutral (IFFN), and battlefield surveillance and situation assessment. The use of data fusion for these applications is appealing. Conceptually, the use of a broad spectrum of sensors should improve system accuracy, decrease uncertainty, and make these systems more robust to changes in the targets and environmental conditions. Techniques for data fusion are drawn from a diverse set of disciplines including signal and image processing, pattern recognition, statistical estimation, and artificial intelligence. Many of these techniques have an extensive history, ranging from Bayesian inference (first published in 1793) to fuzzy logic (originating in the 1920s) to neural nets (developed in the 1940s). In the past two decades an enormous amount of DoD funds have been expended to develop data fusion systems. While there are many successes, there are still a number of challenges and limitations. Indeed, critics of data fusion argue that data fusion technology is disappointing and ask, "why is it that when all is said and done (in data fusion), there is so much more said than done?" This paper presents a summary of the current state and limitations of data fusion. Key issues are identified that limit the ability to implement a successful data fusion system.

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
00012001	N/A	-

Title and Subtitle Dirty Secrets in Multisensor Data Fusion	**Contract Number**
	Grant Number
	Program Element Number
Author(s) Hall, David L.; Steinberg, Alan	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) 1The Pennsylvania State University Applied Research Laboratory P. O. Box 30 State College, PA 16801-0030	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) Director, CECOM RDEC Night Vision and Electronic Sensors Directorate, Security Team 10221 Burbeck Road Ft. Belvoir, VA 22060-5806	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes The original document contains color images.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UNLIMITED
Number of Pages 15	

1.0 Introduction

Data Fusion systems seek to combine information from multiple sensors and sources to achieve improved inferences than those achieved from a single sensor or source. Applications of data fusion related to the Department of Defense (DoD) span a number of areas including automatic target recognition (ATR), identification-friend-foe-neutral (IFFN), *smart w*eapons, battlefield surveillance systems, threat warning systems (TWS), and systems to support precision guided weapons. Waltz and Llinas[1], Hall[2], and Hall and Llinas[3] provide a general introduction to multisensor data fusion. Additional information can be obtained from the texts by Blackman[4], Antony[5], and Hall[6]. Data fusion systems typically use a variety of algorithms and techniques to transform the sensor data (e.g., radar returns, and infrared spectra) to detect, locate, characterize, and identify entities such as aircraft and ground-based vehicles. These techniques include signal and image processing, statistical estimation, pattern recognition, and many others (see Hall and Linn[7]). In addition, the fusion systems may use automated reasoning techniques to understand the context in which the entities are observed (i.e., situation assessment) and to understand the intent and possible threat posed by the observed entities (i.e., threat assessment).

Over the past two decades, an enormous amount of DoD funding has been applied to the problem of data fusion systems, and a large number of prototype systems have been implemented (Hall, Linn, and Llinas[8]). The data fusion community has developed a data fusion process model[9], a data fusion lexicon[10], and engineering guidelines for system development[11]. While a significant amount of progress has been made (Hall and Llinas[12,3]), much work remains to be done. Hall and Garga[13], for example, identified a number of pitfalls or problem areas in implementing data fusion systems. Hall and Llinas[14] described some shortcomings in the use of data fusion systems to support individual soldiers, and M. J. Hall, S. A. Hall and Tate[15] discuss issues related to the effectiveness of human-computer interfaces for data fusion systems.

This paper provides a summary of current progress in multisensor data fusion and identifies areas in which additional research is needed. In addition, the paper describes some issues or *dirty secrets* in the current state of practice of data fusion systems.

2.0 The JDL Data Fusion Process Model

In order to make this paper self-contained, we provide here a brief summary of the Joint Directors of Laboratories (JDL) data fusion process model [9,2,3]. A top-level view of the model is illustrated in Figure 1, and a summary of the processes is shown in Figure 2. This model is commonly used in the data fusion community to assist communications concerning data fusion algorithms, systems, and research issues. It will be used here for the same purpose.

The JDL Data Fusion Working Group was established in 1986 to assist in coordinating DoD activities in data fusion, and to improve communications among different DoD research and development groups. Led by Frank White (NOSC), the JDL working group performed a number of activities including; (1) development of a data fusion process model[9], (2) creation of a lexicon for data fusion[10], (3) development of engineering guidelines for building data fusion systems[11], and (4) organization and sponsorship of the Tri-service Data Fusion Conference from 1987 to 1992. The JDL Data Fusion Working Group has continued to support community efforts in data fusion, leading to the annual National Symposium on Sensor Data Fusion and the initiation of a Fusion Information Analysis Center (FUSIAC[16]).

Figure 1: Top Level View of the JDL Data Fusion Process Model

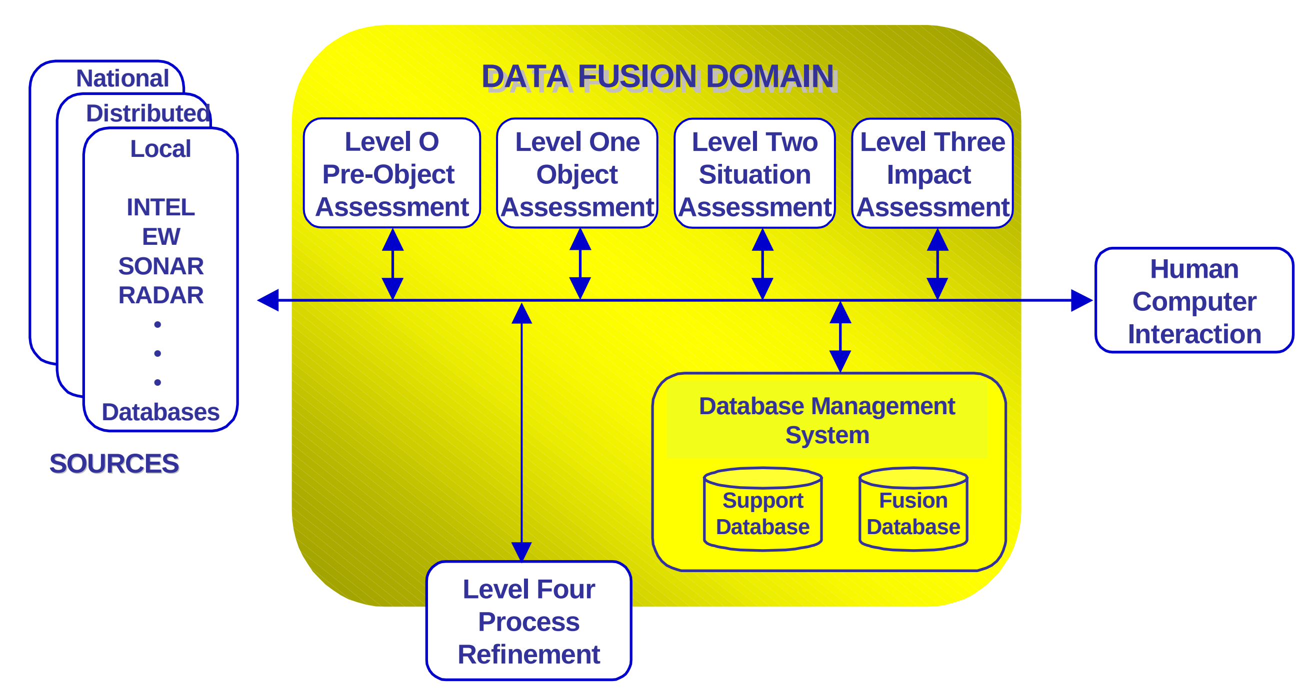


Figure 2: Summary of JDL Processes and Functions		
Process components	**Process Description**	**Functions**
Sources of information	Local and remote sensors accessible to the data fusion system; information from reference systems and human inputs	Local and distributed sensors; external data sources; human inputs
Human Computer Interface (HCI)	Provides an interface to allow a human to interact with the fusion system	Graphical displays; natural language processing
Source Preprocessing	Processing of individual sensor data to extract information, improve signal to noise, and prepare the data for subsequent fusion processing.	Signal and image processing; canonical transformations; feature extraction and data modeling
Level 1 Processing: Object Refinement	Association, correlation, and combination of information to detect, characterize, locate, track, and identify objects (e.g., tanks, aircraft, emitters).	Data alignment; correlation; position, kinematic, attribute estimation; object identity estimation;
Level 2 Processing: Situation Refinement	Development of a description of the current relationships among objects and events in the context of their environment.	Object aggregation; event and activity interpretation; context-based reasoning
Level 3 Processing: Threat Refinement	Projection of the current situation into the future to draw inferences about enemy threats, friendly and enemy vulnerabilities, and opportunities for operations.	Aggregate force estimation; intent prediction; multi-perspective analysis; temporal projections
Level 4 Processing: Process Refinement	A *meta-process* that seeks to optimize the on-going data fusion process (e.g., to improve accuracy of inferences, utilization of communication and computer resources)	Performance evaluation; process control; source requirement determination; mission management
Data Management	Provide access to, and management of, dynamic data fusion data including; sensor data, target state vectors, environmental information, doctrine, physical models, etc.	Data storage and retrieval; data mining; archiving; compression; relational queries and updates

The JDL model is a two layer hierarchical model that identifies fusion processes, processing functions and processing techniques to accomplish the functions. The model was intended for communications among data fusion researchers and implementation engineers, rather than a prescription for implementing a fusion system or an exhaustive enumeration of fusion functions and techniques. The model has evolved since its original exposition to the data fusion community. Steinberg and Bowman[17], for example, have recommended the inclusion of a new *Level zero* processing to account for processing such as pre-detection fusion and coherent signal processing of multi-sensor data. In addition, they suggest a re-naming and re-interpretation of the Level 2 and Level 3 processes to focus on understanding the external world environment (rather than a military-oriented situation and threat focus). C. Morefield[18] has suggested that the distinction between Level 2 and Level 3 is artificial, and that these processes should be considered as a single process. Bowman has suggested that the JDL model can be detrimental to communications if systems engineers focus on the model rather than a systematic architecture analysis and decomposition approach. Many of these comments have merit. However, for the purpose of this paper we will utilize the JDL model for describing the current state of practice and limitations.

3.0 Current Practices and Limitations in Data Fusion

A summary of the current state and limitations of data fusion is provided in Figure 3. This is an update of a similar figure originally introduced by Hall and Llinas in 1993[12] and updated by Hall and Llinas in 1997[3]. For each of the key components of the JDL process, the figure provides a summary of the current practices and limitations. These are summarized below.

Level 1: Object Refinement: Level One processing seeks to combine information about the location and attributes of entities (such as tanks or aircraft) to detect, locate, characterize, track, and identify the entities. Level one processing involves data assignment/correlation, estimation of the state of a entity, and an estimate of the entity's identity. The typical data fusion system partitions the object refinement problem into three basic components; (1) data assignment/correlation, (2) estimate of a state vector (e.g., for target tracking), and (3) estimation of a target's identity. Object refinement is relatively easy when there are a relatively few, widely separated targets moving in predictable paths. Target identity classification can generally be performed when there are observable target attributes (e.g., size, shape, and spectral signature) that can be uniquely mapped to target class or identity. This requires either an accurate model to link attributes with target identity, or a very large set of training data to train a pattern classification algorithm.

When these observing conditions are violated, the problem becomes much more challenging. Closely spaced, rapidly maneuvering targets, for example, are difficult to track because we cannot easily associate the sensor measurements to the appropriate targets. In addition, since acceleration cannot be observed directly, maneuvering targets cause a potential loss of track because we cannot accurately predict the future position of the targets. Complex observing environments, involving multi-path signal propagation clutter, dispersion, or other effects on signal to noise, can cause difficulties in data association and state estimation (because we may lack an accurate model to link the value of a target state vector to predicted observations). It is difficult to combine data from sensors that are co-dependent (viz., for which the sensor data are not statistically independent). Finally, complex targets without distinguishing attributes are difficult to classify or identify.

Figure 3: Summary of Current State of Multisensor Data Fusion		
JDL Process	**Current Practices**	**Limitations & Challenges**
Level 1: Object refinement	▪ Sensor preprocessing using standard signal and image processing methods ▪ Explicit separation of correlation and estimation problem ▪ Multiple target tracking using MHT[4,] JPDA[21], etc. ▪ Use of *ad hoc* maneuver models ▪ Object ID dominated by feature based methods[29] ▪ Pattern recognition using ANN[27] ▪ Emerging guidelines for selection of correlation algorithms[11,41] ▪ Promising work by Poore[18] , Mahler,[19] Barlow, *et al*[40]	▪ Dense target environments ▪ Rapidly maneuvering targets ▪ Complex signal propagation ▪ Co-dependent sensor observations ▪ Background clutter ▪ Context-based reasoning ▪ Integration of identity and kinematic data ▪ Lack of available ANN training data (for target identification)[27] ▪ No true fusion of image and non-image data (at the data level)
Level 2: Situation Refinement	▪ Numerous prototype systems[8] ▪ Dominance of rule-based KBS ▪ Variations include blackboard systems[23], logical templating[22], and case-based reasoning[24] ▪ Emerging use of fuzzy-logic[25] and agent-based systems[26]	▪ Very limited operational systems ▪ No experience in *scaling up* prototypes to operational systems ▪ Very limited cognitive models[15] ▪ Perfunctory test and evaluation against *toy* problems[8] ▪ No proven technique for knowledge engineering[2]
Level 3: Threat Refinement	▪ Same as Level 2 Processing ▪ Limited advisory status ▪ Limited deployment experience ▪ Dominated by *ad hoc* methods ▪ Doctrine-specific, fragile implementations	▪ Same as level 2 ▪ Difficulty to quantify *intent*[6] ▪ Models require established enemy doctrine ▪ Difficult to model rapidly evolving situations
Level 4: Process Refinement	▪ Robust methods for single-sensor systems ▪ Formulations based on operations research[2] ▪ Limited context-based reasoning ▪ Focus on measures of performance (MOP) versus measures of effectiveness (MOE)[1]	▪ Difficult to incorporate mission constraints ▪ Scaling problem when many sensors (10^N) and adaptive systems[36] ▪ Difficult to optimally use non-commensurate sensors ▪ Very difficult to link human information needs to sensor control[28]
Human Computer Interface (HCI)	▪ HCI dominated by the *technology of the week* ▪ Focus on ergonomic versus cognitive-based design ▪ Numerous graphics-based displays and systems[30, 31] ▪ Advanced, 3-D full immersion HCI available[32] and haptic interfaces[33,43]	▪ Very little research has been performed to understand how human analysts' process data and make accurate inferences. ▪ Creative HCI is needed to adapt to individual users and to provide mitigation of known cognitive biases and illusions[15,35]
Data Base Management	▪ Extensive use of 4th and 5th generation COTS DBMS ▪ DBMS individually optimized for text, signal data, imagery, or symbolic information (but not the intersection of any two) ▪ DBMS requires extensive tailoring for individual data fusion systems	▪ Need a generalized DBMS capability for text, signal data, images, and symbolic information ▪ Need a software solution to multi-level security

Current Level One processing is dominated by estimation techniques such as Kalman filters[4], multiple hypothesis tracking (MHT)[4], joint probabilistic data association (JPDA) filters[21], or related techniques. The problem of identity declaration is generally performed using a feature-based, pattern recognition approach[2,29]. This involves representing the sensor data using extracted features (e.g., spectral peaks in a radar cross section observation) and mapping the feature vector to a location in feature-space that can be uniquely identified with a target class or identity. Typical techniques include artificial neural networks (ANN) or cluster algorithms[29]. This identification process works well when there is a unique map between the observed features and the target class, but requires a significant amount of training data. However, the methods fail when training data is lacking[27], or there is ambiguity in the feature-to-target class mapping. Emerging methods include both model-based techniques and syntactic methods that develop descriptions of the makeup of a target in terms of elementary components.

Aubrey Poore[19] and R. Mahler[20] developed two promising methods in Level One fusion. Poore revisited the approach of separating the problems of object correlation, target tracking, and identity estimation. Poore re-links these problems into one single optimization problem having multiple constraints (viz., find the set of state vectors (including the association between observations and tracks) that best fits the observational data). While this larger problem is even more difficult than the original sub-problems, Poore has developed approximation methods to improve the computational feasibility. By contrast, Mahler has developed applications of random set theory to address the joint problem of data association and state estimation. A unified method based on Bayesian inference has been used by Barlow, Stone and Finn[40] to simultaneously estimate target state, identity, and association of the data. Finally, an extensive survey of methods for data correlation has been performed by Llinas *et al*[41].

Level 2: Situation Refinement: Level Two processing seeks to understand the relationships among observed entities and their relationship to the environment. This process involves recognition of patterns, context-based reasoning, and understanding of spatial, temporal, causal, and functional relationships. In general this is a difficult problem. There are numerous prototype systems that have been developed for DoD applications[8]. The predominance of the methods involves knowledge-based systems utilizing production rules[2], fuzzy logic[25], logical templates[22], or case-based reasoning[24]. Emerging systems are beginning to utilize agent-based approaches[26] and blackboard architectures[23].

While this is a very active area of research, the results to date are relatively disappointing. Very few operational systems have been deployed. Many of the prototype systems have addressed limited or *toy* problems with little or no test and evaluation. There is little experience on how to scale these small prototype systems to larger scale operational systems. A key problem for Level Two processing (as well as for Level Three) is the lack of cognitive models for how to perform situation assessment. Current cognitive models can be described as *pathetic*. We simply do not know how to model the reasoning process to perform a *gestalt* type of situation assessment. Numerous *ad hoc* methods (e.g., rules, frames, fuzzy logic, decision trees, scripts, templates, etc.) have been applied. One difficulty involves how to perform the knowledge engineering to identify the key information, inter-relationships, and the associated uncertainty information. Here again, Mahler's random set theory[20] provides a basis for a unified calculus of uncertainty. However, the application to realistic problems is far from routine. A general implementation approach has not yet been developed.

We suggest that improvements to Level Two processing will emerge from an improved understanding of how to select and use existing methods for knowledge representation (e.g., rules, frames, scripts, fuzzy logic), coupled with a better understanding of the strengths and weaknesses of human cognition for these types of tasks. One example would be the incorporation of so-called negative information in reasoning. Negative information involves reasoning about information that has not been observed (but would be expected to for a hypothesized situation). The use of negative reasoning appears to be a key element of successful diagnosis and inference in many areas such as medical diagnosis or diagnosis of mechanical faults[34.] Another promising area for research involves the development of aids for analysts that would address known cognitive biases and shortcomings (e.g., confirmation bias in which humans seek information that confirms a proposed hypothesis rather than evidence that refutes the hypothesis, miss-use of probability, and other biases[15,35]). The original research by J. Wohl[42] and his associates to develop tools for assisting an antisubmarine warfare (ASW) analyst is particularly intriguing. The research suggests that some fairly simple cognitive aids could be developed to significantly improve the data fusion/analysis process.

Level 3: Threat Refinement: Level Three processing involves an interpretation of the situation from a consequences point of view. That is, what is the meaning of the situation in terms of potential opportunities and threats? Alternative hypotheses are generated and projected into the future to determine what are the likely courses of action for engagements, and the consequences of those courses of action. The state of Level Three processing is similar to that of Level Two. There are a number of prototype systems that have been developed, but few deployed systems. The main focus of Level Three processing has been the application of automated reasoning systems and techniques from the discipline of artificial intelligence. A special challenge for Level Three processing is the determination of enemy intent. Conceptually, the determination of an enemy's intent involves a mind-reading exercise; what will the enemy do, under what circumstances, and with what motivation? When a well-known enemy doctrine exists, this can be modeled using a variety of techniques. However, in modern conflict situations this doctrine is often unknown. Hence, it is challenging to automate the process of threat refinement. Another problem for threat refinement is the role of adaptive intelligence opponents. How can engagements be modeled in which an opponent adapts to the actions of a protagonist? Much research has been performed in game theory to address this issue, but there is limited success in applying this work to realistic tactical situations.

Level 4: Process Refinement: The Level Four process is a meta-process; it is a process that monitors the overall data fusion process and seeks to optimize the data fusion within operational and physical constraints[1,2]. Types of functions within Level Four processing include generation of sensor *look angles* (to indicate where to point the sensors to track targets), computation of measures of performance (MOP) and measures of effectiveness (MOE), determination of information needs and sources, and process optimization. Level Four processing is relatively mature for single sensor environments. For single sensors, or a small number of commensurate sensors, Level Four processing becomes a routine problem in multi-objective optimization. This is an area that has received an extensive amount of research, e.g., for applications such as industrial process control.

The Level Four process becomes more challenging under a number of circumstances. These include use of a large number of sensors, use of co-dependent sensors, utilization of non-commensurate sensors (e.g., measuring very diverse physical phenomena on greatly different time scales), and use of sensors in a geographically distributed environment. Modern data fusion systems often involve geographically distributed collection and processing with adaptive systems that self-adjust for system failures and other problems[36]. Under these circumstances it is difficult to develop global MOE and MOP models and to optimize the overall system performance. Another challenge involves modeling sensor performance in realistic data collection environments. Finally, the most effective Level Four process would link the information needs of a human decision-maker to the sensor and source tasking in real-time.

Much research remains to be performed in the Level Four area. However, the improved intelligence and agility of modern sensors makes this an area in which major improvements can be obtained with relatively modest effort. Current research being conducted by M. Nixon[37] using economic theory to model resource utilization is very intriguing.

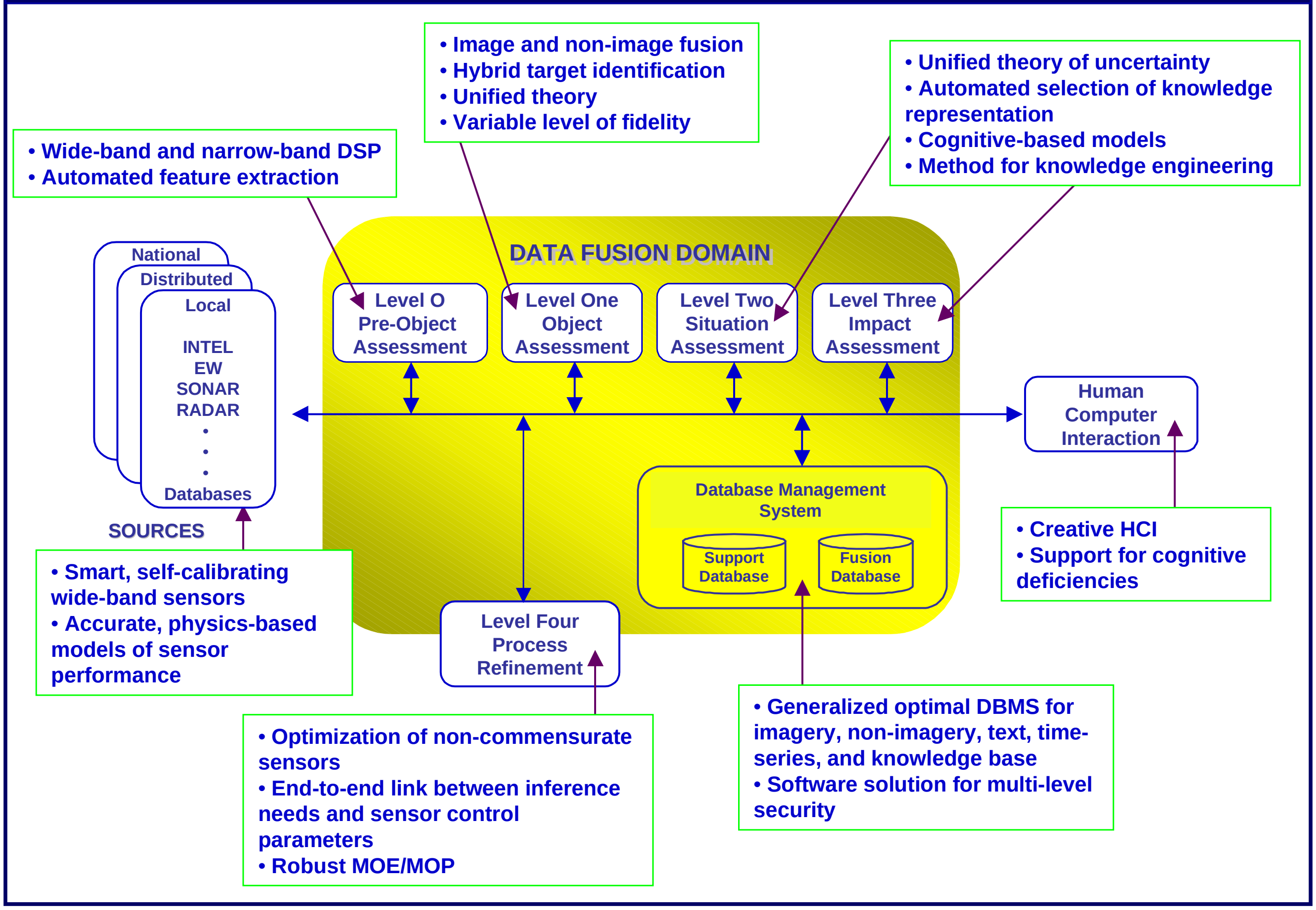
Human Computer Interface: The human-computer interface (HCI) area in data fusion is one that appears to be technology rich and theory poor. M. J. Hall, S. A. Hall, and Tate[15] point out that there is a rapidly evolving capability in HCI technology to provide interfaces such as full-immersion, three dimensional displays[22], haptic interfaces[33,43], three dimensional sound, and other types of interfaces to access and analyze data. However, they note that these interfaces smack of the *technology du jour* and have not been applied with a solid theoretical understanding of how humans access and respond to information displays. Many of the existing HCI for data fusion systems involve geographical information system (GIS) type displays and data access[30,31]. While these are useful, it is not clear that these interfaces truly assist the understanding of information available from a data fusion system, or whether they may actually impede the inference process. B. Feran[38] has argued that the HCI for intelligence systems can actually act as a bottleneck that limits the ability of a user to access and analyze data. Other studies have investigated the issue of trust in decision support systems (Llinas *et*. *al*[39]), and how the HCI affects the extent to which a user believes and trusts the results.

Data Base Management: The final area to be addressed involves data base management for data fusion systems. This is an important area for data fusion systems for several reasons. First, data base management software constitutes the single major amount of software to be developed for a data fusion system (even if one utilizes sophisticated commercial-off-the-shelf (COTS) DBMS packages)[2]. Data required for fusion systems ranges from sensor data (e.g., scalars, vectors, time series, images), to information input by human users, environmental data, textual information, and knowledge such as doctrine. The data base management for a data fusion system must simultaneously accept data at the rate provided by the contributing sensors, and also allow algorithms and users to rapidly retrieve large amounts of data using general Boolean queries. The combination of the complexity of the data sets and need for real-time data storage and retrieval complicates the data base management for data fusion. In addition, the data associated with fusion systems often involves multiple levels of security. Handling multi-level security is currently difficult to do via a software approach. For all of these reasons, extensive special software must be implemented for data fusion systems.

4.0 Research Needs

There are a number of areas of research that could provide value to the data fusion community and improve the ability to develop robust systems. A summary of these research areas is shown in Figure 4 and described below.

Figure 4: Technology Needs in Data Fusion



Data sources: New sensor types and sources are always sought for data fusion applications. The rapid evolution of microprocessors and nano-fabrication techniques provides a basis for rapid evolution of sensors. New, smart, self-calibrating and wide-band sensors would be welcomed for many DoD applications. In addition, accurate, physics-based models of sensor performance could be used to improve the down-stream data fusion processing.

Source Preprocessing: Current advances in digital signal processing and image processing are based on new algorithms and improvements in computer processing speeds and data storage. Advances in source pre-processing will likely come from the application of new wide-band digital signal processing, incorporation of coherent processing (of multi-sensor data), and automated algorithm selection and utilization. For target classification and identification, the ability to perform automated feature extraction would be particularly useful.

Level One Object Refinement: Improvements in level one processing are needed in several areas. These include data level fusion of non-commensurate sensor data (e.g., fusion of image and non-image data) using physics based target and sensor models, and improved target identification using hybrid methods that incorporate target models, human-analyst information, and implicit information learned from the sensor data. It would be very useful to have a better understanding of multiple methods of representing uncertainty, and how to select appropriate ways of representing information. One approach that might be fruitful is to investigate techniques that operate in a hierarchical manner at varying levels of fidelity (e.g., tracking of individual targets, target groups, and general target populations or classification methods that provide varying levels of target identity on demand).

Level Two Situation Refinement and Level Three Threat Refinement: Much work is needed in the Level Two and Level Three areas. Basic cognitive models are needed concerning how to make inferences and decisions about a situation and threat. A unified and practical theory (or calculus) of uncertainty is needed. Automated methods are needed to select appropriate knowledge representation techniques. New methods and tools are required to perform knowledge representation for automated reasoning. Work is required to develop techniques that are more robust (and not as fragile as the current methods). It would be useful to try both a *drill-down* approach as well as a *thin covering* approach. In the *drill-down* method, one might select a very well-bounded problem in situation assessment and attempt to completely solve the problem by a combination of physical models, multiple automated reasoning methods, and *ad hoc* algorithms (i.e., drill down to obtain a complete solution to a narrow problem). In the *thin covering* approach, a broader problem would be selected and addressed. However, the solution would not seek the level of fidelity used for the *drill-down* approach. The results of these approaches could provide valuable insight into how to approach the general Level Two and level three problems.

Human Computer Interface (HCI): The rapid evolution of HCI technologies (e.g., 3-D, haptic interfaces, and natural language processing) should continue to be applied to data fusion systems. However, much more creativity is needed to improve the link between the fusion system and the human. The suggestions by M. J. Hall, S. A. Hall, and Tate[15] (e.g., deliberate synesthesia, time compression/expansion, negative reasoning enhancement, focus/de-focus, pattern morphing, and new uncertainty representation methods) provide an excellent starting point for new HCI research. In addition, more research is needed to understand human cognitive deficiencies and information access preferences. Based on this research, new tools should be developed to enhance the link between a data fusion system and effective human cognition. The focus of this research should be human-centered fusion.

Database Management: New data base management (DBMS) models are needed for data fusion systems. Instead of trying to *cobble together* existing techniques for representing images, signals, text, knowledge, and other data. New models should be developed that *begin with* the requirement for an integrated representation scheme. Software based solutions are also required for multi-level security. On-going research in areas such as distributed data storage and retrieval, data compression, natural-language interfaces to DBMS, improved access and storage schemes, data mining, and related areas should be monitored and applied to the data fusion problem. This is an area in which the commercial market (e.g., for electronic commerce and business) will provide an impetus for significant improvements).

Level Four Processing: Improvements in Level Four processing could have a very large affect on the effectiveness of data fusion systems. The rapid advances in sensors, and the ability to utilize hundreds or thousands of sensors provide both an opportunity and challenge for data fusion systems. New multi-objective, multi-constraint optimization methods are needed to effectively use these sensors. Special areas of research include the effective use of highly non-commensurate sensors (especially those that operate on a greatly different time scale). The link between sensors and the human user needs to be strengthened (to provide an information-based optimization). Research is needed to develop general measures of performance and measures of effectiveness.

Infrastructure Needs: To support the evolving research a strong infrastructure is required for the data fusion community. The data fusion information access center (FUSIAC) could play a strong role for this infrastructure. Key elements include; (1) a set of standard algorithms and software, (2) one or more test-beds to provide a gold standard for algorithm evaluation, (3) warehouses of models for sensors and the environment, and (4) a communication forum. Of particular value would be a universal test case (i.e., a *Lena-world*) for evaluating algorithms. The image processing community, for example, has used a standard picture (of the Playboy model Lena) for evaluating and comparing algorithms. They have also made effective use of a visual programming toolkit (*Khoros*), funded by DARPA, to perform rapid prototyping of image processing techniques. Such a toolkit would be of value to the data fusion community.

5.0 Pitfalls in Data Fusion

The previous part of this paper has provided a broad overview of the state of data fusion technology and identification of potential research issues. A practitioner might well ask the question; so what do I do tomorrow to implement a system? What are some problems and challenges need to be addressed? It is well beyond the scope of this paper to provide a prescription for the implementation of data fusion systems. However, there are several areas worth noting. First, Bowman and Steinberg[11] provide an overview of the general systems engineering approach for implementation of data fusion systems. Engineering guidelines for selection of correlation algorithms are described by Llinas *et al*[41.] Several texts, such as those of Hall[2] and Waltz and Llinas[3] provide detailed information on data fusion algorithms. R. Antony[5] describes issues in data base management systems, and texts are available on specific applications to target tracking (e.g., Blackman[4]) and signal processing techniques[44].

Hall and Garga[13] have discussed the problem of implementing data fusion systems and identified a number of problems or pitfalls. These include the following dictums.

- *There is no substitute for a good sensor:* no amount of data fusion can substitute for a single accurate sensor that measures the phenomena that you want to observe.

- *Downstream processing cannot make up for errors (or failures) in upstream processing:* data fusion processing cannot correct for errors in processing (or lack of pre-processing) of individual sensor data.

- *Sensor fusion can result in poor performance if incorrect information about sensor performance is used:* A common failure in data fusion is to characterize the sensor performance in an *ad hoc* or convenient way. Failure to accurately model sensor performance will result in corruption of the fused results.

- *There is no such thing as a magic or golden data fusion algorithm:* Despite claims to the contrary, there is no perfect algorithm that is optimal under all conditions. Often real applications do not meet the underlying assumptions required by data fusion algorithms (e.g., available prior probabilities or statistically independent sources).

- *There will never be enough training data:* In general there will never be sufficient training data for pattern recognition algorithms used for automatic target recognition or IFFN. Hence, hybrid methods must be used (e.g., model-based methods, syntax representations, or combinations of methods).

- *It is difficult to quantify the value of a data fusion system:* A challenge in data fusion systems is to quantify the utility of the system at a mission level. While measure of performance can be obtained for sensors or processing algorithms, measures of mission effectiveness are difficult to define[1].

- *Fusion is not a static process:* The data fusion process is not static, but rather an iterative dynamic process that seeks to continually refine the estimates about an observed situation or threat environment.

We note that these issues must be addressed for implementation of an effective data fusion system.

6.0 Summary

The technology of multisensor data fusion has made major strides in the past two decades. Extensive research has been performed on data fusion algorithms, distributed architectures, automated reasoning techniques, and new resource allocation and optimization techniques. There is an emerging consensus in the data fusion community concerning basic terminology and engineering guidelines. Recent activities to initiate a data fusion information analysis center (FUSIAC) promise to accelerate the development of data fusion technology by increasing the communications among researchers and system implementers. Despite these rapid advances, however, much research remains to be done. This paper has presented a perspective on the current limitations and challenges in data fusion, and identified recommended areas of research.

7.0 References

[1] E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Inc., Norwood, MA, 1990.

[2] D. L. Hall, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, Inc., Norwood, MA, 1992.

[3] D. L. Hall and J. Llinas, “An introduction to multisensor data fusion,” *Proceedings of the IEEE*, vol. 85, No. 1, January 1997, pp. 6-23.

[4] S. S. Blackman, *Multiple-Target Tracking with Radar Applications*,” Artech House, Inc., Norwood, MA, 1987.

[5] R. T. Antony, *Principles of Data Fusion Automation*, Artech House, Inc., Norwood, MA, 1995.

[6] D. L. Hall, *Lectures in Multisensor Data Fusion*, Artech House, Inc., Norwood, MA, 2000.

[7] D. L. Hall and R. J. Linn, “Algorithm selection for data fusion systems,” *Proceedings of the 1987 Tri-Service Data Fusion Symposium*, APL Johns Hopkins University, Laurel, MD, Vol. 1, pp. 100-110, June 1987.

[8] D. L. Hall, R. J. Linn, and J. Llinas, "A survey of data fusion systems," *Proceedings of the SPIE Conference on Data Structures and Target Classification*, vol. 1470, Orlando, FL, April, 1991, pp. 13-36.

[9] O. Kessler, et al, *Functional Description of the Data Fusion Process*, Technical Report, Office of Naval Technology, Naval Air Development Center, Warminster, PA, January 1992.

[10] F. E. White, *Data Fusion Lexicon*, Data Fusion Sub-Panel of the Joint Directors of Laboratories Technical Panel for C^3 NOSC, San Diego, CA, 1991.

[11] A. N. Steinberg and C. L. Bowman, "Development and application of data fusion system engineering guidelines," *Proceedings of the National Symposium on Sensor and Data Fusion (NSSDF)*, Lexington, MA, May 1998.

[12] D. L. Hall and J. Llinas, "A challenge for the data fusion community I: research imperatives for improved processing," *Proceedings of the 7th National Symposium on Sensor Fusion,* Albuquerque, NM, March 1994.

[13] D. L. Hall and A. K. Garga, "Pitfalls in data fusion (and how to avoid them)," *Proceedings of the 2nd International Conference on Information Fusion – Fusion '99*, vol. 1, pp. 429-436, July 6-8, 1999, Sunnyvale, CA.

[14] D. L. Hall and J. Llinas, "From GI Joe to starship trooper: the evolution of information support for individual soldiers," *Proceedings of the International Conference on Circuits and Systems*, June 1998, Monterey, CA.

[15] M. J. Hall, S. A. Hall, and T. Tate, "Removing the HCI bottleneck: how the human computer interface (HCI) affects the performance of data fusion systems," *Proceedings of the 2000 Meeting of the MSS, National Symposium on Sensor and Data Fusion*, San Antonio, Texas, June 2000.

[16] A. Steinberg, "Standardization in data fusion: new developments," *Proceedings of EuroFusion99: International Conference on Data Fusion,* 5-7 October 1999, Straford-upon-Avon, UK, pp. 269-278.

[17] A. N. Steinberg and C. Bowman, "Revisions to the JDL data fusion process model," *Proceedings of the National Symposium on Sensor and Data Fusion (NSSDF)*, Lexington, MA, May 1998.

[18] C. Morefield, AlphaTech Corporation, private communication to D. L. Hall and A. N. Steinberg, May 18, 2000.

[19] A. Poore, "Multi-dimensional assignment formulation of data association problems arising from multi-target and multi-sensor tracking," *Computational Optimization Applications*, vol. 3, pp. 27-57, 1994.

[20] R. Mahler, "A unified foundation for data fusion," in *Proceedings of the 1994 Data Fusion Systems Conference*, Applied Physics Laboratory, Johns Hopkins University, June 1987.

[21] Y. Bar-Shalom, ed., *Multi-Target, Multi-Sensor Tracking: Applications and Advances*, Artech House, Norwood, MA. 1989.

[22] D. L. Hall and R. J. Linn, "Comments on the use of templating for multi-sensor data fusion," *Proceedings of the 1989 Tri-Service Data Fusion Symposium*, Vol. 1, pp. 345-354, May 1989.

[23] J. Llinas and R. Antony, "Blackboard concepts for data fusion and command and control applications," *International Journal on Pattern Recognition and Artificial Intelligence*, vol. 7, no. 2, April 1993.

[24] I. Watson and F. Marir, "Case-based reasoning: a review," *The Knowledge Engineering Review*, Vol. 9, No. 4, 1994 (http://www.surveying.salford.ac.uk…br-mirror/classroom/cbt-review.htm).

[25] R. Gibson, D. L. Hall and J. Stover, "An autonomous fuzzy logic architecture for multi-sensor data fusion," *Proceedings of the 1994 Conference on Multi-Sensor Fusion and Integration for Intelligent Systems*, Las Vegas, NV. Oct. 1994, pp. 143-150.

[26] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall Series in Artificial Intelligence, Prentice Hall, New Jersey, 1995.

[27] D. Mush and B. Horne, "Progress in supervised neural networks: what's new since Lippman?" *IEEE Signal Processing Magazine*, pp. 8-39, January 1993.

[28] D. L. Hall and A. K. Garga, "New perspectives on level four processing in data fusion systems," *Proceedings of the SPIE AeroSense'99 Conference: Digitization of the Battlefield IV*, April 1999, Orlando, FL.

[29] K. Fukanaga, *Introduction to Statistical Pattern Recognition*, 2nd, ed., Academic Press, NY, 1990.

[30] B. E. Brendle, Jr., "Crewman's associate: interfacing to the digitized battlefield," *Proceedings of the SPIE: Digitization of the Battlefield II*, Vol, 3080, Orlando, Florida, 22-24 April 1997, pp. 195-202.

[31] A. Steele, V. Marzen, B. Corona, "Army Research Laboratory advanced displays and interactive displays, Fedlab technology transitions," *Proceedings of the SPIE: Digitization of the Battlespace IV*, Vol. 3709, Orlando, Florida, 7-8 April 1999, pp. 205-212.

[32] The CAVE at NCSA, http://www.ncsa.uiuc.edu/VEG/ncsaCAVE.html.

[33] R. E. Ellis, O. M. Ismaeil, and M. Lipsett, "Design and evaluation of a high-performance haptic interface," *Robotica*, vol. 14, pp. 321-327, 1996.

[34] D. L. Hall, R. J. Hansen, and D. C. Lang, "The negative information problem in mechanical diagnostics," *Transactions of the ASME*, vol. 119, April 1997, pp. 370-377.

[35] M. Piattelli-Palmarini, *Inevitable Illusions: How Mistakes of Reason Rule over Minds*, John Wiley and Sons, New York, 1994.

[36] D. L. Hall, P. Sexton, M. Warren, and J. Zmyslo, “Dynamo: A tool for modeling integrated air defense systems,” *Proceedings of the 2000 Meeting of the MSS, National Symposium on Sensor and Data Fusion*, San Antonio, Texas, June 2000.

[37] M. Nixon, “Application of economic auction methods to resource allocation and optimization”, presented to an NRO Workshop, Washington, DC., April 17, 2000.

[38] B. Feran, presentation to the National Reconnaissance Office (NRO), spring, 1999 (available on video from the NRO office in Chantilly, Virginia).

[39] J. Llinas, C. Drury, W. Bialas, An-che Chen, *Studies and Analyses of Vulnerabilities in Aided Adversarial Decision-Making*, Technical Report, State University of New York at Buffalo, Dept. of Industrial Engineering, February 1997.

[40] C. A. Barbar, L. D. Stone, and M. V. Hun, “Unified data fusion,” *Proceedings of the 9th National Symposium on Sensor Fusion*, Vol. 1, 12-14 March 1996, pp. 321-330.

[41] J. Llinas, B. Neuenfeldt, L. McConnell, D. Bohney, C. Bowman, D. Hall, J. Lochacki, and P. Applegate, “Studies and analyses within project Correlation: an in-depth assessment of correlation problems and solution techiques,” *Proceedings of the 9th National Symposium on Sensor Fusion*, Vol. 1, 12-14 March 1996, pp. 171-188.

[42] J. G. Wohl, E. E. Entin, D. Serfaty, R. M. James, and J. C. Deckert, “Human cognitive performance in ASW data fusion,” *Proceedings of the 1987 Tri-Service Data Fusion Symposium*, Johns Hopkins University, Applied Physics Laboratory, Laurel MD, 9-11 June 1987, pp. 465-479.

[43] http://haptic.mech.nwu.edu/

[44] D. C. Swanson, *Signal Processing for Intelligent Sensing Systems*, Marcel Dekker, Inc. 2000.