2007

Data Fusion in Defense and National Security: Ubiquitous and Indispensable

Pranab K. Banerjee
Utah State University

Data Fusion in Defense and National Security: Ubiquitous and Indispensable

Pranab K. Banerjee
Space Dynamics Laboratory, 1695 N. Research Park Way, North Logan, Utah 84341, USA
Phone: (435)-797-4207, FAX: (435)-797-4159, Email: Pranab.Banerjee@sdl.usu.edu

1 Overview

Data fusion refers to exploitation of information originating from different sources, homogeneous or heterogeneous, in order to achieve levels and/or quality of knowledge or awareness about the sensed domains that no one source could provide. Fusion of multi-source information can result in global awareness from local awareness, reduced uncertainty or ambiguity, higher information resolution, increased dimensionality, and better reliability through redundancy. These make multi-sensor information fusion a highly desirable and critical technology for use in defense and national security. However, there are quite a few challenges on the way. Continuing advances in sensor hardware present us with increasing diversity of data sources that are non-trivial to align, correlate, and associate in order to produce a unified knowledge product. Also, different sources in a multi-sensor infrastructure can produce conflicting information. This can result from sensor malfunction, local fluctuations in the environment, unreliable communication channel etc. A fusion system need to handle such conflicts in a reasonable manner. In addition, data volume can pose serious challenge to a fusion framework. As the number of information source increases, the amount of data input to the fusion framework goes up as well. This can create computational bottlenecks affecting performance and effectiveness of a fusion system.

2 The enviable fusion engine

One of the most efficient data fusion engines in existence is the human cognitive system. We routinely fuse information from the diverse senses (vision, sound, touch, smell and taste) in real time in order to gain understanding of the environment around us. This is so natural and effortless to us that we often do not realize that multi-sensor heterogeneous data fusion is indeed a non-trivial task. While the human cognitive engine has a highly parallel framework composed of billions of neurons which have been well tuned for such fusion tasks over millions of years of evolutionary processes, the largest parallel computer we have at our disposal today consists of a tiny fraction of computing nodes compared to the human brain, and is quite primitive in comparison. Besides, the inner functions of the human brain are not yet fully understood and hence we do not have a well defined set of fusion algorithms to replicate in order to mimic the functionality of the human cognitive system. However, the data fusion community should keep up with the advances in neuroscience and cognitive modeling, and borrow relevant knowledge that would facilitate evolutionary as well as revolutionary improvements to data fusion algorithms and architectures. In the mean time, continuing advances in computer hardware have made data fusion an ubiquitous and indispensable technology for defense and national security.

3 Data fusion in defense and national security

A critical area of national security is surveillance and intelligence analysis for timely situational awareness. This inherently depends on highly disparate multi-source information analysis. Useful intelligence can be embedded in any modality of information and it is not at all unusual to discover a piece of intelligence in one modality but not in the rest. Hence, an effective data fusion framework is essential in this domain for coherent fusion of diverse intelligence data streams (HUMINT, IMINT, SIGINT, ELINT etc.) for gaining critical situational awareness.

Another critical area of application is the monitoring of health status of remote and space based assets for national security. For example, a surveillance satellite can have a set of sensors to track the status of different critical subsystems. It is of great importance to be able to fuse information from these sensors to create a global picture of the health of the spacecraft which may allow us to predict an impending failure and correct it before it reaches criticality.

In addition, data fusion is essential for guidance of autonomous vehicles, target recognition, battlefield asset allocation, emergency medical treatments of war fighters, and many other areas.

4 Data fusion architectures

Data sources constituting a fusion framework can be classified into three broad categories based on operational modes: (i) *complementary sources* are independent of each other and provide non-overlapping information which are fused to get wider coverage. An example application is wide area persistent surveillance through a set of geo-spatially fixed sensors, each having a narrow surveillance footprint. (ii) *competitive sources* provide highly overlapping information and is useful for providing redundancy in critical asset monitoring. (iii) *cooperative sources* provide parallel information about the same domain to facilitate knowledge augmentation through fusion. For example, an area could be imaged with a radar as well as an infrared imager to track targets based on shape as well thermal profile through fusion of these cooperative sensor data.

Data fusion can be categorized into three main classes based on the level of information abstraction used for fusion: (i) *Raw data fusion* is primarily limited to fusion of homogeneous modalities and requires proper data alignment. (ii) In *feature vector fusion*, a feature extractor is co-located with the sensor and the feature vectors are sent to the fusion node. This is more versatile and can handle disparate data modalities. It also has the advantage of significantly reduced data transfer which is crucial in bandwidth limited environments. (iii) *Decision fusion* combines information at the highest level of abstraction but requires significant computational capabilities, since the sensor nodes need to have classifiers built in. Usually, *feature vector fusion* provides a good middle ground in terms of cost and complexity of the nodes.

The traditional fusion framework consists of a central fusion engine that receives information from individual sources at appropriate levels of abstraction. This has disadvantages such as single point of failure, higher bandwidth requirement for the fusion node etc. Recently, researchers have proposed distributed data fusion architectures where there is no single information aggregator but various nodes can autonomously and dynamically assume this role for an appropriate time span. The nodes can assume other fusion roles as well if needed.

An important component of any heterogeneous data fusion system is the "ontology bridge" that translates a domain specific ontology to a common thesaurus. This is a critical step in mapping disparate sensor specific data and information structures to a uniform representation suitable for a fusion engine.

5 Visualization: keeping the human analyst in mind

The ultimate consumer of any fused information product is the human analyst who needs to gain critical situational awareness in an effective and timely manner. Visualization is key to facilitating this cognitive process since the human faculty of vision is good at comprehending large amount of information from a compact visual representation. Besides, with an appropriately designed visualization scheme, part of the fusion process can be performed by the human brain, thus exploiting the man-machine synergy.

Visualization for presenting a fused knowledge product in the defense and national security domain needs to adopt a different paradigm from most traditional visualization systems, however. This is primarily dictated by the relatively short temporal relevancy of information in the highly dynamic world of surveillance, battlefields, and other national security domains. Traditionally, visualization has been thought of as a post-processing step. However, because of the real time or near-real time requirements in defense and security, visualization needs to become an active and fully interactive component of the real time information flow loop. From a graphics perspective, this is more demanding than the traditional paradigm. However, the tremendous rate of advance in the graphics processing unit (GPU) hardware in recent years fueled by the demands of the computer gaming industry offers many interesting possibilities in this regard. Combination of fast CPUs and GPUs currently available even makes it feasible to adopt a multi-sensory information representation (combination of visual, acoustic, haptic etc.) for enhanced human comprehension through the use of higher cognitive bandwidth.