

Improving the Fusion Process Using Semantic Level *Whole-Brain* Analysis

David L. Hall¹, PhD

M. D. McNeese¹

Eileen Rotthoff²

Tim Shaw²

James Wang¹, PhD

¹The School of Information Sciences and Technology

²The Applied Research Laboratory
The Pennsylvania State University

Abstract

Extensive research in the fusion community has focused on developing techniques for level 1 fusion. The bulk of the literature on multisensor data fusion focuses on the automation of target tracking and automatic target recognition. While such research is needed, current problems involve complexities such as identifying and tracking individual people and groups of people, monitoring global activities and recognition of events that may be a precursor to terrorist activities. The requisite data for this analysis involves sensor data (including signals, images, and vector quantities), textual information (from web sites and human reports), and utilization of models. This analysis process is very human intensive, requiring teams of analysts to search for data, interpret the results, develop alternative hypotheses, and assessment of the consequences of such hypotheses.

This paper describes the development of a set of tools aimed at level-5 fusion to support *whole-brain* analysis. The toolkit encourages analysts to use and integrate both their visual intelligence and their language processing ability. We have developed a set of tools to support multi-INT analysts including tools for; (1) automated semantic labeling of images, (2) collaboration tools, (3) cognitive aids, (4) advanced visualization aids, and (5) context-based interpretation techniques. These tools are being implemented in an open architecture involving a commercial off the shelf geographical information system (GIS). In addition to the tool kit, a "living laboratory environment" has been developed to allow the quantitative evaluation of these aids in a repeatable, experimental environment.

Whole-brain Intelligence Analysis Concept

A fundamental paradox exists in information fusion. Information fusion in this context may be used in traditional areas such as national defense, counter-intelligence, situation assessment for tactical military

applications (Hall and Llinas (2001) and Hall and McMullen (2004)) or non-Department of Defense (DoD) applications such as environmental monitoring, technology assessment for business applications or related areas (Llinas and Hall 1994). The paradox is that information analysts are drowning in a sea of data but unable to obtain the knowledge that they need to address difficult problems. This has often been referred to as the data overload dilemma (Kuperman, 2001) or more recently framed "cogminutia fragmentosa" (McNeese and Vidulich, 2001).

On one hand, an unprecedented capability exists to collect data via distributed sensors, commercial information providers (e.g., AccuWeather, Library Services, and commercial search businesses), human sources, or Internet resources. Smart micro-scale sensors (Jones 1995), wireless communications, and global Internet accessible resources enable the entire earth to be a potential information resource (the I-earth observatory). Such information is available literally at the fingertips of the analysts. In particular, the Internet has exceeded one billion web pages, with a continuing exponential increase. The wealth of data has not produced a commensurate improvement in analyst abilities, however. Analysts are literally swamped with data. They have a wide variety of choices to make as to what is *useful* and *usable*, given the context of what they are trying to *understand* (Woods, 1998).

On the other hand, the glut of data can be overwhelming and may inadvertently promote poor decision processes. Studies of decision-making under stress have shown that too much information can cause ineffective decision styles (Klein et al (1995)). An example is the *hyper-vigilance mode*, in which a decision-maker frantically searches for new information, without taking time for reflection and thoughtful analysis of existing data. The huge glut of rapidly changing data via the Internet may encourage this type of response. Alternatively, a decision-maker may feel overwhelmed with new information

and simply ignore new data. Thus, in a rich atmosphere of data, decision-makers are suffocating for knowledge (McNeese and Vidulich 2001). They may have a large amount of cognitive readiness available to fuse multiple information sources but in fact their meta-cognition (McNeese, 2000) may be very limited. This often makes decisions about "what to do next" daunting.

Through the use of contemporary cognitive systems engineering approaches (e.g., *The Living Lab Framework*, McNeese, 2002), a vision for information-based fusion is to transform the current analyst dilemma to a more productive situation in which the analyst can directly access information.

Our continuing research seeks to develop tools, models and techniques to allow an analyst to effectively use the entire earth's observing resources for situation assessment.

The whole-brain analyst attains maximum value from incoming data inputs if these inputs are formulated in such a way as to extend the analyst's cognitive style. Such an analyst extracts information from both text and visual data, fusing them into a seamless set of knowledge constructs. This enables the analyst to answer questions about logic, feelings and thoughts that do not easily fall either into a single word or phrase, or that can be attributed to a single visual image.

In order to be able to utilize multi-source, multi-INT data, analysts must be able to process both image based data, as well as data based on signals, models, human sources, open text sources, etc. For simplicity we have illustrated in Figure 1 the concept of two basic types of information (image-based information and text-based information) being collected and sent to analysts for evaluation. In this figure, the information is "stove-piped", with different analysts processing each type of data.

The image analyst on the right hand side of Figure 1 conducts classic image processing and interpretation based on internal cognitive processes that convert image information to language constructs and subsequent internal reasoning. The support analyst on the left hand side of Figure 1 processes information from a language or text approach and converts that to use in collaboration with the image analyst.

Ideally, these processes could be conducted by the same analyst (or team of analysts) using a whole-

brain concept. Thus, a single analyst (or team of analysts) should have the capability to:

- Observe and interpret image data directly - observing the image, making internal cognitive assessments using his or her visual skills and language abilities;
- Be provided with a set of candidate semantic labels that describe or characterize the image (to support language based reasoning about the image data);
- Directly view and interpret the text based information; and
- Be provided with the capability to visualize the text based information in image form.

We suggest that analysts should have access both to image data and to semantic labels related to the images, and access to text information and to visualizations of that textual information. Hence, the image data should "move towards language" and the textual (language) data should "move towards imagery". In this way, individual and teams of analysts could simultaneously utilize their visual reasoning skills and their language-based reasoning skills (Pinker, 2000).

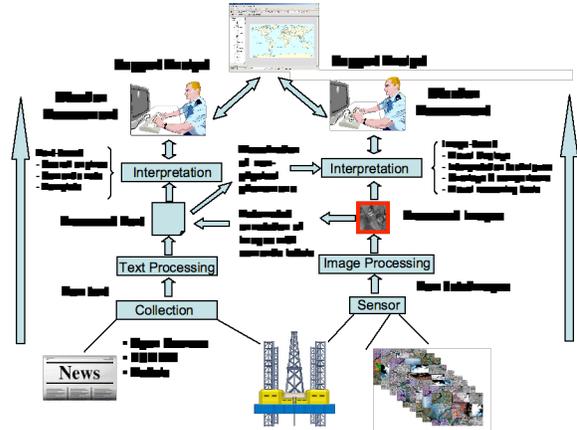


Figure 1: Concept of Whole Brain Analysis

Operational Concept for Information Analysis

A new operational concept was developed to describe how multiple cognitive tools and visualization aids may be used to support geospatial intelligence analysis. The concept is shown in Figure 2. In this concept, an area of interest (potentially the entire earth) is surveyed using multiple heterogeneous sensors. Data may be collected via satellite, airborne vehicles, unattended ground sensors, or from human

observation. The left hand side of the figure shows the concept of data received from multiple sensors, S_1, S_2, \dots , etc. In addition, data may be input or accessible from human observations, open source information (e.g., news reports, etc.). In general, an analyst or team of analysts may be virtually overwhelmed with data that may be of interest. While search engines are available for text-based information (e.g., adapted from the rapid evolution of Internet search engines such as Google™), less flexible techniques are available for images. Image data must be searched using either parametric data associated with the image (e.g., geospatial parameters) or using semantic labels that have been annotated to the image based on human inspection and characterization. The latter requires that a human has inspected each image and performed an initial assessment of image content or context. If thousands of images are available, this is a severe limitation.

- Support for “virtual” collaboration and transactive memory
- Anomaly detection and event prediction, and
- Analyst visualization and information interaction tools.

Each of these is summarized below. Together, these concepts are aimed at providing a basis for improved analysis by “conservation of analyst attention” – that is, by providing support to allow an analyst to focus on analysis and interpretation versus information searching (both of current data and archival information and system “memory”). In addition, improved communications among analysts and new visualization techniques will support increased effectiveness in the analysis process.

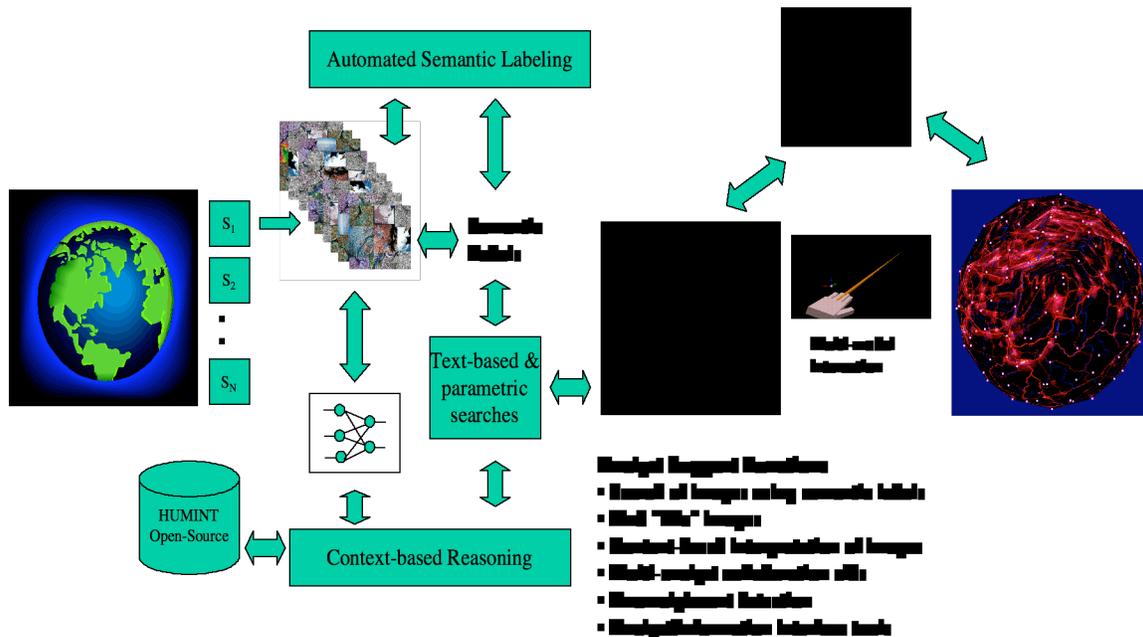


Figure 2: Operational Concept

The operational concept shown in Figure 2 provides a number of elements to support the analytical process. These include:

- Search of images using semantic labels
- Ability to find “like” images based on semantic characteristics
- Context-based interpretation of images
- Multi-analyst collaboration aids

Search for images using semantic labels – The left hand side of Figure 2 shows a function for “automated semantic labeling”. Using a technique developed by Wang and his associates (Wang, et al. 2000), image data can be processed to associate semantic labels for each image. The process involves representing image data using a feature vector involving wavelet coefficients and associating these feature vectors to semantic terms such as “forest”, “urban” or “threat”. In previous research, Wang has demonstrated this concept to characterize over 100,000 pictures from the web, described by 600

semantic terms. After an initial training process, the automated semantic labeling function allows images to be processed and automatically “labeled” with terms selected by the analyst. This will allow analysts to search for images based on both parametric information (e.g., image center, time, etc.) but also by labels. Consequently, searches can be performed using queries such as, “find all the images within the following area of interest, observed within the following time interval, characterized by labels such as “forest”, “threat”, etc. Hence, analysts can use both image and language in the search process. It must be noted that the semantic labeling is NOT equivalent to automatic target recognition, nor does it replace the visual interpretation capabilities of a human analyst. However, this function should allow analysts to sort through very large sets of images to find those of potential interest for further inspection/interpretation. Moreover, labeling images with semantic terms allow searches to be conducted that link textual information (e.g., the contents of news reports, reports by humans, etc.) to images.

Ability to find “like” images using semantic characteristics – Complementary to the image retrieval capability described above, an analyst could use the image characterization feature on an image-by-image basis. That is, an analyst could select an image of interest, use the semantic labeling tool for semantic annotation of the image, and then request retrieval of images that are similar to the image of interest based on similar feature characteristics and semantic labels. For example if an analyst finds something unusual in an image, this capability could be used to help determine whether there are other examples of such an image in the archival data base.

Context-based interpretation of images – This process allows an analyst to further interpret an image based on explicit (subject matter expert) rules as well as implicit information contained in the image. This function involves training a neural network to recognize patterns or image context using both implicit and explicit information. For example, an analyst might be interested in the characteristics of vegetation within an image or in determining the trafficability (e.g., the ability of a ground vehicle to traverse across a certain terrain) or determining the observability of a phenomena based on terrain and vegetation characteristics (e.g., the extent to which a target, activity or event could be observed based on factors such as terrain, vegetation, weather, etc.).

Experiments with this hybrid reasoning approach in various domains have demonstrated it to be very robust and significantly better than either the implicit

or explicit method used by itself (Garga 1996). In particular, research applying this technique to earth science information processing, analysis and understanding showed the capability to fuse non-commensurate, multi-source information and domain knowledge to model, represent, and display complex multi-dimensional terrestrial and atmospheric data (Shaw, et al. 2002).

The advantage of such a context-based reasoning tool for the geospatial intelligence analyst is that it can operate on information contained within an image as well as text-based information from human observers or open sources. In effect, this capability provides an added ability to interpret the context or meaning of information in imagery (augmented by additional source information).

Multi-analyst collaboration aids – It is apparent that the geospatial analyst’s work involves collaboration with other analysts and domain experts. The analyst must have the capability to share information and images with multiple other analysts, co-located and distributed, and with diverse domain experts to facilitate the collaboration necessary for the analysis task. Collaboration aids, including existing commercially available tools as well as tools created specifically for geospatial analyst collaboration can be tested and assessed within the Living Lab framework (McNeese 1996; McNeese et al. 2000). Effective collaboration tools would support sharing of information, sharing of expertise, and facilitation of multi-analyst collaboration to improve the overall accuracy and usefulness of intelligence products.

Support for “virtual” collaboration and transactive memory – In addition to tools to support collaboration among analysts, the assessment of analyst work demonstrated the potential utility for transactive memory tools. These would provide an ability to archive individual and corporate “memory” and expertise for analysis. Such a tool could partially address issues related to training of novice analysts and turnover of analysts. The ability for an analyst to query the computer to ask if the current situation (e.g., image, collection of data, etc.) is similar to previous cases addressed, or to access a “virtual” experienced analyst could improve the functioning of new and novice analysts.

Analyst visualization and information interaction tools – Previous discussions of the whole-brain analysis concept emphasized the need to transform images into semantic labels to allow analysts to utilize their language abilities and reasoning to understand and interpret an observed area of interest.

The converse of this need is also true – that is, the ability to represent information (e.g., non-physical concepts such as “threat”) using visual means. This allows transformation of ancillary data obtained via text or signals to be visualized, especially in relation to collected imagery.

This set of functions would provide novel ways to visualize information of interest to analysts to assist in situation assessment and interpretation. A major challenge in information visualization is to design appropriate visual representations of the non-physical data. An appropriate visual representation enhances the intelligibility of the complex data, enables the user to detect patterns, to perceive data features not easily perceived by looking at the vast amount of input data, to relate apparently unrelated data, and supports individual and collaborative decision-making (Dürsteler, 2000).

An example of such visualization is shown in Figure 3. In this representation, a multi-layer technique provides the complete picture at different levels of abstraction, enables interactive analysis of “raw” and “interpreted” data, and presents the current situation within the context of strategic intelligence and near-term trends (Shaw, et al. 2001).

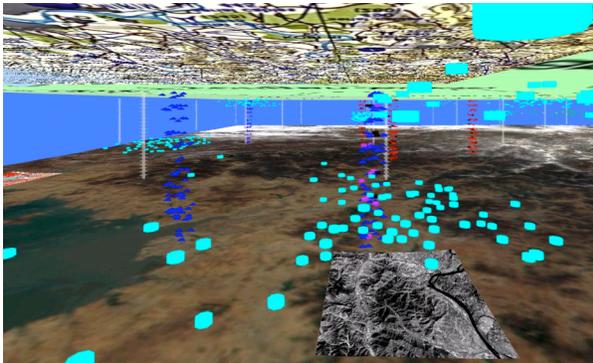


Figure 3: Example of Multi-layer Display to Assist Situation Assessment

Whole-brain Framework Prototype

The framework prototype was designed and implemented using a whole-brain approach to provide an application framework within which geospatial intelligence analysis tools can be integrated. In particular, this approach supports the combination of visually-oriented analysis of images with language-based analysis of text and related information. The framework prototype provides a basis for improved analysis by conservation of analyst attention. It enables the geospatial

intelligence analyst to focus on analysis and interpretation by providing aids for image and text information searching, automatic semantic labeling of images, report generation, and situational visualization. The framework prototype supports the development of a tool set that uses cognitive-based information representation techniques to improve the effectiveness of the integration of intelligence information for increased data understanding. Figure 4 presents the integration concept for this framework.

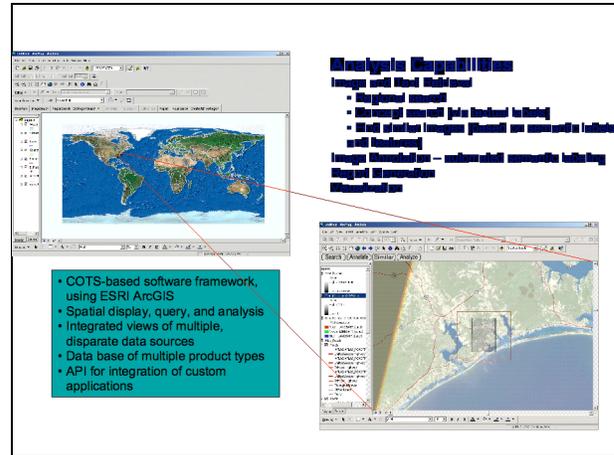


Figure 4: Integration Concept for Geospatial Intelligence Analyst Aids

The whole-brain framework infrastructure was developed using the ESRI (<http://www.esri.com/>) ArcGIS software package as the foundation. A review of the available commercial-off-the-shelf (COTS) geospatial products determined that the ArcGIS package could provide a full-featured basis for the framework prototype. The ArcGIS package furnishes the geographic data management, spatial editing and cartographic visualization functionality as well as a robust application programming interface that can be used to integrate the analysis support tools. The initial set of analysis support tools were integrated in the framework as extensions to the standard functionality accessed through a customized graphical user interface (GUI). The framework can continue to be extended as additional analysis support tools are developed and integrated.

The geospatial analysis aids that have been integrated into the whole-brain framework prototype for this include:

- Region Search – find imagery, pictures, text documents from a specified data base having locations within the selected region
- Image Search – find images based on textual labels
- Report Aids – templates for standardized reporting, free-form reporting, editing of existing reports
- Visualization Definition – select region and data components for automated preparation of 3D visualization.

The framework architecture is extensible for integration of additional analysis tools and for interfaces to standalone tools to provide access to their functionality from within the framework user interface. Transitioning the framework to a useful prototype that can be deployed to an analysis environment requires additional effort in the following areas:

- Assembling the component tools to assist analyst workflow
- Development of a realistic scenario to exercise/demonstrate the tools
- Training of algorithms on selected concepts, test images, etc.
- Experimentation using experienced analysts
- Fine-tuning and adapting the tools based on analyst use.

Summary

An initial prototype system has been developed to demonstrate the whole-brain approach to information analysis. The concept involves developing textual meta-data to support image and signal data. In this way, an analyst can combine his or her visual intelligence along with language capability to improve the overall analysis process. In addition, the generation of meta-data provides an ability to utilize rapid evolution of text-based search capabilities for accessing related textual information from reports and open-source documentation. The prototype developed for this project demonstrates these basic concepts. Future work will seek to augment the tools and concepts described in this paper and evaluate the techniques using large data sets.

Acknowledgements

This work was supported in part by a grant to the Pennsylvania State University via the project, *Innovative Framework for Understanding Integrated Intelligence*, NMA 401-02-9-2001/Task Order No. 0028.

References

- [1] D. L. Hall and J. Llinas, *Handbook of Multi Sensor Data Fusion*, CRC Press, Boca Raton, FL, USA, 2001.
- [2] D. L. Hall, S. A. H. McMullen, *Mathematical Techniques in Multisensor Data Fusion*, Artech House Publishers, 2nd edition, March 1, 2004.
- [3] Llinas, J. and D. Hall, "Dual use of data fusion technology: applying the JDL model to non-DoD applications," Fourth annual IEEE Dual-Use Technology and Applications Conference, SUNY Institute of Technology, Rome, NY, 1994.
- [4] Kuperman, G. G., Cognitive systems engineering for battlespace dominance, in M. D. McNeese and M. Vidulich (editors), "*Cognitive systems engineering in military aviation environments: Avoiding Cogminutia Fragmentosa*" (pp. 230-241), Wright Patterson, Air Force Base, OH, HSIAC Press, 2001.
- [5] McNeese, M. D. and Vidulich, Eds., "*Cognitive systems engineering in military aviation environments: Avoiding Cogminutia Fragmentosa*", Wright Patterson, Air Force Base, OH, CSERIAC Press, 2001.
- [6] Jones, A. Micro-electrical-mechanical Systems (MEMS): A DoD Dual Use Technology Assessment, Washington, DC, Department of Defense, 1995
- [7] Woods, D. D., Designs are hypotheses about how artifacts shape cognition and collaboration. *Ergonomics*, 41,168-173, 1998
- [8] Klein, G. A., Orasanu, J., Calderwood, R., and Zsombok, C. E., editors, *Decision Making in Action: Models and Methods*, Ablex Publishing Corporation, Norwood, NJ, 1995
- [9] McNeese, M. D., "Socio-cognitive factors in the acquisition and transfer of knowledge,"

Cognition, Technology, and Work, 2, 164-177, 2000.

- [10] McNeese, M. D., "Pursuing medical human factors in an information age: The promise of a living lab approach," Proceedings of the Human Factors, Medicine, and Technology Symposium, Baltimore, MD, The University of Maryland Medical Center, 2002.
- [11] Pinker, S., *The Language Instinct: "How the Mind Creates Language"*, Harper Perennial, New York, 2000
- [12] Wang, J.Z., J. Li, and G. Wiederhold, "SIMPLicity: Semantics-sensitive Integrated Matching for Picture Libraries", *IEEE Transactions on PAMI*, 23: p. 1-21. 2000.
- [13] Garga, A. K., "A Hybrid Implicit/Explicit Automated Reasoning Approach for Condition-Based Maintenance", ASNE Smart Ship Symposium II, Philadelphia, PA, November 1996.
- [14] Shaw, T., D. Lang, J. Wise, E. S. Rotthoff, K. L. Slater, M. E. Warren "A Synthetic Multi-Sensory Environment for Time Critical Target Information Analysis", Final Technical Project Report, January 2001.
- [15] McNeese, M. D., An ecological perspective applied to multi-operator systems; in Brown, O. & Hendrick, H. L. (Eds.), *Human factors in organizational design and management—VI* (pp. 365–370). The Netherlands: Elsevier, 1996
- [16] Dowell, J., & Long, J., Conception of the cognitive engineering design problem. *Ergonomics*, 41(2), 126-139, 1998