Wide-Area Motion Imagery (WAMI) Exploitation Tools for Enhanced Situation Awareness

Erik Blasch, Guna Seetharaman
Air Force Research Lab
Rome, NY, 13441
erik.blasch@rl.af.mil
Gunasekaran.Seetharaman@rl.af.mil

Kannappan Palaniappan
Univ. Of Missouri-Columbia
Columbia, MO, 65211
palaniappank@missouri.edu

Haibin Ling
Temple Universtity
Philadelphia, PA 19122
hbling@temple.edu

Genshe Chen
Information Fusion Tech.
Germantown, MD
gchen@intfusiontech.com

Abstract— *The advent of streaming feeds of full-motion video (FMV) and wide-area motion imagery (WAMI) have overloaded an image analyst to detect patterns, movements, and patterns of life. To aid in the process of WAMI exploitation, we explore computer vision and pattern recognition methods to cue the user to salient information. For enhanced exploitation and analysts, there is a need to develop WAMI methods for situation awareness. Computer vision developments in cues, contexts, and communication enhance exploitation capabilities. Multi-data fusion of exploitation context from the video needs to be linked to semantic extraction elements for situation awareness to aid an operator in image analysis. In this paper, we identify (1) opportunities from computer vision techniques to improve WAMI target tracking, (2) relate developments of clustering methods for activity-based intelligence and stochastic context-free grammars for accessing, indexing, and linking relevant information to aid in the processing, and (3) address situation awareness methods of multi-intelligence collaboration for future video techniques. Our example uses the open-source Columbus Large Image Format (CLIF) WAMI data which develops video-based semantic labeling to be connected with other information fusion enterprise capabilities of text-based semantic extraction.*

Keywords: *Wide-Area Motion Imagery, Exploitation, Measures of Effectiveness, Stochistic Context-Free Grammar, Enterprise Fusion*

I. Introduction

Applied Imagery Pattern Recognition seeks to emulate the knowledge and analytical skills of a human towards enhanced perception, motion estimation, and natural language processing. While the ability to completely instantiate these tasks of a human into machines is emerging through computer vision; there are still many functions that require human input. For example, and analyst looking at a full-motion video (FMV) or wide-area motion imagery (WAMI) can see patterns emerging from social, cultural, and behavioral activities, but it is difficult to instantiate these functions by a machine. Thus, we seek methods of enhanced exploitation from video to augment the operator's need for situation awareness. [1]

Fig. 1 demonstrates the functions of applied imagery pattern recognition within an information fusion enterprise [2]. Imagery analysis includes more than exploiting the image; it requires the ability to bring in other sources of sensor and intelligence data, the ability to access, store, and report imagery and the imagery annotated products, as well as collaborate with others users that are processing information (e.g., another imagery data set, text extraction). The use of imagery is only one aspect of coordination as there is a need to (1) use *context* in constrain and optimize the image processing, (2) use of a priori *textual* and verbal requests to coordinate sensor use, and (3) use other sensor modalities to cue sensors. Likewise, the imagery can update all of these issues by outputting terrain information, demonstrating semantic content in the image, and cueing other sensors or imagery tools to exploit the data for further analysis and reporting.

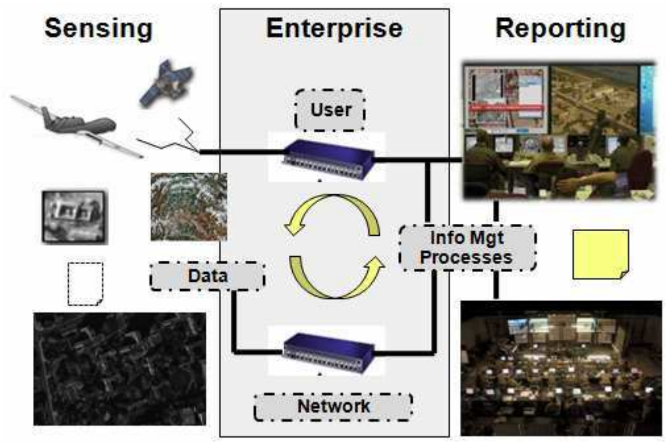

Figure 1 - Information Fusion in the Enterprise.

In this paper, we will highlight current directions of WAMI and FMV that support situational awareness exploitation and analysis. Information fusion of imagery intelligence (IMINT) with non-IMINT intelligence (e.g., human intelligence - HUMINT) requires situation awareness, semantic representation and extraction, and WAMI semantic content outputs to reduce uncertainty [3].

Computer vision has progressed from 2D static image processing, to 3D content, to 4D object dynamic tracking [4]. Recovery of 3D objects from video requires analysis of the sensor motion from dynamic scenes [5, 6, 7], analysis of sensors [8], and processing [9] from which to establish persistent surveillance [10]. These elements of motion processing and the advance of sensor resolutions have resulted in WAMI [11]. Recently, simulations have shown 3D rendering of static scenes [12] to relate to observed imagery.

The rest of the paper is as follows. In Sect. II, we overview situation awareness. Sect. III is recent advances in computer vision to include contextual analysis. Sect. IV is semantic analysis. Sect. V is a challenge problem where Sect. VI provides an example and Sect. VII conclusions.

II. SITUATION AWARENESS

Situation assessment (SA) involves deriving relations among entities, e.g., the aggregation of object states (i.e. classification and location) and Situational Awareness (SAW) is the mental state of a user [1]. While SA/SAW has been recognized in the information fusion and human factors literature, there still exist open questions regarding situation and knowledge representation and theoretical reasoning methods to afford SA/SAW [13]. For instance, while lots of data is collected over a region of interest, how does this information get aggregated and presented to an attention constrained user? Information overload can deteriorate human cognitive reasoning so a pragmatic solution to information representation and semantic constructs are needed for effective and efficient situation understanding.

SAW is a mental state while SA supports (e.g. fusion products) that state. For the computer vision community, there are research developments, but another focus is applications (e.g. military, medical, aviation, security, and environmental). Each might have differences, but the commonality rests in the fact that a multitude of data needs to be efficiently synthesized into a single operating picture (dimensionality reduction) [14], through a user-defined operating picture (UDOP) to assist a user in completing their mission tasks [1].

SAW is an important concept of how people become aware of things happening in their environment. SAW, from a human factor's point of view, determines what effects contribute to a mental picture of action. The *HQ USAF AFISC/SE Safety Investigation Workbook* [15] defines SAW as:

> "keeping track or prioritized significant events and the condition's in one's environment".

One of the premier models for SAW and user decision-making is the extended *Observe-Orient-Decide-Act* (OODA) model which is also termed as Boyd's control loop [16], as shown in Fig. 2. In the "orient" analysis, there is additional information that a user engages in analyzing and synthesizing a situation (e.g. cultural traditions, heritage, and previous experience). Key to the "Orient" phase is the incorporation of new information. New information is combined with previous information to update the situation as well as provide answers to questions posed by the user. The OODA applications for user modeling include information fusion military systems [17], target recognition [18], and recently, semi-automated decision making [19].

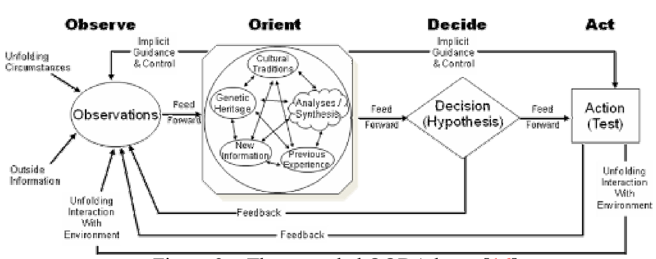

Figure 2 – The extended OODA loop. [16]

However, the classical or extended versions of the OODA loop suffered from a lack of details to sufficiently support the design of systems. Key to the developments and instantiations of the OODA models include: application relevant decision-making based on context, time of analysis, and uncertainty representation. Several developments include the modular M-OODA [20], the Technology, Emotion, Culture, and Knowledge TECK-OODA [21], and the cognitive C-OODA [22]. Current SA/SAW models are natural extensions to the basic tenets, but reflecting the change in technology (e.g. collection and processing observations), the developments of information fusion (e.g. synthesis of observations), and applications to different domains other than military targeting. The Data Fusion Information Group (DFIG) model, shown in Fig. 3, poses different process Levels (L): data characterization (L0), object (L1), situation (L2), and impact (L3) assessment with sensor (L4), user (L5), and mission (L6) (SUM) refinement [23] that extend the OODA loop.

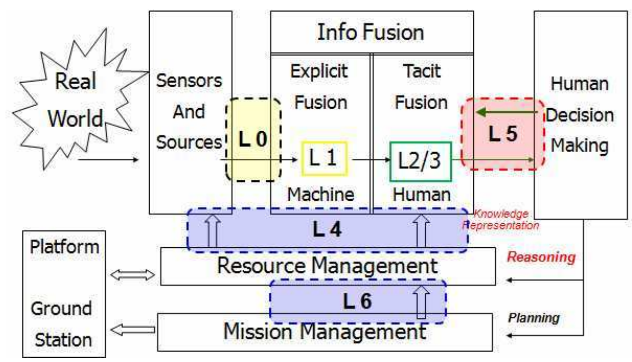

Figure 3 – Data Fusion Information Group Model.

L2 information fusion, SA is the estimation and prediction of relations among entities, to include force structure and force relations, etc. which require adequate user inputs to define entities, outputs from sensors (e..g WAMI) for event detection, and links to other data sources for L3 impact assessment.

The role of SA is to provide an evaluation of the environment and provide information about entities, relations, and objects (e.g. people, assets, or networks for patterns of life). Given a situational picture, there is a need to characterize and display the information to provide awareness. SAW is a mental state about the situation that utilizes the displayed information and affords the user/groups the ability to process, reason, and act on the information. However, computer vision developments for semantic exploitation are needed over context, cues, and communications [24].

III. COMPUTER VISION

Computer vision (CV) supports unmanned aerial vehicles (UAVs) and remotely piloted vehicles (RPVs) that utilize sensor processing of electro-optical sensors (i.e., visual cameras), operate over varying environments, and support user targeting from multimedia sources. Most existing books in CV follow one of three paradigms: biologically inspired CV algorithms [25], signal and image processing [26, 27], and application specific algorithms. There is a much larger system-level consideration that has gone unaddressed for many years. Cameras and modern image manipulation tools have made video a more accessible and ubiquitous medium for communication. The explosive multimedia growth has closed the gap between richness of linguistic communication, once limited to text and print, to video. Therein lays the next set of challenges: How do cultures and 3Cs factor in the way a message is composed in a visual form? Alternatively, given a set of images of forensic nature, how much of can we infer about the context from what is directly seen in the image? Standard physics based and application specific assumptions that were factored in a priori models, which have formally proven useful in vision, will need to be revisited. Those who are practicing in this art will agree that CV has a long way to go, before the already impressive machine vision algorithms are capable of matching parity with human analysts. Foundational ideas needed for multimedia algorithmic frameworks to support system-level processing include 3D analysis, feature extraction, graphical methods, tracking, semantics, and support for user situational awareness, understanding, and assessment), as shown in Table 1.

A. Cues

Cues (*targets*): Image processing for detection, segmentation, classification, and identification require extensions to semantics and SA, as shown in Fig. 4. The cues from images and multimedia content provide a basis for 3D geometry, graphical methods of feature processing, and track analysis. Current methods that build on fundamental approaches include dense data, scene analysis, and semantic content for cause determination and decision making. The extraction of cues over locations supports contextual analysis.

B. Conetxt

Context (*environment*): Three methods for situational awareness contextual processing include: physical (candidate generation), photogrammatic (candidate evaluation), and computational (consistency determination) from which the variable aspects are assessed for enhanced 3C visioning. For example, *context* (Q) is developed from three parameters (variables) of the model set (A), the operator set (O), and the decision policy (I). Direct, recognition-based, and search strategies exist for contextual processing. Instantiating a visual search for context includes: a priori probabilities (P), the operator set (O), and the evaluation metrics (E) of which Q = (A = $\{p\}$, O, I = f(E)). Using background models, syntactic tracking, content-based image retrieval (CBIR), similarity metrics, and moment theory; these concepts support anticipatory autonomy over entities in the environment for sensor and information management through channel analysis.

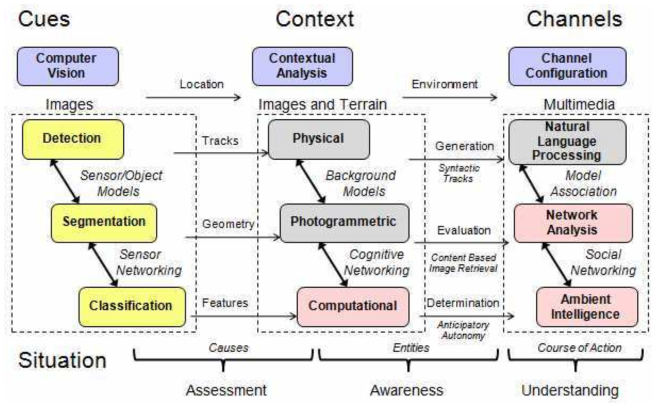

Figure 4 – Concepts introduced in 3C Vision.

C. Channels

Channels (*sensors*): In natural and artificial vision, virtual representations over multimedia analysis support course or action and situational understanding. Methods include (1) use of natural language for explanation, comment, question, and integration and (2) ambient intelligence over embedded, context aware, personalized, adaptive, and anticipatory reasoning. Adding a new channel captures attention

Table 1: Techniques of 3C Vision

	Cues	Contexts	Channels
Vision	Concurrent basic information and relative description	Intended action, a priori knowledge and common culture	Media to convey processed information, including the coding scheme
Analytic Strategy	Capturing current scene (geometry and environment)	Discover relations for scene understanding	Decision making and action planning (inter-human mediation)
3D	Perspective geometry (lighting, local contrast, shape, color, texture)	Image topology (location, foreground/background, pattern or color)	Web-based, mapping, and visualizers, in a web-based virtual reality environment
Features	Feature extraction (region, motion, depth) and decision (computing complexity)	Feature search among: Appearance of scene features, environment considerations, and processing. (Ex: Content based image retrieval)	Linguistic content: document images, textual paragraphs/sentences, and pictorial representations of content
Graphical methods	Bipartite graph matching over feature space	Hypothesis testing using a pattern tree, feature graph, and refinements grammar	Social, cognitive, and sensor networks for information communication extraction and delivery
Tracking	Ego-motion, optical flow, morphological, and spatial tracking	Temporal or circumstantial instances for matching, segmenting, linguistic/syntactic, structuring, and behavior analysis	Information carries (text, sound, images), multimedia which requires content structuring for dynamic applications
Semantics	Tree-scan grammars	Direct (cue, metadata), search, or recognition-based retrieval for linguistic/syntactic applications	Icons, metaphors, and annotations extraction from web content over images/text representative grammars
Situation	Perspective-based analysis for assessment	Knowledge support for awareness based on patterns, retrieval requests, and terrain content.	Interaction through augmented reality, pictorial indexing, and multimedia content for understanding
User	User-directed scene analysis for object and semantic decision making	Attention focusing for perception, action, and cooperative data analysis over locations, activities, and events of interest	Knowledge transfer, social interaction, usability visualizations, annotations to enhance aesthetics, usefulness, and interactivity

(attentative), supports information understanding (explicative), and improves information retention (mnemonic). Use of representation grammar for grouping, detailing, sequencing, comparing, and directing advance traditional image processing methods over social, cognitive, and sensor networks. For grammar, there is a need for semantic exploitation with uncertainty analysis.

IV. Semantic Uncertainty

Semantic ontologies [28] enable a framework for many applications such as command and control, emergency response, and information sharing. Information sharing, and the inherent policies within an architecture, enable data to be fused into actionable knowledge. A key to information fusion is to reduce uncertainty that may come from many sources that require a unified, common, and standardized semantic understanding. Fig. 1 [29], shows the relations between sensed and reported world information from which uncertainty reasoning are required for image processing and user interaction, refinement, and understanding [30, 31, 32, 33].

The evaluation of how uncertainty is processed is dependent on the system-level metrics such as timeliness, accuracy, confidence, throughput, and cost [34], which also are information fusion quality of service (QoS) metrics [35]. Future large complex information fusion systems will require performance evaluation [36] and understanding of the connections between various metrics [37] such as information quality of text-based analysis. It is a goal to formulate, test, and evaluate different methods of a semantic uncertainty ontology that is common, universal, and standardized to link to computer vision techniques. When evaluating both the system's performance as a whole [38] and the specific impact of the uncertainty handling approach, differences arise. For example, when evaluating timeliness (or any other system-level metrics), one will likely find some factors not directly related to the handling of uncertainty itself, such as object tracking and identification report updates (i.e., Level 1 fusion) [39, 40, 41], situation and threat assessment relative to scenario constraints (i.e., Level 2/3 fusion) [42], overall system architectures (e.g. centralized, distributed, etc.), data management processes and feedback / input control processes (i.e., Level 4 fusion considerations) [43], and user-machine coordination based on operating systems (i.e., Level 5 fusion), and others.

Key to the various DFIG [44] levels of information fusion is *evaluation*. For example, there have been efforts in comprehensive tracking [45, 46], object classification [47], and situation awareness evaluation [48] which focus on measures of performance (MOPs). Future evaluations will include high-level information Measures of Effectiveness (MOEs) [49] that include uncertainty characterization [50] and linking to semantic content. One use case is that of WAMI for developments in Level 1 fusion [51, 52, 53]. Other computer vision working groups [54] are exploring semantic technology with datasets that are not necessary focused on uncertainty, but have a rich set of ontologies and datasets for collaboration and comparisons.

As the computer vision community envisions effortless interaction between humans and computers, seamless interoperability and information exchange among applications, and rapid and accurate identification and invocation of appropriate services. As work with semantics and services grows more ambitious, there is increasing appreciation of the need for principled approaches to representing and reasoning under uncertainty. Commonly applied approaches to uncertainty reasoning include probability theory [55], expert systems [56], fuzzy logic, subjective logic [57, 58], Dempster-Shafer theory, DSmT [59], and numerous other techniques. Issues with image processing include:

- *Automated agents* (e.g., to exchange Web information);
- *Uncertainty-laden data*. (e.g., terrain information);
- *Non-sensory collected information* (e.g., human sources);
- *Dynamic composability* (e.g., annotated video); or
- *Information extraction* (e.g., indexing from large databases)

These problems are all related with information fusion, involve both text-based [60] and physics-based [61] data, and can be easily extrapolated to represent the more general classes of problems found in the sensor, data, and information fusion. A recent example of hard-soft fusion uses a controlled natural language (CNL) for data-to-decisions [62].

V. Challenge Problems for Large format Imagery

Video representation and reasoning evaluation framework includes both hard sources (e.g. imaging, radar, video, etc.) and soft sources (e.g., human reports, software alerts, etc.) which requires integration for uncertainty MOEs. Computer vision needs extend to non-saptial data; however, even the spatial data requires new technologies as the size, amount, and speed of data is of the data being collected is increasing. One way to address the problem is to collect imagery data for challenge problem development. The data would enable development of tools needed for data formatting and cross cueing as well as data analysis through algorithm innovation. A general framework for challenge problem testing and evaluation is shown in Fig. 9 that includes the metrics visualization such as receiver operating characteristic curves.

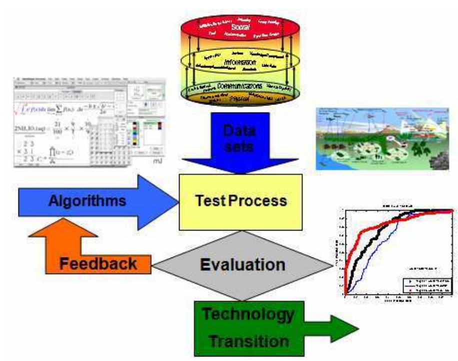

Figure 5 – Evaluation process.

A *challenge problem* includes:

- **Problem Definition**: The scope and significance
- **Data**: Applicable data for the defined problem
  - Tools for reading and processing data
  - Suggestions on training and test sets
  - Characterization of the data
- **Goals**: Research questions and suggested experiments

– **Metrics**: Guidance on reporting results
– **Tools**: Baseline code & results which show reproducible minimum performance standards for the defined problem

Scenarios provide data and support documentation for real world analysis either through analytical, simulated, or empirical results. One example of an open-source WAMI challenge problem is the Columbus Large Image Format (CLIF) collection which includes baseline methods for image registration [63]. Fig. 10 shows the image data set from which results can be compared for infrared, multimodal source, and object tracking and identification [64] over operating conditions solutions [65] event detection.

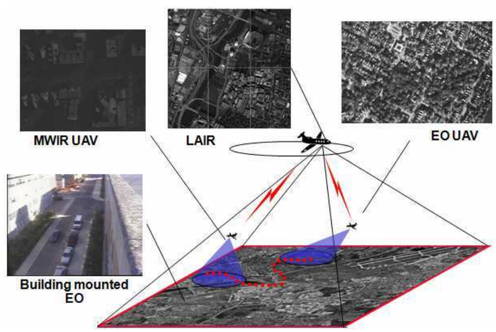

Figure 6 – WAMI data set [X].

Effectiveness relates to a system's capability to produce an effect. Many benefits of fusion include providing locations of events, extending coverage, and reducing ambiguity and false alarms. The goal is to support users in their tasks whether providing refined information, reducing time and workload, or determining completeness, accuracy, and quality in task completion. Effectiveness includes *efficiency*: doing things in the most economical way (good input to output ratio), *efficacy*: getting things done, (i.e., meeting objectives), and *correctness*: doing "right" things, (i.e., setting right thresholds to achieve an overall goal - the *effect*). MOEs support system-level management and design verification, validation, testing, and evaluation. The WAMI output step involves the assessment of how information is presented to the users and, therefore, how it impacts the quality of their decision-making process.

Key aspects of measuring effectiveness come from quality of service (QoS) metrics that can be utilized for hard-soft semantic information fusion [66, 67, 68, 69] Another perspective includes quality of information, or rather information quality (IQ), metrics to combine different types of uncertainty to an established quality. IQ metrics establish user semantic content as a schema or ontology [70] of uncertainty analysis such as a popular method of probabilistic ontologies [71]. Together, these metrics and representations support a formal theory of high-level information fusion [1, 72].

One recent example of algorithm innovation is *context-free grammar* [73] that enables sharing between systems through semantic information [74]. To support geospatial information systems interoperability, the methods of exploitation must switch between pixel-level views (i.e. imagery) versus the graph-level analysis (i.e. social networks) which can be linked between large graphs and images. The distance-level metrics in the graph should quickly be coordinated with the pixel-level information to provide linkage metric.

VI. EXAMPLE – WAMI

Characterizing the semantics in IF processes is not a new research topic. An example is the Semantic Web as part of the Web Ontology Language (OWL) (http://www.w3.org/TR/owl-guide/). OWL operational semantics support message formats (e.g. XML schema) and protocol specifications for an ontology knowledge representation. With a knowledge representation, semantic analysis can be inserted in the message format output from CV semantic analysis.

Semantic analysis is important for activity-based intelligence (ABI) which enables analysts to collect data (e.g., characterize activities), fuse (e.g., locate activities and events), analyze, (e..g., identify and locate actors), and report (e.g. identify and locate networks of these actors and develop patterns of life). Fig. 8 outlines the workflow on an operator doing the OODA "orient" or analyzing the video data. We are interested in image extraction of content, cues and context through ontology text generation. Using the WAMI data, we are interested in behavioral analysis [75], activity recognition [76], and object motion [77].

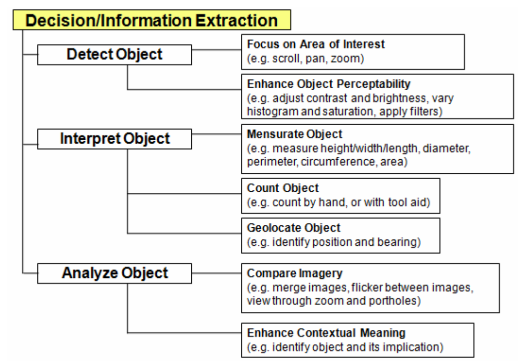

Figure 7 – Analyst Image Processing.

A. Schemas

A *schema* for image processing is shown in Fig. 8 for the Cursor on Target (CoT) program [78]. As detailed, the schema provides target type and identification (ID) allegiance, time stamps, and coordinate locations. While the schema is simple [79], for purposes of information transmission, processing, exploitation, and dissemination, future developments could include semantic fields. It is important to research which semantic content is most relevant for operational information fusion management and systems design.

In order to determine what ontology content can be added to such a message passing schema, there are three issues (1) what, (2) how much, and (3) which ones. For the case of physics-based (video) and textual-based reports, we need to determine what semantic content could be useful. One simple case is that either a human analyst can report a "vehicle" in the uid field, or a machine tracker could extract the information from the video to update the uid field of "vehicle". One example of "vehicle" could be from extracted text and video exploitation of a red vehicle. What is obviously missing from the CoT schema is quality metrics (e.g. confidence, timeliness, and position accuracy) and semantic content (e.g. vehicles).


```
<?xml version='1.0' standalone='yes'?>
<event version="2.0"
    uid="J-01334"
    type="a-h-A-M-F-U-M"
    time="2005-04-05T11:43:38.07Z"
    start="2005-04-05T11:43:38.07Z"
    stale="2005-04-05T11:45:38.07Z" >
  <detail>
  </detail>
  <point lat="30.0090027" lon="-85.9578735" ce="45.3"
            hae="-42.6" le="99.5"  />
</event>
```


Figure 8 – Cursor on Target Schema [79]

B. Wide Area Motion Imagery Tracking

WAMI has gained in popularity as it affords advanced capabilities in persistence, increased track life, and situation awareness, but it also poses new challenges such as tracking evaluation [80] and low frame updates (timeliness) [81, 82]. We utilize the results from a WAMI tracker for track location accuracy and the pixels on target for classification. For example, tracking multiple targets with an on-road analysis from established context, as shown in **Fig. 9**, from the Columbus Large Image Format (CLIF) data set.

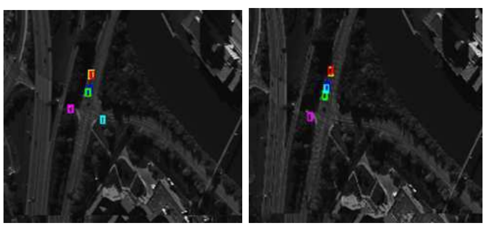

Figure 9 – WAMI Tracking.

Leveraging developments from computer vision [83, 84, 85, 86, 87]. The persistence coverage affords such methods as multiple object and group tracking [88, 89, 90], road assessment and tracking [91, 92], contextual tracking [93], and advances in particle filtering [94]. Because of the numerous objects and their dynamic movements, there are opportunities for linear road tracking, but also there is a need for nonlinear track evaluation [95] such as the randomized unscented transform (RUT) filter [96] for accuracy assessment.

C. Wide Area Motion Imagery Semantic Labeling

As shown in **Fig. 10**, we can detect, cluster information, parse the semantic outputs, and provide probabilities for likelihood analysis based on the assessed semantics.

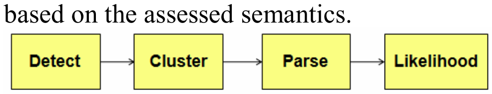

Figure 10 – Analysis of Semantic Analysis

Stochastic Context-free grammar (SCFG) can be expressed as $G=(V, T, S, P, Pr)$:

- V is a finite set of non-terminal variable (symbol) elements;
- T is a finite set of terminals, disjoint from V, which make up the actual content of the sentence;
- $S \in V$ is the start variable (symbol), used to represent the whole sentence;
- P is a finite set of rewrite rules or productions of the grammar.

They are in the form of $X \rightarrow \lambda$, where $X \in V$ and $\lambda \in (V \cup T)$;

- Pr is the set of probabilities for every production rules.

Situation awareness is based on the places, things, and events. Activity-Based Intelligence (ABI) enables analysts to understand what is happening from multi-INT source intelligence. Using SCFG, we develop cognitive semantic representations of the WAMI track outputs using SCFG. Cognitive semantics is the study of meaning applied to words and symbols which can be mapped to image perception. Some of the activity attributes determined from an image could be:

Reference
- **Thing** : actors or entities (vehicle)
- **Place** : locations (event)
- **Time** : reference (day, arrivals)

Correlations (Grammar)
- **Relationship**: a sequence of things (type)
- **Path**: a sequence of places (waypoints)
- **History**: a sequence of time (duration)

Network (Context)
- **Social Network**: a relationship of things (group)
- **Terrain Network**: a relationship of places (road)
- **Time Network**: a relationship of places (order)

Associations (Rules)
- **Action**: behavior or process of things
- **Route**: behavior or process of places (Track)
- **Timeline**: a behavior or process of stamps

Linking (through Probability analysis)
- **Cause**: events that set other things in motion or that constrain motion based on past and present (Markov)
- **Coincidence**: events that happen together
- **Irrelevant** – events that are non-plausible (Non-causal)
- **Causality** : the relationship between cause and effect.

Using the SCFG framework, we can extract semantic content from the image that can be included in the schema, metadata, or reported to an analyst [97]. For example, from **Fig. 11**, we have G = (V, T, S, P, Pr) = (Vehicles, 4, Tracks, Rules, Pr). The things are the vehicle (S), activity context is the rules (P), causal activity is the probability analysis (Pr), and the semantic output (T). For a four word sentence (T) we have $S \rightarrow N\ P\ V\ P$ [1.0], $N\ P \rightarrow Det\ N$ [0.7], $Det \rightarrow red$ [0.4], $N \rightarrow$ car [0.7], $V1 \rightarrow$ start track [0.2], $V2 \rightarrow$ end track [0.8]. The issue here is to link the semantic output from a track before the intersection to after the intersection with the path information. Further exploration and analysis will relate to the semantic analysis for activity assessment.

VII. Conclusions

The paper overviews developments in computer vision for enhanced exploitation to include cues, contexts, and communication for situation awareness. Developing the operational semantics will include issues of representation, reasoning, and policy which need to be considered for command and control [98]. Representing uncertainty has an overall impact on system performance that is hard to quantify or even to assess from a qualitative viewpoint. Operational considerations for WAMI require a common understanding that is achievable by a formal specification of the semantics involved [99, 100].

In the paper, we presented notions of image exploitation for linking to semantic content extraction in relation to schemas and ontologies to support the development for wide-area motion imagery (WAMI) simultaneous tracking and identification. Future work includes group tracking, activity analysis, hard-soft fusion, and contextual understanding and display fusion for enhanced situation awareness.

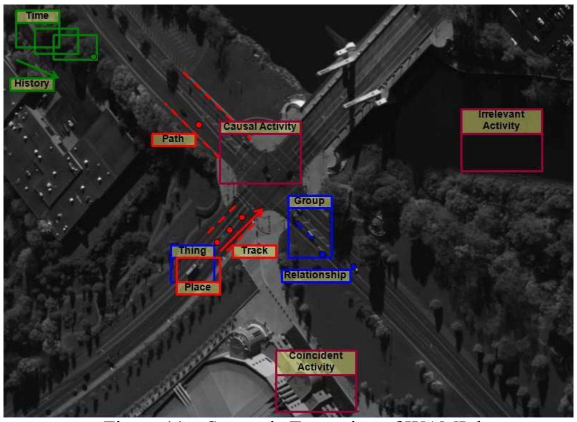

Figure 11 – Semantic Extraction of WAMI data.

REFERENCES

[1] E. P. Blasch, E. Bosse, and D. A. Lambert, *High-Level Information Fusion Management and Systems Design*, Artech House, 2012.

[2] E. Blasch, O. Kessler, J. Morrison, J. F. Tangney *et al*., "Information Fusion Management and Enterprise Processing." *IEEE NAECON*, 2012.

[3] E. Blasch, P. C. G. Costa, K. B. Laskey, H. Ling, and G. Chen, "The URREF Ontology for Semantic Wide Area Motion Imagery Exploitation," *IEEE NAECON*, 2012.

[4] T. Y. Young and G. Seetharaman, "A Regional Approach to Tracking Three Dimensional Motion in an Image Sequence," in, *Advances in Computer Vision and Image Processing,* Vol. 3, (Ed.) Thomas S. Huang, Chapter 3, pp. 63-99, JAI Press, Greenwich, CT, 1987.

[5] G. Seetharaman and T. Y. Young, "A Region Correspondence Approach to the Recovery of 3-Dimensional Motion and Structure in Dynamic Scenes," in *Image Understanding in an Unstructured Environment,* (Ed.) Sushing S. Chen, Chapter 3, pp. 75-124, World Scientific Press, NJ, 1988.

[6] T. Y. Young, G. Seetharaman and Wei-Zao Zhao, "Analysis and Extraction of Three Dimensional Information from An Image Sequence," in *Advances in Artificial Intelligence Research,* Vol. 1, (Ed.) M.B. Fishman, Chapter 8, pp. 209-223, JAI Press, Greenwich, CT, 1989.

[7] G. Seetharaman, "Image Sequence Analysis: Three Dimensional Perception of Dynamic Scenes," in *Handbook of Computer Vision and Image Processing,* Vol. 2, (Ed.) Tzay Y. Young, Chapter 10, pp. 361-403, Academic Press, San Diego, 1994.

[8] G. Seetharaman and H. Le, "Video Assisted GPS for Visual Terrain Navigation using Landmarks," *International Journal of Distributed Sensor Networks,* Vol. 2, No. 2, pp. 103-119, 2006.

[9] T. Rovito, S. Suddarth, J. Layne, K. Priddy, and E. Blasch, "Antenna Aimpoint Integration for Starring-Mode Surveillance (AIMS)," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, July 2008.

[10] E. Blasch, "Emerging Trends in Persistent Surveillance Information Fusion," *IEEE Advanced Video and Signal Based Surveillance*, 2009.

[11] R. Porter, A. M. Fraser, and D. Hush, "Wide Area Motion Imagery: narrowing the semantic Gap," *IEEE Signal Processing Magazine*, Sept 2010, pp. 56-65.

[12] J. M. Ekholrn, K. C. Walli, and J. D. Hendrix, "Optimizing Computer Vision based Scene Reconstruction from Aerial Vehicles" *IEEE Applied Imagery Pattern Recognition*, 2012.

[13] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, S. Das, G. M. Powell, D. D. Corkill, and E. H. Ruspini, "Issues and Challenges in Situation Assessment (Level 2 Fusion)," *J. of Advances in Information Fusion,* Vol. 1, No. 2, pp. 122 - 139, Dec. 2006.

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

[15] *HQ USAF AFISC/SE Safety Investigation Workbook, AFP-1271-Vol. 3*, 1987 (available at: *library.ndmctsgh.edu.tw/milmed/avitation/file-air/AFP-127-1-Vol-3.pdf*)

[16] D. S. Fadok, J. Boyd, and J. Warden, *Air Power's Quest for Strategic Paralysis*, Maxwell Air Force Base AL: Air Univ. Press, 1995.

[17] E. Blasch, "Situation Impact and User Refinement," *Proc. SPIE*, Vol. 5096, 2003.

[18] E. Blasch, "Proactive Decision Fusion for Site Security," *Intl. Conf on Info. Fusion*, 2005.

[19] M. Nilsson, *Capturing semi-automated decision making: the methodology of CASADEMA*. PhD. Thesis, Orebro Univ., SWE, 2010.

[20] R. Rousseau, and R. Breton, "The M-OODA: A Model Incorporating Control Functions And Teamwork In The OODA Loop," *Proc. Command and Control Res, & Tech. Symp*. 2004.

[21] E. Blasch, P. Valin, *et al*., "Implication of Culture: User Roles in Information Fusion for Enhanced Situational Understanding," *Intl. Conf on Info. Fusion*, 2009.

[22] E. Blasch, R. Breton, *et al*., "User Information Fusion Decision Making Analysis with the C-OODA Model," *Int. Conf. on Info Fusion*, 2011.

[23] E. Blasch, "Sensor, User, Mission (SUM) Resource Management and their interaction with Level 2/3 fusion" *Int. Conf. on Info Fusion*, 2006.

[24] V. Cantoni, S. Levialdi, B. Zavidovique, *3C Vision; Cues, Contexts and comunications*, Elsevier, 2011.

[25] V. Cantoni, S. Levialdi, and V. Roberto, *Artificial Vision: Image Description, Recognition and Communication*, Academic Press, 1997.

[26] R.C. Gonzalez, R.E. Woods, and S. L. Eddins, *Digital Image Processing (Second Edition),* Prentice Hall, Upper Saddle River, NJ, 2009.

[27] S. S. Young, R. G. Driggers, and E. L. Jacobs, *Signal Processing and Performance Analysis for Image Systems*, Artech House, 2008.

[28] K.B. Laskey, K.J. Laskey, P.C.G. Costa, M. Kokar, T. Martin, and T. Lukasiewicz (eds.) "Uncertainty Reasoning for the World Wide Web: W3C Incubator Group Report." World Wide Web Consortium. Available at http://www.w3.org/2005/Incubator/urw3/XGR-urw3/.

[29] P. C. G. Costa, K. B. Laskey, E. Blasch and A-L. Jousselme, "Towards Unbiased Evaluation of Uncertainty Reasoning: The URREF Ontology," *Int. Conf. on Info Fusion*, 2012.

[30] E. P. Blasch and P. Hanselman, "Information Fusion for Information Superiority," *IEEE Nat. Aerospace and Electronics Conference,* 2000.

[31] E. Blasch, "Level 5 (User Refinement) issues supporting Information Fusion Management" *Int. Conf. on Info Fusion*, 2006.

[32] E. Blasch, "User refinement in Information Fusion", Chapter 19 in *Handbook of Multisensor Data Fusion 2nd Ed,* (Eds.) D. Hall, and J. Llinas, CRC Press, 2008.

[33] D. L. Hall and J. M. Jordan, *Human Centered Information Fusion*, Artech House, 2010.

[34] E. Blasch, M. Pribilski, B. Roscoe, *et. al.*, "Fusion Metrics for Dynamic Situation Analysis," *Proc SPIE,* Vol. 5429, Aug 2004.

[35] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, S. Das, G. M. Powell, *et al.*, "Issues and Challenges in Situation Assessment (Level 2 Fusion)," *J. of Adv. in Info. Fusion,* Vol. 1, No. 2, pp. 122 - 139, Dec. 2006.

[36] E. Waltz and J. Llinas, "System Modeling and Performance Evaluation," Ch 11 in *Multisensor Data Fusion Systems*, Artech House 1990.

[37] J. Llinas, "Assessing the Performance of Multisensor Fusion Processes," Ch 20 in *Handbook of Multisensor Data Fusion*, (Eds.) D. Hall and J. Llinas, CRC Press, 2001.

[38] P. Hanselman, C. Lawrence, E. Fortunano, B. Tenney, and E. Blasch, "Dynamic Tactical Targeting," *Proc. of SPIE,* Vol. 5441, 2004.

[39] K. C. Chang and Robert M. Fung, "Target Identification with Bayesian Networks in a Multiple Hypothesis Tracking System", *Optical Engineering*, Vol. 36, No. 3, pp.684-691, March, 1997.

[40] E. Blasch, *Derivation of A Belief Filter for High Range Resolution Radar Simultaneous Target Tracking and Identification*, Ph.D. Dissertation, Wright State University, 1999.

[41] C. Yang and E. Blasch, "Pose Angular-Aiding for Maneuvering Target Tracking", *Int. Conf. on Info Fusion*, July 2005.

[42] G. Chen, D. Shen, C. Kwan, J. Cruz, M. Kruger, and E. Blasch, "Game Theoretic Approach to Threat Prediction and Situation Awareness," *J. of Advances in Information Fusion,* Vol. 2, No. 1, 1-14, June 2007.

[43] E. Blasch, I. Kadar, K. Hintz, *et al*., "Resource Management Coordination with Level 2/3 Fusion Issues and Challenges," *IEEE Aero. and Elec. Sys. Mag.,* Vol. 23, No. 3, pp. 32-46, Mar. 2008.

[44] E. Blasch *et al*., "DFIG Level 5 (User Refinement) issues supporting Situational Assessment Reasoning," *Int. Conf. on Info Fusion,* 2005.

[45] K. C. Chang, Z. Tian, S. Mori, and C-Y. Chong, "MAP Track Fusion Performance Evaluation," *Int. Conf. on Information Fusion*, 2002.

[46] E. P. Blasch, O. Straka, J. Duník, and M. Šimandl, "Multitarget Performance Analysis Using the Non-Credibility Index in the Nonlinear Estimation Framework (NEF) Toolbox," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, 2010.

[47] R. Carvalho and KC Chang, "A Performance Evaluation Tool for Multi-Sensor Classification Systems," *Int'l Conf. on Information Fusion*, 2009.

[48] J. Salerno, E. Blasch, *et al*., "Evaluating algorithmic techniques in supporting situation awareness," *Proc. of SPIE,* Vol. 5813, 2005.

[49] E. Blasch, P. Valin, E. Bossé, "Measures of Effectiveness for High-Level Fusion," *Int. Conference on Information Fusion*, 2010.

[50] P.C.G. Costa, R.N. Carvalho, K.B. Laskey, and C.Y.Park, "Evaluating Uncertainty Representation and Reasoning in HLF systems," *Int. Conference on Information Fusion*, 2011.

[51] H. Ling, Y. Wu, E. Blasch, G. Chen, H. Lang, and L. Bai, "Evaluation of visual tracking in extremely low frame rate wide area motion imagery," *Int. Conf. on Information Fusion*, 2011.

[52] R. Pelapur, S. Candemir, F. Bunyak, M. Poostchi, *et. al*, "Persistent Target Tracking Using Likelihood Fusion in Wide-Area and Full Motion Video Sequences," *Int. Conf. on Info. Fusion*, 2012.

[53] P. Liang, *et al*., "Multiple Kernel Learning for Vehicle Detection in Wide Area Motion Imagery," *Int. Conf. on Info Fusion*, 2012.

[54] E. Blasch, H. Ling, Y. Wu, G. Seetharaman, M. Talbert, L. Bai, and G. Chen, "Dismount Tracking and Identification from Electro-Optical Imagery," *Proc. SPIE*, Vol. 8402, 2012.

[55] D. Schum, "The Evidencial Foundations of Probabilistic Reasoning," Northwestern University Press, 1994.

[56] P. Walley, "Measures of Uncertainty In Expert Systems," Artificial Intelligence, 83(1), May 1996, pp. 1-58.

[57] L. M. Kaplan, S. Chakraborty, and C. Bisdikian, "Subjective Logic with Uncertain Partial Observations," *Int. Conf. on Information Fusion*, 2012.

[58] A. J∅sang and R. Hankin, "Interpretation and Fusion of Hyper Opinions in Subjective Logic," *Int. Conf. on Information Fusion*, 2012.

[59] J. Dezert, D. Han, Z-g. L, and J-M. Tacnet, "Hierarchical DSmP transformation for decision-making under uncertainty," *Int. Conf. on Information Fusion*, 2012.

[60] A. Auger and J. Roy, "Expression of Uncertainty in Linguistic Data," *Int. Conference on Information Fusion*, 2008.

[61] H. Ling, L. Bai, E. Blasch, and X. Mei, "Robust infrared vehicle tracking across target pose change using L1 regularization," *Int'l Conf. on Information Fusion*, 2010.

[62] A. Preece, D. Pizzocaro, D. Braines, D. Mott, G. de Mel, and T. Pham, "Integrating hard and soft Information Sources for D2D Using Controlled Natural Language," *Int. Conf. on Information Fusion*, 2012.

[63] O. Mendoza-Schrock, J. A. Patrick, and E. P. Blasch, "Video image registration evaluation for a layered sensing environment." I*EEE Nat. Aerospace Electronics Conf.* (NAECON), 2009.

[64] E. Blasch and B. Kahler, "Multi-resolution EO/IR Tracking and Identification" *Int. Conf. on Info Fusion - Fusion 05*, July 2005.

[65] B. Kahler and E. Blasch, "Sensor Management Fusion Using Operating Conditions," *Proc. IEEE Nat. Aerospace Electronics Conf*, 2008.

[66] W. Perry, D. Signori, and J. Boon, "Exploring the Information Superiority: A Methodology for Measuring the Quality of Information and its Impact on Shared Awareness," *RAND Corporation*, 2004.

[67] M. E. Johnson and K. C. Chang, "Quality of Information for Data Fusion in Net Centric Publish and Subscribe Architectures," *Int. Conf. on Information Fusion*, 2005.

[68] M. A. Hossain, P. K. Atrey, and A. El Saddik, "Modeling Quality of Information in Multi-Sensor Surveillance Systems," *IEEE Int. Conf. on Data Engineering Workshop*, 2007.

[69] C. Bisdikian and L. M. Kaplan, M. B. Srivastava, D. J. Thornley, *et al*., "Building principles for a quality of information specification for sensor fusion," *Int. Conf. on Information Fusion*, 2009.

[70] E. Blasch, "Ontological Issues in Higher Levels of Information Fusion: User Refinement of the Fusion Process," *Int. Conf. on Info Fusion,* 2003.

[71] R. Carvalho, P. Costa, K. Laskey, and KC Chang, "PROGNOS: Predictive Situational Awareness with Probabilistic Ontologies," *Int'l Conf. on Information Fusion*, 2010.

[72] A. Stolcke and J. Segal, "Precise N-gram probabilities from stochastic context-free grammars," *Proc. Comp. Linguistics*, 1994.

[73] J. Gao E. Blasch, and Q. Chen, "Quantitative culturomics using random fractal theory," *IEEE NAECON*, 2012.

[74] G. Levchuk, *et al*, "Adversarial behavior recognition from layered and persistent sensing systems," *Proc. SPIE*, Vol. 7347, 2009.

[75] G. Levchuk, A. Bobick, and E. Jones, E., "Activity and function recognition for moving and static objects in urban environments from wide-area persistent surveillance inputs," *Proc. SPIE*, Vol. 7704, 2010.

[76] Y. Yin, H. Man, J. Wang, and G. Yang, "Human Motion Change Detection By Hierarchical Gaussian Process Dynamical Model With Particle Filter." *IEEE Audio-Vis Sys Surveillance*, 2010.

[77] P.C.G. Costa, KC Chang, *et al*., "High-Level Fusion: Issues in Developing a Formal Theory", *Int. Conf. on Information Fusion*, 2010.

[78] M J. Kristan, J. T. Hamalainen, D. P. Robbins, and P. J. Newer, , "Cursor-on-Target Message Router User's Guide," *MITRE Product – MP090284,* 2009. (accessed on line)

[79] R. A. Shulstad, "Cursor on Target: Inspiring Innovation to Revolutionizing Air Force Command and Control," *Air and Space Power Journal.* Vol. 4, Dec., 2011.

[80] E. Blasch and P. Valin, "Track Purity and Current Assignment Ratio for Target Tracking and Identification Evaluation," *Int. Conf. on Info Fusion*, 2011.

[81] K. Palaniappan, F. Bunyak, P. Kumar, *et al.,* "Efficient feature extraction and likelihood fusion for vehicle tracking in low frame rate airborne video," *Int. Conf. on Information Fusion*, 2010.

[82] K. Palaniappan, R. Rao, G. Seetharaman, "Wide-area persistent airborne video: Architecture and challenges," *Distributed Video Sensor Networks: Research Challenges and Future Directions*, Springer, Part V, Chapter 24, pp. 349 – 371, 2011.

[83] Y. Wu, H. Ling, E. Blasch, G. Chen, and L. Bai, "Visual Tracking based on Log-Euclidean Riemannian Sparse Representation," *Int. Symp. on Adv. in Visual Computing - Lecture Notes in Computer Science*, 2011.

[84] X. Mei, H. Ling, Y. Wu, E. P. Blasch, and L. Bai, "Minimum Error Bounded Efficient L1 Tracker with Occlusion Detection," *IEEE Computer Vision and Pattern Recognition*, 2011.

[85] Y. Wu, E. Blasch, G. Chen, L. Bai, and H. Ling, "Multiple Source Data Fusion via Sparse Representation for Robust Visual Tracking," *Int. Conf. on Info Fusion*, 2011.

[86] Y. Wu, J. Wang, J. Cheng, H. Lu, E. Blasch, L. Bai, and H. Ling, "Real-Time Probabilistic Covariance Tracking with Efficient Model Update," *IEEE Trans. on Image Processing*, 21(5):2824-2837, 2012.

[87] X. Zhang, W. Li, W. Hu, H. Ling, *et al*., "Block covariance based L1 tracker with a subtle template dictionary," *Pattern Recognition*, 2012.

[88] E. P. Blasch and T. Connare, "Improving Track maintenance Through Group Tracking," *Proc of the Workshop on Estimation, Tracking, and Fusion; A Tribute to Yaakov Bar Shalom*, 360 –371, May 2001.

[89] T. Connare, E. Blasch, *et al*., "Group IMM tracking utilizing Track and Identification Fusion," *Proc. of the Workshop on Estimation, Tracking, and Fusion; A Tribute to Yaakov Bar Shalom,* 205 -220, May 2001.

[90] M. Wieneke and W. Koch, "A PMHT Approach for extended Objects and Object Groups," *IEEE T. Aero. and Elec. Syst.*, vol. 48, no. 3, 2012.

[91] C. Yang and E. Blasch, "Fusion of Tracks with Road Constraints," *J. of. Advances in Information Fusion,* Vol. 3, No. 1, 14-32, June 2008.

[92] X. Shi, H. Ling, E. Blasch, and W. Hu, "Context-Driven Moving Vehicle Detection in Wide Area Motion Imagery," *Int'l Conf on Pattern Recognition (ICPR),* 2012.

[93] E. D. Marti, J. Garcia, and J. L Crassidis, "Improving Multiple-Model Context-Aided Tracking through an Autocorrelation Approach," *Int. Conf. on Info Fusion*, 2012.

[94] F. Papi. M. Podt, Y. Boers, *et al.,* "Constraints Exploitation for Particle Filtering based Tagret Tracking," *Int. Conf. on Info Fusion*, 2012.

[95] Z. Duan, X. Rong Li, *et al.,* "Design and Analysis of Linear Equality Constrained Dynamic Systems," *Int. Conf. on Info. Fusion*, 2012.

[96] O. Straka, J. Duník, and M. Šimandl, "Randomized Unscented Kalman Filter in Target Tracking," *Int. Conf. on Information Fusion*, 2012.

[97] Z. Yao, X. Yang, L. Lin, M. W. Lee, et al., "I2T: Image Parsing to Text Description," *Proc. of IEEE*, Vol. 98, no.8, pp 1485-1508, 2010.

[98] L. Scholz, D. Lambert, D. Gossink, *et al*., "A Blueprint for Command and Control: Automation and Interface," *Int. Conf. on Info Fusion*, 2012.

[99] E. Blasch, P. C. G. Costa, K. B. Laskey, D. Stampouli, G. W. Ng, *et al*., "Issues of Uncertainty Analysis in High-Level Information Fusion – Fusion2012 Panel Discussion," *Int. Conf. on Info Fusion*, 2012.
[100] P. C. G. Costa, *et al*., "Uncertainty Evaluation: Current Status and Major Challenges –Panel Discussion," *Int. Conf. on Info Fusion*, 2012.