This document is published in:


© 2013. Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.
Overview of Contextual Tracking approaches in Information Fusion

Erik Blasch a, Jesus Garcia Herrero b, Lauro Snidaro c, James Llinas d,
Guna Seetharaman a, Kannappan Palaniappan e

a Air Force Research Laboratory, Information Directorate, Rome, NY, 13441
b University of Madrid, Spain
c University of Udine, Udine, Italy
d University of Buffalo, Buffalo, NY, 14214
e Univ. Of Missouri-Columbia, Columbia, MO, 65211

ABSTRACT

Many information fusion solutions work well in the intended scenarios; but the applications, supporting data, and capabilities change over varying contexts. One example is weather data for electro-optical target trackers of which standards have evolved over decades. The operating conditions of: technology changes, sensor/target variations, and the contextual environment can inhibit performance if not included in the initial systems design. In this paper, we seek to define and categorize different types of contextual information. We describe five contextual information categories that support target tracking: (1) domain knowledge from a user to aid the information fusion process through selection, cueing, and analysis, (2) environment-to-hardware processing for sensor management, (3) known distribution of entities for situation/threat assessment, (4) historical traffic behavior for situation awareness patterns of life (POL), and (5) road information for target tracking and identification. Appropriate characterization and representation of contextual information is needed for future high-level information fusion systems design to take advantage of the large data content available for a priori knowledge target tracking algorithm construction, implementation, and application.

Keywords: Contextual Tracking, Sensor management, Patterns of Life, Traffic Behavior, Road Information, Situation Awareness, Group Tracking, Behavior Analysis, Activity-based Intelligence, Information Fusion, WAMI

1. INTRODUCTION

Target tracking has matured to include non-Gaussian nonlinear tracking methods to detect numerous targets over varying terrain. However, there are needed external sources for robust target tracking solutions such as map updates, roads and road conflation, detection of patterns of life, trafficability maps, and other products that relate to context information. Fig. 1 highlights some issues where (1) cues relate to standard situation assessment of target tracking and classification, (2) context supports situation awareness over physical, computational, and environment issues, and (3) channels for situation understanding over ambient intelligence, language, and networks.

![Fig. 1. Contextual Awareness and Understanding [1]](image-url)
Target tracking is a subset of information fusion which supports many applications [2]. One commonly accepted information fusion model is the Data Fusion Information Group (DFIG) model [3] (shown in Fig. 2) originally developed for military systems, but used by many in the International Society of Information Fusion (www.isif.org) as a common processing framework. The levels (L) determine the processing in the system such as L0 data registration, L1 object tracking and ATR assessment [4] (shown in Fig. 3), L2 situation awareness [5, 6] and L3 impact assessment [7]. The complementary control levels are: L4 sensor management, L5 user refinement [8], and L6 mission management. Fig. 3 highlights that target tracking composes many aspects of which all the models (e.g., sensor, target, environment, behavior, and performance) inherently provide context to improve target tracking performance.

Fig. 2. Data Fusion Information Group Model (L = Level)  Fig. 3. Context Modeling to support target tracking.

Context has well been explored in the last decade for target tracking and information fusion issues especially for group tracking [9]. Using a group tracking example [10], as adapted from Steinberg [11], context supports all levels of information fusion and management, shown in Fig. 4. For example, context has been reported for user refinement (L5) [12], situation assessment (L2) [13], resource management (L4) [14], and threat/scenario assessment (L3) [15, 16] of targets. Current needs of information fusion and target tracking models include information exploitation and management to take advantage of all the contextual information [17]. Context can therefore be exploited as binding element for the synergic interaction of techniques residing at different DFIG levels.

Context is a widely used term by philosophers and scientists with many different definitions and connotations [18, 19]. Nevertheless, there are very general definitions where context is a subset of a physical or conceptual state, which is related to a specific entity, as in Dey and Abowd [20] where an application-driven definition is reported. Context can be framed as "any information that can be used to characterize the situation of entities that are considered relevant to the interaction
between a user and an application”, or more general, “between the operator and the system” [21]. It is generally an addition to sensing devices data, and “surrounds” a situation of interest, aiding comprehension and interaction [22].

2. CONTEXTUAL TRACKING METHODS (REVIEW)

Context in support of tracking has well been explored to support or enhance target state estimation and display the results of simultaneous tracking and classification/identification to aid user contextual analysis [23]. Of the many examples explored over the years, context can be based on three aspects: sensor information, target capabilities, and environmental constraints [24]. It is the environmental constraints that affords context-enhanced target tracking (DFIG Level 1), wherein improved performance is reported. While the advent of big data machine analytics (Section 4), due to high-performance computers and cloud computing, is new, one of the first context tracking papers uses video results to index context [25].

For target tracking, other mediums such as radar-based group tracking sought methods combined the use of group target associations as well as these groups of targets constrained on roads [26, 27]. In this sense, context was used in the form of both the targets themselves as well as the environment. Then, researchers applied methods for convoy tracking [28], improvement of track association [29], use of speed for kinematic target selection [30], and map information to improve target location accuracy [31]. Other examples included feature selection based on context [32] and learning spatial relations for moving target assessment [33].

As contextual information from roads added to tracking algorithms, new nonlinear tracking methods were demonstrated to show improvement not only to capture non-linear motion estimation, but also how contextual information from roads support the motion estimation [34]. In specific domains such as airport traffic monitoring, the knowledge of taxiway layout and motion rules at surface allowed significant improvement in tracking results [35]. Movement of vehicles expanded to include developments in people tracking [36], and methods to improve driver’s ability to stay on roads [37, 38]. Similarly, methods of mathematical performance analysis were developed for maneuvering target pose [39], maneuvering target state estimation [40] and tracking robustness for spatial-temporal context [41].

In a related approach, improvements in audio tracking utilized context. For audio tracks, methods were developed for phonemics [42] and use of semantic context for modeling audio tracking [43]. Other methods included context to constrain Hidden Markov Models [44] and beat tracking [45]. These methods could be useful for determining the tracks associated with targets on roads with GPS capabilities and cell phones [46].

In 2007, there were multiple applications of using context to support video tracking. Sanchez et al. [47], demonstrated tracking for video by using both a general tracking layer and a context layer. Specifically, the context layer supports track initiation, maintenance, and updates of the tracks. Other methods included appearance context to re-acquire lost targets [48] which can be used to restart lost tracks or initiate new tracks. Other prominent methods included pedestrian systems [49] and methods to improve target search for video-based simultaneous tracking and classification [50] (as an extension by Nguyen to [41]).

Further refinements of the use of contextual information supported the estimation algorithms for sensor management (DFIG, Level 4 Fusion) such as matrix refinement [51] and sensor models [52]. Contextual estimation constraints included road networks [53] and track states [54]. Likewise, “context-aware visual tracking,” [55] was coined to demonstrate that tracking multiple objects in the scene provides context for the designated target. New developments also supported the use of tracking to update other contexts such as ontologies for situation assessment (DFIG Level 2) video tracking [56] (as an extension by Sanchez et al. to [47]), surveillance [57] and threat assessment (DFIG Level 3) [58]. Finally, using context supports route planning [59], scene context [60], and multimodal fusion [61]. As an example, context tracking issues were summarized in a Bayesian Tracking tutorial by Koch [62].

The representation of contextual domain knowledge with ontologies allowed the integration of data from available hard/soft sources (surveillance sensors, human reports, databases, etc.) in the context of a situation [63]. Gomez-Romero, et al., [64] utilized ontology-based models to specify concepts and relationships of contextual knowledge by separation of context reasoning and feedback to support track management. Contextual information is represented with ontologies, which are used to model the heterogeneous entities present in the domain (the abstract scene model) and the normal situations (the normale model). The normale model contains axioms and rules to classify vessel behavior as compliant to the operational rules or not in a maritime domain awareness scenario. The harbor domain brings challenges
to advanced fusion systems [65]. Also, the representation of context with ontologies has been used to classify harbor objects and basic situations by deductive reasoning according to the harbor regulations (e.g. navigation rules) [66].

In 2011, there were many tracking solutions using tracking as method of dealing with environment context for such aspects of occlusions [67, 68] and feature adaptation for target signature variation [69, 70]. A novel method by Rice and Vasquez [71] demonstrated the use of context for hyperspectral sensor-based target tracking. Other developments included methods for effective computing [72], hybridized methods for group tracking [73], multimodal fusion from electro-optical and infrared sensing [74], anomaly detection from known path of travels [75], and track-segment association [76]. Finally, another application versus ground target tracking, was for underwater tracking [77, 78].

Contextual tracking, aided by situational awareness, has been explored through logical methods. Visentini and Snidaro [21] explored context of natural language expression of physical entities in environments to reduce the ambiguity and uncertainty of measurements through likelihood maps to constrain the location estimate of a target in a building. They followed up [79] by introducing domain expert knowledge as context through Markov Logical Networks (MLN) which are a form of Statistical Relational Learning (SRL). MLNs combine first-order logic and graphical models (e.g. Markov). First-order logic, in contrast to propositional logic, represents complex environments in a concise way. Contextual information was developed for a maritime domain tracking example by using the “isA” formulation over locations, tracks, situations, and threats. An operator supplied given evidence which provided contextual information to refine the sensor observed evidence (DFIG Level 5). As related to ontological methods [56] for insertion of contextual information into tracking systems, logical methods support situation awareness for high-level information fusion.

Continuations of ground target tracking methods included methods for mobile tracking [80], context-aided tracking [81], and occlusions [82]. Other methods included learning target labeling [83], vehicle detection [84], and background context [85]. Currently methods include using context for sensor management and placement [86, 87], anomaly detection from stochastic free grammars [88], occlusion detection [89], and finally towards the growth of machine analytics in dictionary learning to support target identification context [90].

One final area of discussion of context is in moving target classification assessment. This has been coined as “context enhancement” [91] as improvements on tracking and classification/recognition/identification [92, 93]. Context awareness including: environments, sensors, and targets, improved multi-source robustness [94] and clutter suppression [95]. Currently, context-enhancement is being mapped to common qualitative and quantitative results for user assessment (DFIG Level 5) and refinement of context moving and stationary target data (DFIG Level 1) [96]. Next we present an overview of the categories of contextual analysis for target tracking based on the literature review.

### 3. CONTEXTUAL TRACKING METHODS

From the review of the many methods in contextual tracking, we sought to organize the previous methods as well as look to the future of future needs. Our preliminary categorization includes:

1. Road information for target tracking and identification, group detection, and context-aided tracking;
2. Environment-to-hardware processing for sensor modeling/management, and context enhancement;
3. Historical traffic behavior for situation awareness patterns of life (POL) for context awareness;
4. Known distribution of entities for situation/threat/scenario assessment for context inference; and
5. Domain knowledge from a user to aid tracking through selection and analysis for context cognition.

Categories (1) and (2) have well been vetted in the literature. Category (3) relates to the use of tracking as well as the constraints available to provide context awareness. For example, in video tracking, many examples track the scene to infer behavior. Over the many issues for contextual target tracking, we looked at the variations in the themes and while there is discussion on context awareness (as situation awareness using ontologies and logical networks), there was limited analysis of connections to threat assessment [97] for context inference. Ideas exist for game-theoretic modeling of multiple affiliation entities being tracked [98]; however, there is a need for context-based human, social, cultural, and behavior (HSCB) modeling and assessment [99]. For example, HSCB can be used can be used with road information to isolate which pedestrians, how fast cars are moving on roads, and clutter mitigation that does not conform to social, cultural and behavioral norms which leads to human, animal, vehicle and clutter (HVAC) target categorization. As detailed in [22], Garcia, Snidaro, and Visentini advocate the need for context cognition by the user (DFIG L5 fusion).
4. MACHINE ANALYTICS FOR CONTEXTUAL TRACKING

With the enormous amount of data types, distributed locations, and various connections to different applications (e.g., finance to surveillance) resulting from the expansion of the World-Wide Web, new techniques are needed to exploit context. Related concepts recently emerging are context awareness (Fig. 1) and machine, descriptive, prescriptive, predictive, visual, and other analytics, shown in Fig. 5. There are three issues of importance: hardware (e.g., Apache Hadoop data intensive distributed architecture), software (e.g., machine analytics), and user/domain applications (e.g., visual analytics, text analytics). Data, process, and visual analytics pave the wave for big-data processing to utilize more contextual information in target tracking [17].

![Machine Analytics Diagram]

Fig. 5. Big Data Analysis

5. CONTEXTUAL TRACKING EXAMPLE

Using a multimodal cooperative sensing example, we are interested in simultaneous target tracking and identification (STID) [101]. Multi-modal measurements could be from infrared, visual, and/or wide-area motion imagery (WAMI). Together, contextual environmental modeling of the weather for the aerial sensors and the terrain information for the ground sensors would aid in the analysis of the complete system for accurate automatic target recognition (ATR) (i.e., high probability of detection with low false alarms [102]). One recent complex challenge that requires contextual information is WAMI target tracking [103] with information management that supports moving intelligence [104]. Complications of real-time WAMI sensor processing include a low frame rates [105] and mappings to geospatial intelligence systems [106] such as environmental (e.g., terrain modeling). Together, machine analytics and contextual tracking support enhanced situation awareness [107] for cooperative control of multimodal sensors as depicted in Fig. 6.

![Fig. 6. Wide Area Motion Imagery (WAMI) data.]

Fig. 6. Wide Area Motion Imagery (WAMI) data.
Accurate contextual statistical modeling is needed of the environment, sensor, and target data (i.e., operating conditions) to support the mathematical algorithms as shown in Table 1. Note that identification includes composing ATR and kinematic data for threat assessment of friend, foe, and neutral (FFN) affiliation [108].

Table 1. Contextual Analysis over the levels of information fusion

<table>
<thead>
<tr>
<th>Info Fusion</th>
<th>Measurement</th>
<th>Model</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0 – Data Registration</td>
<td>Pixels</td>
<td>Terrain</td>
<td>Road-assisted</td>
</tr>
<tr>
<td>Level 1 – Object Assessment</td>
<td>Kinematic/Features</td>
<td>Kinematic/Target</td>
<td>Context Aware (Assisted) Tracking</td>
</tr>
<tr>
<td>Level 2 – Situation Assessment</td>
<td>Object Groups</td>
<td>Behavioral</td>
<td>Group Tracking</td>
</tr>
<tr>
<td>Level 3 – Impact Assessment</td>
<td>Threat (FFN)</td>
<td>Intent/Allegiance</td>
<td>Anomaly detection</td>
</tr>
<tr>
<td>Level 4 – Sensor Management</td>
<td>Sensor Type</td>
<td>Camera</td>
<td>Sensor, appearance, models</td>
</tr>
<tr>
<td>Level 5 – User Refinement</td>
<td>POL</td>
<td>Cognitive</td>
<td>Activity, Behavioral Analysis</td>
</tr>
<tr>
<td>Level 6 – Mission Management</td>
<td>Objectives</td>
<td>Goal-driven</td>
<td>Social, Cultural modeling</td>
</tr>
</tbody>
</table>

Managed layers of sensors offer capabilities to robustly track targets over various operating conditions of differing targets, sensors, and environments. Numerous advances in algorithms, database methods, and sensors offer opportunities for future capabilities. Inherent in the analysis are three techniques: (1) feature extraction, processing, and tracking for targeting [109], (2) common data sets for analysis and algorithm comparison over environmental conditions, and (3) persistent wide-area motion imagery for long-term consistent sensing.

2.1 Feature Tracking and Identification (Targets)

For tracking targets, various features are important to determine the automatic target recognition/classification/identification, behavior, and location [110]. These features (such as group dynamics) would aid in tracking through occlusions, illumination changes, and links to common data bases [111, 112]. The behavioral features are those related to the global movement, relation of vehicles to the center motion for local movements, and common attributes that aid in affiliation/association of members in group movement. Features can also be elements of graphical methods, Markov Logic Networks, ontologies, and situation attributes that can be continuous or discrete to constrain target state estimates (e.g., recognition/classification/identification, behavior, and location).

2.2 Wide Area Motion Imagery (Sensors)

Wide Area Motion Imagery (WAMI) is an emerging capability that affords large spatial coverage, constant monitoring over time, and potential for diverse frequency sensing (e.g. EO/Radar) for tracking and identification wide area surveillance [113]. Since the WAMI data covers a city (see Fig. 7), the ability to maintain track (after initiation) is increased as the objects are within the sensed region of interest for potentially an extended duration [114], activity analysis [115, 116], and occlusion detection [117]. Likewise, with constant staring, there is the increased advantage of extending track lifetime by linking associated tracks, projecting tracks onto road networks, recovering from lost tracks, and coordinating hand-off to ground sensors. Finally, with the advent of WAMI, there are other modalities emerging for electro-optical visual cameras, moving synthetic aperture radar (SAR), and hyperspectral (HSI) methods. Together, these sensors provide a rich new set of contextual data that needs to be exploited using novel fusion (e.g. ATR) methods over small pixel resolutions in addition to the traditional vehicle tracking.

2.3 Situations and Scenarios (Environment)

WAMI data provides new opportunities that relate to targets and environments [118], increasingly so when combined with other sensors such as ground-based detectors. WAMI data sets cover a broad range of environmental conditions and various target behaviors. Using contextual information for target tracking algorithm development, the basic techniques such as tracking and behavioral semantic labels can be applied over a larger spatial distances and temporal intervals. As an example, using in the Columbus Large Image Format (CLIF) data set, identified contextual conditions include sensor system performance (camera motion and frame rate, contrast, and camera model fidelity), targets (turning, type, and speed), and environments (shade, occlusion, on and off roads), as shown in Fig. 7. The environment can be used to constrain target motion (e.g., roads), but also the tracking measurements can be used as history of normalcy and abnormal behavior. For example, the normalcy models, shown in Fig. 7, of tracking behavior details the patterns of life (POL) of normal activities of the vehicles on roads including direction, speed, and numbers of vehicles.
6. DISCUSSION

In Table 2, we list the issues associated with contextual information over the sensors, targets, and the environment.

**Table 2: Issues for Contextual Tracking**

<table>
<thead>
<tr>
<th>Concept</th>
<th>Advantage</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAMI</td>
<td>Wide area scene understanding</td>
<td>Increased computations from environment (roads) for context-aided tracking with limited bandwidth</td>
</tr>
<tr>
<td>Multimodal</td>
<td>Can combine HSI, radar, and EO</td>
<td>Inconsistent geo-registration of multiple mixed resolution sensors hinders real-time context analysis</td>
</tr>
<tr>
<td>Ground sensors</td>
<td>Ability to get high-resolution pixels for classification/ recognition/identification</td>
<td>Accuracy of reported and project information of change appearance for context enhancement</td>
</tr>
<tr>
<td><strong>Targets</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pedestrians/ Vehicles</td>
<td>Can track through various dynamic changes (e.g. on-off roads)</td>
<td>Prediction of movement through occlusions requires predicted (versus estimated) context awareness</td>
</tr>
<tr>
<td>Tracking</td>
<td>Increased track lifetime from extended spatial coverage</td>
<td>Increased number of confuser objects that require rudimentary context scene analysis</td>
</tr>
<tr>
<td>Group Association</td>
<td>Maintain database of associations of common movements, affiliations, and reduced state estimates</td>
<td>Separating activities of interest for group association versus that of routine and independent activities for verifiable context awareness</td>
</tr>
<tr>
<td>Patterns of life</td>
<td>Can determine normalcy modeling of behaviors and activities</td>
<td>Determining the unknown actions resulting from sparse activities for context inference/cognition</td>
</tr>
<tr>
<td>Intent</td>
<td>Can link person to a priori known places of activity to help in tracking, and/or build up notion of behavioral intent</td>
<td>Determining the social/cultural norms of various groups that have yet to be identified or actions which are routine but require context cognition</td>
</tr>
<tr>
<td><strong>Environment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roads</td>
<td>Context aiding through realizable vehicle/pedestrian paths of travel</td>
<td>Requires machine analytics to maintain known road networks and travel for context awareness</td>
</tr>
<tr>
<td>Weather</td>
<td>With different modalities, have the opportunity for distance and weather invariant observations.</td>
<td>Some sensors need to detect the variations in features due to changes in weather (e.g. illumination) requiring context enhancement</td>
</tr>
<tr>
<td>Terrain</td>
<td>Can observe through varying conditions (e.g. occlusions, obscuration)</td>
<td>Need to detect change in conditions linking HCSB with routes for context awareness</td>
</tr>
<tr>
<td><strong>Use</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysts</td>
<td>Provide context information to operators</td>
<td>Cueing of information to many users who all have different roles and functions of context cognition</td>
</tr>
<tr>
<td>Social Networks</td>
<td>Link to available machines analytics databases for social networks</td>
<td>Requires context based indexing for efficient retrieval to support context cognition/awareness</td>
</tr>
<tr>
<td>Cultural Networks</td>
<td>Determine activities of activities (e.g. credit card transactions, police records, vehicle registration records, etc.)</td>
<td>Delay of information to support context tracking needs, including time latency and low confidence correlations impeding context cognition</td>
</tr>
<tr>
<td>Sensor Management</td>
<td>Increased correlation of features for tracking and performance models</td>
<td>Real-time model updates for changing targets; high demand for limited assets requiring context assessment</td>
</tr>
</tbody>
</table>
7. CONCLUSIONS

This paper overviewed the many discussions of context in the domain of target tracking. Through explorations of the subject and multiple discussions, we presented various themes on the subject and provided an example for WAMI tracking where context aids wide-area surveillance, group tracking, maritime-domain awareness, and patterns-of-life estimation. Our current categories include:

1. **Context-aided tracking**: road information for target tracking and identification, group detection, and targeting;
2. **Context enhancement**: environment-to-hardware processing for sensor modeling/management;
3. **Context awareness**: historical traffic behavior for situation awareness patterns of life (POL) and scene analysis;
4. **Context inference**: known distribution of entities for situation/threat/scenario assessment; and
5. **Context cognition**: domain knowledge from a user to aid tracking through selection and analysis of objectives.

To further target tracking techniques, we envision that machine analytics and human social, cultural, and behavioral modeling will be incorporated into future context aided-tracking, enhancement, awareness, inferencing, and cognition.

REFERENCES


