A Conceptual Framework for Automatic Situation Assessment

Y. Fischer, A. Bauer, and J. Beyerer

Abstract—**Humans make decisions on the basis of their situation awareness and it is well-known that insufficient situation awareness leads to incorrect decisions. The challenge of an advanced surveillance system for supporting situation awareness of a human decision maker is therefore to detect and assess complex situations that evolve over time. In this article, we present a conceptual framework for automatic situation assessment that consists of four parts, namely the situation characterization, the situation abstraction, the situation recognition and the situation projection. The situation itself can be described at several different levels of abstraction. The proposed framework can be used as a guideline when designing automatic situation assessment processes.**

Index Terms—**High-level data fusion, situation assessment, situation awareness, situational modeling, surveillance system**

I. Introduction

During the operation of complex systems that include human decision making, acquiring and interpreting information from the environment forms the basis for the state of knowledge of a decision maker. This state is often referred to as *situation awareness*. The most commonly used definition of situation awareness was provided by Endsley in [1]:

> "Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future."

Due to this definition, situation awareness consists of three levels, namely perception, comprehension, and projection, as depicted in Figure 1. The first level of situation awareness includes the detection of relevant elements and its characteristics in the environment. The elements are of course domain specific and their status, their attributes and their dynamics have to be observable by sensorial means. The second level of situation awareness is based on first level elements and includes the understanding of the significance of these elements in relation to the operator's goals. The third and highest level of situation awareness is again based on the lower levels 1 and 2 and deals with the ability to project future actions of elements in the environment.

Thus, a high level of situation awareness consists of much more than simply collecting information about elements in the environment. It is furthermore a result from the comprehension of its meaning and the projection of future states in order to make decisions on the most favorable actions. Situation awareness is therefore referred to as a mental state or a state of knowledge, whereas the processes to achieve and maintain that state are referred to as *situation assessment*. As a high level of situation awareness provides the complete knowledge which is necessary for effective decision making, the decision process itself and the performance of actions are separate stages of the dynamic decision making process as illustrated in Figure 1.

Endsley described several factors that have a major influence on the decision making process. First, individual factors influence the situation assessment process, for example the operator's abilities, experience, and training. But individuals do not only vary in their information processing mechanisms but also in their expectations and objectives. Other influencing factors can be summarized as system factors which include the system capabilities or the interface design, and also some features of the task environment like workload, stress or complexity.

The concept of situation awareness established by Endsley is applicable in many different domains and it can also be used for advanced surveillance systems. Especially in security-related tasks, like the surveillance of specific areas, decision makers should always have a high level of situation awareness. Situations of interest that take place in surveyed areas are often of a high complexity and dynamic, because they consist of multiple different objects that interact with each other and their activities evolve over time. In such a complex and dynamic environment, the limited capacity of a person's attention is quickly exhausted. The focus of attention is therefore a major limit on situation awareness.

In today's surveillance system, level 1 situation awareness is highly supported through the various heterogeneous sensors

Manuscript received October 22, 2010.

Yvonne Fischer is with the Institute for Anthropomatics - Vision and Fusion Laboratory (IES) at Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany. (phone: +49 721 6091-571; e-mail: yvonne.fischer@kit.edu).

Alexander Bauer is with the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, 76131 Karlsruhe, Germany (e-mail: alexander.bauer@iosb.fraunhofer.de).

Jürgen Beyerer is with the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, 76131 Karlsruhe, Germany (e-mail: juergen.beyerer@iosb.fraunhofer.de). He is also with the Institute for Anthropomatics - Vision and Fusion Laboratory (IES) at Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany.

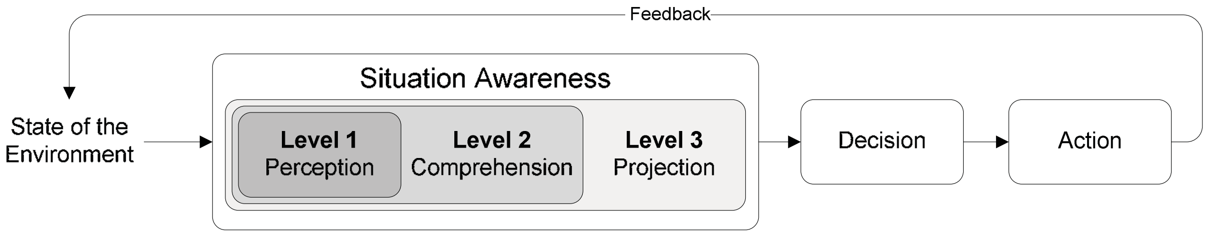

Figure 1. The process of dynamic decision making (adopted from [1]).

and appropriate signal-processing methods for extracting as much information as possible about the surveyed environment and its elements. The challenge of advanced surveillance systems is therefore not only to collect as much sensor data as possible, but also to process and present them in an intelligent and meaningful way to give a sufficient information support to some decision maker. Or, in other words, to detect and assess complex situations that evolve over time as an automatic support to an operator's situation assessment process. The information overload is then reduced by providing only relevant or task-oriented information, which can be used to guide the focus of attention of a decision maker and allow him to decide and react in a timely and effective manner. However, there is still a need for concepts and methods supporting higher level situation awareness (level 2 and 3) that are able to infer real situations from observed elements in the environment and to project their status in the near future. Especially, there is no framework that addresses the problems when assessing situations during their development and not only when they are finished. Therefore, the proposed framework was designed with the focus on this.

The paper is structured as follows. The next section gives an overview of the related work in situation awareness, data fusion, world modeling and more specific, situation modeling. Section III deals with a discussion on situational abstraction levels and tries to come up with a definition of the term situation. In Section IV, several problems for automatic situation assessment in surveillance systems are identified and a conceptual framework for automatic situation assessment that tries to deal with these problems is presented. The paper finishes with a conclusion and outlook in Section V.

II. Related Work

As mentioned in the introduction, Endsley provided in [1] a theoretical model of situation awareness and it is the most widely used model today. She developed the model based on an analysis of dynamic human decision making and stated that individuals with good situation awareness will have a greater likelihood of making appropriate decisions in complex environments. Working with heterogeneous sensors, the theories of multi-sensor data fusion [2], [3] offer a powerful technique for supporting situation awareness. A lot of data fusion models have been developed and compared to Endsley's situation awareness model [4], whereas the most dominant model is the JDL (Joint Directors of Laboratories) data fusion process model [5].

Regarding data fusion in surveillance systems, the *object-oriented world model (OOWM)* is an approach to represent the relevant information extracted from sensor signals, fused into a single comprehensive, dynamic model of the monitored area. It was developed in [6], whereas the basic ideas have been published in [7]. A detailed description of the architecture can be found in [8] and an application of the OOWM for wide area maritime surveillance is proposed in [9].

More advanced systems also support high-level functions such as situation assessment as described in general in [10]. Probabilistic methods like hidden Markov models can be used for situation recognition [11], but are strongly dependent on training data. Therefore, several other approaches have been proposed, for example grammar-parsing for detection of abnormal behavior of a person's movement in an indoor surveillance [12] or logic based approaches for the recognition of human activities [13]. In [14] a heuristic graph matching approach to identify meaningful patterns in large volumes of data have been proposed as an enhancement to existing situation assessment methods. In [15], Markov random fields are used to model contextual relationships and maximum a posteriori labeling is used to infer intentions of observed elements.

Regarding situational modeling, several concepts exist in literature. Roy proposed in [16] the concept of situation analysis as a process to provide and maintain a state of situation awareness. He also proposed definitions of situational elements like entities, events and activities. Another refinement of the situational terminology with respect to the JDL data fusion model is given in [17]. The concept of situation management in dynamic systems proposed by Jakobson [18] includes not only the processes of perceiving and recognizing situations, but also the analysis of past situations and the prediction of future situations. In [19], a rough taxonomy of functions related to situation assessment is proposed and a general overview of current approaches to automating this process is given.

III. Situational Abstraction

In the revised version of the JDL data fusion model [4], situation assessment (JDL-level 2) is defined as the estimation and prediction of relations among entities. The resulting network of relations among its elements is then referred to as the state of aggregation or the estimated situation. However, there is no formal representation of a situation, as the JDL definition admits any variety of relations to be considered. Types of relations exist at many different levels of abstraction, ranging from quantitative to highly abstract qualitative statements. Therefore, one formal representation of a situation, which fulfills several requirements in various application areas, does not exist. Situations are characterized mainly by their qualitative statements and their representation is therefore strongly dependent on the application domain.

Figure 2 shows a general decomposition of a situational description with respect to different abstraction levels. The level of abstraction is determined by the quantity of context knowledge added to the observed element, whereas only relevant context information is used. The context information consists of knowledge that is not directly observable, for example some expert knowledge. Its content and relevancy is in turn determined by the application domain and the task that an operator has to solve. The higher the level of abstraction, the lower is the level of detailed knowledge of a single, observed element. In the following, we will explain the decomposition in detail and give examples for each level of abstraction.

With the focus on surveillance systems, the perception stage includes the acquisition of object information by means of various sensors. Time invariant attributes about an object are summarized as properties and time variant attributes are summarized as the object's state. When observing for example human beings in a surveillance system, the result from the perception stage is therefore the person's position and velocity as states and the height as its property. This information is the input for the next stage, the comprehension of a situation.

At the lowest level of abstraction, a scene includes all observed objects at a point of time. A scene can therefore be interpreted as a snapshot or as a spatial subset of the world's observable objects at a point of time, whereas an episode includes also the time-dimension. An episode is the recording of all observed objects in a period of time (either discrete or continuous) and can therefore be defined as a spatio-temporal subset of the world's observable objects. Note, that at this level of abstraction, no relational aspects between objects are regarded.

The next level of abstraction deals with the description of quantitative relations that can be extracted directly from the information content of a scene or of an episode. Quantitative relations are statements about two or more relevant information values, mostly about the attribute values of some objects. The spatial distance measured in meter between two objects is for example a quantitative relation. Note that quantitative relations do not assume that the information values are derived from different objects. Another example of a quantitative relation is therefore the distance that an object has passed between two time points.

In Figure 2, special placements between quantitative and

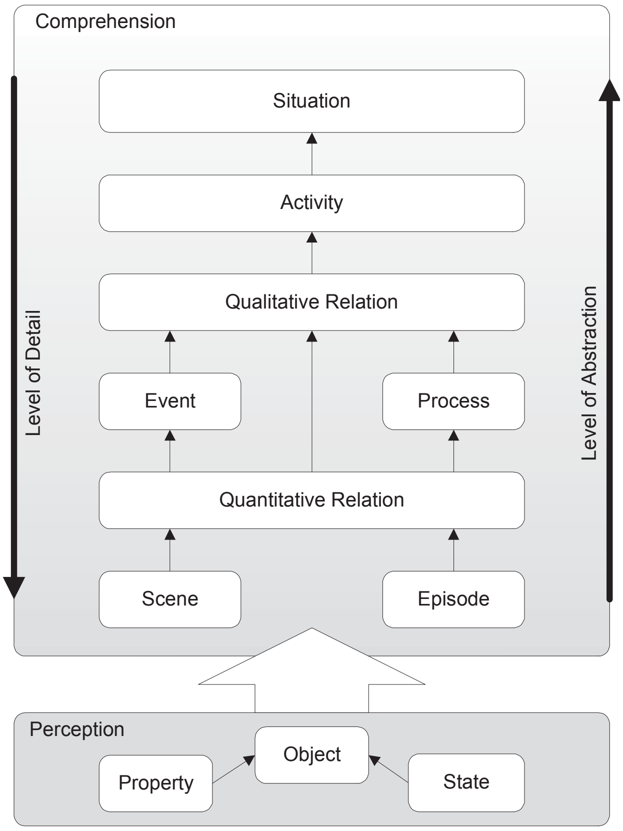

Figure 2. Situational abstraction levels.

qualitative relations are given to events and processes. They can be interpreted as special cases of quantitative relations. An event is defined as the change of relevant object information at a point of time and a process describes the behavior of relevant object information during a time period. For example, the disappearance of an observed person could be tagged as an event or that a person's attribute value, indicating its speed, has changed to zero. A process would be the person's speed value or the direction of movement over a time period. Events and processes are not limited to a single object. A process between two objects could be the decreasing distance between them or an event could be that the distance value of the quantitative relation changed to zero.

On the next higher level of abstraction, events, processes and quantitative relations can be summarized to qualitative relations. Detailed knowledge of attribute values of the observed objects goes lost at this level. A qualitative relation is therefore an interpretation of the underlying events, processes and quantitative relations. Examples for qualitative relations are therefore a person that is walking, a person that stops its movement, a person that is moving towards another object, or a person that meets another person.

Qualitative relations are strongly connected to activities. However, we state that activities take place in a longer period of time and are more complex in their construction. As qualitative relations can be interpreted as single and nondecomposable structures, an activity includes also the temporal relationships between these components. An activity is therefore a sequence of qualitative events, processes and relationships. Temporal relationships of overlapping processes can for example be expressed by Allen's temporal interval logic [20]. An example for this level of abstraction is a fighting activity between two human beings. The term behavior is often used if the focus is on activities conducted by humans or only by a single object. However, we will use the term activity for this level because it has a broader meaning.

At the highest level of abstraction, there is the situation itself. The human comprehension of a situation can be interpreted as the knowledge of everything of relevance that is going on. Therefore, based on our discussion so far, we come up with the following definition of the term situation:

> A situation at time t is defined as a world state, which is characterized by the collection of relevant activities up to the time t and their interpretation with respect to the context knowledge.

As the world evolves over time, it changes from one state to another. Therefore, the change from one situation to another is due to the change of any activity that is going on or due to a change of the context. As an example, we assume a fighting activity between two humans that is going on so far. Regarding the context, the situation is completely different if the fighting takes place on the street or inside a boxing ring, although the underlying activity is the same.

The situation assessment process can therefore be described as the estimation of the state of the world, which however does not only consist of the recognition of all activities that are going on. Moreover, it also includes contextual conditions like the environment in which the activity is taking place and its aim is to reduce the quantity of information with respect to its relevance.

IV. CONCEPTUAL FRAMEWORK

The conceptual framework proposed in this Section for supporting situation awareness by automatic situation assessment was developed with the focus on advanced surveillance systems. In such systems, a lot of heterogeneous sensors are deployed on various positions for observing a specific environment. The type of such a surveyed environment can vary from rather small areas when thinking of indoor surveillance of a building, to extremely wide areas, for example the surveillance of maritime areas. Furthermore, the sensors can be deployed stationary or on a mobile platform like an unmanned aerial vehicle (UAV). We assume that for each sensor type, there are appropriate signal-processing methods, such that each sensor is able to produce observations on object level. Furthermore, we assume that there are several low-level fusion methods like data association and tracking algorithms for generating an object-level representation of the surveyed area. This representation is called the object-oriented world model (OOWM) [6]. However, as mentioned above, simply collecting all the information of the observed elements does not automatically lead to a high level of situation awareness in an advanced surveillance system. There are several challenges to solve if a surveillance system should be able to support these higher levels:

- Mostly, there is a lack of training data, especially for critical situations that an operator wants to detect.
- For interventional reasons, critical situations have to be detected timely, which means during their development and not only when they are already finished.
- The system should be able to deal with uncertain observations, as signal processing methods usually provide estimated feature values and also false detections.
- The system should be able to detect incomplete observations, whereas the incompleteness can be of spatial and of temporal type. Spatial incompleteness follows from sensor coverage, as for example in wide areas it is not possible to continuously observe every part of the environment. Temporal incompleteness follows from spatial incompleteness in the past. As situations evolve over time, it is possible that the beginning of a situation was not observed.
- Furthermore, the system should be able to predict the situation state in the near future and give a clue to the question: What might happen next?

The process of automatic situation assessment in an advanced surveillance system should be established for refining and reducing the quantity of information that a decision maker needs to examine. Object observations are building the basis for the assessment process and the aim is to guide the focus of attention of the decision maker to relevant higher-level activities or situations.

The conceptual framework for supporting situation awareness by automatic situation assessment that tries to cope with the above listed challenges is illustrated in Figure 3. It is originally inspired by the taxonomy of functions related to situation assessment that was proposed in [19]. The proposed framework consists of four major process parts (situation characterization, situation abstraction, situation recognition, and situation projection) and the associated results of the processes. The four process parts and its connection will be explained in the following.

A. *Situation Characterization*

The first process part of the situation assessment framework is the characterization of relevant situations. As learning-based methods for situation recognition are often not realizable due to the lack of training data, this process has to be conducted by human experts. The experts provide descriptions of relevant situations, including their salient features. Such situations of interest determined by the experts are then tried to assess

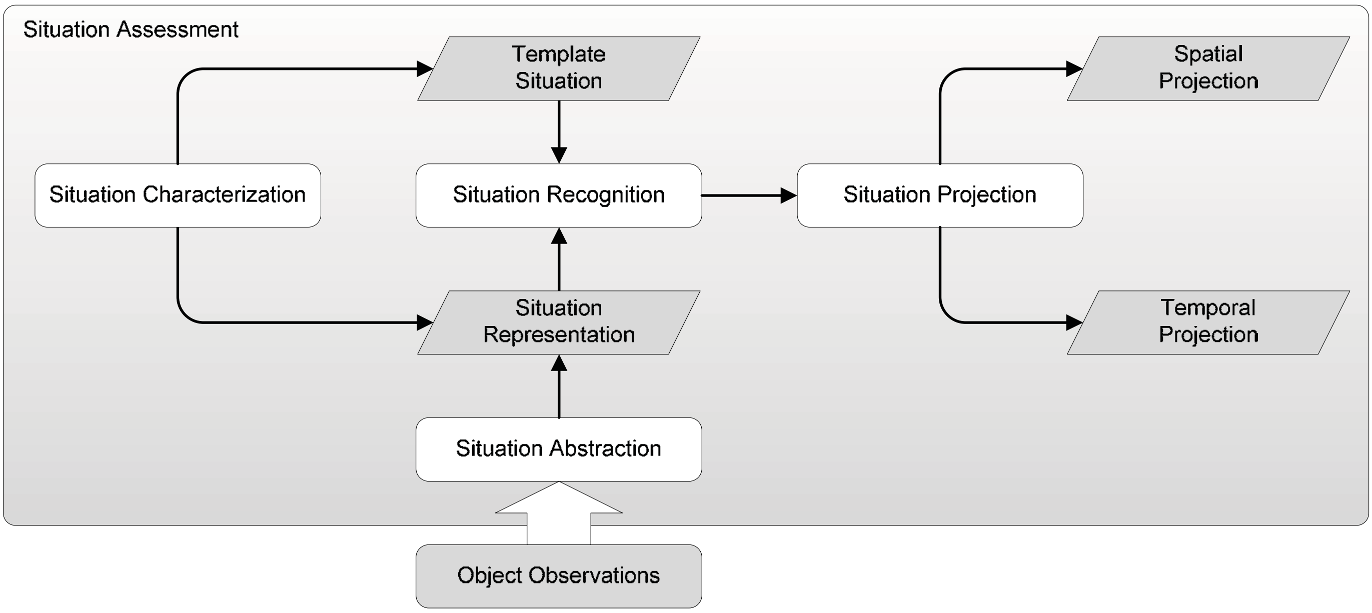

Figure 3. The Conceptual Framework for automatic situation assessment

during surveillance operation. However, the description has to be transformed into a formalized representation, namely the template situation. Especially, it has to be determined on which situational abstraction level (see Section III) the template situation is established. The template situation is not fixed at one single abstraction level, but due to performance issues and data overload, it should not vary over several levels.

B. Situation Abstraction

The second process part is the situational abstraction of the observed objects. The aim of this process is to determine the level of abstraction of the currently observed situation. Object observations are building the basis for the situation abstraction process, which finally results in the state estimation of quantitative relations, events, processes, qualitative relations or activities, depending on the level of abstraction. This result is then the state estimate of the current situation, namely the situation representation. The level of abstraction of the situation representation is again dependent on the situation characterization process. In order to support the following situation recognition process, the level of abstraction of the situation representation has to be exactly the same as the level of abstraction of the situation template. For a low level of abstraction, the situation abstraction process could be quite simple, for example if it includes only the extraction of some relevant quantitative relations. However, the achievement of a high level of abstraction could include very complex inference processes, for example when a fighting activity has to be inferred from observations of persons.

C. Situation Recognition

The third process part of the framework is the situation recognition, which deals with matching the situation representation to the situation template. However, the main challenge in situation recognition during operation is to find the correlating template to the current representation, or in other words, which observed elements in a situational representation belong to which template situation. On the one hand, there are several template situations and not only one, and on the other hand, there may be a large amount of situation representations that do not belong to any of the predefined templates, for example false detections or ongoing activities that are not relevant. The situation recognition can also be viewed as a classification problem, whereas the different template situations define the different classes.

However, methods chosen for situation recognition have to deal with several problems:

- Situation recognition should be able to detect time segments of situation templates, as it could have been missed to observe the beginning of a template situation. Furthermore, as situations should be detected while they are currently ongoing, they are not yet completed in time.
- It should be able to detect only a spatial part of a situation, as it is possible that at the point of observation, the sensors cover only a part of the environment which is under surveillance. Another reason for the spatial reduction of the situation representation could be that the inference process for generating a high level of abstraction has failed.
- Situation recognition should be able to deal with situational representations that do not match to any template situation, as mentioned above.

Due to these challenges, the result of the situation recognition should not be a binary decision if the situation is recognized or not. The result should be a degree of belief for each template situation, indicating the existence of the underlying and ongoing situation representation.

D. Situation Projection

The last process part of our framework is the projection of situations. Due to the environment, only partial information about an ongoing situation is known, whereas it can be partial in a temporal and in a spatial sense. Reasons for spatial and temporal reduction of the situation representation have been discussed in the situation recognition part. However, for a decision maker it is not only relevant to recognize a situation, or, in other words, to have a degree of belief that a certain situation is ongoing, but also to infer about unobserved elements of the situation representation. These unobserved elements could lie in the past, in the present or in the future. Results of inferring missing elements in present are called spatial projections and results of missing elements in the past or in the future are called temporal projections. The spatial and temporal projection is in general possible if the situation template is assumed to be true. Missing elements in the situation representation, either spatial or temporal, are therefore elements of the situation template that do not have a correlating element in the situation representation.

V. Conclusion and Outlook

In this article, we introduced a conceptual framework for automatic situation assessment supporting situation awareness of a decision maker in advanced surveillance systems. The framework consists of four parts. The first part, situation characterization, deals with defining situations of interest at a certain level of abstraction and generating so-called situation templates. Situational abstraction levels have been discussed in detail. The second part deals with generating a situation abstraction out of observed object information. The resulting situation representation and the situation template are matched during the third process part, the situation recognition. The result of the situation recognition is a degree of belief of every template situation, indicating the existence of the underlying situation. Furthermore, the concept differentiates between spatial and temporal projections. They are results of the situation projection and should guide the focus of attention of the decision maker on what he might have missed or what might happen in the near future. In summary, a system that regards all these parts of the proposed situation assessment framework would cover a lot of aspects that enhance the situation awareness of a decision maker in a surveillance system.

Further directions are the establishment of such a framework. First effort will be given to select several template situations at different levels of abstraction and to test several methods for the situation recognition with respect to the challenges highlighted in this article.

References

[1] M. R. Endsley, “Towards a theory of situation awareness in dynamic systems,” in *Human Factors*, vol. 37, no. 11, pp. 32-64, 1995.

[2] H. B. Mitchell, *Multi-Sensor Data Fusion: An Introduction*, Springer, 2007.

[3] D. L. Hall and S. A. H. McMullen, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, Inc., 2004.

[4] É. Bossé, J. Roy, and S. Wark, *Concepts, Models, and Tools for Information Fusion*, Artech House, Inc., 2007.

[5] A. N. Steinberg, C. L. Bowman, and F. E. White, “Revisions to the JDL Data Fusion Model,” in *Sensor Fusion: Architectures, Algorithms, and Applications, Proceedings of the SPIE*, vol. 3719, 1999.

[6] A. Bauer, T. Emter, H. Vagts, and J. Beyerer, “Object oriented world model for surveillance systems,” in *Future Security: 4^{th} Security Research Conference*, pp. 339–345, Fraunhofer Press, 2009.

[7] T. Emter, I. Gheta, and J. Beyerer, “Object oriented world model for video surveillance systems,“ in *Future Security: 3^{rd} Security Research Conference*, pp. 315-320, Fraunhofer Press, 2008.

[8] J. Moßgraber, F. Reinert, and H. Vagts, “An architecture for a task-oriented surveillance system: a service and event based approach,” in *Proc. Fifth International Conference on Systems*, ICONS, pp. 146-151 2010.

[9] Y. Fischer and A. Bauer, “Object-oriented sensor data fusion for wide maritime surveillance,” 2^{nd} International Conference on Waterside Security, WSS, 2010, to be published.

[10] S. Das, *High-Level Data Fusion*, Artech House, Inc.; 2008.

[11] D. Meyer-Delius, C. Plagemann, and W. Burgard, “Probabilistic Situation Recognition for Vehicular Traffic Scenarios,” in *Proceedings of the 2009 IEEE international conference on Robotics and Automation*, pp. 4161-4166, 2009.

[12] A. Bauer and Y. Fischer, “Task-oriented situation recognition,” in *Proceedings of SPIE*, vol. 7709, 2010.

[13] J. Ijsselmuiden and R. Stiefelhagen, “Towards high-level human activity recognition through computer vision and temporal logic,” in *KI 2010: Advances in Artificial Intelligence*, Lecture Notes in Computer Science, vol. 6359, pp. 426-435, 2010.

[14] K. Sambhoos, R. Nagi, M. Sudit, and A. Stotz, “Enhancements to high level data fusion using graph matching and state space search,” in *Information Fusion*, vol. 11, no. 4, October 2010, pp. 351-364.

[15] R. Glinton, J. Giampapa, and K. Sycara, “A Markov Random Field Model of Context for High-Level Information Fusion,” in *9th International Conference on Information Fusion*, 2006.

[16] J. Roy, “From Data Fusion to Situation Analysis,” in *Proceedings of the Fourth International Conference on Information Fusion*, 2001.

[17] J. J. Salerno, “Where's level 2/3 fusion - a look back over the past 10 years,” in *10th International Conference on Information Fusion*, IEEE, 2007.

[18] G. Jakobson, J. Buford, and L. Lewis, “Situation Management: Basic Concepts and Approaches” in *Information Fusion and Geographic Information Systems*, 2007, pp. 18-33.

[19] A. N. Steinberg, “Foundations of Situation and Threat Assessment,” in *Handbook of Multisensor Data Fusion: Theory and Practice*, 2nd Edition, edited by M. E. Liggins , D. L. Hall , and J. Llinas, CRC Press, 2008, pp. 437–501.

[20] J. F. Allen, “Maintaining Knowledge about Temporal Intervals,” in *Communications of the ACM*, vol. 26, no. 11, 1983.