Appears in the 8th *International Conference on Information Fusion*, Philadelphia, PA, USA 2005.

Comparing Future Situation Pictures

Hedvig Sidenbladh
Department of IR Systems
Division of Sensor Technology
Swedish Defence Research Agency
SE–164 90 Stockholm, Sweden
hedvig@foi.se

Pontus Svenson   Johan Schubert
Department of Data and Information Fusion
Division of Command and Control Systems
Swedish Defence Research Agency
SE–164 90 Stockholm, Sweden
ponsve,schubert@foi.se
http://www.foi.se/fusion/

Abstract – *In this paper, we suggest a practical method for measuring the difference between situation pictures obtained independently from each other. In previous work, we have considered the output of two independent tracking methods, while we here consider the output from a threat analysis/prediction system at two time instances. At two instances in time, a hypothesized threat analysis/prediction system generates future situation pictures, describing the world at the same future point in time. Qualitative and quantitative differences between the two future situation pictures are then detected. The detected difference map serves as an indication of the amount of unpredictability in the world; or alternatively, of the ability of the threat analysis/prediction system to model different areas in the future situation picture accurately. This could either be used as feedback in training a learning threat analysis/prediction system, or to inform a human user about the reliability of the future situation picture.*

Keywords: Performance evaluation, Quality estimation, Situation difference measure, Situation assessment, Threat assessment

1 Introduction

Modern data and information fusion systems and algorithms aim to speed up the processing in the OODA loop. This leads to the possibility of "getting inside the enemy's OODA loop", giving us decision superiority.

The situation pictures and threat analyses that are the outputs of fusion algorithms give commanders a better starting-point when planning own forces' future actions. Such planning is also helped by many other kinds of decision support systems, ranging from geographical databases to logistical support systems and tools for making simulations of the enemy's expected future behavior.

Even the best laid plans, however, often go awry. When unexpected things occur, it is important to be able to quickly adapt to a new situation. The planning tools mentioned before can be used to change the current plan and adapt it to new circumstances. Is is also important to be able to discover when a current situation picture is different from an anticipated one. Today, this discovery needs to be done by human operators. Since they might be preoccupied with

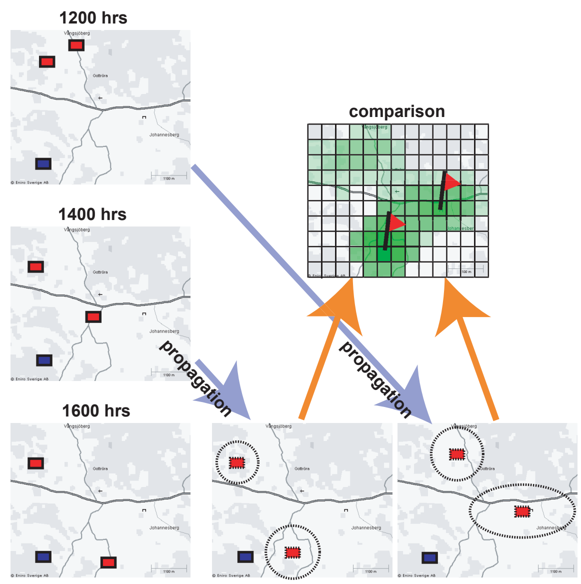

Figure 1: An information fusion system gives a situation picture at 1200 hrs. A method for propagation of situation pictures in time is used to predict the situation at 1600 hrs. At 1400 hrs, the information fusion system has an updated situation picture. The same method of propagation is again used to predict the situation at 1600 hrs. We present a method to measure the difference between the two predictions. This difference gives a measure of the reliability of the prediction method. Furthermore, it gives an indication to a operator of the information fusion system as to which areas of the predicted situation pictures to trust.

making plans or analyses for other parts of the battlefield, it would be useful to have an automatic system that alerts the human operator when re-planning needs to be done. In this paper, we describe how such a capability could work, and how it can be implemented for a simple example.

The idea is to compare threat analyses or future situation pictures made at different times (Figure 1). Given a situation picture consisting of the location of, e.g., hostile units or mobs, we hypothesize a method for predicting their future position. In a decision support system, this prediction is used to initiate resource allocations and planning to meet the threat posed by the enemy. After some time has passed, more information about the enemy is received, allowing us to update the prediction. By comparing the new prediction with the old, we can automatically detect if the enemy is following the behavior suggested by the prediction method. There can be several reasons for a discrepancy: our first prediction might be wrong because is was based on too little information, the enemy might have been forced to change their plans, or they might have succeeded in deceiving us at first.

When the divergence in behavior is discovered, the system notifies the human operator of the situation. The operator can then analyze the situation picture and determine whether replanning needs to be done or not.

The work presented here is an extension of previous work [11], where we suggested that a useful command and control system should have several subsystems that implement the same functionality. The subsystems could use different algorithms to answer the same questions, one method could for instance be slow but exact while another is fast but approximate. Another scenario is the case where new modules are integrated into existing systems — here it will ease the transition for operators if they can continue to use the old, well-known subsystem in conjunction with the new, hopefully improved version: by showing where the methods differ, the operators will learn the new capabilities of the system quickly.

The rest of this paper is outlined as follows. In Section 2, we briefly describe some related work. Section 3 describes a scenario and the hypothesized methods for prediction and threat analysis, while Section 4 reviews the mathematics of comparing situation pictures and threat maps. The paper is concluded with a discussion about potential applications of the methods, followed by a summary in Section 6.

2 Related work

With the growing amount of information flowing into command and control systems, automatic performance evaluation is an extremely important tool to help a human operator decide on what information to attend to, and when replanning is needed. We present a practical method to help the operator make such decisions. Replanning is a vital field in artificial intelligence, but most of its applications have so far been rather simple systems, such as path planning for robots (e.g., [12]). For military applications, discovering when replanning needs to be done currently relies on manual observation.

In our view, a random set [2] is the most general way to mathematically describe a situation picture. Using the random set representation, Mahler [5], Zajic and Mahler [13], and Hoffman et al. [3] have employed generalized Kullback-Leibler difference and Csiszár metrics to measure the efficiency and correctness of fusion methods in a number of papers. Although we have a similar goal, there are several differences between the approach taken in these papers and our work.

Firstly, the data compared in papers [5, 13, 3] is on the form of full random sets, while our data is on the form of probability density functions (PDF:s) or probability hypothesis density (PHD) functions [7]. Obviously, the problem of defining a distance between two multi-target probability density functions over random sets is quite different from the problem of defining a distance metric between two PDF:s or PHD functions.

Secondly, our goal is to obtain an online quality measure based on the difference between the outputs of two propagation methods, while the goal of Mahler et al. is to obtain a measure of the information content, i.e., an entropy measure, in the multi-target probability density function, or a measure of the difference between the obtained function and a known ground truth.

Thirdly, the measures in [3, 5, 13] are global, while we are interested in local measures. The desired output from our comparison method is a "difference map" over the state-space, that can be presented to a human operator as a basis for decisions on replanning.

3 System

For the rest of this paper, we will let the term situation picture refer to a list of observed and/or inferred task forces. We will assume that each task force has certain information associated with it. A minimum of information in a geographical situation picture is the position and speed of the task force. However, most of the discussion applies equally to more advanced representations, such as intent, mission, or composition of the task force. We intend to study such problems in more detail in the future.

An example of how the information fusion system could work is shown in Figure 2. The left column of the figure shows situation pictures that are obtained from a set of data fusion methods. In this example, only one task force is present, for clarity of presentation. It should be noted that this is not a limitation of the comparison algorithm.

At time t_1, the situation picture S_1 shows that an enemy task force has been observed in the upper part of the map. A commander determines that it is vital to infer where this task force will be located at time t_3. The com-

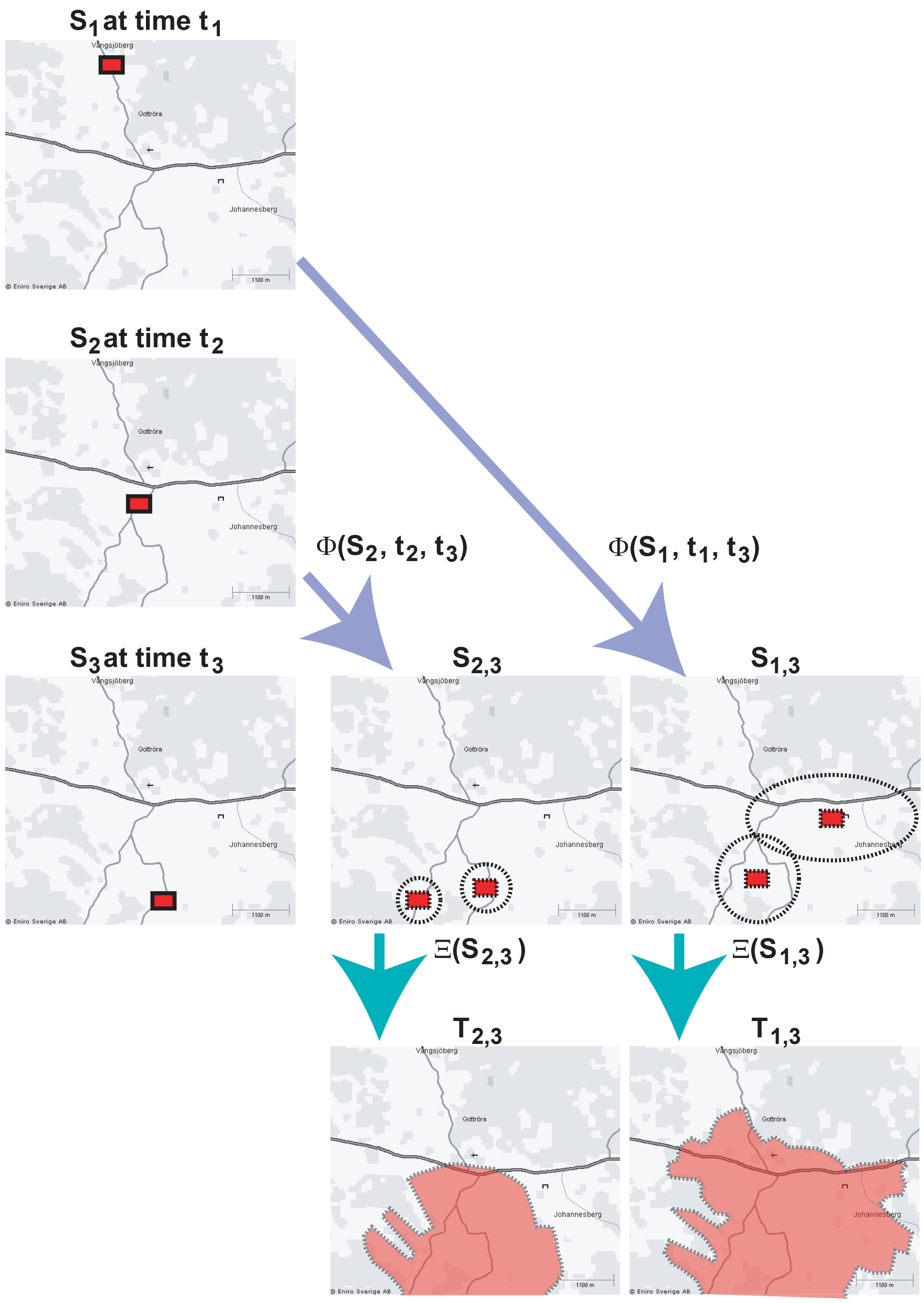

Figure 2: A hypothesized scenario. In the left column, the situation picture from an information fusion algorithm is shown. At times t_1 and t_2 respectively, two independent predictions over time of the situation picture are made. A threat analysis algorithm is used to analyze each predicted situation picture in terms of threatened areas and range of fire of hostile units.

mand and control system has access to a prediction method Φ, which gives as output the estimated situation picture $S_{1,3} = \Phi(S_1, t_1, t_3)$ based on the known state at time t_1.[1] This prediction is used by the commander to plan ahead.

The estimated situation picture $S_{1,3}$ is a probability density function over possible states with two peaks, here indicated by the task force symbols with ellipses around them. In the case of more than one task force, the estimated situation picture is represented by a random set [2, 6] structure, e.g., a PHD [7].

A threat analysis module Ξ is also available, which computes areas of visibility and range of fire [9] for enemy forces, based on the estimated situation picture. From $S_{1,3}$, a threat map $T_{1,3} = \Xi(S_{1,3})$ is computed using this module. The commander might use this map along with $S_{1,3}$ for planning.

At time t_2, the situation picture S_2 is computed. The task force has moved onto one of the smaller roads in the lower part of the map. The propagation of this picture to time t_3 gives the estimated situation picture $S_{2,3} = \Phi(S_2, t_2, t_3)$, which is based on more evidence and propagated over a shorter time span than $S_{1,3}$, thus with sharper peaks, which is indicated with smaller ellipses. The threat map $T_{2,3} = \Xi(S_{2,3})$ has a smaller threatened area than $T_{1,3}$, for the same reasons.

The scope of this paper is methods for measuring the differences in $S_{2,3}$ and $T_{2,3}$, compared to $S_{1,3}$ and $T_{1,3}$. The observed differences are due to two factors. The first source of difference is noise introduced by the methods Φ and Ξ. This part of the difference does not give the commander any vital information about the situation picture or the behavior of the enemy. The second source of differences is actual change in the behavior of the enemy between times t_1 and t_2; we term this part of the observed difference *qualitative*, since it tells the commander about qualitative changes in the planning, intent or behavior of the enemy.

Which differences are considered qualitative depends on the context in which the method is employed. This is further addressed in Section 5.

The key question for a commander using the system is:

> Is there a qualitative difference in $S_{2,3}$ and $T_{2,3}$, compared to $S_{1,3}$ and $T_{1,3}$? That is, does the commander have to replan after receiving the second prediction?

The key question for the designer of the system is:

> How does the difference measure separate between qualitative difference and difference due to method noise?

[1]The reason for including t_1 and t_3 as individual parameters, as opposed to the time of propagation $t_3 - t_1$ only, is that the propagation might be dependent on the time of day or on the season when the propagation takes place, not only on the situation picture S_1.

In the next section, a qualitative difference measure is presented, which partly compensates for method noise.

4 Measures of Comparison

In this section, we describe methods for comparing predicted situation pictures and predicted threat maps, to obtain a measure of the *qualitative* difference over time.

4.1 Comparing Predicted Situation Pictures

The first thing that can be noted about the two propagated situation pictures (see Figure 3) is that they have been propagated over different amounts of time; $S_{1,3}$ has been propagated $t_3 - t_1$ time units, while $S_{2,3}$ only $t_3 - t_2$ time units. Now imagine that the prediction method Φ adds noise in proportion to the propagation time. Addition of noise to a probability density function or PHD can be expressed mathematically as the convolution of the function with a kernel[2] of standard deviation proportional to the noise level – in other words, "smoothing" or "blurring" the function, making its peaks less pronounced. The noise added by the prediction method Φ in rendering $S_{1,3}$ can be expressed as a kernel $K_{(t_3-t_1)\sigma}$ with standard deviation $(t_3 - t_1)\sigma$, while the noise added to $S_{2,3}$ is equivalent to $K_{(t_3-t_2)\sigma}$ with standard deviation $(t_3 - t_2)\sigma$. The noise level per time step, σ, is a constant determined in Φ.

Thus, in order for an informative comparison between $S_{1,3}$ and $S_{2,3}$ to take place, the noise $(t_2 - t_1)\sigma$ must be added to $S_{2,3}$. If the distributions are expressed analytically, this can be achieved by convolving $S_{2,3}$ with a kernel $K_{(t_2-t_1)\sigma}$ of standard deviation $(t_2 - t_1)\sigma$:[3]

$$S^*_{1,3} = K_{(t_2-t_1)\sigma} \otimes S_{2,3} \qquad (1)$$

If the distributions $S_{1,3}$ and $S_{2,3}$ are instead expressed as particle clouds [10], the diffused distribution $S^*_{1,3}$ can be generated from $S_{2,3}$ by adding a random noise term sampled from $K_{(t_2-t_1)\sigma}$ to each particle in $S_{2,3}$.

A common measure for comparison of two probability density functions f and g is the Kullback-Leibler divergence [1, 4]

$$K(f, g) = \int f(\mathbf{x}) \log \frac{f(\mathbf{x})}{g(\mathbf{x})} d\mathbf{x} \, . \qquad (2)$$

However, this measure does not suit our purposes. There are two reasons for this. Firstly, the measure is only well defined when $\int f(\mathbf{x})d\mathbf{x} = \int g(\mathbf{x})d\mathbf{x}$, which is the case for probability density functions, but not for PHD functions. Secondly, it is not a proper metric, since $K(f, g) \neq K(g, f)$ in the general case. The divergence was defined

[2]A kernel is here a zero-mean function that integrates to 1. In the case of Gaussian noise, the kernel is a zero-mean Gaussian distribution. In the case of no noise, the kernel is a zero-mean Dirac function.

[3]This corresponds to the convolution with the sub-unit kernel $D_{doctrine}$ in [11].

to measure the difference between the ground truth f and an approximation g of the ground truth, or the entropy of a distribution f compared to a non-informative, or prior, distribution g - not the difference between two different approximations. Thirdly, the Kullback-Leibler divergence is a global measure. We are instead interested in a measure that is defined over a small part of the state space. In Figure 3, the comparison is made locally for each part of the state space defined by a square in the grid $\Delta S_{1,3}$.

Therefore, we use standard norms as a distance metric. The L_p norm of a function f over a sub-area A in the state space is defined as

$$\parallel f \parallel_{p,A} \equiv \left(\int_A |f(\mathbf{x})|^p d\mathbf{x}\right)^{1/p} \tag{3}$$

For the special case of $p = \infty$, we define

$$\parallel f \parallel_{\infty,A} \equiv \max_{\mathbf{x} \in A} f \ . \tag{4}$$

Using this norm, the distance function $\Delta S_{1,3}(A)$ between a certain area A in the two situation pictures $S_{1,3}$ and $S^*_{1,3}$ is defined as

$$\Delta S_{1,3}(A) = \parallel S_{1,3} - S^*_{1,3} \parallel_{p,A} \ . \tag{5}$$

This function computed for each bin in the grid (see Figure 3) gives the difference map $\Delta S_{1,3}$, a measure of the *qualitative* difference between the two predicted situation pictures $S_{1,3}$ and $S_{2,3}$.

4.2 Comparing Predicted Threat Maps

In addition to the predicted situation pictures displayed in Figure 3, more advanced representations can be compared in the same manner. One such representation is the threat maps $T_{1,3}$ and $T_{2,3}$ (see Figure 4). The threat difference $\Delta T_{1,3}$ can be presented to the operator of the information system to indicate areas where, e.g., placement of own troops needs to be replanned.

5 Discussion

We see a number of different applications of this method in military command and control systems:

Alerting the user. With the increasing amounts of information an operator of a command and control system is presented with, focus of attention is an important issue. The difference measure can draw the operator's attention to the areas in the situation picture where the prediction has changed. For example, in areas where a qualitative difference (see Section 3) is detected, a "red flag" can be raised (as in Figure 1).

As discussed in Section 3, whether the difference is qualitative is dependent on the context in which the method is employed. One alternative could be to learn [8] rules for

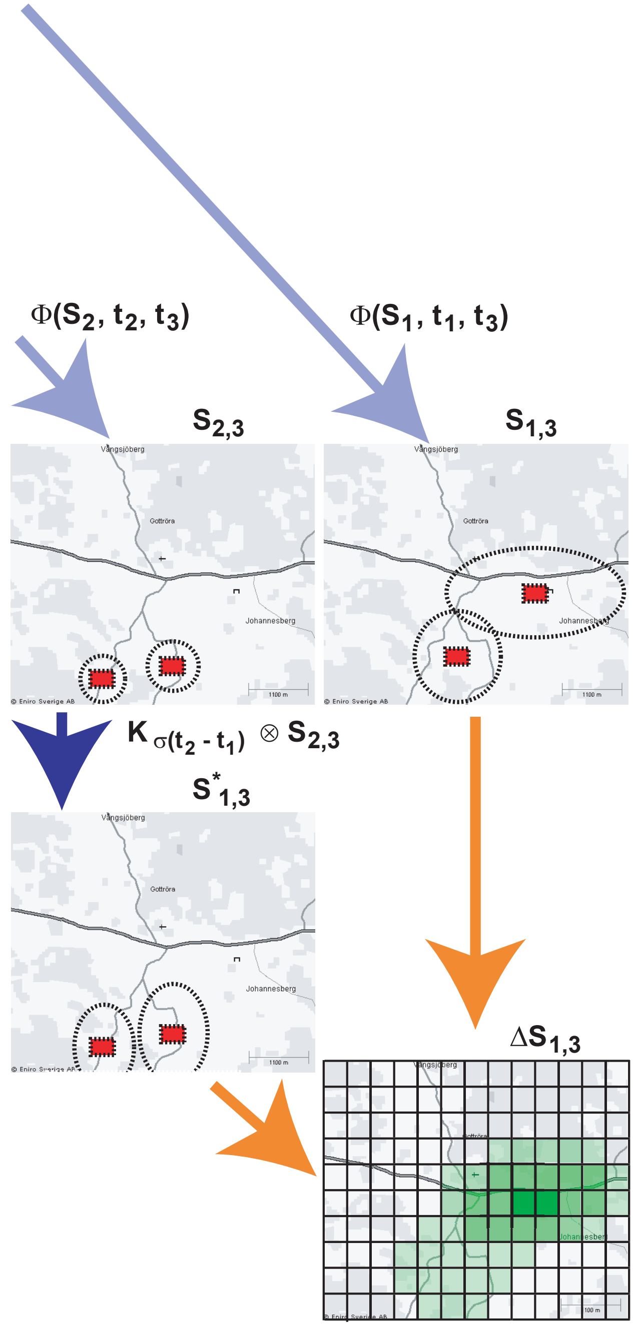

Figure 3: Comparison of predicted situation pictures with compensation for prediction noise.

the shape and level of qualitative differences, versus differences due to method noise, possibly in interaction with a human operator.

The potential reasons for qualitative differences in the prediction are many, such as an erroneous motion model in the prediction model, the fact that the enemy has changed their plans, or that the enemy has actively deceived us. A difference module can of course not give information on

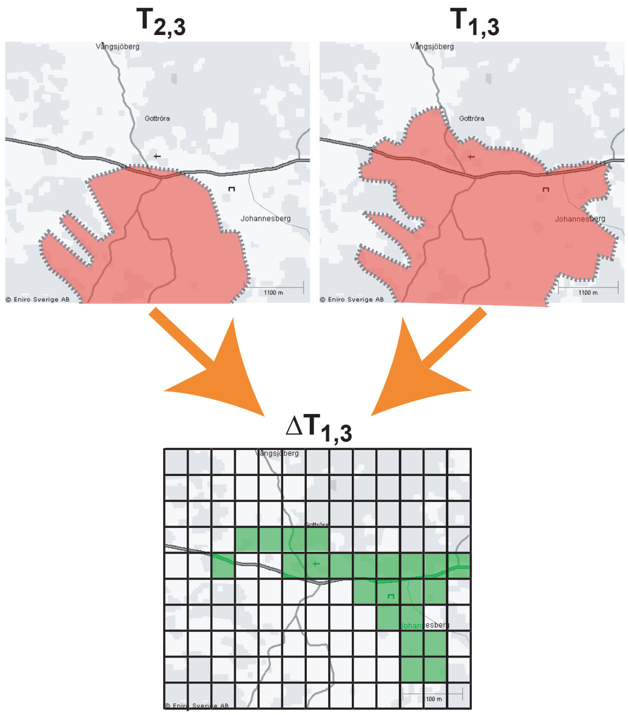

Figure 4: Comparison of threat maps.

what has happened, only alert the operator when replanning is needed, indicating in which areas of the situation picture the qualitative difference was detected.

Learning systems. It is also possible to use the difference map to automatically improve the information fusion methods in the system. It could, e.g., be used as input to an automatic replanning tool or for reinforcement learning [8] in an adaptive prediction method.

Operator education. In a training or war-game situation, the method could be used to illustrate the effect of actions taken by training operators of a command and control system. By simulating the outcomes of multiple choices the pupil could have made, the system or supervisor can later show the pupil exactly what were the consequences of his or her decisions.

Alternatively, if the pupil intervenes in the fusion process and changes an intermediate result, the system can process both the operator-changed and the old information and maintain two different end-results. A comparison module can then check the two situation pictures and the resulting differences be presented to the operator.

6 Conclusion

A method for measuring the qualitative difference between predicted situation pictures, made with the same system at different times, is presented. The goal is to help a human operator attend to events in the predicted situation pictures requiring replanning or other forms of action.

We use the L_p norm as a measure of difference between the two situation pictures, expressed as probability density or PHD functions. The measure is local, and computed over a grid in the state-space. The method also compensates for differences in the amount of noise in the predicted situation pictures. This is done by convolving the least noisy picture with a kernel with a standard deviation dependent on the difference in time span over which the predictions were made. The method can also be used to measure the difference between, e.g., threat maps evaluated from the predicted situation pictures.

References

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 1991.

[2] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen. *Mathematics of Data Fusion*. Kluwer Academic Publishers, 1997.

[3] J. R. Hoffman, R. P. S. Mahler, and T. Zajic. User-defined information and scientific performance evaluation. In Proc. SPIE Vol. 4380 *Signal Processing, Sensor Fusion, and Target Recognition X*, pages 300–311, 2001.

[4] S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.

[5] R. P. S. Mahler. Information for fusion management and performance estimation. In Proc. SPIE Vol. 3374 *Signal Processing, Sensor Fusion, and Target Recognition VII*, pages 64–75, 1998.

[6] R. P. S. Mahler. *An Introduction to Multisource-Multitarget Statistics and its Applications*. Lockheed Martin Technical Monograph, 2000.

[7] R. P. S. Mahler and T. Zajic. Multitarget filtering using a multitarget first-order moment statistic. In Proc. SPIE Vol. 4380 *Signal Processing, Sensor Fusion, and Target Recognition X*, pages 184–195, 2001.

[8] T. Mitchell. *Machine Learning*. McGraw Hill, 1997.

[9] R. Richbourg and W. K. Olson. A hybrid expert system that combines technologies to address the problem of military terrain analysis. *Expert Systems with Applications*, 11(2):207–225, 1996.

[10] H. Sidenbladh. Multi-target particle filtering for the probability hypothesis density. In *Proc. Sixth International Conference on Information Fusion*, pages 1110–1117, 2003.

[11] H. Sidenbladh, P. Svenson, and J. Schubert. Comparing multi-target trackers on different force unit levels. In Proc. SPIE Vol. 5429 *Signal Processing, Sensor Fusion, and Target Recognition XIII*, pages 306–314, 2004.

[12] A. Stentz. The focussed D* algorithm for real-time replanning. In *Proc. Fourteenth International Joint Conference on Artificial Intelligence*, 1995.

[13] T. Zajic and R. P. S. Mahler. Practical information-based data fusion performance evaluation. In Proc. SPIE Vol. 3720 *Signal Processing, Sensor Fusion, and Target Recognition VIII*, pages 92–103, 1999.