

Research Article

Naval Target Classification by Fusion of Multiple Imaging Sensors Based on the Confusion Matrix

S. Giompapa,¹ A. Farina,¹ F. Gini,² A. Graziano,¹ R. Croci,¹ and R. Di Stefano¹

¹ *Analysis of Integrated Systems Unit, SELEX Sistemi Integrati, via Tiburtina Km 12,400, 00131 Roma, Italy*

² *Department of "Ingegneria dell'Informazione", University of Pisa, via G. Caruso 16, 56122 Pisa, Italy*

Correspondence should be addressed to F. Gini, f.gini@iet.unipi.it

Received 14 July 2009; Accepted 12 December 2009

Recommended by Aleksandar Dogandzic

This paper presents an algorithm for the classification of targets based on the fusion of the class information provided by different imaging sensors. The outputs of the different sensors are combined to obtain an accurate estimate of the target class. The performance of each imaging sensor is modelled by means of its confusion matrix (CM), whose elements are the conditional error probabilities in the classification and the conditional correct classification probabilities. These probabilities are used by each sensor to make a decision on the target class. Then, a final decision on the class is made using a suitable fusion rule in order to combine the local decisions provided by the sensors. The overall performance of the classification process is evaluated by means of the "fused" confusion matrix, i.e. the CM pertinent to the final decision on the target class. Two fusion rules are considered: a majority voting (MV) rule and a maximum likelihood (ML) rule. A case study is then presented, where the developed algorithm is applied to three imaging sensors located on a generic air platform: a video camera, an infrared camera (IR), and a spotlight Synthetic Aperture Radar (SAR).

Copyright © 2009 S. Giompapa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The ability to quickly and reliably recognize non cooperative targets is of primary importance for surveillance operations in Homeland Security (HS) applications. The development of efficient fusion strategies and the improvements in the design of more reliable sensors have increased the research interest in the classification techniques in many fields. In particular, automatic surveillance systems based on imaging sensors are gaining significant interest, as proved by recent research work [1–3] addressed to improve the quality of image-based surveillance systems. The recognition techniques can include approaches based either on the human interpretation of the data provided by a sensor system or on automatic methods. The Automatic Target Classification (ATC) techniques can use data coming from sensors of different nature, such as infrared, electro-optic cameras, and radar systems. As described in [4], the process of target recognition can be conceptualized as composed by five levels or subprocesses: the detection, that is, the process of distinguishing the target from thermal noise; the

discrimination, that is, the capability to extract potential targets from surrounding clutter; the preclassification, that is, a sort of prescreen in order to exclude targets that are not of interest from further processing; the classification, that is, the process during which the targets are characterized as belonging to a specific class according to some particular features; and, finally, the identification, that is, a more sophisticated process which may refer to the individuation of the target cooperativeness or to the extrapolation of more specific features, for example, in a maritime environment the type or the name of a ship previously associated to a naval class. This work assumes that the first three processes [5] have occurred and is only concerned with the classification process. In particular, we investigate how data coming from sensors of different nature can be combined to improve the classification task. In ATC the classification task can be accomplished using several approaches. A model-based technique uses a model of the target, obtained, for example, by a Computer-Aided Design (CAD) or an Electro-Magnetic (EM) simulation [6], to compare the simulated models with the signature of the target under test. The computational

load of this methodology can be very high, especially when more than one sensor are used. Another methodology can consist of collecting many real versions of the target signature and of comparing them with the signature of the current target under test, but in this case a very large database is required and if the target (or observing environment) changes significantly the classification process may be unsuccessful.

In this work a classification algorithm is developed, which uses a different approach, where the information on the target class is provided by imaging sensors of different nature and it is expressed by means of a confusion matrix (CM). This approach allows us to overcome the difficulties related to the high computational load of the methodologies described above and to insert the classification task in the analysis and simulation of a large and complex system. The CM is analytically computed for each imaging sensor and it models the performance of the sensor during the classification process. The entries of this matrix are the conditional error probabilities in the classification and the conditional correct classification probabilities. These probabilities amount to the target class likelihood functions and are used to make the decision on the target class by each sensor. The sensor CM is analytically computed as a function of its sensitivity features, its resolution, and using a database of prestored reference images. Then a final decision on the class is made, using a suitable fusion rule in order to combine the decisions coming from the different sensors. The overall performance of the classification process is evaluated by means of the “*fused*” confusion matrix, that is, the CM pertinent to the final decision on the target class. Two fusion rules are considered to combine the class information coming from the different sensors: a majority voting (MV) rule and a maximum likelihood (ML) rule. The ultimate purpose of this fusion process is to combine the outputs of the different imaging sensors to obtain an accurate and reliable estimate of the target class. This analytical approach is then applied to a case study, where three imaging sensors located on a generic platform, performing in coastal surveillance, are considered: a video camera, an infrared (IR) camera, and a spotlight Synthetic Aperture Radar (SAR). The final information on the class, considered by means of the “*fused*” CM, could then be exploited by the system where the sensors are located to perform other surveillance operations, such as the evaluation of the threat level of a noncooperative target.

In [7], different levels of abstraction in the fusion of data coming from different imaging sensors are described: the signal-level fusion is the combination of signals from different sensors, performed before the production of images; the pixel-level fusion consists of merging different digital images; and the feature-level fusion extracts specific features from different image and combines them. The approach developed in this work, using the CM to model the classification capability of the imaging sensor, refers to a higher level of abstraction. A similar approach, where the CM is used to model the classification capability of the sensor, is also used in [8], but there the classification process is used to support the data association and to improve the tracking, especially

in presence of association uncertainty in kinematic measurements. In the literature many applications are proposed where radar images are combined with images from different kind of sensors [7, 9–11] or where heterogeneous data sets coming from dissimilar imaging sensors are combined at an information fusion level [12]. In [13] we have described the classification algorithm based on the CM applied to visible and infrared images. In the present work three sensors are considered instead of two. In particular, we have considered electromagnetic simulated images from a spotlight SAR, in addition to those from the visible camera and from the IR sensor. The results of a similar case study have already been presented in [14], in relation to three sensors. In the work presented herein we present a more complete and methodical description of the algorithm, we show more details about the numerical case study considered and the figures of the simulated images, and we report in the appendix the entire mathematical details of the analytical computation of the CM.

The main contribution of the proposed classification algorithm is the development of a methodology that allows us to emulate and incorporate the classification process in the study and simulation of a complex multisensor system, without increasing the computational load of the overall simulation. In fact, in [15] the proposed algorithm has been inserted into the simulation of a multisensor system for coastal border surveillance, without increasing the computational load of the whole simulator.

The paper is organized as follows. Section 2 describes the classification algorithm, based on the analytical computation of the CM. The fusion of the decisions on the target class coming from different imaging sensors is presented in Section 3. Two decision rules are considered, that is, a majority voting (MV) rule and a maximum likelihood (ML) rule. The performance of the decision rule is described in Section 4. In Section 5 a case study is illustrated, where the developed algorithm is applied to three imaging sensors located on a platform performing in a maritime surveillance scenario: a video camera, an IR camera, and a spotlight SAR. The numerical results for this case study are presented. Finally, in Section 6 some conclusions are drawn. The analytical details of the computation of the CM are reported in the appendix.

2. The Classification Algorithm

The generic entry of the CM of a classifier is the probability that a target belonging to the class i is misclassified as belonging to class j :

$$\begin{aligned} c_{ij}^{(k)} &= \Pr\{\text{the } k\text{th sensor decides for } H_j \text{ when } H_i \text{ is true}\} \\ &= \Pr\{\hat{d}_k = j \mid H_i\}, \end{aligned} \quad (1)$$

where H_i represents the hypothesis that the target belongs to class i and $\{\hat{d}_k = j\}$ represents the event {the k th sensor decides for H_j }. Thus the i th row of the CM represents the event {the true class of the target is i } and the class

likelihood function for the sensor output j is the j th column of \mathbf{C} [8]. The off-diagonal elements of the CM represent the conditional error probability during the classification and the diagonal elements are the conditional correct classification probabilities for a given sensor, under the hypothesis H_i :

$$\begin{aligned} P_{CC|H_i}^{(k)} &= \Pr\{\text{correct classification for the } k\text{th sensor} \mid H_i\} \quad (2) \\ &= \Pr\{\text{the } k\text{th sensor decides for } H_i \text{ when } H_i \text{ is true}\}. \end{aligned}$$

Then the correct classification probability for the k th sensor is

$$P_{cc}^{(k)} = \sum_{i=1}^M P_{CC|H_i}^{(k)} \cdot \Pr\{H_i\}, \quad (3)$$

where M is the number of hypotheses (i.e., the number of classes considered), the term $P_{CC|H_i}^{(k)}$ is the conditional correct classification probability given by the diagonal element $c_{ii}^{(k)}$ of the CM, and $\Pr\{H_i\}$ is the probability that the i -th hypothesis is true. The error probability conditioned to the i th class, for the k th sensor, is

$$P_{ERR|H_i}^{(k)} = \sum_{j=1, j \neq i}^M c_{ij}^{(k)}. \quad (4)$$

The entries of the CM are used to model the performance of each sensor during the classification and to make the decision on the target class. This means that a target detected by the system is declared as belonging to class j with a probability equal to $p_{ij}^{(k)}$ derived from the elements of the CM and this probability is used by the sensor as a threshold for the decision on the class. More specifically, in order to associate a class to an incoming target, a random variable uniformly distributed in the interval $[0, 1]$ is generated and it is compared with the threshold given by the entries of the sensor CM:

$$\begin{aligned} 0 \leq u < c_{ii}^{(k)} & \text{ then the } k\text{th sensor decides for } H_i, \\ \sum_{j=1}^{i-1} c_{ij}^{(k)} \leq u < \sum_{j=1}^i c_{ij}^{(k)} & \text{ then the } k\text{th sensor decides for } H_i \\ & \text{(for } 1 < i \leq M). \end{aligned} \quad (5)$$

This is done to simulate the classification event without generating the data. The simulation of the classification event based on the elements of the CM is shown in Figure 1.

In the classification approach described here, the entries of the CM are computed in an approximated closed-form by means of an analytical evaluation, whose details are described in the appendix. The parameters required for the analytical evaluation for each sensor are: (i) the signal-to-noise ratio (SNR) at the output of the sensor; (ii) the sensor resolution; (iii) a set of reference images stored in a database;

and (iv) the cross-correlation between the images of the database. The CM can be expressed as the following function:

$$\mathbf{C} = [c_{ij}] = f(SNR, N_H, N_V, M), \quad (6)$$

where SNR is the signal-to-noise ratio, N_H and N_V represent the sensor resolution in terms of number of pixels on the horizontal and vertical planes, respectively, and M is the dimension of the reference database. In order to simplify the analysis, the following assumptions are made:

- (1) the detection of the target has already occurred (e.g., it has been performed by a radar system);
- (2) the image database for each sensor is exhaustive, that is, the possibility that the image of the target under test is not contained in the database is not considered;
- (3) the reference images of each database do not contain noise, but this is added during the analytical computation of the CM;
- (4) the noise added on each image is Gaussian and independent from pixel to pixel.

As described in more details in the appendix, the computation of the entries of the CM in the i th row is derived from the computation of the classification error probability for the i th class. The error probability is computed in an incremental way, by defining the elementary error event in the space composed by all the possible hypotheses (H_1, \dots, H_M) and by adding the contribution of this event to the overall error probability. The partial contributions for the i th class are assigned to the off-diagonal elements $c_{ij}^{(k)}$. The diagonal elements can be computed as

$$c_{ii}^{(k)} = 1 - P_{ERR|H_i}^{(k)} = 1 - \sum_{j=1, j \neq i}^M c_{ij}^{(k)}. \quad (7)$$

In the case considered here, the dimension of the reference database is equal to the number of classes considered, since the hypothesis of exhaustive database is made. The images of the reference database for each sensor can be derived from a CAD model of the target. The algorithm for the computation of the CM is schematically represented in Figure 2. An example of database construction is mentioned in the case study of Section 5 and described in [13].

3. Fusion of the Decisions on Target Class

The purpose of the fusion process is to combine the outputs of all the imaging sensors in the system to obtain an accurate and reliable estimate of the target class. As stated before, the performance of each imaging sensor during the classification process is modelled by means of its confusion matrix $\mathbf{C}^{(k)} = [c_{ij}^{(k)}]$.

The fusion process is described in Figure 3 in the case of K imaging sensors. For simplicity, let us consider $K = 3$ sensors. For each imaging sensor, the CM matrix is analytically computed as described in Section 2 and its entries are used to make a local decision on the class, that is,

d_1 , d_2 , and d_3 . Then these local decisions are combined using a suitable decision rule. Thus the observed data is a three-dimensional vector $\mathbf{d} = (d_1, d_2, d_3)$ whose elements $\{d_k\}$ are discrete random variables (r.v.) that take values in the set $I_S = \{1, 2, \dots, M\}$, where M is the number of classes considered, and represent the decision on the target class coming from each imaging sensor. Moreover, we assume that the elements of \mathbf{d} are mutually independent, that is, the decisions made by different sensors are independent.

Let us consider the set $\mathbf{I}_d = \{\mathbf{d}^{[m]}, m = 1 \dots M^K\}$, that is, the set of all the observable sequences of $K = 3$ elements, that can be constructed with the M elements of the set I_S . The dimension of \mathbf{I}_d is M^3 . Our purpose is to map the three-dimensional vectors \mathbf{d} into a scalar value belonging to the set I_S and representing the final estimated class of the target, that is, d_f in Figure 3. This means that there are M possible hypotheses $\{H_1, H_2, \dots, H_M\}$. We assume that these hypotheses have the same a priori probability:

$$\Pr\{H_i\} = \frac{1}{M}, \quad \text{for } i = 1, \dots, M. \quad (8)$$

We indicate with $g(\mathbf{d})$ the fusion rule, that is, the function that maps the observed vector \mathbf{d} into a final decision in favour of one of the M hypotheses:

$$d_f = g(\mathbf{d}), \quad (9)$$

where $\{d_f = j\}$ represents the event (we decide in favour of H_j). This approach, based on the fusion of the decisions made by each sensor through the CM entries, allows us to manage the combination of information coming from very dissimilar imaging sensors and to compensate for the sensor parameter differences, such as the fields of view, the resolutions, and the noise features. The overall performance of the fusion process can be expressed by means of a “fused” confusion matrix, that is, the matrix pertinent to the final decision on the target class d_f . Two fusion strategies are investigated and compared in this work: one based on the majority voting decision rule and the other based on the maximum likelihood decision rule.

3.1. Majority Voting Decision Rule (MV). The majority voting decision rule consists in choosing for target class that occurs more times in the observed sequences. In the case of the three-dimensional sequences $\mathbf{d} = (d_1, d_2, d_3)$ considered here, the MV rule can be analytically expressed as follow

$$\begin{aligned} d_f = q & \quad \text{if } q : \{L_q(\mathbf{d}) \geq 2\} \\ & \equiv \{[L_q(\mathbf{d}) = 2] \cup [L_q(\mathbf{d}) = 3]\}, \end{aligned} \quad (10)$$

where $L_q(d_1, d_2, d_3)$ is the number of times the value q appears in the sequence $\mathbf{d} = (d_1, d_2, d_3)$, that is, the number of occurrences of the q th class in the observed sequence. When $L_q(\mathbf{d}) = 1$ for $q = 1, 2$, and 3 , that is, $\{(d_1 \neq d_2) \cap (d_1 \neq d_3) \cap (d_2 \neq d_3)\}$, the MV rule is not applicable. In these cases, we choose the class in favour of which the “more reliable sensor” has decided. Note that the more reliable sensor is the sensor for which the conditional correct

classification probability, given by the diagonal elements of the CM, is higher. For instance, if the sequence $\mathbf{d}^{[8]} = (1, 2, 4)$ is observed, the final class will be $d_f = q$ for which $\Pr\{d_k = q \mid H_q\}$ is maximum, for $k = 1, 2, 3$. In this example we have: $d_1 = 1$, $d_2 = 2$, and $d_3 = 4$; then we consider the diagonal elements $c_{11}^{(1)}$ for the first sensor, $c_{22}^{(2)}$ for the second sensor and $c_{44}^{(3)}$ for the third one and we decide

$$d_f = \begin{cases} 1 & \text{if } \max(c_{11}^{(1)}, c_{22}^{(2)}, c_{44}^{(3)}) = c_{11}^{(1)}, \\ 2 & \text{if } \max(c_{11}^{(1)}, c_{22}^{(2)}, c_{44}^{(3)}) = c_{22}^{(2)}, \\ 4 & \text{if } \max(c_{11}^{(1)}, c_{22}^{(2)}, c_{44}^{(3)}) = c_{44}^{(3)}. \end{cases} \quad (11)$$

The observable three-dimensional vectors $\mathbf{d}^{[m]}$ are all the possible sequences of $K = 3$ objects $\mathbf{d} = (d_1, d_2, d_3)$ that can be made with the $M = 4$ elements of the previously defined set \mathbf{I}_d . Table 1 shows all the 64 observable sequences. According to the decision rule described by (10), we can construct a fusion table for MV decision rule, as shown in Table 2. The last column of the table contains the final decision on the target class made according to the MV rule. Using this table, we can construct the “fused” matrix \mathbf{F} , after the fusion of the information on the target class:

$$\mathbf{F} = [f_{ij}]. \quad (12)$$

The entries of this matrix are

$$f_{ij} = \Pr\{d_f = j \mid H_i\} = \sum_{m \in D_j} \Pr(\mathbf{d} = \mathbf{d}^{[m]} \mid H_i), \quad (13)$$

where D_j is the decision zone of H_j , that is, the set of m 's for which we decide in favour of H_j . It is defined as

$$D_j = \{m : g(\mathbf{d}^{[m]}) = j\}. \quad (14)$$

The elements of the sum $\Pr(\mathbf{d} = \mathbf{d}^{[m]} \mid H_i)$ can be computed as:

$$\begin{aligned} \Pr(\mathbf{d} = \mathbf{d}^{[m]} \mid H_i) & \\ & = \Pr(d_1 = d_1^{[m]} \mid H_i) \cdot \Pr(d_2 = d_2^{[m]} \mid H_i) \\ & \quad \cdot \Pr(d_3 = d_3^{[m]} \mid H_i) \end{aligned} \quad (15)$$

that represents the product of the entries of the CMs of the three sensors:

$$\begin{aligned} \Pr(\mathbf{d} \mid H_i) & = \Pr\{d_1 = j, d_2 = k, d_3 = n \mid H_i\} \\ & = c_{ij}^{(1)} \cdot c_{ik}^{(2)} \cdot c_{in}^{(3)} \end{aligned} \quad (16)$$

with $j, k, n, i = 1, \dots, 4$.

Note that the sum with respect to j of the elements in (13), that is, the sum of the elements of each row of the fused CM, is equal to 1 by construction. In fact, the combination of elementary events (i.e., single-sensor decision events) belonging to three distinct probability sets, where probability sums to 1, will provide a set of probabilities whose sum will be again equal to 1.

TABLE 1: Observable sequences

111	121	131	141
112	122	132	142
113	123	133	143
114	124	134	144
211	221	131	141
212	222	232	242
213	223	233	243
214	224	234	244
311	321	331	341
312	322	332	342
313	323	333	343
314	324	334	344
411	421	431	441
412	422	432	442
413	423	433	443
414	424	434	444

TABLE 2: Fusion table for the MV rule.

	\mathbf{d}	# occurrences of 1	# occurrences of 2	# occurrences of 3	# occurrences of 4	d_f
$\mathbf{d}^{[1]}$	111	3	0	0	0	1
$\mathbf{d}^{[2]}$	112	2	1	0	0	1
$\mathbf{d}^{[3]}$	113	2	0	1	0	1
$\mathbf{d}^{[4]}$	114	2	0	0	1	1
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$\mathbf{d}^{[8]}$	124	1	1	0	1	♦
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$\mathbf{d}^{[64]}$	444	0	0	0	3	4

♦ Choose class coming from the more reliable sensor.

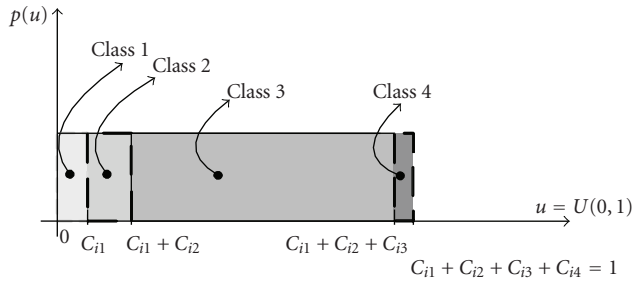


FIGURE 1: Simulation of the classification event based on the elements of the CM.

3.2. Maximum Likelihood Decision Rule. In many applications, the most common approach utilized to distinguish between two or more hypotheses is based on the *Bayes rule*, that assume a priori knowledge of the probabilities of the hypotheses under test. The Bayes rule is based on the minimization of the expectation of the cost function C_{ij} , defined as the cost assigned to the decision to choose in favour of H_j when H_i is true [16]. The analytical formulation

of the Bayesian approach applied to the decision on the target class is

$$d_f = q : q = \arg \max_j \{ \Pr(H_j | \mathbf{d}) \} \quad \text{for } j = 1, \dots, M. \quad (17)$$

The rule expressed by (17) is called *M-ary maximum a posteriori probability* (MAP) decision rule, since $\Pr(H_j | \mathbf{d})$ is the probability that the hypothesis H_j is the true one after the observation of the data \mathbf{d} , thus it is an a posteriori probability. As stated before, this decision rule assumes prior knowledge of the likelihoods of the hypotheses.

According to the Bayes theorem, the a posteriori probability $\Pr(H_j | \mathbf{d})$ can be expressed as

$$\Pr(H_j | \mathbf{d}) = \frac{p(\mathbf{d} | H_j) \cdot \Pr(H_j)}{p(\mathbf{d})}, \quad (18)$$

where $\Pr(H_j)$ is the prior probability of the j th hypothesis and $p(\mathbf{d})$ is the probability mass function (pmf) of the discrete data \mathbf{d} . Since $p(\mathbf{d})$ is a positive function that does not depend on the hypothesis, it does not affect the maximization of $\Pr(H_j | \mathbf{d})$. When the assumption of equal

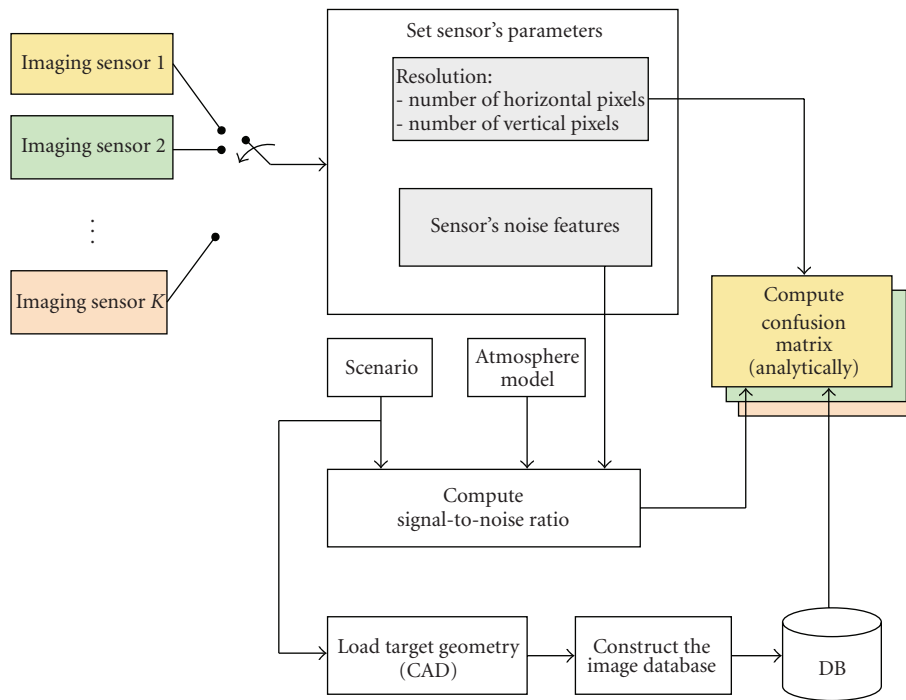


FIGURE 2: Computation of the CM.

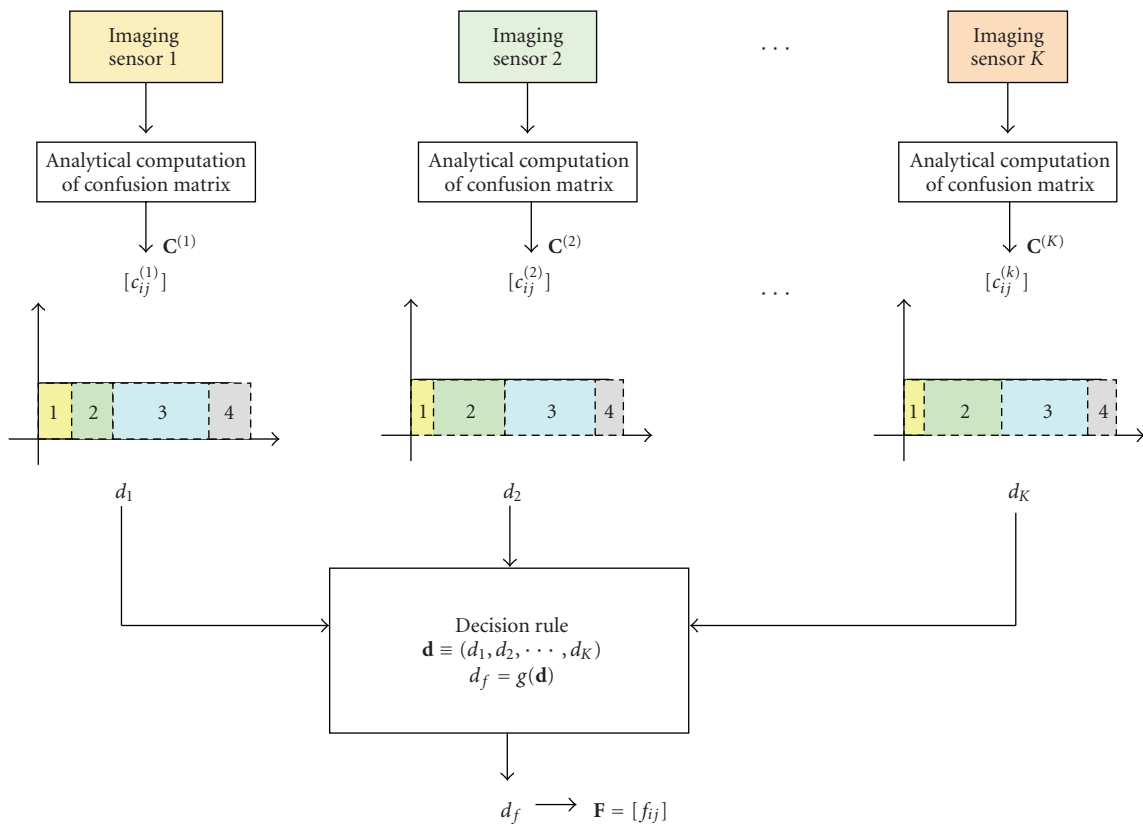


FIGURE 3: Fusion of the decision on the target class.

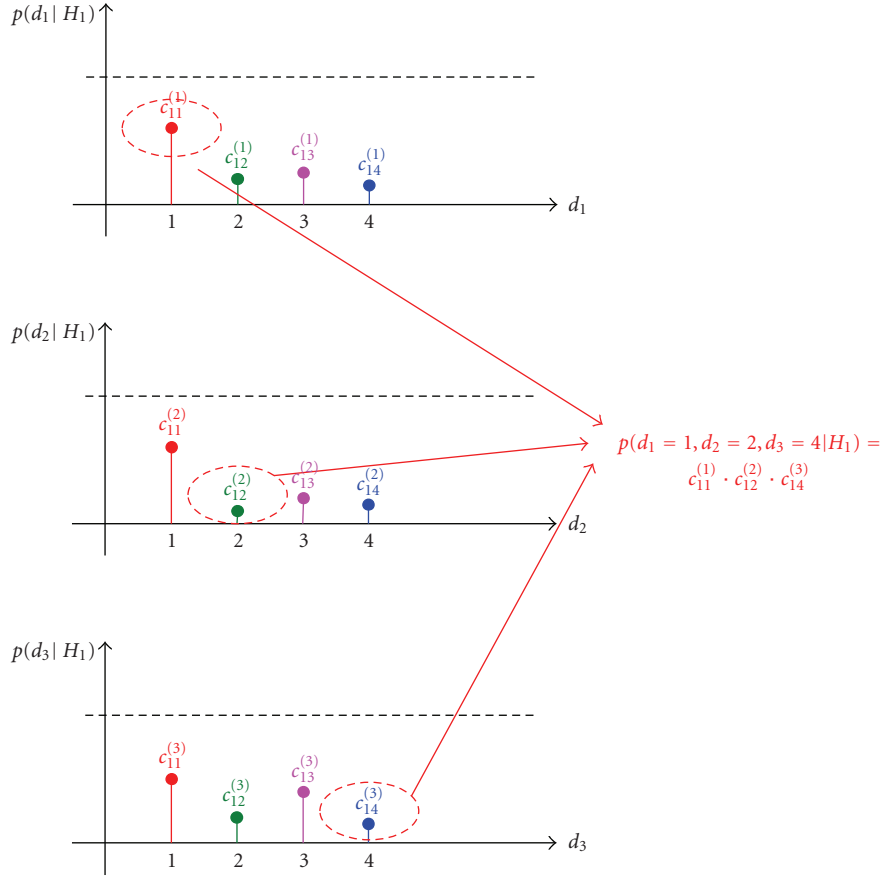


FIGURE 4: Example of the pmf, for the sequence $\mathbf{d}^{[8]} = (1, 2, 4)$ when hypothesis H_1 is true.

prior probability of the hypotheses can be done, the decision rule of (17) can be expressed as

$$d_f = q : q = \arg \max_j p(\mathbf{d} | H_j) \quad \text{for } j, q = 1, \dots, M, \quad j \neq q. \quad (19)$$

This is called *M-ary maximum likelihood* (ML) decision rule, since $p(\mathbf{d} | H_j)$ is the likelihood function of the j th hypothesis. Note that the decision rule (19) provides the minimum error probability only when the prior probabilities $\Pr(H_q)$ are all equal.

According to the ML rule, in order to decide the final class of the target using the observed data sequences \mathbf{d} , we have to choose for the hypothesis H_q that maximizes the following probability mass functions:

$$p(\mathbf{d} = \mathbf{d}^{[m]} | H_q), \quad \text{for } \mathbf{m} = 1, \dots, M^K, \quad (20)$$

for $q = 1, \dots, M,$

where

$$\begin{aligned} p(\mathbf{d} = \mathbf{d}^{[m]} | H_q) \\ = p(d_1 = d_1^{[m]} | H_q) \cdot p(d_2 = d_2^{[m]} | H_q) \\ \cdot p(d_3 = d_3^{[m]} | H_q). \end{aligned} \quad (21)$$

The elements of the product in (21) are the entries of the confusion matrices $\mathbf{C}^{(1)}$, $\mathbf{C}^{(2)}$, and $\mathbf{C}^{(3)}$, respectively. The joint conditional probability mass function of \mathbf{d} can be expressed as follows:

$$\begin{aligned} p(\mathbf{d} | H_q) &= \Pr\{d_1 = i, d_2 = j, d_3 = n | H_q\} \\ &= c_{qi}^{(1)} \cdot c_{qj}^{(2)} \cdot c_{qn}^{(3)}. \end{aligned} \quad (22)$$

This is shown in Figure 4 for the sequence $\mathbf{d}^{[8]} = (1, 2, 4)$. According to the ML decision rule described above, we can derive a fusion table for all the observable sequences, as shown in Table 3. The last column of the table contains the final decision on the target class made according to (19). Similarly to the case of the MV rule, from this table we can evaluate the fused confusion matrix \mathbf{F} , by using (13) and (14).

4. Performance Analysis

The performance of the decision rule can be expressed in terms of closeness of the fused confusion matrix \mathbf{F} to the identity matrix, which represents the ideal case. In fact, an ideal classification process is characterized by probabilities of error (off-diagonal elements of the CM) equal to zero and

by probabilities of correct classification (diagonal elements) equal to one, that is,

$$\mathbf{F} = \mathbf{I}_M, \quad (23)$$

where \mathbf{I}_M is the identity matrix of order M . The conditional correct classification probabilities for the fused matrix \mathbf{F} can be expressed from its diagonal elements, similarly to those of the CMs of the K sensors, $\mathbf{C}^{(1)}, \mathbf{C}^{(2)} \dots \mathbf{C}^{(K)}$:

$$P_{CC|H_i} = \Pr\{\text{correct classification} \mid H_i\} = f_{ii}. \quad (24)$$

The probability of correct classification is then

$$P_{CC} = \sum_{i=1}^M P_{CC|H_i} \cdot \Pr\{H_i\} = \frac{1}{M} \cdot \sum_{i=1}^M P_{CC|H_i}, \quad (25)$$

where in the last part of the equality, we have used the assumption (8) of equal a priori probability for the hypotheses. By replacing (24) in (25), we obtain

$$P_{CC} = \frac{1}{M} \cdot \sum_{i=1}^M P_{CC|H_i} = \frac{1}{M} \cdot \sum_{i=1}^M f_{ii} = \frac{\text{tr}(\mathbf{F})}{M}, \quad (26)$$

where we have considered that the sum of the diagonal elements represents the trace of matrix \mathbf{F} . To evaluate the performance of the fusion process, we consider the probability of correct classification expressed in (26) and we select as the best performing matrix the one for which the probability of correct classification, and then the ratio $\text{tr}(\mathbf{F})/M$, is the nearest to 1, that is, its maximum value. This occurs when the trace of matrix \mathbf{F} at the numerator is close to M , which is the trace of the identity matrix. From this point of view, the correct classification probability is an indication of the closeness of matrix \mathbf{F} to the identity \mathbf{I}_M .

The same performance criterion can be explained by considering an alternative interpretation. In order to evaluate the closeness of the fused matrix \mathbf{F} to the identity, we can define the following quality factor [14]:

$$\begin{aligned} Q &= 1 - \frac{\text{tr}\{\mathbf{I}_M - \mathbf{F}\}}{\text{tr}(\mathbf{I}_M)} = 1 - \frac{\text{tr}(\mathbf{I}_M) - \text{tr}(\mathbf{F})}{\text{tr}(\mathbf{I}_M)} \\ &= \frac{\text{tr}(\mathbf{F})}{\text{tr}(\mathbf{I}_M)} = \frac{\text{tr}(\mathbf{F})}{M}. \end{aligned} \quad (27)$$

The parameter Q belongs to the interval $[0, 1]$ and it is close to 1 when the fused matrix \mathbf{F} is very close to identity and it is close to 0 when \mathbf{F} is significantly different from identity. As we can see by comparing (26) and (27), the parameter Q is equivalent to the probability of correct classification. Thus, the best fused matrix is the one for which this quality factor is nearest to 1, that is also the maximum value of the correct classification probability. The difference $\text{tr}(\mathbf{I}_M) - \text{tr}(\mathbf{F})$ at the numerator of the first side of expression (27) represents the sum of the off-diagonal elements of \mathbf{F} :

$$\text{tr}(\mathbf{I}_M) - \text{tr}(\mathbf{F}) = M - \text{tr}(\mathbf{F}) = \sum_{\substack{j=1 \\ j \neq i}}^M \sum_{i=1}^M f_{ij}. \quad (28)$$

This property is due to the fact that the sum of all the elements of the matrix \mathbf{F} is equal to M , that is due to the fact that the sum of the elements in each row of the CM is equal to 1.

5. Numerical Case Study

In this section a case study is presented, where the developed algorithm is applied to three imaging sensors located on a generic air platform: a video camera, an infrared camera (IR), and a spotlight Synthetic Aperture Radar (SAR). A numerical example, concerning the classification process performed by the three imaging sensors, is provided, for four classes of naval targets. This case study allowed us to include and test the algorithm proposed in this work inside the simulation of a complex multisensor system, which performs its operations in a realistic scenario for maritime border surveillance. The considered system is notional. The numerical values considered in this example reflect a typical maritime situation, with standard environmental conditions.

5.1. The System and the Scenario. The background of this case study is represented by an integrated multisensor system for the coastal surveillance. The focus herein is on the classification process, in particular on the fusion of the target class data coming from different imaging sensors. This is a part of a research activity whose final goal is the development of a computer simulator that emulates the main functions of the integrated multisensor system for coastal surveillance (see [15, 17, 18]). The integrated multisensor system is composed of two platforms of multiple sensors: a land-based platform, located on the coast, and an air platform, moving in front of the coast. The land platform is equipped with a Vessel Traffic Service (VTS) radar, an infrared camera (IR), and a station belonging to an Automatic Identification System (AIS) that provides an information on the target cooperativeness. The air platform carries an Airborne Early Warning Radar (AEWR), which can operate in a spotlight SAR mode, a video camera, and a second IR camera. The mission of the system is the detection, the tracking, the identification and the classification of multiple targets that enter a sea region, the assessment of their threat level and the selection of a suitable intervention on them. The threat evaluation and the selection of the intervention are performed by a command and control centre (C^2), which coordinates all the operations of the multisensor system. The threat evaluation logic is based on a deterministic comparison between the target kinematical parameters detected by the two radars and some tolerance thresholds on the speed, on the distance between the target and the coast, and on the direction. This logic also takes into account for the class information provided by the imaging sensors inside the system. The three imaging sensors of the air platform are considered herein. After that the decision on the target class is made by each imaging sensor, according to the algorithm described in Section 2, the fusion of the decisions is performed as described in Section 3. The information on the class is generally not very reliable for long

TABLE 3: Fusion table for the ML rule.

	\mathbf{d}	$p(\mathbf{d} H_1)$	$p(\mathbf{d} H_2)$	$p(\mathbf{d} H_3)$	$p(\mathbf{d} H_4)$	d_f
$\mathbf{d}^{[1]}$	[111]	$c_{11}^{(1)} \cdot c_{11}^{(2)} \cdot c_{11}^{(3)}$	$c_{21}^{(1)} \cdot c_{21}^{(2)} \cdot c_{21}^{(3)}$	$c_{31}^{(1)} \cdot c_{31}^{(2)} \cdot c_{31}^{(3)}$	$c_{41}^{(1)} \cdot c_{41}^{(2)} \cdot c_{41}^{(3)}$	1
$\mathbf{d}^{[2]}$	[112]	$c_{11}^{(1)} \cdot c_{11}^{(2)} \cdot c_{12}^{(3)}$	$c_{21}^{(1)} \cdot c_{21}^{(2)} \cdot c_{22}^{(3)}$	$c_{31}^{(1)} \cdot c_{31}^{(2)} \cdot c_{32}^{(3)}$	$c_{41}^{(1)} \cdot c_{41}^{(2)} \cdot c_{42}^{(3)}$	1
$\mathbf{d}^{[3]}$	[113]	$c_{11}^{(1)} \cdot c_{11}^{(2)} \cdot c_{13}^{(3)}$	$c_{21}^{(1)} \cdot c_{21}^{(2)} \cdot c_{23}^{(3)}$	$c_{31}^{(1)} \cdot c_{31}^{(2)} \cdot c_{33}^{(3)}$	$c_{41}^{(1)} \cdot c_{41}^{(2)} \cdot c_{43}^{(3)}$	1
$\mathbf{d}^{[4]}$	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$\mathbf{d}^{[5]}$	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$\mathbf{d}^{[8]}$	[124]	$c_{11}^{(1)} \cdot c_{12}^{(2)} \cdot c_{14}^{(3)}$	$c_{21}^{(1)} \cdot c_{22}^{(2)} \cdot c_{24}^{(3)}$	$c_{31}^{(1)} \cdot c_{32}^{(2)} \cdot c_{34}^{(3)}$	$c_{41}^{(1)} \cdot c_{42}^{(2)} \cdot c_{44}^{(3)}$	
$\mathbf{d}^{[64]}$	[444]	$c_{14}^{(1)} \cdot c_{14}^{(2)} \cdot c_{14}^{(3)}$	$c_{24}^{(1)} \cdot c_{24}^{(2)} \cdot c_{24}^{(3)}$	$c_{34}^{(1)} \cdot c_{34}^{(2)} \cdot c_{34}^{(3)}$	$c_{44}^{(1)} \cdot c_{44}^{(2)} \cdot c_{44}^{(3)}$	4

TABLE 4: CM of the video camera, $\mathbf{C}^{(1)}$.

output \rightarrow input \downarrow	<i>Dinghy</i>	<i>Motor boat</i>	<i>Fishing boat</i>	<i>Oil tanker</i>
<i>Dinghy</i>	0.9209	0.0648	0.0142	0.0000
<i>Motor boat</i>	0.0648	0.8490	0.0862	0.0000
<i>Fishing boat</i>	0.0142	0.0862	0.8996	0.0000
<i>Oil tanker</i>	0.0000	0.0000	0.0000	1.0000

distance from the air platform, when only the spotlight SAR is active, and it becomes more reliable when the target is closer to the platform, when also the video camera and IR camera are active.

According to the evaluation logic a *Threat Level* is assigned to the non-cooperative targets in the set (TL0, TL1, TL2), where TL0 indicates a neutral target, TL1 a suspect target, and TL2 a threat target. The intervention is only selected for the targets assessed as threat and it consists in the allocation of a system resource in order to inspect the nature of the target. Two types of resources are considered here: a helicopter and a patrol boat; both the resources are used only for the target inspection [15, 18]. The architecture of the surveillance multisensor system we refer to is shown in Figure 5. The simulated scenario is composed of: the geographical area considered, the position of the sensors in this area, the multiple naval targets that enter the scene, and the resources of the system. Four classes of naval targets are considered in this scenario.

- (i) *Class 1*: high speed dinghy;
- (ii) *Class 2*: motor boat;
- (iii) *Class 3*: fishing boat;
- (iv) *Class 4*: oil tanker.

5.2. Numerical Results. The CMs of the three sensors have been computed by considering the analytical algorithm described above, for the four classes of naval targets. The analytical computation of the CM requires the setup of a database of reference images. In this numerical example the reference database for each sensor is composed of simulated images, no real data have been considered for now. This

database has been constructed using a three-dimensional (3D) CAD model of the naval targets considered in the scenario. The same CAD models have been exploited for the construction of the reference database for the video camera, for the IR camera, and for the spotlight SAR. The sizes considered for the naval targets reported are: (10 × 4.6 × 3) m for the dinghy; (15 × 4.7 × 5.3) m for the motor boat; (16 × 5.3 × 7) m for the fishing boat; (100 × 33.5 × 37.6) m for the oil tanker. For the video camera the image generation is simply obtained by the projection of the 3D CAD on the camera focal plane. For the IR camera, the images are simulated using a specific simulation software, the Open-source Software for Modelling and Simulation of Infrared Signatures (OSMOSIS) [19], developed at the Military Royal Academy of Brussels. For the spotlight SAR the CADs have been processed by a software for the simulation of electromagnetic (EM) images. An example of the simulated images for the dinghy for a view angle equal to 45° and is shown in Figures 6(a)–6(c), respectively, for the video camera, the IR camera, and the spotlight SAR. The distance between the sensor and the target is 5 Km for the video camera and 1 Km for the IR camera, where the temperature information is represented by the gray scale of the images.

The SNR over the single pixel of the reference images has been evaluated by considering the noise level of each imaging sensor. As it concerns the electro-optical sensors (EO) we have considered the Noise Equivalent Illuminance (NEIL) for the video camera and the Noise Equivalent Temperature Difference (NETD) for the IR camera. In both cases we have taken into account that the SNR decreases with the distance between the target and the sensor because of the atmosphere attenuation. For the video camera we have analytically computed the atmosphere extinction coefficient

TABLE 5: CM of the IR camera, $\mathbf{C}^{(2)}$.

output → input ↓	<i>Dinghy</i>	<i>Motor boat</i>	<i>Fishing boat</i>	<i>Oil tanker</i>
<i>Dinghy</i>	0.8076	0.0289	0.1635	0.0000
<i>Motor boat</i>	0.0289	0.9529	0.0182	0.0000
<i>Fishing boat</i>	0.1635	0.0182	0.8182	0.0000
<i>Oil tanker</i>	0.0000	0.0000	0.0000	1.0000

TABLE 6: CM of the spotlight SAR, $\mathbf{C}^{(3)}$.

output → input ↓	<i>Dinghy</i>	<i>Motor boat</i>	<i>Fishing boat</i>	<i>Oil tanker</i>
<i>Dinghy</i>	0.5995	0.2306	0.1698	0.0000
<i>Motor boat</i>	0.2306	0.7167	0.0526	0.0000
<i>Fishing boat</i>	0.1698	0.0526	0.7775	0.0000
<i>Oil tanker</i>	0.0000	0.0000	0.0000	1.0000

[20], assuming a wavelength value of 550 nm. For the IR camera the extinction coefficient has been computed by LOWTRAN (LOW resolution TRANsmission model) for standard weather conditions: a temperature equal to 30°C, a relative humidity equal to 43%, and a sea state equal to 0. For the evaluation of the SNR in the case of the spotlight SAR, we have considered the radar equation, revisited in order to take into account for the SAR geometry [21, 22]. The simulated images of the three sensors are referred to the same geometrical and environmental conditions, but the SNR value can be different from one sensor to the other, due to the different nature of the sensors. The classification approach described in this work allows compensating for the sensor parameter differences, such as the fields of view, the resolutions and the noise features.

In this case study we have assumed that the decisions coming from the three sensors are aligned in time. In a future development of this work we will also consider the time misalignment in the decisions, by introducing a delay in the fusion process in order to take into account for the sampling rate of the slower sensor.

The CMs of the three imaging sensors considered in the case study are $\mathbf{C}^{(1)} = [c_{ij}^{(1)}]$ for the video camera, $\mathbf{C}^{(2)} = [c_{ij}^{(2)}]$ for the IR camera, and $\mathbf{C}^{(3)} = [c_{ij}^{(3)}]$ for the spotlight SAR. According to the definition given in (1), the generic entry of the matrix $\mathbf{C}^{(1)}$ is the following probability:

$$c_{ij}^{(1)} = \Pr\{\text{video camera decides for } H_j \text{ when } H_i \text{ is true}\} \\ = \Pr\{d_1 = j \mid H_i\}. \quad (29)$$

Similarly, the entries of matrices $\mathbf{C}^{(2)}$ and $\mathbf{C}^{(3)}$ are defined as

$$c_{ij}^{(2)} = \Pr\{\text{IR camera decides for } H_j \text{ when } H_i \text{ is true}\} \\ = \Pr\{d_2 = j \mid H_i\}, \quad (30)$$

$$c_{ij}^{(3)} = \Pr\{\text{Spotlight SAR decides for } H_j \text{ when } H_i \text{ is true}\} \\ = \Pr\{d_3 = j \mid H_i\}. \quad (31)$$

Moreover we have:

- (i) $\{d_1 = j\} \equiv \{\text{Video camera decides for } H_j\}$;
- (ii) $\{d_2 = j\} \equiv \{\text{IR camera decides for } H_j\}$;
- (iii) $\{d_3 = j\} \equiv \{\text{Spotlight SAR decides for } H_j\}$.

Thus, the conditional correct classification probabilities are $c_{ii}^{(1)} = \Pr\{d_1 = i \mid H_i\}$ for the video camera; and $c_{ii}^{(2)} = \Pr\{d_2 = i \mid H_i\}$ for the IR camera; $c_{ii}^{(3)} = \Pr\{d_3 = i \mid H_i\}$ for the spotlight SAR.

In the case study a distance between the target and the sensor equal to 10 Km and a view angle equal to 180° have been considered. The resulting CMs for the video camera, the IR camera, and the spotlight SAR are reported, respectively, in Tables 4, 5, and 6. These tables show that the less reliable sensor, as it concerns the classification of the four targets considered, is the spotlight SAR. On the other side this sensor has a major coverage with respect to the other two sensors. The correct classification probability conditioned to *Class 4* (oil tanker) is always $P_{CC|H_4} = 1$, due to the fact that the size of this class of target (100 m) is significantly different from the size of the other targets considered. The fused CMs obtained by the majority voting rule, \mathbf{F}_{MV} , and by the maximum likelihood rule, \mathbf{F}_{ML} , are shown in Tables 7 and 8, respectively. From these tables we can observe that the best performing fused matrix, that is, the nearest to \mathbf{I}_4 , is the one obtained by the ML decision rule, \mathbf{F}_{ML} . The goodness of the CMs in Tables 4–8 is expressed by means of the probability of correct classification P_{CC} that is equal to the quality factor Q defined in (27). As shown in Table 9, the value of P_{CC} nearest to 1 is that corresponding to the fused matrix \mathbf{F}_{ML} . The fusion process can provide an improvement in the correct classification probability equal to 3.59% for the MV rule and

TABLE 7: Fused CM by using the MV rule, F_{MV} .

output → input ↓	<i>Dinghy</i>	<i>Motor boat</i>	<i>Fishing boat</i>	<i>Oil tanker</i>
<i>Dinghy</i>	0.9455	0.0226	0.0317	0.0000
<i>Motor boat</i>	0.0523	0.9407	0.0069	0.0000
<i>Fishing boat</i>	0.0660	0.0069	0.9269	0.0000
<i>Oil tanker</i>	0.0000	0.0000	0.0000	1.0000

TABLE 8: Fused CM by using the ML rule, F_{ML} .

output → input ↓	<i>Dinghy</i>	<i>Motor boat</i>	<i>Fishing boat</i>	<i>Oil tanker</i>
<i>Dinghy</i>	0.9460	0.0229	0.0309	0.0000
<i>Motor boat</i>	0.0262	0.9597	0.0140	0.0000
<i>Fishing boat</i>	0.0166	0.0097	0.9735	0.0000
<i>Oil tanker</i>	0.0000	0.0000	0.0000	1.0000

equal to 5.24% for the ML rule, with respect to the most reliable sensor, that is, the video camera in the numerical example considered here. The value $P_{CC} \equiv Q$ for all the CMs considered in this numerical example is also graphically shown in Figure 7.

6. Conclusions

This work describes a classification algorithm based on the fusion of the class information provided by multiple imaging sensors. The classification algorithm automatically exploits the a priori knowledge provided by the sensor CM, which is used to model the sensor performance during the classification process. The entries of the CM are the conditional error probabilities in the classification and the conditional correct classification probabilities, and they are used to make the decision on the target class by each sensor. The CM is analytically computed as a function of the sensor SNR, the sensor resolution, a set of simulated reference images stored in a database, and the cross-correlation between the reference images. Then a final decision on the class is made, using a suitable fusion rule, in order to combine the decisions coming from the three sensors. The fusion, operated on the single decisions, allows us to manage the combination of information coming from very dissimilar imaging sensors and to compensate for the sensor parameter differences. The overall performance of the classification process is evaluated by means of the fused CM, that is, the matrix pertinent to the final decision on the target class. Two decision rules are described in the paper: a majority voting (MV) rule and a maximum likelihood (ML) rule. A numerical example is finally proposed where the described classification algorithm is applied to a case study where three imaging sensors are located on a generic platform. The three imaging sensors are a video camera, an IR camera, and a spotlight SAR, and they operate into a multisensor system for coastal surveillance. The final information on the class is used in the multisensor system, as a support to other processes required during the surveillance operation. This methodology allowed us to

include the classification process inside the simulation of a complex multisensor surveillance system, without increasing the overall computational load [18].

As a final remark, we note that in this analysis we have assumed that a recognition process always occurs. Future developments of the described approach are expected to refine the model, by considering the possibility that the image under test is not contained in the image database and by evaluating the performance of the joint process of recognition and classification.

Appendix

Analytical Computation of the Confusion Matrix

The generic entry of the CM of a sensor is the probability that a target belonging to the class i is misclassified as belonging to class j :

$$c_{ij} = \Pr\{\text{the sensor decides for } H_j \text{ when } H_i \text{ is true}\}, \quad (\text{A.1})$$

where H_i represents the hypothesis that the target belongs to class i . The computation of the entries of the CM in the i th row is derived from the computation of the classification error probability for the i th class. The error probability is computed in an incremental way by adding the contribution of the generic elemental error event to the overall error probability: these partial contributions to the error probability for the i th class are assigned to the off-diagonal elements c_{ij} of the of the CM. The diagonal elements, representing the conditional correct classification probabilities, can be consequently computed as

$$c_{ii} = 1 - P_{\text{ERR}|H_i} = 1 - \sum_{i=1, i \neq j}^M c_{ij}, \quad (\text{A.2})$$

where M is the number of classes considered. The elemental error event in the classification of an image belonging to

TABLE 9: Performance of the CMs.

	$C^{(1)}$	$C^{(2)}$	$C^{(3)}$	F_{MV}	F_{ML}
$Q \equiv P_{CC}$	0.9174	0.8947	0.7734	0.9533	0.9698

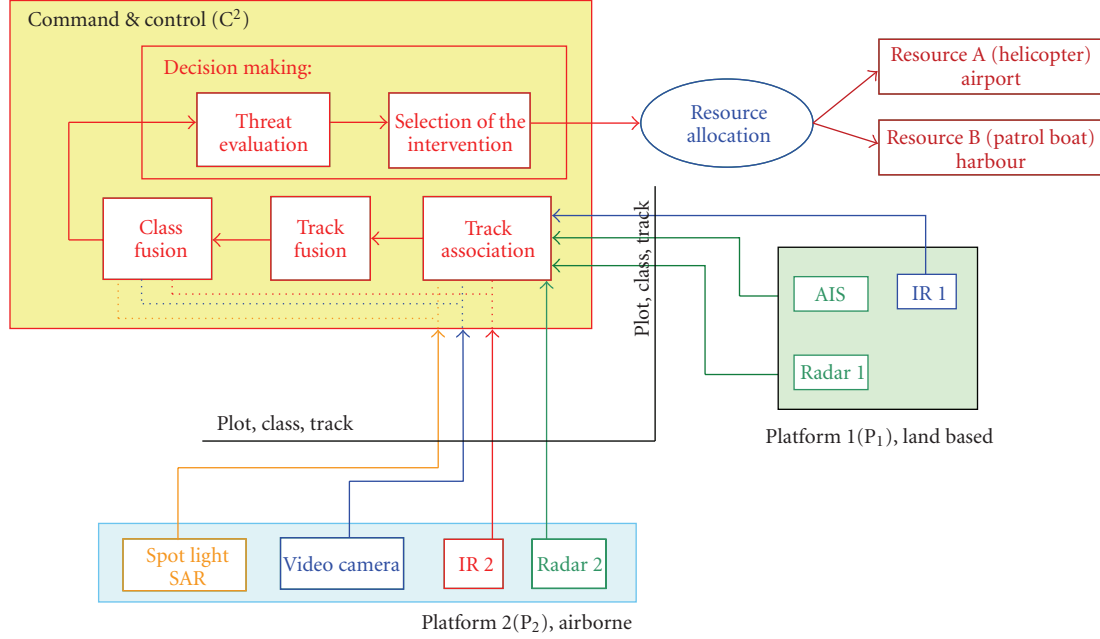


FIGURE 5: Architecture of the integrated multisensor system for coastal surveillance.

a given class is defined through the correlation between the reference images and through their energy differences, compared to the variance of the noise over the single pixel. The error probability can be defined as the average of the conditional error probabilities:

$$P_{ERR} = \sum_{i=1}^M P_{ERR|H_i} \cdot P(H_i). \quad (A.3)$$

Consider a database of M reference images $\{\mathbf{I}_1, \dots, \mathbf{I}_M\}$, one for each class, and a received image \mathbf{I} . These images are represented by matrices whose dimension depends on the sensor resolution, in terms of pixels on the horizontal and vertical planes. In general, these matrices are dependent on the target coordinate along the azimuth and the elevation. In the rest of this appendix this dependence is omitted to simplify the notation. The elements of these matrices are intensities proportional to the optical power received from the target for the video camera and the IR camera, and to the target RCS for the spotlight SAR. In this analysis the following assumptions are made.

- (i) The image database is exhaustive, that is, the possibility that the image of the target under test is not contained in the database is not considered.
- (ii) The reference images of each database do not contain any source of noise, but this is added during the analytical computation of the CM.

- (iii) The noise added over each image is additive, Gaussian, and independent of pixel to pixel.

Let us indicate with Y the observation space and let us divide this space in M decision zone $\{Y_1, \dots, Y_M\}$, such that if the image \mathbf{I} belongs to the zone Y_k then the hypothesis H_k is true. The error probability is

$$P_{ERR} = \sum_{i=1}^M P(H_i) \cdot \sum_{k=1, k \neq i}^M \int_{Y_k} p(\mathbf{I} | H_i) dY, \quad (A.4)$$

$$P_{ERR} = \sum_{k=1}^M \int_{Y_k} \left(\sum_{i=1, i \neq k}^M p(\mathbf{I} | H_i) \cdot P(H_i) \right) dY, \quad (A.5)$$

where the term $\sum_{k=1, k \neq i}^M \int_{Y_k} p(\mathbf{I} | H_i) dY$ in (A.4) represents the probability that an image \mathbf{I} generated by the i th class belongs to the decision zone Y_k , thus generating an error in the classification, and the term $\int_{Y_k} (\sum_{i=1, i \neq k}^M p(\mathbf{I} | H_i) \cdot P(H_i)) dY$ in (A.5) represents the probability that an image \mathbf{I} belonging to the decision zone Y_k is generated by the i th class, thus generating an error in the classification.

A given image, \mathbf{I} , belongs to the decision zone Y_k when

$$\sum_{i=1, i \neq k}^M p(\mathbf{I} | H_i) \cdot P(H_i) < \sum_{i=1, i \neq m}^M p(\mathbf{I} | H_i) \cdot P(H_i) \quad \forall m \neq k, \quad (A.6)$$

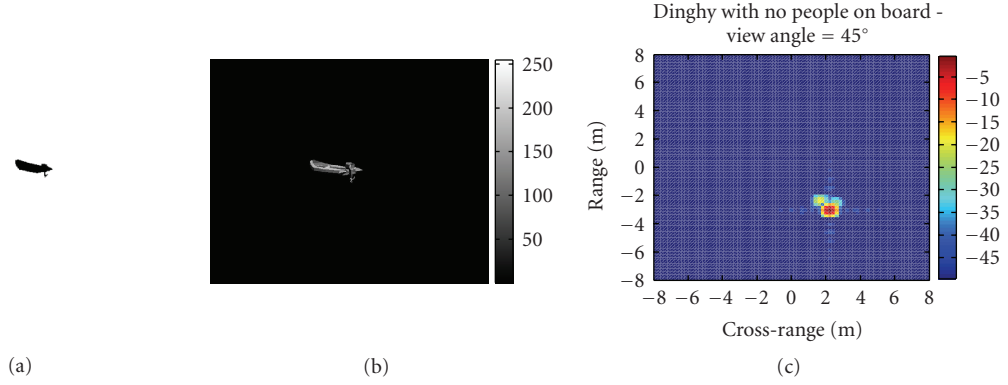


FIGURE 6: Example of simulated images for the dinghy, view angle 45°: (a) video camera; (b) IR camera (by courtesy of F. Lapiere); (c) spotlight SAR (by courtesy of U. D'Elia, M.G. Del Gaudio, F. Prodi).

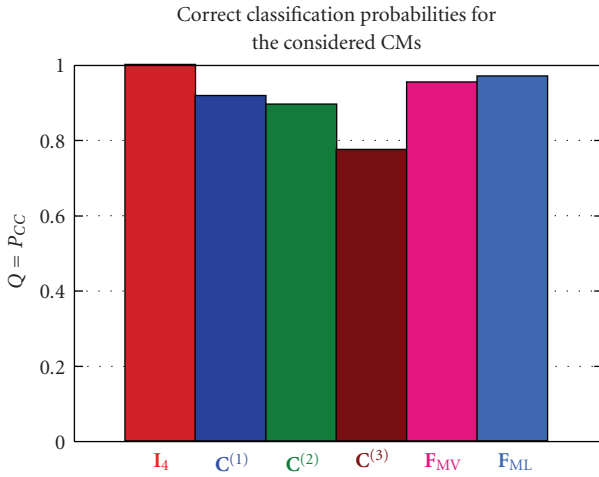


FIGURE 7: Probability of correct classification equal to the factor Q .

that is, when the error probability conditioned to the hypothesis H_k is minimum [23]. Equation (A.6) can be written as follows:

$$\sum_{i=1}^M p(\mathbf{I} | H_i) \cdot P(H_i) - p(\mathbf{I} | H_k) \cdot P(H_k) < \sum_{i=1}^M p(\mathbf{I} | H_i) \cdot P(H_i) - p(\mathbf{I} | H_m) \cdot P(H_m) \quad (\text{A.7})$$

so that the image \mathbf{I} belongs to the decision zone Y_k when

$$p(\mathbf{I} | H_m) \cdot P(H_m) < p(\mathbf{I} | H_k) \cdot P(H_k) \quad \forall m \neq k. \quad (\text{A.8})$$

The set of inequalities in (A.8) defines the boundaries between the decision zones.

Since the reference images have finite energy over their range of definition, that is, the sensor field of view, the set of the reference images $\{\mathbf{I}_1, \dots, \mathbf{I}_M\}$ can be seen as a vector space where we can define a scalar product, by using the correlation function between a generic pair of elements inside the set. Using the scalar product, we can describe the elements of

this vector space by means of their coordinates with respect to an orthonormal base, constructed by a Gram-Schmidt orthonormalization. This representation for the reference images can be used to express the error probability given a certain hypothesis, $P_{\text{ERR}|H_k}$.

In the vector space each image \mathbf{I}_i , for $i = 1, \dots, M$, is represented by an M -dimensional vector \mathbf{y}_i , and the energy difference between two images represented by the vectors \mathbf{y}_1 and \mathbf{y}_2 is

$$\begin{aligned} \Delta E = E_1 - E_2 &= \sum_{k=1}^M (y_{1,k}^2 - y_{2,k}^2) \\ &= \sum_{k=1}^M (y_{1,k} + y_{2,k}) \cdot (y_{1,k} - y_{2,k}) \\ &= (\mathbf{y}_1 + \mathbf{y}_2)^T \cdot (\mathbf{y}_1 - \mathbf{y}_2), \end{aligned} \quad (\text{A.9})$$

where $y_{i,k}$ represents the k th component of the vector \mathbf{y}_i .

Let us indicate with \mathbf{n} the vector representing the noise, assumed to be zero-mean, additive, Gaussian, and independent of pixel to pixel with variance σ^2 , and with ζ the elements of the orthonormal Gram-Schmidt base. The statistics of the noise vector \mathbf{n} are related to the sensor signal-to-noise ratio and they do not depend on the class of the image under test.

Assuming that the noise is Gaussian and that the M hypotheses have the same a priori probabilities, the decision criterion of (A.8) can be expressed as

$$\mathbf{y}^T \cdot \mathbf{y}_k - \left(\frac{1}{2}\right) \cdot (\mathbf{y}_k^T \cdot \mathbf{y}_k) \stackrel{k}{>} \mathbf{y}^T \cdot \mathbf{y}_j - \left(\frac{1}{2}\right) \cdot (\mathbf{y}_j^T \cdot \mathbf{y}_j), \quad (\text{A.10})$$

where \mathbf{y} is the vector representing the received image in the vector space. Equation (A.10) can be written as:

$$\left[\mathbf{y}^T - \frac{1}{2} (\mathbf{y}_k^T + \mathbf{y}_j^T) \right] \cdot (\mathbf{y}_k - \mathbf{y}_j) \stackrel{k}{>} 0 \quad (\text{A.11})$$

since the scalar quantities $\mathbf{y}_k^T \cdot \mathbf{y}_j$ and $\mathbf{y}_j^T \cdot \mathbf{y}_k$ are equal. Then the error event can be characterized as follows:

(i) the inequality

$$\left[\mathbf{y}^T - \frac{1}{2}(\mathbf{y}_k^T + \mathbf{y}_j^T) \right] \cdot (\mathbf{y}_k - \mathbf{y}_j) < 0 \quad (\text{A.12})$$

represents the elemental error event in the classification of an image belonging to the k th class;

(ii) by defining the normalized vector

$$\boldsymbol{\nu}_{jk} = \frac{(\mathbf{y}_k - \mathbf{y}_j)}{\sqrt{E_{jk}}}, \quad (\text{A.13})$$

where $E_{jk} = \sum_{i=1}^M (y_{k,i} - y_{j,i})^2$ is the cross energy between an image belonging to the k th class and another image belonging to the j th class, the conditional error probability can be expressed as

$$P_{\text{ERR}|H_k} = \Pr \left\{ \bigcup_{j=1, j \neq k}^M \left[\left(\mathbf{n} + \frac{1}{2}(\mathbf{y}_k - \mathbf{y}_j) \right)^T \cdot \boldsymbol{\nu}_{jk} < 0 \right] \right\}; \quad (\text{A.14})$$

(iii) Considering that

$$\frac{1}{2}(\mathbf{y}_k - \mathbf{y}_j)^T \cdot \boldsymbol{\nu}_{jk} = \frac{1}{2}\sqrt{E_{jk}} \quad (\text{A.15})$$

the conditional error probability becomes

$$P_{\text{ERR}|H_k} = \Pr \left\{ \bigcup_{j=1, j \neq k}^M \left[\mathbf{n}^T \cdot \boldsymbol{\nu}_{jk} < -\frac{1}{2}\sqrt{E_{jk}} \right] \right\} \quad (\text{A.16})$$

and it can be divided in the following contributions:

$$\begin{aligned} P_{\text{ERR}|H_k} &= \Pr \left\{ \bigcup_{j=1, j \neq k}^M \left(\mathbf{n}^T \cdot \boldsymbol{\nu}_{jk} < -\frac{1}{2}\sqrt{E_{jk}} \right) \right\} \\ &= \Pr \left\{ \bigcup_{j=1, j \neq k}^{M-1} \left(\mathbf{n}^T \cdot \boldsymbol{\nu}_{jk} < -\frac{1}{2}\sqrt{E_{jk}} \right) \right\} \\ &\quad + \Pr \left\{ \mathbf{n}^T \cdot \boldsymbol{\nu}_{Mk} < -\frac{1}{2}\sqrt{E_{Mk}} \right\} \\ &\quad - \Pr \left\{ \left[\bigcup_{j=1, j \neq k}^{M-1} \left(\mathbf{n}^T \cdot \boldsymbol{\nu}_{jk} < -\frac{1}{2}\sqrt{E_{jk}} \right) \right] \right. \\ &\quad \left. \cap \left[\mathbf{n}^T \cdot \boldsymbol{\nu}_{Mk} < -\frac{1}{2}\sqrt{E_{Mk}} \right] \right\} \end{aligned} \quad (\text{A.17})$$

such that the conditional error probability $P_{\text{ERR}|H_k}$ can be computed by considering two elements in the union and by reiterating $M-3$ times the operation expressed in (A.17).

More details about this analytical computation can be found in [18].

Nomenclature

AEWR:	Airborne Early Warning Radar
AIS:	Automatic Identification System
ATC:	Automatic Target Classification
C ² :	Command and Control
CAD:	Computer-Aided Design
CM:	Confusion Matrix
EM:	Electromagnetic
EO:	Electro-Optical
HS:	Homeland Security
NEIL:	Noise Equivalent Luminance
NETD:	Noise Equivalent Temperature Difference
IR:	Infrared
ML:	Maximum Likelihood
MV:	Majority Voting
SAR:	Synthetic Aperture Radar
TL:	Threat Level
VTS:	Vessel Traffic Service.

Acknowledgments

The authors would like to acknowledge Fabian D. Lapierre (Royal Military Academy, Brussels, Belgium) for kindly providing the temperature file exploited for the simulation of the infrared images; Ugo D'Elia and Maria Grazia Del Gaudio (MBDA, Rome, Italy), and Francesco Prodi (SELEX Sistemi Integrati, La Spezia, Italy) for kindly providing the simulated electromagnetic images exploited in this work; Paolo Marrucci (SELEX Galileo, Pomezia, Italy) for his help with the meteorological vectors for IR simulation.

References

- [1] G. L. Foresti, C. Micheloni, L. Snidaro, P. Remagnino, and T. Ellis, "Active video-based surveillance system," *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 25–37, 2005.
- [2] H.-M. Chen, S. Lee, R. M. Rao, M.-A. Slamani, and P. K. Varshney, "Imaging for concealed weapon detection: a tutorial overview of development in imaging sensors and processing," *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 52–61, 2005.
- [3] L. Snidaro, R. Niu, G. L. Foresti, and P. K. Varshney, "Quality-based fusion of multiple video sensors for video surveillance," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 37, no. 4, pp. 1044–1051, 2007.
- [4] M. N. Cohen, "An overview of radar-based, automatic, noncooperative target recognition techniques," in *Proceedings of the IEEE International Conference on Systems Engineering*, pp. 29–34, Fairborn, Ohio, USA, August 1991.
- [5] F. Gini, A. Farina, and M. S. Greco, "Radar detection and preclassification based on multiple hypothesis testing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 3, pp. 1046–1059, 2004.
- [6] C. J. Baker, H. D. Griffiths, and M. Vespe, "Multi-perspective imaging and image interpretation," in *Proceedings of the NATO ASI Lectures on Imaging for Detection and Identification*, July 2006.
- [7] M. Costantini, A. Farina, and F. Zirilli, "The fusion of different resolution SAR images," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 139–146, 1997.

- [8] Y. Bar-Shalom, T. Kirubarajan, and C. Gokberk, "Tracking with classification-aided multiframe data association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 3, pp. 868–878, 2005.
- [9] A. Farina, P. Lombardo, and M. Marsella, "Joint tracking and identification algorithms for multisensor data," *IEE Proceedings: Radar, Sonar and Navigation*, vol. 149, no. 6, pp. 271–280, 2002.
- [10] G. Davidson, S. J. Symons, and M. Everett, "Sensor fusion system for infrared and radar," in *Proceedings of the IEE Seminar on Signal Processing Solutions for Homeland Security*, October 2005.
- [11] S. J. Symons, R. Philpott, D. Manson, and M. Everett, "MFR-IRST integration in the naval domain," in *Proceedings of the IEE Colloquium on the Use of Radars with Others Sensing Types*, pp. 4/1–4/6, October 2000.
- [12] A. Howard and H. Seraji, "Multi-sensor terrain classification for safe spacecraft landing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 4, pp. 1122–1131, 2004.
- [13] S. Giompapa, R. Croci, R. Di Stefano, et al., "Naval target classification by fusion of IR and EO sensors," in *Electro-Optical and Infrared Systems: Technology and Applications IV*, vol. 6737 of *Proceedings of SPIE*, Florence, France, September 2007.
- [14] S. Giompapa, A. Farina, F. Gini, A. Graziano, R. Croci, and R. Di Stefano, "Naval target classification based on the confusion matrix," in *Proceedings of the IEEE Aerospace Conference*, Big Sky, Mont, USA, March 2008.
- [15] S. Giompapa, A. Farina, F. Gini, A. Graziano, R. Croci, and R. Di Stefano, "Study of the classification task into an integrated multisensor system for maritime border control," in *Proceedings of the IEEE Radar Conference*, Rome, Italy, May 2008.
- [16] S. M. Kay, *Fundamental of Statistical Signal Processing, Volume 2: Detection Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 1998.
- [17] S. Giompapa, A. Farina, F. Gini, A. Graziano, and R. Di Stefano, "Computer simulation of an integrated multi-sensor system for maritime border control," in *Proceedings of the IEEE Radar Conference*, pp. 308–313, Boston, Mass, USA, April 2007.
- [18] S. Giompapa, *Analysis, modelling, and simulation of an integrated multisensor system for maritime border control*, Ph.D. dissertation, University of Pisa, Pisa, Italy, 2008.
- [19] F. D. Lapierre, R. Dumont, A. Borghgraef, J.-P. Marcel, and M. Acheroy, "OSMOSIS: an open-source software for modelling and simulation of infrared signatures using general emissivity profiles," in *Proceedings of the 3rd International IR Target and Background Modelling and Simulation Workshop (ITBMS '07)*, Toulouse, France, June 2007.
- [20] I. I. Kim, B. McArthur, and E. Korevaar, "Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications," in *Optical Wireless Communications III*, vol. 4214 of *Proceedings of SPIE*, pp. 26–37, Boston, Mass, USA, November 2001.
- [21] W. G. Carrara, R. S. Goodman, and R. M. Majewski, *Spotlight SAR Signal Processing Algorithms*, Artech House, Boston, Mass, USA, 1995.
- [22] G. Franceschetti and R. Lanari, *Synthetic Aperture Radar Processing*, CRC Press, Boca Raton, Fla, USA, 1999.
- [23] J. L. Melsa and D. L. Cohn, *Decision and Estimation Theory*, McGraw-Hill, New York, NY, USA, 1978.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

