Feature-based target recognition with a Bayesian network

Jun Liu
Kuo-Chu Chang
George Mason University
School of Information Technology and Engineering
Center of Excellence in Command, Control, Communications, and Intelligence (C^3I)
Fairfax, Virginia 22030
E-mail: jliu@c3i.gmu.edu

Abstract. The problem of target classification with high-resolution, fully polarimetric, synthetic aperture radar (SAR) imagery is considered. We propose a framework of using a Bayesian network for feature fusion to deal with the difficult problem of SAR target classification. One difficult problem in SAR feature identification and fusion for target classification is that the features identified may not be independent and that it is not easy to find the ''right'' fusion rule to combine them. The Bayesian network model when constructed properly can explicitly represent the conditional independence and dependence between various features and therefore provide a sound and natural framework for feature fusion. This paper summarizes our recent work in SAR target recognition using a feature-based Bayesian inference approach. The approach works on the selected features which are chosen so that the separability of the original data are well maintained for later classification. Once the original data are mapped into feature space, the probabilistic model between features and the target is estimated and represented by a Bayesian network, which is then used to calculate the probabilities that a target belongs to one of the given classes based on the observed features. A comparison between the above technique and the traditional statistical approaches such as nearest mean and Fisher pairwise is illustrated based upon performance on a fully polarimetric ISAR (inverse SAR) image data set. Note that although the feature set used in the paper is obtained from the same sensor, the concepts of feature selection and Bayesian network formulation discussed in the paper are not restricted to this case only. They can be applied for multisensor feature-level fusion as well.

Subject terms: sensor fusion; feature level fusion; Bayesian network; synthetic aperture radar (SAR) imagery; target recognition.

Paper SF-002 received July 3, 1995; revised manuscript received Sep. 6, 1995; accepted for publication Oct. 3, 1995.

1 Introduction

The objective of an automatic target recognition (ATR) system is to detect and recognize targets from sensor data. One of the important components of an ATR system is its classifier. The function of the classifier is to categorize input measurements that represent detected targets according to target type. The classifier output corresponding to each input is an estimate of correct category label, based on the observable characteristics of the input.

In general, a feature-based classifier consists of two major parts, a feature selection and a classification mechanism. For the purpose of target classification, the features selected do not necessarily have physical meaning. The only goal in designing features is to preserve class discriminant information of the data while ignoring information that is irrelevant to the discrimination task. Once a feature is identified, it will define a transformation to map input measurements into feature space, which usually has a much lower dimension than that of the input space. This will greatly simplify the classification problem.

A number of attributes that are present in ISAR images can be exploited to discriminate between targets and clutter false alarms. They are size, shape, signal strength, polarimetric properties, spatial distribution of reflected signal, and so on. However, only a few of them can be used to discriminate among classes of targets. It is very difficult to develop discrimination features for exploiting these attributes in any optimal fashion. Furthermore, the features identified may not be independent and it is not easy to find the ''right'' fusion rule. In fact, past experimental results[1] showed that adding features does not necessarily improve performance if they are not handled correctly. In this study, in the first stage, we examined up to twelve features against the data; some were studied before,[1] and others are new. Out of the twelve features, we then selected the best discrimination features for classification. The main idea in this stage is to select the most useful information or processed results, while ignoring the irrelevant or bad ones. Although the feature set used in the paper is computed from the data of the same sensor, the concepts of feature selection and decision making discussed here are not restricted to this case only.

Research was supported by ARPA under Contract No. N66001-93-C-6010.

In the second stage, a Bayesian network is used for classification. A Bayesian network is a directed, acyclic graph in which the nodes represent random variables, and the arcs between the nodes represent probabilistic dependence between the variables. The Bayesian network model when constructed properly can explicitly represent the conditional independence and dependence between various features and therefore provide a sound and natural framework for feature fusion. Much attention has been drawn to this technology in the past few years and it has been successfully applied both to tasks of assessment under uncertainty and tasks of decision-making under uncertainty.[2–4] Recently it has also been applied to multisource intelligence fusion.[5] In this paper, for feature-level fusion, we first identify the network topology for various target parameters such as class and orientation and sensor data features. Instead of working directly on the input measurements, the transformed data on the chosen feature spaces are used as input. With the selected topology, we then learn (estimate) the probabilistic relationship between various variables in the Bayesian network. The conditional probability that a target belongs to a class given observed features is then computed based on a probabilistic inference algorithm using the network. Finally, the target is assigned to the class with the highest probability. Note that the idea proposed here is general and can be applied to multisensor domain directly. In fact, the idea of applying a Bayesian network to multisensor fusion has become more and more popular in the fusion community.[6]

This paper is organized as follows: Section 2 describes target sets and radar image data, Sec. 3 introduces the network algorithm and the classification process, Sec. 4 presents the performance results, and finally, Sec. 5 contains some concluding remarks.

2 Data Description

The database used for this study was obtained from MIT's Lincoln Laboratory; it consists of four representative targets: a Dodge van, a Chevrolet Camaro, a Dodge pickup truck, and an International Harvester bulldozer. Each of the four targets was put on a platform, and the millimeter-wave radar image data collected using a 35-GHz normal frequency at a fixed 5.5-deg depression angle while the platform turned over a complete 360-deg azimuth. These inverse synthetic aperture radar (ISAR) images are 1-ft. range-processed with full polarizations, namely, horizontal transmit, horizontal receive (HH); horizontal transmit, vertical receive (HV); and vertical transmit, vertical receive (VV). The images of these vehicles are available at 0.04-deg azimuth intervals and each is associated with a viewing angle. The original image of the vehicles has the size of 32×20 pixels. Figure 1 shows an example of four target images using single-channel HH at a 0.4-deg azimuth.

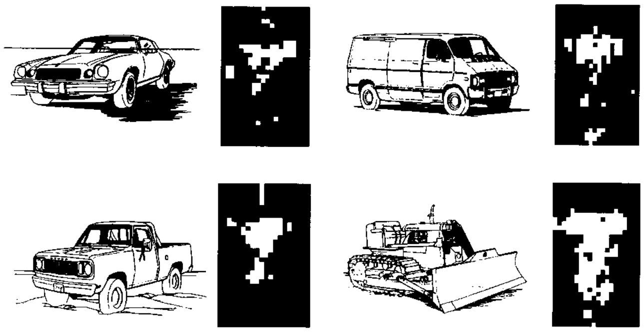

Fig. 1 ISAR images of an HH channel for four targets at 0.4 azimuth with a 1-ft.×1-ft. resolution.

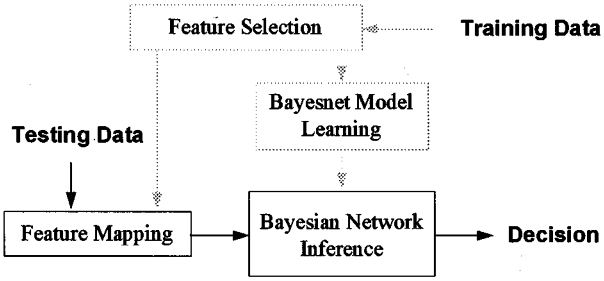

Fig. 2 Feature-based classifier with a Bayesian network.

The fully polarimetric ISAR data were first filtered by a polarimetric whitening filter (PWF),[7] and then normalized and compressed by a window slicing technique to a 15×9 dimension.[8] In this study, 5280 processed images were picked up from the database for each target. It should be mentioned that the targets look confusing from different angles. In particular, an image for a target at one angle may look like the image of another target at the same or a different angle, while some images from adjacent angles for the same target may look much different. This made the classification task very difficult.

3 Feature-Based Classification Procedure

Our feature-based classifier is composed of the following stages. As shown in Fig. 2, the observed measurement is first transformed into feature space based on preselected features. Then the transformed data are input into the Bayesian network for probabilistic inference. The result is a set of estimated conditional probabilities that the observed target is from one of the classes given the observed features. Finally, the decision-making procedure simply compares these estimated conditional probabilities, and the observed target is assigned to the class with the highest conditional probability. Note that the feature selection and Bayesian network model learning modules are based on the training data and are done a priori off-line.

The most difficult part to build into this classifier is feature selection. Once features are chosen, the second step is to learn the probabilistic models between features and the targets and represent the model by a Bayesian network. For simplicity, we have assumed a simple two-level network topology where the observed features are assumed to be conditionally independent given the target class and image azimuth angle[*] (see Fig. 3). Given the network topology, the first task is to estimate the conditional probabilities of the observed features given the target parameters. These

---

[*] Note that the observed features are not conditionally independent given only the target class.

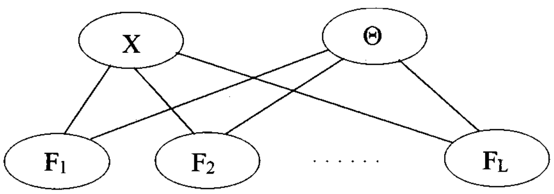

Fig. 3 A probabilistic model between features and target states.

estimated conditional probability distributions are used to describe the variable relationships in the Bayesian network. In the rest of this section, we will describe feature selection, Bayesian network modeling and decision making in detail.

3.1 *Feature Selection*

As mentioned before, the useful features here are those that preserve class discriminant information on the data while ignoring information that is irrelevant to the discrimination task. Since, at present, no method exists for developing discrimination features in any optimal fashion, one way is to test all the proposed features on the data set to see how well they can separate the targets of different classes. The good features, which have a better ability to separate different target classes, are maintained for the further use of classification, and the poor ones are discarded.

In this study, a total of twelve features were examined, of which three (standard deviation, fractal dimension, and weighted-rank fill ratio) were developed by MIT's Lincoln Laboratory to discriminate targets from nature-clutter false alarms in their ATR system,[9] and the rest are new features. Some features are contrast-based, and the other spatial-distribution related. They are

(i) The standard deviation (SD) feature is a measure of the fluctuation in intensity in an image. It is computed from the typical estimator for the standard deviation by using the power (expressed in decibels) of all the pixels in an image. If the radar image in power is denoted by $P(r,a)$, then the log standard deviation σ can be estimated as:

$$\hat{\sigma}=\left(\frac{S_2-S_1^2/N}{N-1}\right)^{1/2}, \tag{1}$$

where

$$S_1=\sum_{r,a} 10\log_{10} P(r,a), \tag{2}$$

and

$$S_2=\sum_{r,a} [10\log_{10} P(r,a)]^2. \tag{3}$$

N is the total number of pixels in the image, which is $9\times 15=135$.

(ii) The *fractal dimension* (FD) feature provides a measure of the spatial distribution of the brightest scatterers in the images. To calculate the FD feature of an image, the first step is to convert the image to a binary one. To do so, the K brightest pixels in the image are selected and their values are converted to 1, while the rest of the pixel values are converted to 0. Then the fractal dimension can be estimated as:

$$\dim=-\frac{\log M_1-\log M_2}{\log 1-\log 2}=\frac{\log M_1-\log M_2}{\log 2}, \tag{4}$$

where M_1 is the number of 1-pixel-by-1-pixel boxes needed to cover the image, and M_2 is the number of 2-pixel-by-2-pixel boxes needed to cover the image. Obviously, $M_1=K$.

(iii) The *weighted-rank fill ratio* (WRFR) feature measures the percentage of the total energy contained in the brightest scatterers of an image. Using the notation of Eq. (2), this feature is defined as follows:

$$\eta=\frac{\sum_{\text{k brightest pixels}} P(r,a)}{\sum_{\text{all pixels}} P(r,a)}. \tag{5}$$

(iv) The *counting* (CNT) feature is obtained by counting the number of pixels in the image that exceed a specific threshold, then dividing by the total number of pixels of the image.

(v)–(xii) The following eight features are designed to measure the spatial distribution of the brightest pixels in the images. First we convert the image to binary by using *amplitude thresholding,* in which all pixel values exceeding a specified threshold are converted to 1, and the remaining pixel values are converted to 0. Assuming the converted binary image is denoted as $B(i,j)$, then these features are defined as:

$$M_X=\frac{1}{N}\sum_{i=1}^{9}\sum_{j=1}^{15} i\times B(i,j), \tag{6}$$

$$M_Y=\frac{1}{N}\sum_{i=1}^{9}\sum_{j=1}^{15} j\times B(i,j), \tag{7}$$

$$S_{XX}^2=\frac{1}{N-1}\sum_{i=1}^{9}\sum_{j=1}^{15} (i-M_X)^2\times B(i,j), \tag{8}$$

$$S_{YY}^2=\frac{1}{N-1}\sum_{i=1}^{9}\sum_{j=1}^{15} (j-M_Y)^2\times B(i,j), \tag{9}$$

$$S_{XY}^2=\frac{1}{N-1}\sum_{i=1}^{9}\sum_{j=1}^{15} (i-M_X)(j-M_Y)\times B(i,j), \tag{10}$$

$$W_X=\max\{i:B(i,j)=1\}-\min\{i:B(i,j)=1\}, \tag{11}$$

$$W_Y=\max\{j:B(i,j)=1\}-\min\{j:B(i,j)=1\}, \tag{12}$$

$$W_{XY}=W_X\times W_Y, \tag{13}$$

where N is the total number of pixels in the image.

Features need to be evaluated since using similar features does not guarantee a better discrimination performance, and sometimes adding features can even degrade performance. One way to evaluate whether a feature is good for discriminating targets of different classes is to test it on the training data set in a heuristic manner. The other way is to use the optimal feature set selection approach,[10] which uses a Bhattacharrya upper bound[11] on the probability of classification error to determine which subset of any l features of the available L features has the lowest upper bound. It can be seen in Sec. 4 that this latter approach is an effective way to select features and its performance is very close to what we obtained using an ''optimal manually chosen'' feature set.

3.2 *Probabilistic Modeling of a Bayesian Network*

Our objective is to recognize targets from the observed data. Bayesian networks show great promise for performing this function since they can be used to represent complicated probabilistic relationships among variables of interest. Furthermore, many efficient algorithms have been developed for drawing inferences from the evidence.[12–15] For the current Bayesian network model, let the class status of an observed target be a random variable X, and assume the target belongs to one and only one of K classes; then, obviously, X is a discrete random variable, and without loss of generality, we can assume it takes on a value 1,2,...,K (in our case $K=4$). In a radar image, an important factor that greatly affects the appearance of the target is the target orientation. Let Θ denote the azimuth angle when a target is imaged, which can have values from 0 to 360 deg. If we discretize the 360-deg space into M small sectors and each has an equal interval, Θ can be treated as a discrete random variable. Finally, based on a set of selected features, denoted as random variables $F_1,F_2,\ldots,F_L$, discrete or continuous, a simple two-level probabilistic model between features and target states (class and azimuth) can be obtained as in Fig. 3.

As shown in Figure 3, each node represents a random variable, and each line indicates the conditional probabilistic relationship between the connected nodes. Note here that the network topology implicitly assumes that the observed features are conditionally independent given the target class and orientation. However, in general, this is not the case, and a more complicated network topology is needed to model the problem. In a Bayesian network, the conditional probability distribution of a child given all of its parents is assumed to be given before any probabilistic reasoning can be drawn. In our case, this is to say that given target type X and radar azimuth angle Θ, the feature F_l is distributed with the known distribution $P(F_l|X,\Theta)$, $l=1,2,\ldots,L$. In reality, the conditional distributions $P(F_l|X,\Theta)$ need to be elicited by expert knowledge or physical models or estimated with the training data. It should be mentioned that for some continuously distributed features, since their distributions are hardly close to any of the well-known parameterized probability distributions, they must be estimated with nonparametric methods. We will discuss this in more detail in the next section.

With the Bayesian network, the class probabilities of observed targets can be computed with any probabilistic inference algorithm.[12–15] Basically, the question becomes one of how to calculate the conditional probability that an observed target is from a class at an azimuth angle given the observed features, e.g., $P(X,\Theta|F_1,F_2,\ldots,F_L)$. For the current simplified model, since the features $F_1,F_2,\ldots,F_L$ are conditionally independent given X and Θ, it can be shown that the required conditional probability can be obtained as:[†]

$$P(X,\Theta|F_1,F_2,\ldots,F_L)=\frac{1}{C}\prod_{l=1}^{L}P(F_l|X,\Theta), \tag{14}$$

where C is a normalizing constant. If there are K classes of targets and M sectors of angles, based on Eq. (14), we can obtain a total of $K\times M$ probability estimations. These are then used to make a decision.

3.3 *Decision Making*

In this step, the observed target is assigned to an appropriate class. What we obtained from Eq. (14) is a set of $K\times M$ probabilities, e.g., how likely an observed target is from class k and at azimuth sector m, $k=1,2,\ldots,K$, and $m=1,2,\ldots,M$. To make a decision, there are two basic decision rules.

Decision rule 1. Among $K\times M$ estimations, find the one with the highest probability value, then assign the target to the class associated with this estimate. It can be seen that, at the time the target class is determined, so can the target azimuth angle. However, this may not be necessary, and hence we have the second decision rule.

Decision rule 2. Instead of estimating $P(X,\Theta|F_1,F_2,\ldots,F_L)$ using Eq. (14), we estimate $P(X|F_1,F_2,\ldots,F_L)$. This can be obtained by the following equation:

$$\begin{aligned}P(X|F_1,F_2,\ldots,F_L)&=\sum_{\Theta}P(X,\Theta|F_1,F_2,\ldots F_L)\\&=\frac{1}{C}\sum_{\Theta}\prod_{l=1}^{L}P(F_l|X,\Theta).\end{aligned} \tag{15}$$

From Eq. (15), we can obtain K probability estimates. Again, we choose the one with the highest value, then assign the target to the class associated with this estimate.

When making a probabilistic inference, an interesting consideration is to treat the radar azimuth angle as known. This may happen if the radar azimuth angle at which the target is imaged can be determined by other sources of information. If this is the case, namely the radar azimuth angle is known to be in the m'th sector, the conditional probability that the target is from a specific class given the observed features and $\Theta=m$ can be obtained using the following equation:

$$P(X|F_1,F_2,\ldots,F_L,\Theta=m)=\frac{1}{C}\prod_{l=1}^{L}P(F_l|X,\Theta=m), \tag{16}$$

where C is a normalization constant. Again, the decision making is based on K calculated probability estimates.

[†] Again, in general, the network is more complicated and an efficient algorithm such as SPI [16] can be used to compute the conditional probability distributions.

Table 1 Averaged correct classification rates (ACCR) in % using OLPARS.

Optimal feature set	Nearest mean classifier			Fisher pairwise classifier	
	weight	training	testing	training	testing
all twelve features	E	39.7	25.0	86.2	68.3
	var.	66.8	66.2	rej. 20	rej. 830
	cov.	80.2	64.2		
1,3,6,7,12	E	38.3	25.0	83.8	72.4
	var.	66.0	67.9	rej. 20	rej. 44
	cov.	81.5	76.0		
1,3,6,7	E	71.7	71.0	81.9	82.0
	var.	66.9	67.1	rej. 17	rej. 49
	cov.	80.4	80.2		
1,3,6	E	69.1	67.8	78.4	78.4
	var.	69.0	68.5	rej. 26	rej. 26
	cov.	76.7	76.9		
1,3	E	63.1	62.9	72.8	72.2
	var.	65.0	65.0	rej. 59	rej. 141
	cov.	72.4	72.1		

4 Performance Evaluation

The data set used for this study is first randomly and uniformly split into two data sets, the training data set (1728×4) and the testing data set (3552×4). Using the training data set, the conditional probability distributions $P(F_l|X,\Theta)$ are first estimated. They are then used to define the Bayesian network and later to calculate conditional probabilities $P(X,\Theta|F_1,F_2,\ldots,F_L)$ for classification. The testing data set is input to the system to examine the classification performance, and the results are reported in terms of the averaged correct classification rate (ACCR), which is defined as the ratio of the number of correctly classified observations to the total number of observations in the testing data set.

4.1 *Estimation of Conditional Probability Distributions*

In this approach, the conditional probability distributions are estimated by a smoothing kernel approach, which is the most thoroughly developed approach in literature. Assuming we have K classes of targets ($K=4$ in our problem), L features, and the angle space is decomposed into M sectors, there will be a total of $K\times L\times M$ distribution functions to be estimated. For the l'th feature F_l, k'th class, and m'th sector of angles, $P(F_l|X=k,\Theta=m)$ is estimated by using only that image data from the k'th class of target and m'th sector of angles. The data are first transformed based on the feature F_l. If we consider $F_l^1,F_l^2,\ldots,F_l^n$ as n-transformed observations based on the feature F_l, and they are from the k'th class and m'th sector of angles, the kernel estimate has the form

$$\hat{P}(F_l|X,\Theta)=\frac{1}{n}\sum_{i=1}^{n}k(F_l-F_l^i), \tag{17}$$

where $k(.)$ is a probability density function symmetric about the origin. Here we use a Gaussian density with variance σ^2, which is the only parameter. The parameter should be chosen so that the ACCR is maximized.

4.2 *Classification Results*

In this section, for the purpose of comparison, we first present some test results by using traditional classifiers—nearest mean and Fisher pairwise. We then show some test results to illustrate the key issues discussed in previous sections. The first part of the results (Table 1) is obtained by means of a software package called OLPARS (On-Line Pattern Analysis and Recognition System).[10] The system provides users with a convenient tool to realize a variety of traditional pattern recognition methods, especially optimal feature set selection, as mentioned in the previous section. The traditional classifiers such as nearest mean, Fisher pairwise and so on can also be designed and evaluated easily.

In Table 1, the left-most column refers to the optimal feature set selected in terms of the Bhattacharrya upper bound. The number of features in the feature set should be given in order to do the feature selection. For example, if we decide to use only two features, OLPARS reports that the best feature set is {1,3}, i.e., the first feature SD and the third feature WRFR as defined in Sec. 3.1. Corresponding to each of the optimal feature sets, Table 1 presents the classification results using the nearest mean and Fisher pairwise classifiers for both the training and testing data sets (note: the classifier parameters are estimated by the training data set only). In the table, "E," "var," and "cov." represent the different distance measures, Euclidean, weighted by a diagonal variance matrix, and weighted by a covariance matrix, respectively, and "rej." refers to rejection. As can be seen from the table, the feature set {1,3,6,7} has the overall best performance.

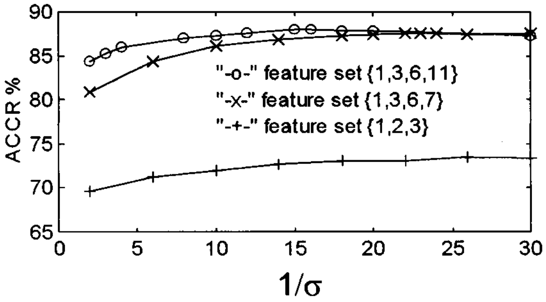

Fig. 4 ACCR vs. $1/\sigma$ for {1,3,6,11} {1,3,6,7}, and {1,2,3}, $M=48$, and using decision rule 2.

The second part of the results is obtained using feature-based classification with a Bayesian network. Figure 4 shows a comparison of classification results among the optimal feature sets, i.e., {1,3,6,7} selected by OLPARS, the manually selected set {1,3,6,11} and the previously proposed feature set {1,2,3}. The best ACCRs are 87.55% for {1,3,6,7} and 88.00% for {1,3,6,11} respectively, and the latter feature set performs slightly better than the former one. The two feature sets selected by the two different approaches have only one differing feature. However, the optimal feature selection approach is more computationally efficient. In Fig. 4, it also can be seen that the overall performance of feature sets {1,3,6,7} and {1,3,6,11} is much better than that of the previously proposed set {1,2,3}.

Figure 5 gives a comparison between decision rules 1 and 2. We find the second rule is better than the first. However, using the first rule, the target azimuth angles can be predicted simultaneously.

Figure 6 displays the impact of the number M when the azimuth angle space is partitioned. It can be seen that when M is increased, performance improves as well.

The last result, shown in Fig. 7, is obtained using Eq. 16, assuming the radar azimuth angle can be determined from other information sources. In this model, performance is about 5% more accurate than that of decision rules 1 and 2. The best ACCR rate is 93.3%. This result is not surprising because more information is assumed to be available in this model.

By comparing the two experimental results, it can be seen that the feature-based Bayesian net classifier performs noticeably better than the traditional classifiers—nearest mean and Fisher pairwise.

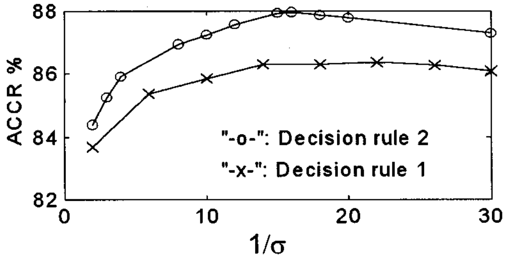

Fig. 5 ACCR vs. $1/\sigma$ for decision rule 1 and 2, $M=48$, and using {1,3,6,11}

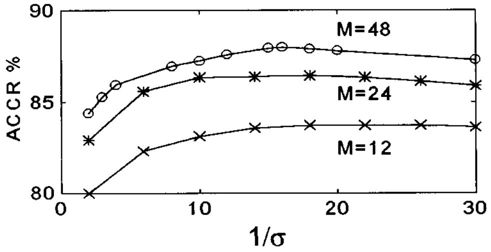

Fig. 6 The impact of M, using decision rule 2 and {1,3,6,11}.

5 Conclusions

In this paper, we applied the Bayesian network technology for feature-level fusion. Bayesian networks show great promise for multiattribute fusion since they can be used to represent complicated probabilistic relationships among variables of interest. We first identify the network topology for various target parameters such as class and orientation and sensor data features. The network topology explicitly represents the conditionally independent or dependent relationships among various features. Instead of working directly on the input measurements, the transformed data on the chosen feature spaces are used as input. With the selected topology, we then learn (estimate) the probabilistic relationship among various variables in the Bayesian network. The conditional probability that a target belongs to a class given observed features is then computed based on a probabilistic inference algorithm using the network. The network model used here is relatively simple. In a separate but related research,[16] we also studied the problem of Bayesian network construction using neural learning techniques where the network is more general and complicated.

In the current approach, performance depends on a number of factors, including selection of a set of workable features, choosing an appropriate probabilistic model to describe the observations, handling approximation of the required conditional probability distributions, and so on. When classifying the SAR image, not only does the Bayesian network model lead to a better performance than certain types of traditional classifiers, for example, nearest mean and Fisher pairwise, it also possesses a certain degree of flexibility to handle other target parameters such as orientation. The orientation can be predicted at the time the target is classified. In the case where the orientation is

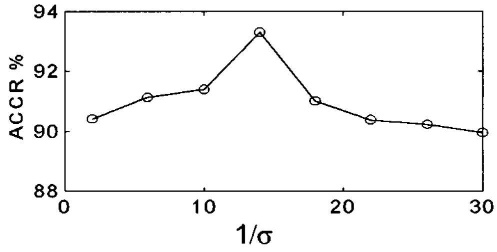

Fig. 7 ACCR vs. $1/\sigma$, assuming Θ is known, using {1,3,6,11}, and $M=48$.

known, the classification accuracy is improved significantly. The decomposition of azimuth angle space and the parameter of the conditional probability distribution estimates are two main factors that could be adjusted to improve performance.

Although the feature set used in the example is obtained from the same sensor, the concepts of feature selection and decision making discussed in the paper are not restricted to this case. The idea of applying a Bayesian network to feature-level fusion can also be applied to a multisensor domain directly. In fact, the idea of applying a Bayesian network to multisensor fusion has become more and more popular in the fusion community. The evaluation of our results demonstrates the usefulness of the proposed approach.

Acknowledgment

This research was supported by ARPA under Contract No. N66001-93-C-6010.

References

1. L. M. Novak, G. J. Owirka, and C. M. Netishen, ‘‘Performance of a high-resolution polarimetric SAR automatic target recognition system,’’ *Lincoln Laboratory J.* **6**(1), 11–24 (1993).
2. R. A. Howard and J. E. Matheson, ‘‘Influence diagrams,’’ in *The Principles and Applications of Decision Analysis*, Vol. II, R. A. Howard and J. E. Matheson, Eds., Menlo Park: Strategic Decisions Group, Menlo Park, CA (1981).
3. S. L. Lauritzen and D. J. Spiegelhalter, ‘‘Local computations with probabilities on graphical structures and their application in expert systems,’’ *J. Roy. Stat. Soc. B* **50** (1988).
4. R. D. Shachter, ‘‘Intelligent probabilistic inference,’’ in *Uncertainty in Artificial Intelligence*, L. N. Kanal and J. F. Lemmer, Eds., North-Holland, Amsterdam (1986).
5. K. C. Chang, ‘‘Multiple intelligence correlation and fusion with Bayesian networks,’’ *The Seventh Joint Service Data Fusion Symposium* (1994).
6. K. C. Chang and Clayton Stewart, ‘‘Application of Bayes nets in sensor fusion,’’ SPIE Euro OPTO, Oct. (1993).
7. L. Novak and C. Netishen, ‘‘Polarimetric synthetic aperture radar imaging,’’ *Int. J. Imaging System Tech.* **4**, 306–318 (1992).
8. K. C. Chang and Y. C. Lu, ‘‘High resolution polarimetric SAR target classification with vector quantization,’’ to appear in *Proc. FUZZ-IEEE/IFCS ’95*, Japan.
9. M. C. Burl, G. J. Owirka, and L. M. Novak, ‘‘Texture discrimination in synthetic aperture radar imagery,’’ *Proc. IEEE 23rd Asilomar Conf. on Circuits, Systems, and Computers*, p. 399 (1989).
10. *OLPARS User’s Manual, Version 7.3,* PAR Government System, Inc., 1010 Prospect St., Suite 200, La Jolla, CA (1995).
11. P. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis,* Wiley, New York (1973).
12. R. Shachter, A. Del Favero, and B. D’Ambrosio, ‘‘Symbolic probabilistic inference: A probabilistic perspective,’’ *Proc. AAAI* (1990).
13. M. Henrion, ‘‘Propagating uncertainty by logic sampling in Bayes’ networks,’’ Technical Report, Department of Engineering and Public Policy, Carnegie-Mellon University, Pittsburgh, PA (1986).
14. J. H. Kim and J. Pearl, ‘‘A computational model for combined causal and diagnostic reasoning in inference systems,’’ *Proc. 8th Internat. Joint Conf. Artificial Intelligence* (1985).
15. K. C. Chang and Robert Fung, ‘‘Symbolic probabilistic inference with both discrete and continuous variables,’’ *IEEE Trans. SMC* (1995).
16. K. C. Chang and Sulin Yang, ‘‘Bayesian network construction with neural learning,’’ submitted for publication.
17. D. E. Kreithen, S. D. Halversen, and G. J. Owirka, ‘‘Discriminating targets from clutter,’’ *Lincoln Laboratory J.* **6**(1), 25–51 (1993).

Jun Liu received his BS in applied mathematics from the Beijing Polytechnic University in 1981, and an MS in remote sensing and cartography from the Graduate College of Chinese Science and Technology University, Beijing, in 1985. From 1985 to 1992 he was an assistant researcher at the Institute of Remote Sensing Application, Chinese Academy of Sciences, working in the areas of information system and field spectral analysis for remote sensing, image processing, and application and evaluation of inertial navigation systems in aerial photography. Since 1993 he has been a graduate research assistant with the Center of Command, Control, Communications and Intelligence at George Mason University, Fairfax, VA where he is a PhD candidate. His current research interests include image processing, image feature detection, and target recognition and classification as well as Bayesian networks.

Kuo-Chu Chang received a BS in communication engineering from the National Chiao-Tung University, Taiwan, in 1979 and MS and PhD degrees in electrical engineering from the University of Connecticut in 1983 and 1986, respectively. From 1983 to 1992, he was a senior research scientist in the Advanced Decision Systems (ADS) division, Booz-Allen & Hamilton, Mountain View, California. He joined the Systems Engineering Department of George Mason University in 1992 as an associate professor. His research interests include estimation theory, optimization, signal processing, and data fusion. He is particularly interested in applying unconventional techniques to conventional decision and control systems. He has published more than 50 papers in the areas of multitarget tracking, distributed sensor fusion, and Bayesian probabilistic inference. He is currently an editor on navigation/tracking systems for *IEEE Transactions on Aerospace and Electronic Systems.* Dr. Chang is a member of Etta Kappa Nu and Tau Beta Pi.