Target Classification Using PAS and Evidence Theory

Jia Li
School of Information System and Management
National University of Defense Technology
Changsha 410073, China
jialiasus@sina.com

Li Guohui
School of Information System and Management
National University of Defense Technology
Changsha 410073, China
guohli@nudt.edu.cn

Yan Hai
School of Information System and Management
National University of Defense Technology
Changsha 410073, China
yvrw@sina.com

Zhang Hui
School of Information System and Management
National University of Defense Technology
Changsha 410073, China
zhanghui_nudt09@yahoo.com.cn

Abstract—**This paper presents a novel Dempster-Shafer evidence construction approach for aircraft aim recognition. The prior-probability of the properties of aircraft was used for establishing a probabilistic argumentation system. Dempster-Shafer evidence was constructed by assumption- based reasoning. Therefore, additional information could be provided to the classification of the data fusion system. Experiments on artificial and real data demonstrated that the proposed method could improve the classification results.**

Keywords; Dempster-Shafer theory, probabilistic argumentation system (PAS), target classification

I. INTRODUCTION

Data fusion is a very important approach for aircraft target recognition because of the imprecision of single sensor. The radar sensor can track multiple targets at one time, but with large intersection between two measurements and low recognition ability. The IR image sensor can recognize the aim property, but with short sphere of action and sensitivity with weather and other environmental factors. A system with multi-sensors can make use of the complementary and redundancy of different sensors to improve the effects of the recognition and tracking. Fusion of the different data improves the recognition precision and achieves more credible results.

Dempster-Shafer theory[1] is one of the most important data fusion theories[2-4]. It can adapt most situations in real application, and can be integrated with imprecision reasoning theories. It improves the expression of void belief assignment. In the use of Dempster-Shafer theory, the construction of the evidence from different sensors is an important step. In traditional data fusion systems, image properties are usually used for target classification. The numerical properties such as velocity and acceleration are ignored. In fact, the distributions of the numerical properties are also important information for target classification.

The PAS theory[5-7,13] is a reasoning system that combines both qualitative approach and quantitative approach. In this method, we used PAS with numerical properties for target classification. By estimating the priority probabilities of the properties, the knowledge base for reasoning was established. According to the special form of knowledge base in this method, the approach for calculating the support degree of a hypothesis was presented. Because the result measurement is in the form of belief function, our method can be integrated in a data fusion system by using Dempster-Shafer theory, or directly do the classification in the use of radar or other measure sensors.

The rest of the paper is organized as follows. Section 2 describes the fundamental theories that used in our method. Section 3 presents our method for knowledge base construction and evidence reasoning. Section 4 presents the experiments on artificial and real data. Section 5 concludes the paper.

II. FUNDAMENTAL THEORIES

A. Evidence Theory

The Dempster-Shafer theory of evidence, also referred to as evidence theory, is a theoretical for reasoning with uncertain and imprecision information and data.

Let $\Omega = \{\varpi_1, \dots, \varpi_T\}$ be a finite set of exclusive and exhaustive classes called the frame of discernment. A basic belief assignment (BBA) or mass function is a function $m: 2^{\Omega} \rightarrow [0,1]$. It satisfies the two following conditions:

$$m(\emptyset) = 0 \quad (1)$$

$$\sum_{A\subseteq\Omega} m(A) = 1 \quad (2)$$

where $\emptyset$ means empty set. Such a BBA that satisfies $m(\phi) = 0$ is called normal[8-10].

Let m_1 and m_2 induced by two independent items of evidence. Then the Dempster's rule of combination is

$$m(C) = \frac{\sum_{A\cap B=C} m_1(A)\cdot m_2(B)}{1-\sum_{A\cap B=\emptyset} m_1(A)\cdot m_2(B)} \quad (3)$$

After combining all pieces of evidence, a decision has to be made using the final belief function. A method called

Published by Atlantis Press, Paris, France.

Pignistic Transformation[11,12] is used for decision-making. from the BBA, a pignistic probability distribution is defined as

$$BetP(\omega) = \sum_{A\subseteq\Omega,\omega\in A} \frac{1}{|A|} \cdot \frac{m(A)}{1-m(\emptyset)} \quad (4)$$

B. Probabilistic Argumentation System

The Probabilistic Argumentation System is an alternative approach for non-monotonic reasoning under uncertainty. It allows to judge open questions (hypothesis) about the unknown in the use of the given knowledge.

A probabilistic argumentation system usually to be defined as a quadruple $PAS_P = (\xi, P, A, \Pi)$ [6], where ξ denote the knowledge base, P denote the question variables, A denote the assumptions and Π denote the probability distribution of assumptions.

An assignment of truth values to the elements of set A is called interpretation. Towards set variables to corresponding truth values according to an interpretation, a sentence's value can be evaluated. An interpretation x is called a model of a sentence γ, if γ evaluates to 1 according to proportional logic.

At last, a scenario s is called a quasi-support scenario for h relative to ξ, if h is a logical consequence of ξ when the scenario s happened[7]. The sets of all quasi-support scenarios for h relative to ξ is denoted by $QS_A(h,\xi)$.

Let $s = (x_1, \dots, x_m)$ be a scenario, and then the prior probability of s is determined by

$$p(s) = \prod_{i=1}^{m} p(x_i) \quad (5)$$

Therefore, if S is an arbitrary set of scenarios, then the probability of S is simply the sum of the probabilities of its elements

$$p(S) = \sum_{s\in S} p(s) \quad (6)$$

If $h \in \mathcal{L}_{A\cup P}$ is a hypothesis that we want to judge, then $dqs(h,\xi) = p(QS_A(h,\xi))$ is called degree of quasi-support of h relative to ξ. This measure corresponds to unnormalized belief in the Dempster-Shafer theory of evidence.

III. EVIDENCE CONSTRUCTION

In this section we present an approach to construct the evidences for target recognition.

A. Knowledge Base Construction

Suppose there are T different types of targets, and K properties that can be measured. Let C be the entire set of target types

$$C = \{C_1, C_1, \dots, C_T\} \quad (7)$$

And let c denotes the correct unknown class of the target. Then obviously there must have $c \in C$.

For a given property f (such as velocity, acceleration, etc.) of aircraft aim, and its type $c = C_j$, then its probability density function (PDF) f_p can be denoted as

$$P(f < x | c = C_j) = \int_{-\infty}^{x} f_p(t) d_t \quad (8)$$

We can estimate the empirical distribution $\widehat{f_p}$ as the approximation of f_p from the history data.

Suppose that property f belongs to a region R and R is divided into several intervals $I_i (i = 1,2, \dots, n)$ which form a partition

$$I_i \cap I_j = \emptyset \text{ for } i \neq j,\ \cup_{i=1}^{n} I_i = R \quad (9)$$

To make things easy, the event that f belongs to an interval I_i is denoted as a random variable set to a certainty value, which is expressed as

$$(d_f = I_i) \Leftrightarrow (f \in I_i) \quad (10)$$

For a certain interval I_i, the probability that f belongs to it can be calculated with

$$P(d_f = I_i | c = C_j) = \int_{I_i} f_p(t) d_t \approx \int_{I_i} \widehat{f_p}(t) d_t \quad (11)$$

That means if we know the target type c, then we can get the probability that its value belongs to each interval. But this rule seems difficult to be used in PAS.

To get a rule for classifier reasoning in the form of PAS, we can denote T multi-valued random variables $f_j (j = 1,2, \dots, T)$, corresponding to T different types of targets. Each variable has the value range which is $F_j = \{f_{j1}, f_{j2}, \dots, f_{jn}\}$. The distribution of each random variable is set to

$$P(f_j = f_{ji}) = P(d_f = I_i | c = C_j) \quad (12)$$

which corresponds to the probability that the property f belongs to a segment. Then f_j can be seen as the factors (such as environment, mission requirement, etc.) that influence on the target, to make the property f in such certainty value range. Therefore, a rule that used to form knowledge base can be derived as

$$c = C_j \rightarrow (f_j = f_{ji} \rightarrow (d_f = I_i)) \quad (13)$$

This rule is obvious, since that when the target type is confirmed, the PDF of property f is certainly confirmed. Then, the approximate probability that $f \in I_i$ can be derived by the empirical distribution $\widehat{f_p}$, which means when some affections whose probability equal to the probability of $f \in I_i$ happen, the value range of f can be determined.

The rule (13) can also be written as:

$$(c \neq C_j) \vee (f_j \neq f_{ji}) \vee (d_f = I_i) \quad (14)$$

Let K_{fji} denotes rule (14). This rule is for property f, target type C_j and value interval I_i. In general, for T different types of targets, K properties and n intervals of each property, there will be total $T \cdot K \cdot n$ rules. Then knowledge base can be constructed by the conjunction of all such rules.

B. PAS Reasoing

In PAS, two sets of proportions must be distinguished, hypothesis and scenarios. In our case, Knowledge base is the conjunction of all the rules of K_{fji}. The target type is the issue which wants be figured out. So the variable *c* is the hypothesis. The priority probabilities of the properties of the certain target type can be estimated with historical data, and the distributions of the target properties can be measured by the sensors such as radar. Thus, the variables d_f and f_j build up the scenarios. Then the inference principles of probabilistic assumption based reasoning can be applied to process the target classification. By definition in section 2, for a hypothesis h to the target type, the quasi-support of h is the set of assumption scenarios that allow us to prove h or make knowledge base contradictory. Let $QS(h)$ denotes the quasi-support of h and $QS(\perp)$ denotes the quasi-support of the contradiction, i.e. the set of scenarios that make the knowledge base contradictory. Then the degree of $QS(h)$ is denoted as

$$qs(h) = P(QS(h)) \tag{15}$$

and the degree of contradiction is denoted as

$$qs(\perp) = P(QS(\perp)) \tag{16}$$

where *P* is the joint probability measure over the quasi-support, i.e. the sum of probabilities of the scenarios in the quasi-support. Then the degree of the support of the hypothesis h by the knowledge base can be derived as the quantity

$$sp(h) = \frac{qs(h)-qs(\perp)}{1-qs(\perp)} \tag{17}$$

In other words, the probability of the hypothesis h is estimated from the knowledge base and the probability density of the scenario variables.

Since the rules in the knowledge base have a special form same to formula (14), then we can achieve $sp(h)$ in a special way. In the enumeration of all the scenarios, we can firstly calculate the probability of the scenario, and then get its support hypothesis. In such way, the difficulty of searching the quasi-support is avoided. The calculation of the probability of the scenario is simple, because it is just the multiplication of every variable's probability. To determine the support hypothesis, the proportion that knowledge base proves must be found.

For a given scenario, every variable in it (such as d_f and f_j) has the fixed value. Because the knowledge base has a conjunctive normal form, the result must be the conjunction of every clause. A clause is true when any literal in it is true. That means when ($f_j = f_{ji}$) and ($d_f \neq I_i$), there will be

$$\begin{gathered}(c \neq C_j) \vee (f_j \neq f_{ji}) \vee \left(d_f = I_i\right) \\ \Rightarrow (c \neq C_j) \rightarrow c \in C/C_j\end{gathered} \tag{18}$$

because the disjunction will remove the negative literals in a formula.

Therefore, the knowledge base in such a scenario will prove a result from the clauses that satisfy ($f_j = f_{ji}$) and ($d_f \neq I_i$). Get the intersection of such clauses in equation (14), then the set that contains the intersection will be supported by this scenario. In other words, the quasi-support to such hypothesis includes the scenario, then the probability of the scenario can be added to $qs(h)$. After the enumeration of all scenarios, the function *sp* is determined.

According to [1], the function *sp* is also a belief function in the sense of Dempster-Shafer theory. Therefore, such result can be used in data fusion, or directly used in target classification by pignistic decision.

IV. EXPERIMENTS

To prove the effects of our method, experiments on artificial and real data were carried out.

A. Artificial Data

In artificial data experiments, we set 3 types of aircraft with 2 main properties f and g, which denote velocity and acceleration respectively. Each property was divided into 6 intervals. The priority probabilities of the properties are shown in the tables below.

TABLE I. VELOCITY SETTING IN ARTIFICIAL DATA

class	I_1	I_2	I_3	I_4	I_5	I_6
C_1	0	0.2	0.4	0.2	0.1	0.1
C_2	0.2	0.4	0.1	0.1	0.1	0.1
C_3	0	0.1	0.2	0.4	0.2	0.1

TABLE II. ACCELERATION SETTING IN ARTIFICIAL DATA

class	K_1	K_2	K_3	K_4	K_5	K_6
C_1	0	0.1	0.7	0.1	0	0.1
C_2	0.1	0.2	0.3	0.2	0.1	0.1
C_3	0.2	0.0	0.3	0.0	0.4	0.1

Then we randomly created 90 artificial targets of three classes (each class has 30 samples). For each target, we created *K* random measure points according to the priority probabilities of the properties. Then the approximate property distributions were estimated. As a result, the recognition rate come from the pignistic transformation is given in the following table according to different values of *K*.

TABLE III. RECOGNITION RATE

K	10	30	90
rate	85.56%	95.56%	97.78%

The reason that recognition rate is low when *K* is small is that the data measured on each target is rarely little and the estimate of the target property distributions are imprecise.

The result at K=30 is acceptable, and can make contribution to the data fusion system.

To analyze the performance of our method, the confusion matrix at K=30 is given by the following table.

TABLE IV. CONFUSION MATRIX ($K = 30$)

	C_1	C_2	C_3
C_1	30	0	0
C_2	3	27	0
C_3	1	0	29

It can be seen that all the mistakes are judged to belong to C_1. It seems that the knowledge base prefers to C_1. Similar situation happened when K was set to other values, such as 10 or 90.

An extreme result can be derived when set $K = 1$, which means there is only one measurement on the target. In such case the confusion matrix is shown as below.

TABLE V. CONFUSION MATRIX ($K = 1$)

	C_1	C_2	C_3
C_1	23	5	2
C_2	9	20	1
C_3	14	1	15

There are 46 times judged to C_1 in 90, and only 18 times judged to C_3. This result shows that the knowledge base has its tendency.

B. Real Data

In the real data experiments, there are three types of target that is UAV, delta winged aircraft and airship which are denoted as C_1, C_2 and C_3 respectively. For each aircraft, a tracking path of 2 hours long was observed by radar. Then we cut each path into 100 pieces which are called measure points and confuse the sequence. Half of the measure points were used to estimate the empirical distribution of the properties. By cutting each property's value range into 6 intervals, the knowledge base is constructed as below.

TABLE VI. VELOCITY DISTRIBUTION IN REAL DATA

class	I_1	I_2	I_3	I_4	I_5	I_6
range	[0,10]	[10,20]	[20,25]	[25,30]	[30,35]	[35,40]
C_1	0	0.0007	0.1087	0.3739	0.4499	0.0668
C_2	0.0011	0.45	0.1529	0.1647	0.0987	0.1327
C_3	0.5481	0.4519	0	0	0	0

TABLE VII. ACCELERATION DISTRIBUTION IN REAL DATA

class	K_1	K_2	K_3	K_4	K_5	K_6
range	[-10,-6]	[-6,-3]	[-3,0]	[0,3]	[3,6]	[6,10]
C_1	0.0007	0.0029	0.4806	0.5124	0.0026	0.0008
C_2	0	0.0133	0.4653	0.5118	0.0090	0.0005
C_3	0	0.0021	0.5043	0.4902	0.0030	0.0005

The other 50 measure points were used to test our method, and the recognition rate achieves 98.67%, the confusion matrix is shown in table (VIII).

TABLE VIII. CONFUSION MATRIX IN REAL DATA

	C_1	C_2	C_3
C_1	50	0	0
C_2	2	48	0
C_3	0	0	50

V. CONCLUSION

In this paper, a method for target classification in use of numerical properties was presented. By estimating the priority probability distributions of each class, knowledge base was built by our method, then the use of PAS to the results suit for Dempster-Shafer theory. Because of the special form of knowledge base in the problem of classification, an approach for reasoning is established and makes the processing simple. It also can be used for classification directly. Experiments on artificial and real data show the performance of our method.

REFERENCES

[1] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1967.

[2] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster–Shafer theory", IEEE Transactions on Systems, Man and Cybernetics , vol. 25, 1995, pp. 804–813.

[3] T. Denoeux, "A neural network classifier based on Dempster–Shafer theory", IEEE Transactions on Systems, Man and Cybernetics A, vol. 30, 2000, pp. 131–150.

[4] T. Denoeux, L.M. Zouhal, "Handling possibilistic labels in pattern classification using evidential reasoning", Fuzzy Sets and Systems, vol. 122, 2001, pp. 47–62.

[5] R. Haenni, J. Kohlas, N. Lehmann, Probabilistic argumentation systems, in: J. Kohlas, S. Moral (Eds.), Algorithms for Uncertainty and Defeasible Reasoning, Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 5, Kluwer, Dordrecht, 2000.

[6] J. Kohlas, R. Haenni, S. Moral, "Propositional information systems", Journal of Logic and Computation, vol. 9, 1999, pp. 651–681.

[7] R. Haenni, "Probabilistic argumentation", Journal of Applied Logic, vol. 7, 2009, pp. 155–176.

[8] P. Smets, R. Kennes, "The transferable belief model", Art. Intell, vol. 66, 1994, pp. 191–234.

[9] P. Smets, Practical uses of belief functions, in: K.B. Laskey, H. Prade (Eds.), Uncertainty in Artificial Intelligence, vol. 15, Stockholm, Sweden, 1999.

[10] P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem", International Journal of Approximate Reasoning, vol. 9, 1993, pp. 1–35.

[11] B.R. Cobb, P.P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models", Int. J. Approx.Reason, vol. 41, 2006, pp. 314–330.

[12] P. Smets, "Decision making in the TBM: the necessity of the pignistic transformation", Int. J. Approx. Reason, vol. 38, 2005, pp. 133–147.

[13] J. Kohlas, P.A. Monney, "A Mathematical Theory of Hints", Lecture Notes in Economics and Mathematical Systems, vol. 425, Springer, Berlin, 1995.