A New Method to Determine BPA in Evidence Theory

Wen Jiang[*]
School of Electronics and Information, Northwestern Polytechnical University,
Xi'an, 710072, China
Email: jiangwen@nwpu.edu.cn

Yong Deng[1,2,3,4]
1.Hangzhou Key Lab of E-Business and Information Security, Hangzhou Normal University,
Hanzhou, Zhejiang, 310036, China
2.Key Laboratory of Intelligent Computing & Information Processing of Ministry of Education,
Xiangtan University, Xiangtan, Hunan, 411105, China
3. School of Computer and Information Sciences, Southwest University,
Chongqing, 400715, China
4.School of Electronics and Information Technology, Shanghai Jiao Tong University,
Shanghai, 200240, China
Email: dengyong@sjtu.edu.cn

Jinye Peng
School of Electronics and Information, Northwestern Polytechnical University,
Xi'an, 710072, China
Email: jinyepeng@nwpu.edu.cn

Abstract—**Dempster Shafer theory of evidence has been widely used in many data fusion application systems. However, how to determine basic probability assignment, which is the main and the first step in evidence theory, is still an open issue. In this paper, based on the distance measure between the sample data under test and the model of attribute of species, a new method to obtain BPA is proposed. A numerical example is used to illustrate the efficiency of the proposed method.**

Index Terms—**Sensor fusion, Dempster Shafer theory of evidence, basic probability assignment, normal distribution, distance measure**

I. INTRODUCTION

Data fusion is a technique that enables us to simultaneously take into account heterogeneous data coming from different sources to get an optimal estimation of objects under investigation. Whatever the application, the goal of data fusion is to reduce uncertainty and imprecision by combining both redundant and complementary data. A few mathematical formalisms (probability, fuzzy logic, possibilities, evidence theory, etc.) are available to perform a measure of the uncertainty and the imprecision. Among these,

evidence theory, also called Dempster-Shafer (DS) evidence theory, is a powerful and flexible mathematical tool for handling uncertain, imprecise, and incomplete information. That is so for at least for three reasons. First, by representing the uncertainty and the imprecision of a body of knowledge via the notion of evidence, belief can be committed to a single hypothesis or a composite. Second, the evidence combination rule of the DS theory provides an interesting operator to integrate multiple information from different sources. Finally, the decision on the optimal hypothesis choice can be made in a flexible and rational manner [1-4].

In real data fusion application systems based on DS theory, the basic probability assignment function should be given so that the combined BPA can be obtained through Dempster's rule of combination. However, how to determine basic probability assignment, which is the main and the first step in evidence theory, is still an open issue. Many authors have addressed this problem using different approaches. Yager associated the DS belief structure with a whole class of fuzzy measures, and discussed the entropy of a fuzzy measure [5]. Zhu et.al. proposed a method to derive mass values from fuzzy membership degrees. To do this, fuzzy c-means (FCM) clustering is used to represent the grey levels as fuzzy sets [6]. Bloch used the distance to cluster centers [7]. In this work mass functions are estimated in unsupervised way and using fuzzy membership functions to take into account the ambiguity within pixels. This ambiguity is due to inherent vagueness rather than randomness. Even

if these approaches do not assume any particular probability distribution for the grey level values, they suffer from the fact that mass values are derived in a more or less heuristic manner. The mass functions may also be derived from probabilities. Le Hegarat-Mascle et al. [4] and Salzenstein et al. [8] modeled the knowledge provided by each information source using probability density functions (pdfs). A subtractive scheme is then used to convert these pdfs into confidence levels. Indeed, the mass value associated with each compound hypothesis is obtained by subtracting the mass values of the involved individual hypotheses. Bendjebbour et al. [9] proposed a probabilistic model where the frame of discernment contains the individual clusters and the ignorance that is modeled by the union of all individual clusters. In this work, the authors derived the mass value of ignorance from the mixture of distributions of the individual clusters composing it. We use random sets to obtain the BPA in target recognition [10], which to some extend to be seen as generalized fuzzy Bayesian method.

In this paper, a new method to obtain BPA is proposed. The new method uses the existing sample data to build a normal distribution described model of attribute categories. Based on the distance measure between the sample data under test and the model of attribute, BPA is obtained. The paper is organized as follows: In section 2, the preliminaries of evidence theory are briefly introduced. In section 3, a new method to determine BPA is proposed. An experiment of Iris data classification is used to illustrate the efficiency of our method in section 4. Conclusions are made in Section 5.

II. DEMPSTER SHAFER EVIDENCE THEORY

Evidence theory first supposes the definition of a set of hypotheses Θ called the frame of discernment, defined as follows:

$$\Theta = \{H_1, H_2, \ldots, H_N\}. \tag{1}$$

It is composed of N exhaustive and exclusive hypotheses. Form the frame of discernment Θ, let us denote $P(\Theta)$, the power set composed with the 2^N propostions A of Θ:

$$P(\Theta) = \{\varnothing, \{H_1\}, \{H_2\}, \ldots, \{H_N\}, \\ \{H_1 \cup H_2\}, \{H_1 \cup H_3\} \ldots, \Theta\} \tag{2}$$

where $\varnothing$ denotes the empty set. The N subsets containing only one element are called singletons. A key point of evidence theory is the basic probability assignment (BPA). A BPA is a function from $P(\Theta)$ to $[0,1]$ defined by :

$$m: \quad P(\Theta) \to [0,1] \\ \quad A \mapsto m(A) \tag{3}$$

and which satisfies the following conditions:

$$\sum_{A \in P(\Theta)} m(A) = 1; m(\varnothing) = 0. \tag{4}$$

Dempster's rule of combination (also called orthogonal sum), noted by $m = m_1 \oplus m_2$, is the first one within the framework of evidence theory which can combine two BPAs m_1 and m_2 to yield a new BPA :

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - k}. \tag{5}$$

With

$$k = \sum_{B \cap C = \varnothing} m_1(B) m_2(C). \tag{6}$$

where k is a normalization constant, called conflict because it measures the degree of conflict between m_1 and m_2, $k = 0$ corresponds to the absence of conflict between m_1 and m_2, whereas $k = 1$ implies complete contradiction between m_1 and m_2. The belief function resulting from the combination of J information sources S_J defined as

$$m = m_1 \oplus m_2 \cdots \oplus m_j \cdots\cdots \oplus m_J. \tag{7}$$

As can be seen from above, multi source information can be easily fused in the framework of evidence theory, if we can obtain the BPA functions. We present our method in next section.

III. PROPOSED METHOD TO DETERMINE BPA

In fact, some sample are existed in many system, which often approximatively submit the normal distribution. So we use the existing sample data to build a normal distribution described model of attribute categories, then generate the function of BPA based on the relationship between the test data and the normal distribution model. In order to be understood easily, the following Iris data classification problem shows the detailed approach of the proposed method.

In the Iris data, there are 3 species of Iris flower, i.e., *Setosa*, *Versicolor*, and *Virginica* [11]**.** The Iris data contains 150 instances, and each species contains 50 instances. There are four attributes, i.e., Sepal Length (SL), Sepal Width(SW), Petal Length(PL), and Petal Width(PW) in the Iris data.

We randomly choose 40 instances from Setosa and the mean μ and the standard deviation σ of the Sepal Length (SL) of Setosa are obtained as follows: $\mu = 5.0375$, $\sigma = 0.3621$. Hence, we can construct the normal distribution described model of SL attribute of Setosa in Fig. 1.

In the same way, we randomly choose 40 instances from Versicolor and Virginica. the mean μ and the standard deviation σ of the Sepal Length (SL) of each species can be obtained, shown in Table 1. The normal distribution described model of SL attribute of each species, shown in Fig. 2, can be constructed based on the data of Table I.
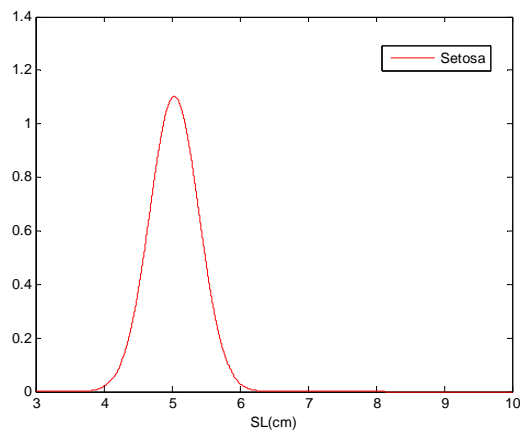
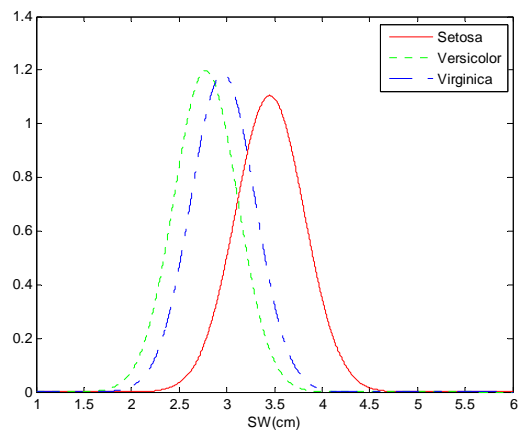Figure 1.   The normal distribution described model of SL attribute of Setosa

TABLE I.

The mean and the standard deviation of SL attribute of each species

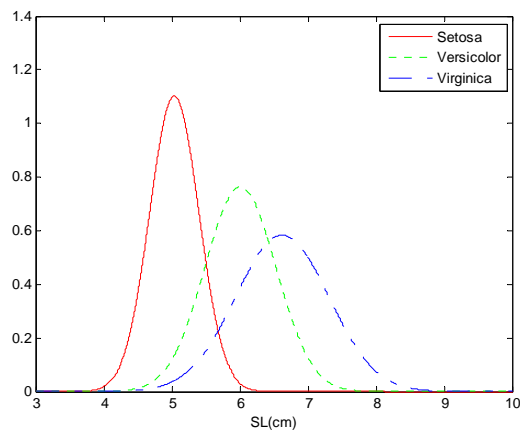| Setosa(S)	Versicolor(C)	Virginica(V)
$\mu = 5.0375$	$\mu = 6.0100$	$\mu = 6.6225$
$\sigma = 0.3621$	$\sigma = 0.5232$	$\sigma = 0.6841$

Figure 2.   The normal distribution described model of SL attribute of each species

In a similar way, the normal distribution described models of Sepal Width(SW) attribute, Petal Length(PL) attribute, and Petal Width(PW) attribute of each species can be constructed, shown in Fig. 3, Fig. 4, Fig. 5.

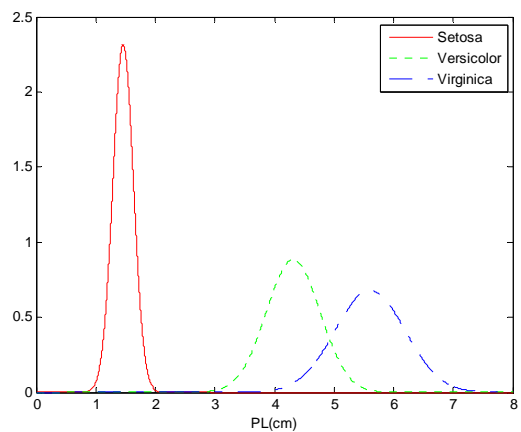Figure 3.   The normal distribution described model of SW attribute of each species

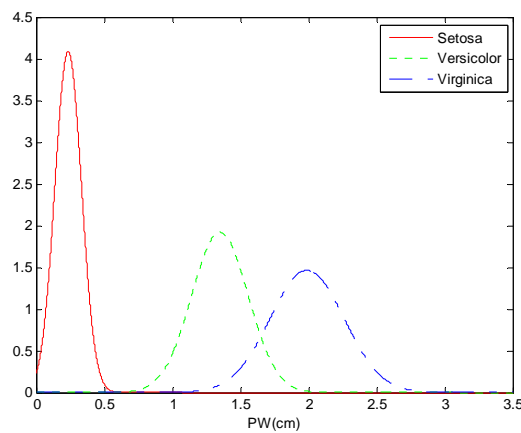Figure 4.   The normal distribution described model of PL attribute of each species

Figure 5.   The normal distribution described model of PW attribute of each species

We randomly choose a datum from Iris source, for example, a new instance (NI) can be shown as (5.0cm, 3.5cm, 1.3cm, 0.3cm). In Fig. 6, the relation between SL attribute of NI and the normal distribution described model of SL attribute of each species are distinctly showed.
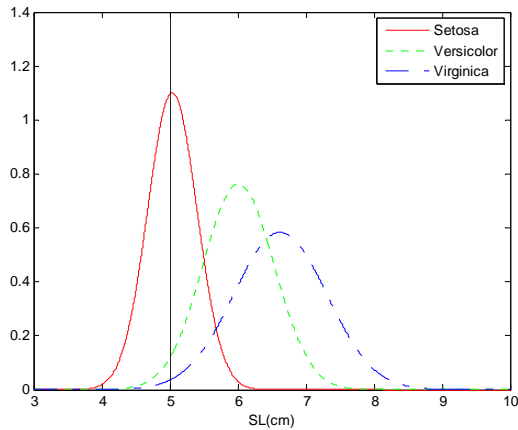

Figure 6.   The relation between the new instance and the model of SL attribute of each species

Now a distance measure [12] between the sample data x_i and the normal distribution described model of attribute category A is introduced as follows:

$$d(A) = 2\int_{\mu}^{x_i} p[x \mid \mu]dx . \tag{8}$$

Where $p[x \mid \mu]$ is probability density function (PDF) of the normal distribution described model of attribute category A, defined as follows:

$$p[x \mid \mu] = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2}[\frac{x-\mu}{\sigma}]^2\} . \tag{9}$$

Where μ is the mean of the normal distribution described model of attribute category A, and σ is the corresponding standard deviation.

Furthermore, error function $erf[x]$ [12] is defined as

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt . \tag{10}$$

Therefore, the value of the distance measure $d(A)$ can be easily calculated by error function $erf[x]$ [12], we have

$$d(A) = erf[\frac{x_i - \mu}{\sqrt{2}\sigma}] . \tag{11}$$

The value of the distance measure $d(A)$ can express the deviation degree between the sample data x_i and the normal distribution described model of attribute category A. So we can easily conclude that the more is the value of the distance measure, the smaller is the probability

which x_i belongs to category A. Furthermore, we define $p(A) = 1 - d(A)$ to express the probability which x_i belong to category A.

First, we apply Eq. (11) to calculate the value of the distance measure between the SL attribute of NI and the normal distribution described model of SL attribute of each species，shown as follows.

$$d(\text{Setosa}, \text{S}) = 0.0825$$

$$d(\text{Versicolor}, \text{C}) = 0.9464$$

$$d(\text{Virginica}, \text{V}) = 0.9823$$

Accordingly, the probability p can be calculated as follows.

$$p(S) = 1 - d(\text{S}) = 0.9175$$

$$p(C) = 1 - d(\text{C}) = 0.0536$$

$$p(V) = 1 - d(\text{V}) = 0.0177$$

We define

$$p(S, C, V) = 1 - \max\{p(S), p(C), p(V)\}$$
$$= 1 - 0.9175 = 0.0825$$

Finally, normalized the obtained probability p, the BPA can be obtained as follows.

TABLE II.

The BPA of an instance attribute SL

	m(S)	m(C)	m(V)	m(S,V,C)
SL	0.8565	0.0500	0.0165	0.0770

In the same way, the BPA of an instance attribute SW, PL, PW can be calculated, shown in Table III.

TABLE III.

The BPA of an instance

	m(S)	m(C)	m(V)	m(S,V,C)
SL	0.8565	0.0500	0.0165	0.0770
SW	0.7859	0.0269	0.0953	0.0919
PL	0.3527	0.0000	0.0000	0.6473
PW	0.5051	0.0000	0.0000	0.4949

Using the combination rule in Eq. (5), the fusion results can be shown as follows.

$$m(S) = 0.9901 \quad m(C) = 0.0031$$
$$m(V) = 0.0041 \quad m(S,C,V) = 0.0027$$

Hence, the instance can be classed as Setosa. The result is consistent with the actual situation.

The algorithm of our proposed method can be listed step by step as follows.

Step 1: Use the existing sample data to build the normal distribution, which describe the model attributes of instances.

Step 2: Calculate the distance measure between the collected attribute of the sample data and the model attribute.

Step 3: Calculate the probability which the sample data belongs to the model category.

Step 4: Normalized the probability measure to obtain the BPA function.

IV. EXPERIMENTS

We randomly select 120 instances, each 40 instances for 3 species, to construct species model. The remaining 30 instances, each 10 instances for 3 species is used as collected instances whose class is not known. By applying the method proposed in above section to classify the 30 instances, the results is showed in Table IV.

TABLE IV.

The result of simulation

The number of experiments	The number of Right results	The number of Wrong results	Correct rate
30	28	2	93.33%

The results show that the BPA are generated by the proposed method, the correct rate of Iris data classification is up to 93.33%. It can been seen that our proposed method has good results in classification problem.

V. CONCLUSIONS

In real data fusion application systems based on DS theory, how to determine basic probability assignment is still an open issue. In this paper, a new method to obtain BPA is proposed based on the distance measure between the sample data under test and the model of attribute of species. The proposed method, which effectively overcome the problem of subjectivity, have strong Generality. A numerical example is used to illustrate the efficiency and the low computational complexity of the proposed method.

ACKNOWLEDGMENT

REFERENCES

[1] D. L. Hall, *Mathematical techniques in multisensor data fusion.* Artech House, MA, 1992.

[2] A.Dempster, "Upper and lower probabilities induced by multivalued mapping," *Ann. Math. Stat.*, vol. 38, pp. 325-339, 1967.

[3] G. Shafer, *A Mathematical theory of evidence*, Princeton University Press, 1976.

[4] S. Le Hegarat-Mascle, I. Bloch, and D. Vidal-Madjar, "Application of DS evidence theory to unsupervised classification in multiple remote sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1018-1031, 1997.

[5] R. R. Yager, "A class of fuzzy measures generated from a Dempster-Shafer belief structure," *Int. J. Intel. Syst.* vol. 14, pp. 1239-1247, 1999.

[6] Y. M. Zhu, L. Bentabet, O. Dupuis, V. Kaftandjian, D. Babot, and M. Rombaut, "Automatic determination of mass functions in DS theory using FCM and spatial neighbourhood information for image segmentation," *Opt. Eng.* vol. 41, pp. 760-770, 2002.

[7] I. Bloch, "Some aspect of DS evidence theory for classification of multimodality medical images taking partial volume effect into account," *Pat. Recog.* vol. 17, pp. 905-916, 1996.

[8] F. Salzenstein, and A. O. Boudraa, "Unsupervised multisensor data fusion approach," *IEEE Proc. ISSPA, KualaLumpur*, vol. 1, pp. 152-155, 2001.

[9] A. Bendjebbour, Y. Delignon, L. Fouque, V. Samson, and W. Pieczynski, "Multisensor image segmentation using DS fusion in Markov fields context," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 1789-1798, 2001.

[10] Y. Deng, Z. F. Zhu, and S. Zhong, "Fuzzy Information Fusion Based on Evidence Theory and Its Appl ication in Target Recognition," *ACTA AERONAU TICA ET ASTRONAU TICA SINICA*, vol. 26, no. 6, pp. 754-758, 2005.

[11] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179-188, 1936.

[12] K. Fukunaga, T. E. Flick, "A test of the Gaussian-ness of a data set using clustering," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 2, pp. 240-247, 1986.

Wen Jiang was born in Hunan, P. R. China, in 1974. She received the Bachelor degree in signal and system from Information Engineering University, Zhengzhou, China, in 1994, the Master degree in image processing from Information Engineering University, Zhengzhou, China, in 1997, the Ph. D. degree in systems engineering from Northwestern Polytechnical University, Xi'an, China, in 2009.

She is currently an associate professor in school of electronics & information, Northwestern Polytechnical University. Her research interests are in the areas of information fusion and intelligent information processing.

Yong Deng was born in Hunan, P. R. China, in 1975. He received the Bachelor degree in physics from Shanxi Normal University, Xi'an, China, in 1997, the Master degree in instrument engineering from Hunan University, Changsha, China, in 2000, the Ph. D. degree in precise instrument and mechanic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2003.

He is currently a professor in the School of Computer and Information Sciences, Southwest University. His research interests are in the areas of information fusion, fuzzy logic and decision analysis.

Jinye Peng received his MS degree in Radio Electronics from Northwest University, Xi'an, China in 1996 and his PhD degree in signal and information processing from Northwestern Polytechnical University, Xi'an, China, in 2002. He was nominated as one of the "New Century Excellent Talents" and sponsored by the "Program for New Century Excellent Talents in University" by the Ministry of Education in 2007. He has published two books and more than 80 scientific papers in journals and proceedings, including the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, ACM TRANSACTIONS ON MULTIMEDIA COMPUTING COMMUNICATIONS AND APPLICATIONS, etc. He has become a professor since 2003.

He is currently a professor at the School of Electronics and Information, Northwestern Polytechnical University, and also acts as director of the department of electronic science and technology. His research interests include image/video analysis and retrieval, quantum information processing.