

# Multisensor Data Fusion Based on Modified D-S Evidence Theory

Yingming Zhou, Hongji Xu\*, Junfeng Sun, Lingling Pan, Baozhen Du and Min Chen

School of Information Science and Engineering, Shandong University, Jinan, China

\*Corresponding author

**Abstract**—Dempster-Shafer (D-S) evidence theory has been widely used in multisensor data fusion to deal with uncertain information. But unreasonable results may be produced by using D-S combination rule in the case of that data are conflicting with each other. This paper proposes a modified evidence combination method based on information gain and fuzzy preference relations. This method takes account of both historical data and real-time data by introducing the concepts of historical support and real-time support, so it can obtain more accurate results by using more effective information. In order to evaluate the performance of the proposed evidence combination method, an example of classifying the patient's state by five vital signs is given in this paper. The simulation experiment shows that the proposed modified method achieves higher classification accuracy compared with other three data fusion methods.

**Keywords**—Dempster-Shafer (D-S) evidence theory; multisensor data fusion; information gain; fuzzy preference relations

## I. INTRODUCTION

Multisensor data fusion is a subset of data fusion that is used to combine data and information from multiple sources to obtain more accurate results. The main idea of data fusion is to make more accurate decisions in consideration of the imperfections (incomplete data, uncertainty, inconsistency, etc.) of the data from a single data source. In 1986, the joint directors of laboratories (JDL) of the US Defense Department proposed the term “data fusion” [1]. Since then data fusion has developed rapidly.

Data fusion algorithms play an important role in the efficiency and accuracy of the processed information. Some algorithms, such as fuzzy set theory [2], neural networks [3], Dempster-Shafer (D-S) evidence theory [4], support vector machines (SVM) [5], etc., are all important algorithms in data fusion field. Compared with other methods, D-S evidence theory proposes a mathematical framework to handle uncertainty problems, so it is widely used in data fusion field.

Unfortunately, there are still some insufficiencies in D-S theory that are mainly focused on the combination rule. Typically, the results may be inaccurate or unreasonable when highly conflict evidences are synthesized by D-S combination rule.

In response to the above problems, scholars have put forward a lot of improvement methods which can be divided into two kinds of categories. The methods in the first category

are to modify the D-S combination rule: Yager [6] proposed a conservative method assigning the conflict factors to the recognition framework. Li, et al. [7] and Pan, et al. [8] assigned conflict factor to all focal elements by a certain proportion, getting better results than that of Yager. The methods in the second category are to process the original evidence: Murphy, et al. [9] assigned equal weight to each evidence by an average way. Deng, et al. [10] and Liu, et al. [11] proposed a weighted average method according to the interrelationships of evidences. Most of the methods require the support or weight of evidence, but the calculation of support is based on the evidence itself without taking into account the impact of the source of evidence which can be quantified by historical data.

In this paper, a modified evidence combination method is proposed for multisensor data fusion which is based on information gain and fuzzy preference relations. It can help data fusion systems make more accurate decisions by rational use of both historical data and real-time data.

## II. MODIFIED EVIDENCE COMBINATION METHOD

In the practical application of data fusion, it usually contains vast amounts of historical data and real-time data. However, most of the improved algorithms for D-S evidence theory cannot combine the two kinds of data effectively to form a reasonable fusion result. In this section, a new evidence combination method based on information gain and fuzzy preference relations is proposed to solve the above problem.

Suppose the detection system has  $n$  sensors with  $m$  results and each decision result can be represented by a single element  $t_j$ . The recognition framework [4] is  $\Theta = \{t_1, t_2, \dots, t_m\}$ , and  $n$  basic belief assignments (BBAs) that are expressed as  $m_1, m_2, \dots, m_n$  can be obtained from  $n$  sensors. When the belief is only assigned to the single element  $t_j$ , the entropy of  $m_i$  is consistent with the Shannon entropy [12].

### A. Proposed Historical Support Based on Information Gain

Historical support proposed in this paper is calculated by historical data. Historical data includes not only the sensor data, but also the true verdicts, so it is obvious to see which sensor has higher accuracy. The historical support of evidences can be measured by the information gain. The greater the historical support is, the higher the credibility of the evidence is.

The information entropy  $H(d)$  of decision variable  $d$  can be expressed as

$$H(d) = - \sum_{t_j \in \Theta} P(d = t_j) \log_2(P(d = t_j)) \quad (1)$$

where  $P(d = t_j)$  is the probability of event  $d = t_j$ .

Then the conditional information entropy of  $m_i$  can be calculated by

$$\begin{cases} H(d | m_i = t_j) = - \sum_{t_j \in \Theta} P(d = t_j | m_i = t_j) \log_2(P(d = t_j | m_i = t_j)) \\ H(d | m_i) = \sum_{t_j \in \Theta} P(m_i = t_j) H(d | m_i = t_j) \end{cases} \quad (2)$$

The information gain of  $m_i$  can be expressed as

$$G(d, m_i) = H(d) - H(d | m_i). \quad (3)$$

Finally, normalize the information gain and get the historical support by

$$Sup_{his}(m_i) = \frac{G(d, m_i)}{\sum_{i=1}^n G(d, m_i)}. \quad (4)$$

#### B. Proposed Real-time Support Based on Fuzzy Preference Relations

The real-time support proposed in this paper is calculated by real-time data. The distance between two evidences can be used to show the degree of similarity. The farther the distance is, the lower the similarity is. The real-time support based on fuzzy preference relations is proposed to measure the credibility of the evidence.

The Jousselme's distance  $d_{ij}$  between evidences  $m_i$  and  $m_j$  can be calculated as in [13]. Then calculate the average distance between  $m_i$  and other evidences by

$$\bar{d}_i = \frac{\sum_{j=1, j \neq i}^n d_{ij}}{n-1}. \quad (5)$$

The logsig function ( $\log sig(x) = \frac{1}{1+e^{-kx}}$ ) is widely used in neural networks as the transfer function. The reference [14] used the function to calculate the fuzzy preference relations between samples to characterize the orderly structure between samples. We extend the application to characterize the support

of evidences. The raw real-time support of  $m_i$  can be calculated by

$$Sup_{raw}(m_i) = \frac{1}{1+e^{k\bar{d}_i}} \quad (6)$$

where  $k \in N_+$ . Different values of the parameter  $k$  reflect different quantitative criterion for the same distance  $\bar{d}_i$  between evidences, the appropriate value can be selected by experiments or expert opinions.

Then get the real-time support by normalizing the raw real-time support expressed as

$$Sup_{ret}(m_i) = \frac{Sup_{raw}(m_i)}{\sum_{i=1}^n Sup_{raw}(m_i)}. \quad (7)$$

#### C. Modified Evidence Combination Method

Assume that  $m_1$  and  $m_2$  are BBAs of the recognition framework  $\Theta$  for two evidences  $E_1$  and  $E_2$ , the focal elements of  $m_1$  are  $A_i (i=1, 2, \dots)$  and the focal elements of  $m_2$  are  $B_j (j=1, 2, \dots)$ .

The classical D-S combination method [4] can be expressed as

$$m(X) = \begin{cases} \frac{1}{1-K} \sum_{A_i \cap B_j = X} m_1(A_i) m_2(B_j) & X \neq \emptyset \\ 0 & X = \emptyset \end{cases} \quad (8)$$

with

$$K = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j). \quad (9)$$

The proposed combination method in this paper is expressed as

$$m(X) = \begin{cases} m_{\cap}(X) + m_h(X) + m_r(X) & X \neq \emptyset \\ 0 & X = \emptyset \end{cases} \quad (10)$$

where  $m_{\cap}(X) = \sum_{A_i \cap B_j = X} m_1(A_i) m_2(B_j)$ ,  $m_h(X)$  is expressed as

$$m_h(X) = \alpha \frac{\prod_{A_i \subseteq X} m_1(A_i)^{Sup_{his}(m_1)} \cdot \prod_{B_j \subseteq X} m_2(B_j)^{Sup_{his}(m_2)}}{\sum_{X \subseteq \Theta} \left( \prod_{A_i \subseteq X} m_1(A_i)^{Sup_{his}(m_1)} \cdot \prod_{B_j \subseteq X} m_2(B_j)^{Sup_{his}(m_2)} \right)} K \quad (11)$$

and  $m_r(X)$  is expressed as

$$m_r(X) = \beta \frac{\prod_{A_i \subseteq X} m_1(A_i)^{Sup_{ret}(m_1)} \cdot \prod_{B_j \subseteq X} m_2(B_j)^{Sup_{ret}(m_2)}}{\sum_{X \subseteq \Theta} (\prod_{A_i \subseteq X} m_1(A_i)^{Sup_{ret}(m_1)} \cdot \prod_{B_j \subseteq X} m_2(B_j)^{Sup_{ret}(m_2)})} K \quad (12)$$

where  $\alpha + \beta = 1$ ,  $\alpha$  and  $\beta$  are adjustment factors to control the proportion of the historical part and the real-time part respectively. The factor  $\alpha$  can be calculated by  $\alpha = \frac{num_{his}}{num_{tot}}$ ,

where  $num_{his}$  is the number of historical records and  $num_{tot}$  is the number of total records. The parameter  $K$  is calculated in the same way as the D-S evidence theory by (9).

#### D. Procedures of Modified Evidence Combination Method

BBAs are formed from the original data firstly and then the BBAs are combined according to the proposed method in this paper as shown in Figure 1. The key procedures of the method are as follows.

Step1: Obtain BBAs through membership function from the original data.

Step2: Calculate historical support and real-time support.

Step3: Fuse BBAs by the modified evidence combination method and select the decision corresponding to the maximum reliability value as the final result.

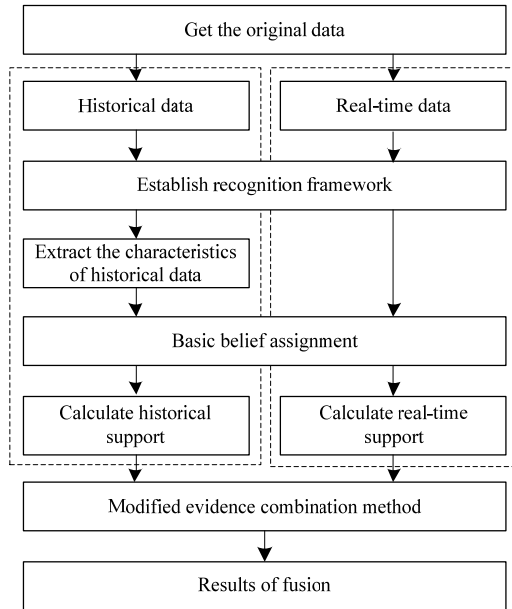


FIGURE 1. THE PROCEDURES OF MODIFIED EVIDENCE COMBINATION METHOD

### III. CASE ANALYSIS

Real medical readings collected from the multiparameter intelligent monitoring in intensive care (MIMIC) database [15]

are used in our simulations. 1800 records of the patient numbered 055n are extracted to be divided into three different categories represented by A, B, and C respectively. Each category consists of 600 records and each record contains five vital signs: blood pressure (BP), heart rate (HR), pulmonary artery pressure (PAP), respiratory rate (RESP) and the blood temperature (Tblood).

The extracted dataset is divided into two parts, where one part is taken as the historical data (1200 records) and another part is taken as the real-time data (600 records). There are five vital signs in one record. Five vital signs correspond to five evidences whose BBAs can be expressed as  $m_{BP}$ ,  $m_{HR}$ ,  $m_{PAP}$ ,  $m_{RESP}$  and  $m_{Tblood}$  respectively.

The recognition framework is set to  $\Theta = \{A, B, C\}$  because our goal is to determine the category based on the vital signs. The characteristics of historical data are extracted to determine original BBAs of evidences. The definition of an original BBA can be expressed as

$$\begin{cases} R_i = \frac{1}{e^{\frac{(x-\mu_i)^2}{\sigma_i^2}}} & i = 1, 2, 3 \\ m_{BP}(A) = \frac{R_1}{R_1 + R_2 + R_3} \end{cases} \quad (13)$$

where  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  are the mean values of the three categories from historical data,  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$  are standard deviations of the three categories from historical data.

TABLE I. ORIGINAL HISTORICAL DATA IN CATEGORY A

No.	BP (mmHg)	HR (bpm)	PAP (mmHg)	RESP (bpm)	Tblood (degC)	Category
1	82	88	20	15	37.1	A
2	82	88	19	15	37.1	A
3	83	88	20	15	37.1	A
...	...	...	...	...	...	...
398	97	109	26	22	36.7	A
399	96	109	25	22	36.7	A
400	96	109	25	23	36.7	A

Table I shows the part of the sample data of original historical data in category A, BBAs can be determined by (13). In the same way, we can get the BBAs of historical data and real-time data in other categories. After the determination of BBAs, we can continue to fuse them following the procedures in Figure 1. Based on the BBAs, the results of proposed modified combination method are compared with that of D-S evidence theory, Murphy's method in [9] and Yu's method in [16].

Figure 2 shows the influence of parameter  $k$  of membership function in (6). As mentioned in section III, the appropriate parameter  $k$  can be selected by experiments or expert opinions. Here we choose the appropriate value by experimenting with historical data. From Figure 2, it is obvious to see that the proposed modified method in this paper works best when  $k =$

105 and the correct rate tends to be stable as  $k$  increases. So we set  $k = 105$  in the next simulation.

Figure 3 illustrates the performance of four different combination methods. There are 600 real-time records to be classified in this experiment. The changing trends of these algorithms are similar, but the modified combination method in this paper obtains the highest correct rate, whose correct rate is 0.9450, when all of the samples are classified. The correct rate of Murphy's method is similar to that of Yu's method, and they are all a little lower than 0.94. The correct rate of classical D-S evidence theory is worse than other methods which is 0.9067. It can be seen that Murphy's method and Yu's method have a certain effect by improving the D-S evidence theory, but the modified method in this paper works better.

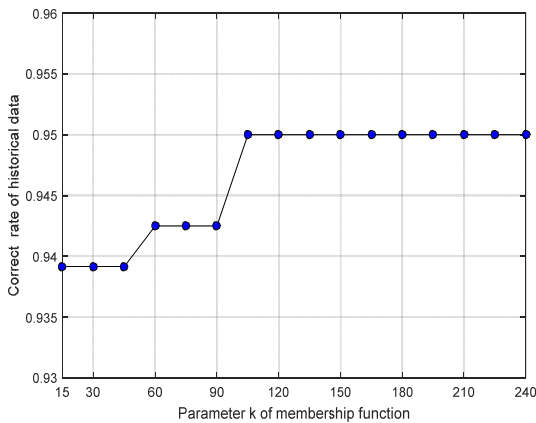


FIGURE II. THE INFLUENCE OF PARAMETER  $K$

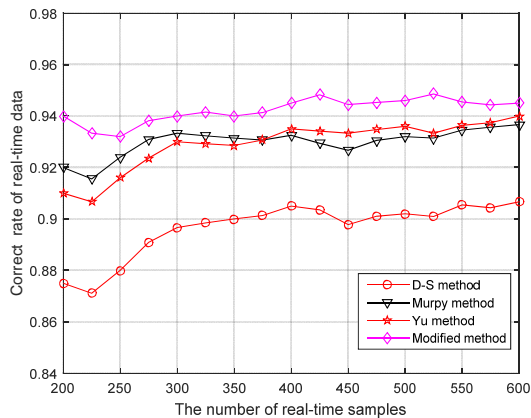


FIGURE III. THE CORRECT RATE OF DIFFERENT METHODS

#### IV. CONCLUSIONS

Different sensors always produce conflicting data in a multisensor data fusion system which contains vast amounts of historical data and real-time data. D-S evidence theory cannot deal with problems effectively in such a situation. With analysis of different kinds of improved methods, a new modified evidence combination method based on information gain and fuzzy preference relations is proposed. The proposed method can not only obtain more accurate results with the

effective combination of historical data and real-time data but also handle the conflict validly between different sensors as shown in the above experiment.

#### ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China (61771292), the National Key Research and Development Program of China (2017YFC0803403), and the Natural Science Foundation of Shandong Province of China (ZR2016FM29).

#### REFERENCES

- [1] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6-23, 2002.
- [2] K. Kolomvatsos, C. Anagnostopoulos and S. Hadjiefthymiades, "Data fusion and type-2 fuzzy inference in contextual data stream monitoring," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 47, no. 8, pp. 1839-1853, 2017.
- [3] S. Sapna, "Fusion of big data and neural networks for predicting thyroid," *Proceedings of the International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques*, pp. 243-247, 2016.
- [4] A. P. Dempster, "Upper and lower probabilities induces by a multivalued mapping," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325-339, 1967.
- [5] H. Wang, Y. Xiao and Y. Long, "Research of intrusion detection algorithm based on parallel SVM on spark," *Proceedings of IEEE International Conference on Electronics Information and Emergency Communication*, pp. 153-156, 2017.
- [6] R. R. Yager, "On the Dempster-Shafer framework and new combination rules," *Proceedings of International Forum on Information Technology and Applications*, IEEE Computer Society, pp. 563-565, 2009.
- [7] B. Li, B. Wang, J. Wei, et al., "Efficient combination rule of evidence theory," *Journal of Data Acquisition & Processing*, vol. 17, no. 1, pp. 237-240, 2002.
- [8] G. Pan and L. Wu, "An improved D-S evidence combination rule and its application," *Systems Engineering*, vol. 32, no. 5, pp. 150-153, 2014.
- [9] C. K. Murphy, "Combining belief functions when evidence conflicts," *Decision Support Systems*, vol. 29, no. 1, pp. 1-9, 2000.
- [10] Y. Deng, W. Shi, Z. Zhu, et al., "Combining belief functions based on distance of evidence," *Decision Support Systems*, vol. 38, no. 3, pp. 489-493, 2004.
- [11] H. Liu, Z. Zhao and X. Liu, "Combination of conflict evidences in D-S theory," *Journal of University of Electronic Science & Technology of China*, vol. 37, no. 5, pp. 701-704, 2008.
- [12] Y. Deng, "Deng entropy: A generalized Shannon entropy to measure uncertainty," in press.
- [13] A. L. Jousselme, D. Grenier and É. Bossé, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, no. 2, pp. 91-101, 2001.
- [14] Q. Hu, W. Pan, L. Zhang, et al. "Feature selection for monotonic Classification," *IEEE Transactions on Fuzzy Systems*, vol.20, no.1, pp.69-81, 2012.
- [15] A. L. Goldberger, L. A. N. Amaral, L. Glass, et al. "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215-220 2000.
- [16] Y. Yu, C. Yang, H. Jiang, et al. "An improved evidence combination approach based on credibility of evidence," *Proceedings of Chinese Control Conference*, pp. 5518-5522, 2017.