PCARST: a method of weakening conflict evidence based on principal component analysis and relatively similar transformation

Xuan Zhao[a], Huizi Cui[b], Zilong Xiao[c], Bingyi Kang[a,b,c,*]

[a] *College of Information Engineering, Northwest A&F University, Yangling, Shaanxi, 712100, China.*
[b] *Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China*
[c] *Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China*

Abstract

How to deal with conflict is a significant issue in Dempster-Shafer evidence theory (DST). In the Dempster combination rule, conflicts will produce counter-intuitive phenomena. Therefore, many effective conflict handling methods have been presented. This paper proposes a new framework for reducing conflict based on principal component analysis and relatively similar transformation (P-CARST), which can better reduce the impact of conflict evidence on the results, and has more reasonable results than existing methods. The main characteristic feature of the BPAs is maintained while the conflict evidence is regarded as a noise signal to be weakened. A numerical example is used to illustrate the effectiveness of the proposed method. Results show that a higher belief degree of the correct proposition is obtained comparing previous methods.

Keywords: Conflict management, Dempster-Shafer evidence theory, Relatively similar transformation, Principal component analysis, Information fusion, Decision making

---

*Corresponding author: B. Kang, College of Information Engineering, Northwest A&F University, Yangling, Shaanxi, 712100, China. Email address: bingyi.kang@nwsuaf.edu.cn; bingyi.kang@hotmail.com

1. Introduction

Dempster-Shafer theory (DST) is an effective model to deal with imprecise, vague, partially uncertain information [1, 2]. The uncertain information from multi-source sensors or experts, which is always described by basic probability assignments (BPAs), can be aggregated by the combining rule, e.g. Dempster combination rule to obtain more accurate results. Therefore, DST has been widely used to address plenty of scientific applications, such as information fusion [3, 4, 5], pattern recognition [6], risk analysis [7, 8], multiple attribution decision making (MADM) [9, 10, 11].

However, there are some shortcomings in classical DST proposed by Dempster and Shafer [1, 2]. The main disadvantage is that aggregating extreme conflicting evidence may lead to counterintuitive results. The abnormal phenomenon caused by highly conflict evidence in DST is firstly discussed by Prof. Zadeh [12]. When there is a big conflict between the evidence participating in the synthesis, the result of the fusion may be contrary to common sense. Various effective methods have been developed in the evidence theory to solve the compatibility and difference of multi-source information. Briefly, there are two strategies to deal with conflict management in DST according to the reasons discussed above: (1) one is to improve the combination rule. The focus of this direction lies in how to reallocate the conflict. For instance, Yager thinks that the conflict should be assigned to the unknown information, i.e. the full set [13], Smets thinks that the conflict should be assigned to the empty set since the conflict may be caused by the object outside the FoD. Some extension versions of DST are also developed, e.g. Dezert-Smarandache theory (DSmT) [14], transferable belief model (TBM) [15], evidential reasoning rule (ER) [16], generalized evidence theory (GET) [17]. (2) the other strategy is modifying the measure of the body of evidence, there are now a variety of measures for the compatibility and difference of evidence bodies, such as Murphy's [18] average measure, Deng's [19] similarity measure,Jiang's [20] information entropy measure, Jiang's [21] correlation coefficient measure, Xiao's [22] divergence measure, Mi and Kang's [23] evidence gravity measurement, Shang and Deng's [24] AE-K-Means measurement, etc. Giving reasonable weight to the evidence to be fused, which can make DST converge faster, thereby enhancing the reliability of the fusion result.

In work-related to machine learning, Wen proposed a concept of relative transformation for machine learning [25]. Relative Transformation (RT) is used to simulate the relative perception of humans on the focused object. Relative perception transforms the original data space into a relative space, which means that the information of the focused object depends not only on its characteristics but also on the characteristics of other objects. The relative transformation makes the isolated point far away from the normal point to suppress noise. However, the result of the relative transformation is not stable, which is not conducive to improving the robustness of the fusion system.

Principal component analysis (PCA) uses coordinate rotation to maximize the variance of the data in the principal direction, that is, the projection of the entire data in this direction is the most scattered, other directions are often related to noise, and abandoning them can reduce noise to a certain extent. PCA is often used to process high-dimensional data, and has good applications in dimensionality reduction and noise reduction [26, 27, 28, 29], image analysis [30, 31, 32, 33, 34], and text mining [35, 36, 37]. Its characteristics are: determine the potential dimensionality of the data, remove unimportant data, and reconstruct the original data set.

In this article, a new method of weakening conflict evidence combination, called the principal component analysis of relative similarity transformation (PCARST), is proposed. The proposed method analyzes the main features from the perspective of relatively similar transformation, to achieve the effect of noise reduction. Furthermore, numerical examples show that the new method can better distinguish between trust evidence and conflict evidence than do other existing conflicting evidence weakens methods.

The main contributions of this article are as follows:

(1) This is the first study in evidence theory that combines relative similarity conversion and principal component analysis, and measures conflict evidence and trust evidence from a new perspective.

(2) PCARST is a universal framework that can weaken the impact of conflict on the final result.

(3) Compared with the previous method, PCARST has a better ability to weaken collisions.

The structure of this article is organized as follows. The second part reviews basic knowledge and related work. The third part proposes a new method to weaken the conflict evidence, and gives a numerical example. The fourth part compares and discusses with previous methods. The fifth part introduces an application example and discusses the rationality and effectiveness of this method. Finally, the sixth part summarizes this article.

2. Preliminaries

In this section, some preliminaries are briefly introduced.

2.1. Dempster-Shafer (DS) theory of evidence

The Dempster-Shafer theory (DST) of evidence, which was first proposed by Dempster [1] and then developed by Shafer [2], is regarded as a generalization of the Bayesian theory of probability. Due to its ability to handle uncertainty or imprecision embedded in the evidence, the DS theory has increasingly been applied in recent years [38, 39, 40], and applied to multiple attribute decision analysis problems [9, 41], etc.

The introduction of DS theory are briefly summarized as follows:

(1) “Frame of discernment” [2]:

Let $\Theta = \{H_1, H_2, \ldots, H_N\}$ be a finite set of n elements, and $P(\Theta)$ denote the power set composed of 2^N elements of Θ.

$$P(\Theta) = \{\emptyset, \{H_1\}, \{H_2\}, \ldots, \{H_N\}, \{H_1 \cup H_2\}, \{H_1 \cup H_3\}, \ldots, \Theta\} \quad (1)$$

(2) “Basic probability assignment (BPA)” [2]:

The BPA function or mass function is defined as a mapping of the power set $P(\Theta)$ to a number between 0 and 1.

$$m : P(\Theta) \to [0, 1] \quad (2)$$

and which satisfies the following conditions:

$$m(\emptyset) = 0, \sum_{A \subseteq P(\Theta)} m(A) = 1 \quad (3)$$

The mass $m(A)$ represents how strongly the evidence supports A.

(3) “Belief and plausibility functions” [2]:

The belief function Bel is defined as

$$Bel : P(\Theta) \to [0, 1] \textit{ and } Bel(A) = \sum_{B \subseteq A} m(B) \quad (4)$$

and the plausibility function Pl is defined as

$$Pl : P(\Theta) \to [0, 1] \textit{ and } Pl(A) = 1 - Bel(\bar{A}) = \sum_{B \cap A \neq \emptyset} m(B) \quad (5)$$

Bel(A) and *Pl(A)* are the lower limit and the upper limit, respectively, of the belief level of hypothesis A which is illustrated in Figure 1. Both imprecision and uncertainty can be represented by them.

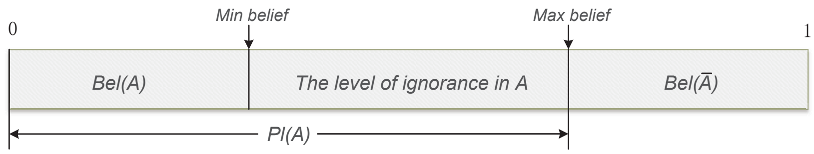

Figure 1: The relation between Bel and Pl.

(4) "Dempster's combination rule":[42]

Two bodies of evidence X and Y regarding Θ can be used to calculate the belief level for some new hypothesis C as follows:

The measure of conflict K is given as

$$K = \sum_{X\cap Y=\emptyset, \forall X,Y\subseteq\Theta} m_i(X)\times m_{i'}(Y) \tag{6}$$

and the mass function after combination is

$$\begin{aligned} m(C) &= m_i(X)\oplus m_{i'}(Y) \\ &= \begin{cases} 0, & if\ X\cap Y=\emptyset, \\ \frac{\sum_{X\cap Y=C,\forall X,Y\subseteq\Theta} m_i(X)\times m_{i'}(Y)}{1-K}, & if\ X\cap Y\neq\emptyset. \end{cases} \end{aligned} \tag{7}$$

2.2. Relative transformation (RT)

To simulate the effect of the relative perception, Wen propose a notion of relative transformation (RT) [25]. Suppose an original n-dimensional space X^n is defined, m samples in original space X^n is denoted by $x^i=\left(x_1^i,x_2^i,\ldots,x_n^i\right)$, $i=1,2,...,m$. The relative space is defined as Y^m, the samples in relative space is denoted by $y^j=\left(y_1^j,y_2^j,\ldots,y_m^j\right)$, $j=1,2,...m$. The relative transform (RT) f is defined as

$$Y^m = f\left(X^n\right) \tag{8}$$

The element y_i^j of the sample y^j in relative space can be represented as

$$y_i^j = f\left(x^i, x^j\right) = d\left(x^i, x^j\right) \tag{9}$$

where the function d is the distance function, e.g., Euclidean distance. Some special properties of the RT have been given in Ref.[25], e.g., the ability to handle noisy data, fault detect.

Especially, RT has a good performance to cope with the noisy data because RT has the augmenting effect on the distance between the samples from the original space to relative space, i.e., if the samples behave coherence, they will be gathered more after RT, and if a sample is an outlier, it will be further away from other samples.

2.3. Principal components analysis (PCA)

Principal component analysis (PCA) was introduced by Karl Pearson [43] and developed by Harold Hotelling [44], which is regarded as an orthogonal linear transformation that transforms the original data into a new coordinate system to obtain the greatest variance. The data in the new system assigns the data on the first coordinate (namely the first principal component) with the biggest variance, and one the second coordinate with the second biggest variance, etc. PCA has been used in many scopes, e.g., dimensionality reduction [45, 46], noise signal detections [27].

Assume we have m samples $x(i)$, $i = 1, \ldots, m$, each sample $x(i)$ has n features, the original data is described by matrix $D = [d_{i,j}]_{mn}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. The procedure of the PCA is briefly summarized as follows:

Step 1: Shift the sample mean of each column to zero using $x\left(i\right) = x\left(i\right) - \frac{1}{m}\sum\limits_{j=1}^{m} x\left(j\right)$.

Step 2: Obtain the covariances matrix $X^T X$.

Step 3: Obtain the eigenvalues $\lambda = [\lambda_1, \ldots, \lambda_m]$ and the corresponding eigenvectors $v = [v_1, \ldots, v_m]$ of the covariances matrix $X^T X$.

Step 4: Select the corresponding eigenvectors to assemble a new eigenvector matrix $E_{n,k}$ according to the k greatest eigenvalues.

Step 5: Project the data to the new coordinate system using $DE_{n,k}$.

Inspired by RT and PCA, we proposed a new management plan. The following section will discuss the BPA conflict in DST.

2.4. Previous framework to modify the conflict evidence

In the complex information world, the sources of information are diverse. The first screening and then fusion of different information can make the final result more reliable. Murphy [18] proposed the idea of the average weight of evidence, that is, to give each piece of evidence the same weight. Deng [19] used the concept of evidence distance to determine the new weight with the sum of the similarity between himself and all other evidence and assigned a lower weight to conflicting evidence. The main methods are as follows:

Step 1: Construct a similarity measure matrix (SMM). S_{ij}=1-d(m_i,m_j) is the similarity measure.

$$SMM = \begin{bmatrix} 1 & S_{12} & & \cdots & & S_{1n} \\ S_{21} & 1 & & \cdots & & S_{2n} \\ \vdots & \vdots & \ddots & & & \vdots \\ & & & S_{ij} & & \\ \vdots & \vdots & & & \ddots & \vdots \\ S_{n1} & S_{n2} & & \cdots & & 1 \end{bmatrix}$$

Step 2: Calculate the support degree of each piece of evidence.

$$Sup(m_i) = \sum_{j=1, j\neq i}^{n} S(m_i, m_j) \tag{10}$$

Step 3: Calculate weights and normalize.

$$Crd_i = \frac{Sup(m_i)}{\sum\limits_{i=1}^{n} Sup(m_i)} \tag{11}$$

$$MAE(m) = \sum_{i=1}^{n} (Crd_i \times m_i) \tag{12}$$

Since then, many researchers have assigned various weights based on this framework. The most important one is to modify the matrix value directly. For example, Jiang [21] uses correlation coefficient to replace similarity, Mi and Kang [23] use evidence theory gravitational value instead of similarity.

3. Proposed method to weaken the conflict evidence using PCAoRST

In this section, a generalized method, called principal component analysis based on relative similarity transformation, is proposed for conflict evidence.

Model multiple sources of evidence as a BPA model and transform it into a relatively similar model. Currently, methods to measure BPA relationships include similarity, correlation coefficient, RB divergence measure, evidence gravity, etc. However, these standards can be divided into the relative Similarity Model (RSM) and the Relative Gap Model (RGM). We transform the RGM into an RSM for a definition for unified processing. Finally, we perform PCA processing on the relative similarity model based on relative similarity transformation.

Definition 1. *Let the FoD be Θ, suppose we have t BPAs m_i, $i = 1, \ldots, t$. The relative similarity transformation(RST) of BPAs is defined as follows:*

$$X_i : m_i \xrightarrow{RST} Y_i : RSM_i = (RSM_{i1}, \cdots, RSM_{in}) \tag{13}$$

Then the matrix after relative similarity transformation is defined as follows:

$$RSM = \begin{bmatrix} 1 & RSM_{12} & & \cdots & RSM_{1n} \\ RSM_{21} & 1 & & \cdots & RSM_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ & & & RSM_{ij} & \\ \vdots & \vdots & & \ddots & \vdots \\ RSM_{n1} & RSM_{n2} & & \cdots & 1 \end{bmatrix} \tag{14}$$

The RSM_{ij} represents the degree of the relative similarity between evidence i and evidence j.

Definition 2. *The definition of the relative similarity transformation of the relative gap model is as follows:*

$$RSM(i,j) = \begin{cases} 1 - RGM(i,j), 0 \leq RGM(i,j) \leq 1 \\ \dfrac{\sum\limits_{i=1}^{t} (RGM(i,t) \times RGM(j,t))}{\sqrt{\sum\limits_{i=1}^{t} RGM(i,t)} \times \sqrt{\sum\limits_{j=1}^{t} RGM(j,t)}}.RGM(i,j) \geq 0 \end{cases} \tag{15}$$

3.1. Procedure of the proposed method

In this method, we regard the degree of the relative similarity associated with a certain evidence body as several characteristics of this evidence body. The improved method maximizes the variance on the first principal component of the data by rotating projection. The most critical features of low-dimensionality are retained to show the data's variability better. That is to say, all characteristics in this direction are the most scattered, which means that more information is retained, which is equivalent to a projection of the feature set on the angle with the most information. After PCARST projection processing, the length of the projection intercept is regarded as the new weight source. The complete PCARST process is shown in Figure 2.

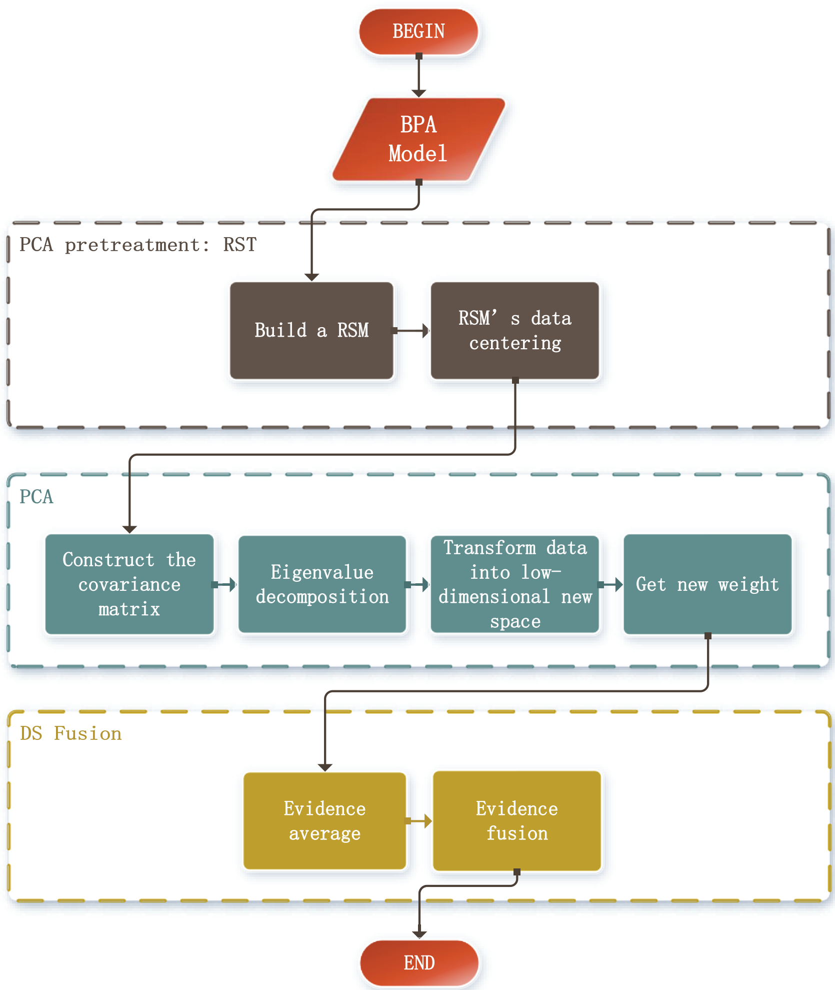

Figure 2: The process of PCARST.

The steps of using PCARST to weaken conflict evidence are as follows:

Step 1 (Build a relatively similar model): The original BPA data set is mapped by a relatively similar transformation, and construct a relatively similar model $RSM = [RSM_{i,j}]_{tt}$, $i = 1, \ldots, t$, $j = 1, \ldots, t$. If it is a relative gap model, use the Eq.(15) to transform it into a relative similar model.

Step 2 (RSM's data centering): Shift the sample mean of each column of the relative similarity matrix to zero using $RSM\left(i\right) = RSM\left(i\right) - \frac{1}{t}\sum\limits_{j=1}^{t} x\left(j\right)$. The centralized matrix of RSM is regarded as X.

Step 3 (Construct the covariance matrix): Obtain the covariances matrix $X^T X$. The elements of matrix $X^T X$ express the degree of difference between the two characteristics.

Step 4 (Eigenvalue decomposition of covariance matrix): Obtain the eigenvalues $\lambda = [\lambda_1, \ldots, \lambda_t]$ and the corresponding eigenvectors $v = [v_1, \ldots, v_t]$ of the covariances matrix $X^T X$.

Step 5 (Transform data into low-dimensional new space): Select the corresponding eigenvectors $v_{max(\lambda)}$ according to the greatest eigenvalues, and project the data S to the new coordinate system using $Sv_{max(\lambda)}$.

One of the most critical issues is selecting several eigenvalues in this step. If the dimension is too large, the noise reduction rate is not high, and the data error will be enormous if the dimension is too small. In response to this problem, a classic way is to consider the percentages left by different feature values. The processed data represents the original data as much as possible. In this article, our principle is $\frac{\lambda_i}{\sum \lambda_i} \geq 90\%$. The information represented by the eigenvalues is shown in Figure 3.

Figure 3: Graphical eigenvalue information.

The depth of the color in the figure represents the amount of information contained in the feature value of the corresponding location.

Step 6 (Get new weight): Obtain the PCA weight W of each BPA using $W = \frac{\left|Sv_{\max(\lambda)}\right|}{\sum \left|Sv_{\max(\lambda)}\right|}$.

Step 7 (Evidence average): Average the BPAs to one BPA by the weight W, i.e. $\bar{m} = \sum m_i W(j)$.

Step 8 (Evidence fusion): Aggregate the average BPA $\bar{m}$ $t-1$ times using the Dempster combination rule.

According to intuition, conflicting evidence should be given a smaller weight, so that it has a smaller impact on the result in the process of evidence fusion. This will be illustrated in the numerical examples in the next section.

3.2. Numerical Example

To illustrate the superiority of the proposed method, some numerical calculations will be constructed.

Example 1. *An example in Ref.[19] is used to describe the procedure of the proposed method. The example includes five BPAs, which is shown as follows:*

$$\begin{aligned}
&m_1 : m_1(A) = 0.5, m_1(B) = 0.2, m_1(C) = 0.3 \\
&m_2 : m_2(A) = 0, m_2(B) = 0.9, m_2(C) = 0.1 \\
&m_3 : m_3(A) = 0.55, m_3(B) = 0.1, m_3(A, C) = 0.35 \\
&m_4 : m_4(A) = 0.55, m_4(B) = 0.1, m_4(A, C) = 0.35 \\
&m_5 : m_5(A) = 0.6, m_5(B) = 0.1, m_5(A, C) = 0.3
\end{aligned}$$

Step 1, we can calculate the correlation coefficients in Ref.[21] to obtain the similarity matrix RSM, which is shown as follows

$$RSM = \begin{bmatrix} 1 & 0.3762 & 0.8908 & 0.8908 & 0.8922 \\ 0.3762 & 1 & 0.1499 & 0.1499 & 0.1449 \\ 0.8908 & 0.1499 & 1 & 1 & 0.9981 \\ 0.8908 & 0.1499 & 1 & 1 & 0.9981 \\ 0.8922 & 0.1449 & 0.9981 & 0.9981 & 1 \end{bmatrix}$$

Step 2, we shift the sample mean of each column to zero. The centralized matrix of RSM using $RSM(i) = RSM(i) - \frac{1}{t}\sum_{j=1}^{t} x(j)$ is regarded as X.

$$X = \begin{bmatrix} 0.1900 & 0.0120 & 0.0831 & 0.0831 & 0.0856 \\ -0.4338 & 0.6358 & -0.6579 & -0.6579 & -0.6617 \\ 0.0808 & -0.2143 & 0.1922 & 0.1922 & 0.1914 \\ 0.0808 & -0.2143 & 0.1922 & 0.1922 & 0.1914 \\ 0.0822 & -0.2192 & 0.1903 & 0.1903 & 0.1933 \end{bmatrix}$$

Step 3, we obtain the covariances matrix $X^T X$, which is denoted by,

$$X^T X = \begin{bmatrix} 0.2441 & -0.3262 & 0.3479 & 0.3479 & 0.3501 \\ -0.3262 & 0.5443 & -0.5414 & -0.5414 & -0.5441 \\ 0.3479 & -0.5414 & 0.5499 & 0.5499 & 0.5528 \\ 0.3479 & -0.5414 & 0.5499 & 0.5499 & 0.5528 \\ 0.3501 & -0.5441 & 0.5528 & 0.5528 & 0.5559 \end{bmatrix}$$

Step 4, we obtain the eigenvalues $\lambda = [0, 0, 0, 0.0352, 2.4088]$ and the corresponding eigenvectors $v = [v_1, \ldots, v_t]$ of the covariances matrix $X^T X$, which are shown as follows:

$$v = \begin{bmatrix} -0.0063 & -0.2720 & 0.3844 & 0.8288 & 0.3022 \\ 0.0088 & 0.3799 & -0.5669 & 0.5593 & -0.4705 \\ 0.7033 & -0.1733 & -0.4971 & 0.0048 & 0.4778 \\ -0.7106 & -0.1406 & -0.4971 & 0.0048 & 0.4778 \\ 0.0198 & 0.8555 & 0.1916 & 0.0169 & 0.4804 \end{bmatrix}$$

Step 5, we select the corresponding eigenvectors $v_{max(\lambda)}$ according to the greatest eigenvalues, and project the data S to the new coordinate system using $Sv_{max(\lambda)}$.

$$Sv_{max(\lambda)} = \begin{bmatrix} 1.4050 \\ -0.1440 \\ 1.6337 \\ 1.6337 \\ 1.6355 \end{bmatrix}$$

Step 6, we obtain the weight W of each BPA using $W = \frac{|Sv_{\max(\lambda)}|}{\sum |Sv_{\max(\lambda)}|}$

$$W = \frac{|Sv_{\max(\lambda)}|}{\sum |Sv_{\max(\lambda)}|} = \begin{bmatrix} 0.2178 \\ 0.0223 \\ 0.2532 \\ 0.2532 \\ 0.2535 \end{bmatrix}$$

In this article, we take the feature direction corresponding to the maximum feature value, because the maximum feature transformation can retain the internal information of the data to the greatest extent, and it is expected to have the largest dispersion in the projected dimension.

4. Comparing with previous methods

This section uses the example in Section 4 and compares the improved method with the existing conflict measurement methods. The results are as follows:

Table 1: Fusion results produced by various methods

Method	m_1,m_2	m_1,m_2,m_3	m_1,m_2,m_3,m_4	m_1,m_2,m_3,m_4,m_5
Dempster [42]	$m(A) = 0$	$m(A) = 0$	$m(A) = 0$	$m(A) = 0$
	$m(B) = 0.8571$	$m(B) = 0.6316$	$m(B) = 0.3288$	$m(B) = 0.1228$
	$m(C) = 0.1429$	$m(C) = 0.3684$	$m(C) = 0.6712$	$m(C) = 0.8772$
Murphy [18]	$m(A) = 0.1543$	$m(A) = 0.3500$	$m(A) = 0.6027$	$m(A) = 0.7958$
	$m(B) = 0.7469$	$m(B) = 0.5224$	$m(B) = 0.2627$	$m(B) = 0.0932$
	$m(C) = 0.0988$	$m(C) = 0.1276$	$m(C) = 0.1346$	$m(C) = 0.1110$
Deng et al. [19]	$m(A) = 0.1543$	$m(A) = 0.4861$	$m(A) = 0.7773$	$m(A) = 0.8909$
	$m(B) = 0.7469$	$m(B) = 0.3481$	$m(B) = 0.0628$	$m(B) = 0.0086$
	$m(C) = 0.0988$	$m(C) = 0.1657$	$m(C) = 0.1600$	$m(C) = 0.1005$
Proposed	$m(A) = 0.1543$	$m(A) = 0.7923$	$m(A) = 0.9627$	$m(A) = 0.9896$
	$m(B) = 0.7469$	$m(B) = 0.1130$	$m(B) = 0.0030$	$m(B) = 0.0001$
	$m(C) = 0.0988$	$m(C) = 0.0157$	$m(C) = 0.0239$	$m(C) = 0.0066$
	$m(A, C) = 0$	$m(A, C) = 0.0138$	$m(A, C) = 0.0105$	$m(A, C) = 0.0036$
Jiang [21]	$m(A) = 0.1543$	$m(A) = 0.7839$	$m(A) = 0.9528$	$m(A) = 0.9873$
	$m(B) = 0.7469$	$m(B) = 0.1131$	$m(B) = 0.0078$	$m(B) = 0.0006$
	$m(C) = 0.0988$	$m(C) = 0.0929$	$m(C) = 0.0313$	$m(C) = 0.0090$
	$m(A, C) = 0$	$m(A, C) = 0.0101$	$m(A, C) = 0.0080$	$m(A, C) = 0.0030$
Proposed	$m(A) = 0.1543$	$m(A) = 0.8286$	$m(A) = 0.9615$	$m(A) = 0.9891$
	$m(B) = 0.7469$	$m(B) = 0.0762$	$m(B) = 0.0031$	$m(B) = 0.0002$
	$m(C) = 0.0988$	$m(C) = 0.0793$	$m(C) = 0.0259$	$m(C) = 0.0075$
	$m(A, C) = 0$	$m(A, C) = 0.0159$	$m(A, C) = 0.0095$	$m(A, C) = 0.0033$
Xiao [22]	$m(A) = 0.1543$	$m(A) = 0.6722$	$m(A) = 0.9413$	$m(A) = 0.9871$
	$m(B) = 0.7469$	$m(B) = 0.2271$	$m(B) = 0.0177$	$m(B) = 0.0011$
	$m(C) = 0.0988$	$m(C) = 0.0929$	$m(C) = 0.0327$	$m(C) = 0.0083$
	$m(A, C) = 0$	$m(A, C) = 0.0078$	$m(A, C) = 0.0083$	$m(A, C) = 0.0035$
Proposed	$m(A) = 0.1543$	$m(A) = 0.8235$	$m(A) = 0.9610$	$m(A) = 0.9897$
	$m(B) = 0.7469$	$m(B) = 0.0810$	$m(B) = 0.0048$	$m(B) = 0.0003$
	$m(C) = 0.0988$	$m(C) = 0.0799$	$m(C) = 0.0226$	$m(C) = 0.0058$
	$m(A, C) = 0$	$m(A, C) = 0.0156$	$m(A, C) = 0.0115$	$m(A, C) = 0.0042$
Mi and Kang [23]	$m(A) = 0.1543$	$m(A) = 0.8648$	$m(A) = 0.8648$	$m(A) = 0.9591$
	$m(B) = 0.7469$	$m(B) = 0.0382$	$m(B) = 0.0382$	$m(B) = 0.0238$
	$m(C) = 0.0988$	$m(C) = 0.0833$	$m(C) = 0.0833$	$m(C) = 0.0145$
	$m(A, C) = 0$	$m(A, C) = 0.0137$	$m(A, C) = 0.0137$	$m(A, C) = 0.0026$
Proposed	$m(A) = 0.1543$	$m(A) = 0.8991$	$m(A) = 0.8991$	$m(A) = 0.9513$
	$m(B) = 0.7469$	$m(B) = 0.0093$	$m(B) = 0.0093$	$m(B) = 0.0035$
	$m(C) = 0.0988$	$m(C) = 0.0774$	$m(C) = 0.0774$	$m(C) = 0.0412$
	$m(A, C) = 0$	$m(A, C) = 0.0142$	$m(A, C) = 0.0142$	$m(A, C) = 0.0039$

[1] In Mi and Kang's method, the value of the adjustable parameter δ is set to 1/4.

[2] In addition, since the belief universal gravitation characterizes the magnitude of the measure between different bodies of evidence, for the two identical bodies of evidence, m_3 and m_4, their values before and after fusion are the same.

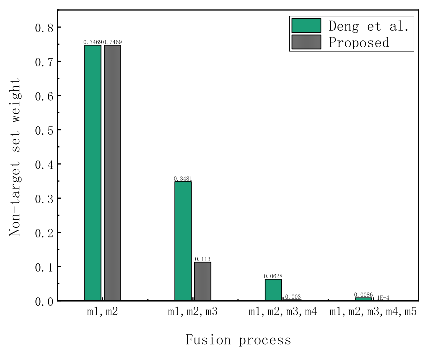

Figure 5: The weight of the non-target set (noise) is compared between Deng et al. and the proposed method

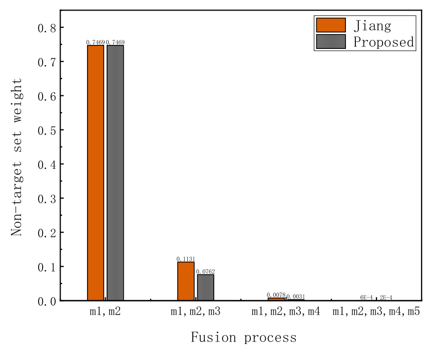

Figure 6: The weight of the non-target set (noise) is compared between Jiang and the proposed method

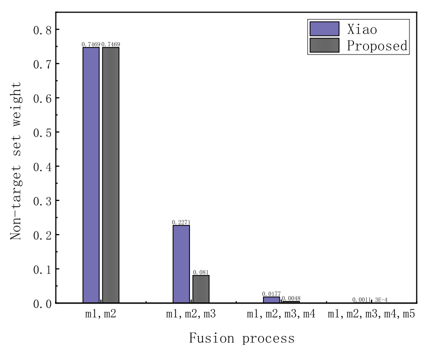

Figure 7: The weight of the non-target set (noise) is compared between Xiao and the proposed method

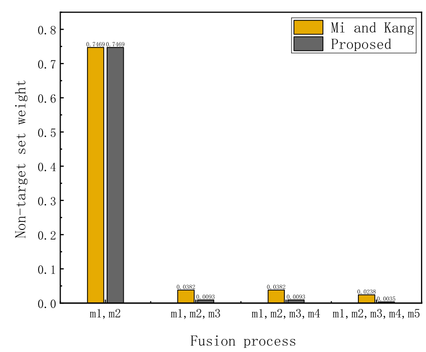

Figure 8: The weight of the non-target set (noise) is compared between MiKang and the proposed method

In Table 1, non-target set B (noise) is supported by a piece of evidence. From Figure 5 to Figure 8, the PCARST method we proposed has the best effect in reducing the weight of non-target set B (noise), which is better than other methods. Therefore, the PCARST method has certain advantages.

5. Application

5.1. Fault Diagnosis

The practical application of fault diagnosis illustrates the feasibility and effectiveness of PCARST. Generally, an autonomous vehicle needs to use multiple sensors at the same time, such as radar sensors and sonar that can detect obstacles by implementing obstacles. Multi-sensor fusion methods are often used to control an autonomous car more safely and reliably. However, no matter which type of sensor has its advantages and disadvantages, according to the record, three possible faults make the car unable to make a judgment: low oil pressure, ABS failure, and solenoid valve stuck, which are recorded as F1, F2, and F3. Five sensors, denoted as S1, S2, S3, S4, S5, measure them and model the results into BPA. The corresponding models are shown in Table 2. It can be seen from the table that m5 has given the wrong information due to sensor failure. The above methods can identify and weaken fault information, but the PCARST way we proposed has the highest ability to weaken conflicts and is superior to other methods. When conflicting evidence occurs, its weight decreases and the method suffers the most negligible impact. In practical applications, example of fault diagnosis [47] is used to illustrate the superiority of the PCARST method.

Table 2: BPAs representation of sensor failure.

	m(F1)	m(F2)	m(F3)	m(Θ)
m1	0.65	0.15	0	0.2
m2	0.7	0.1	0	0.2
m3	0.7	0	0	0.3
m4	0.75	0.05	0	0.2
m5	0	0.2	0.8	0

Table 3: The fusion results.

	F1	F2	F3	Θ
Deng et al.	0.9918	0.0051	0.0022	0.0009
Proposed	0.9950	0.0039	**0.0002**	0.0009
Jiang	0.9937	0.0045	0.0009	0.0009
Proposed	0.9942	0.0043	**0.0006**	0.0009
Xiao	0.9908	0.0055	0.0028	0.0009
Proposed	0.9951	0.0039	**0.0001**	0.0009
Mi and Kang	0.9950	0.0042	0.00002	0.0008
Proposed	0.9956	0.0036	**0**	0.0008

The results in Table 3 show that when conflict evidence m5 appears, we can see that the improved method recognizes the fault better and reduces the weight of the fault. This application shows that our method has certain advantages.

6. Conclusion

Reducing the impact of conflicting evidence is a long-term problem that needs to be resolved in evidence fusion. Based on PCA and relative transformation, we propose a new method of weakening conflict evidence, a principal component analysis method of relatively similar transformation, thereby effectively reducing the pollution of evidence noise to the final result. Its characteristics are: (1) PCARST maps the original evidence to a relatively similar model. This relatively similar modification process is the preprocessing of P-CA. (2) PCARST uses the principle of maximum variance to find the long axis of the direction with the largest difference in each position of a group of sample points and uses a small amount of data to represent the original data as much as possible to achieve reduction the purpose of noise. This article illustrates the method and application of the proposed scheme with numerical examples. In the future, weakening the conflict of evidence is still a problem that needs to be solved continuously.

Acknowledgment

The work is partially supported by the Fund of the National Natural Science Foundation of China (Grant No. 61903307), China Postdoctoral Science Foundation (Grant No. 2020M683575) and Chinese Universities Scientific Fund (Grant No. 2452018066). Thanks for the anonymous reviewers for their valuable suggestions and comments.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, in: Classic works of the Dempster-Shafer theory of belief functions, Springer, 2008, pp. 57–72.

[2] G. Shafer, A mathematical theory of evidence, Princeton university press, 1976.

[3] O. Basir, X. Yuan, Engine fault diagnosis based on multi-sensor information fusion using dempster–shafer evidence theory, Information fusion 8 (4) (2007) 379–386.

[4] F. Smarandache, J. Dezert, Information fusion based on new proportional conflict redistribution rules, in: 2005 7th International Conference on Information Fusion, Vol. 2, IEEE, 2005, pp. 8–pp.

[5] E. Lefevre, O. Colot, P. Vannoorenberghe, Belief function combination and conflict management, Information fusion 3 (2) (2002) 149–162.

[6] T. Denoeux, Analysis of evidence-theoretic decision rules for pattern classification, Pattern recognition 30 (7) (1997) 1095–1107.

[7] J. Yang, H.-Z. Huang, L.-P. He, S.-P. Zhu, D. Wen, Risk evaluation in failure mode and effects analysis of aircraft turbine rotor blades using dempster–shafer evidence theory under uncertainty, Engineering Failure Analysis 18 (8) (2011) 2084–2092.

[8] L. Sun, R. P. Srivastava, T. J. Mock, An information systems security risk assessment model under the dempster-shafer theory of belief functions, Journal of Management Information Systems 22 (4) (2006) 109–142.

[9] J.-B. Yang, M. G. Singh, An evidential reasoning approach for multiple-attribute decision making with uncertainty, IEEE Transactions on systems, Man, and Cybernetics 24 (1) (1994) 1–18.

[10] V.-N. Huynh, Y. Nakamori, T.-B. Ho, T. Murai, Multiple-attribute decision making under uncertainty: the evidential reasoning approach revisited, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 36 (4) (2006) 804–822.

[11] J.-B. Yang, P. Sen, A general multi-level evaluation process for hybrid madm with uncertainty, IEEE Transactions on Systems, Man, and Cybernetics 24 (10) (1994) 1458–1473.

[12] L. A. Zadeh, Review of a mathematical theory of evidence, AI magazine 5 (3) (1984) 81–81.

[13] R. R. Yager, On the dempster-shafer framework and new combination rules, Information sciences 41 (2) (1987) 93–137.

[14] F. Smarandache, J. Dezert, Advances and Applications of DSmT for Information Fusion, Vol. 3: Collected Works, Vol. 3, Infinite Study, 2004.

[15] P. Smets, R. Kennes, The transferable belief model, Artificial intelligence 66 (2) (1994) 191–234.

[16] J.-B. Yang, D.-L. Xu, Evidential reasoning rule for evidence combination, Artificial Intelligence 205 (2013) 1–29.

[17] Y. Deng, Generalized evidence theory, Applied Intelligence 43 (3) (2015) 530–543.

[18] C. K. Murphy, Combining belief functions when evidence conflicts, Decision support systems 29 (1) (2000) 1–9.

[19] D. Yong, S. WenKang, Z. ZhenFu, L. Qi, Combining belief functions based on distance of evidence, Decision support systems 38 (3) (2004) 489–493.

[20] W. Jiang, B. Wei, C. Xie, D. Zhou, An evidential sensor fusion method in fault diagnosis, Advances in Mechanical Engineering 8 (3) (2016) 1687814016641820.

[21] W. Jiang, A correlation coefficient for belief functions, International Journal of Approximate Reasoning 103 (2018) 94–106.

[22] F. Xiao, A new divergence measure for belief functions in ds evidence theory for multisensor data fusion, Information Sciences 514 (2020) 462–483.

[23] X. Mi, B. Kang, On the belief universal gravitation (bug), Computers & Industrial Engineering 148 (2020) 106685.

[24] Q. Shang, H. Li, Y. Deng, K. H. Cheong, Compound credibility for conflicting evidence combination: An autoencoder-k-means approach, IEEE Transactions on Systems, Man, and Cybernetics: Systems.

[25] W. Guihua, Relative transformation for machine learning, Journal of Computer Research and Development 45 (4) (2008) 612–618.

[26] H. G. Gauch Jr, Noise reduction by eigenvector ordinations, Ecology 63 (6) (1982) 1643–1649.

[27] J. Salmon, Z. Harmany, C.-A. Deledalle, R. Willett, Poisson noise reduction with non-local pca, Journal of mathematical imaging and vision 48 (2) (2014) 279–294.

[28] K. Yata, M. Aoshima, Effective pca for high-dimension, low-sample-size data with noise reduction via geometric representations, Journal of multivariate analysis 105 (1) (2012) 193–215.

[29] A. De Cheveigné, J. Z. Simon, Denoising based on time-shift pca, Journal of neuroscience methods 165 (2) (2007) 297–305.

[30] J. Yang, J.-y. Yang, From image vector to matrix: A straightforward image projection techniqueimpca vs. pca, Pattern Recognition 35 (9) (2002) 1997–1999.

[31] H. Jégou, O. Chum, Negative evidences and co-occurences in image retrieval: The benefit of pca and whitening, in: European conference on computer vision, Springer, 2012, pp. 774–787.

[32] C. E. Thomaz, G. A. Giraldi, A new ranking method for principal components analysis and its application to face image analysis, Image and vision computing 28 (6) (2010) 902–913.

[33] J. Wang, C.-I. Chang, Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis, IEEE transactions on geoscience and remote sensing 44 (6) (2006) 1586–1600.

[34] M. A. O. Vasilescu, D. Terzopoulos, Multilinear image analysis for facial recognition, in: Object recognition supported by user interaction for service robots, Vol. 2, IEEE, 2002, pp. 511–514.

[35] M. W. Berry, M. Castellanos, Survey of text mining, Computing Reviews 45 (9) (2004) 548.

[36] R. Albright, Taming text with the svd, SAS Institute Inc.

[37] S. Lee, B. Yoon, Y. Park, An approach to discovering new technology opportunities: Keyword-based patent map approach, Technovation 29 (6-7) (2009) 481–497.

[38] L. Dymova, P. Sevastianov, P. Bartosiewicz, A new approach to the rule-base evidential reasoning: Stock trading expert system application, Expert Systems with Applications 37 (8) (2010) 5564–5576.

[39] B. Kang, Y. Deng, R. Sadiq, S. Mahadevan, Evidential cognitive maps, Knowledge-Based Systems 35 (2012) 77–86.

[40] B. Kang, P. Zhang, Z. Gao, G. Chhipi-Shrestha, K. Hewage, R. Sadiq, Environmental assessment under uncertainty using dempster–shafer theory and z-numbers, Journal of Ambient Intelligence and Humanized Computing 11 (5) (2020) 2041–2060.

[41] J.-B. Yang, D.-L. Xu, On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 32 (3) (2002) 289–304.

[42] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics 38 (2) (1967) 325–339.

[43] K. Pearson, Liii. on lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin philosophical magazine and journal of science 2 (11) (1901) 559–572.

[44] H. Hotelling, Analysis of a complex of statistical variables into principal components., Journal of educational psychology 24 (6) (1933) 417.

[45] L. Cao, K. S. Chua, W. Chong, H. Lee, Q. Gu, A comparison of pca, kpca and ica for dimensionality reduction in support vector machine, Neurocomputing 55 (1-2) (2003) 321–336.

[46] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, R. Kaluri, D. S. Rajput, G. Srivastava, T. Baker, Analysis of dimensionality reduction techniques on big data, IEEE Access 8 (2020) 54776–54788.

[47] Z. Wang, F. Xiao, An improved multi-source data fusion method based on the belief entropy and divergence measure, Entropy 21 (6) (2019) 611.