Adaptive Weighted Multi-view Evidential Clustering

Zhe Liu[1]([✉]) [iD], Haojian Huang[2], and Sukumar Letchmunan[1]

[1] Universiti Sains Malaysia, Penang 11800, Malaysia
liuzhe921@gmail.com, zheliu@ieee.org
[2] Harbin Engineering University, Harbin 150001, China

Abstract. Multi-view clustering, which integrates information from different views for better performance, has gained widespread attention. However, the existing clustering methods cannot represent the uncertainty and imprecision in cluster assignment caused by the tendency of clusters to overlap and the diversity among views. To surmount this issue, in this paper, we propose an adaptive weighted multi-view evidential clustering (WMVWC) method based on the framework of belief functions. The proposed WMVEC can be viewed as a multi-view version of conventional evidential c-means clustering. We construct the objective function of WMVEC by integrating the learning of view weights and credal partition into a unified framework, and design an optimization scheme to obtain the optimal results of WMVEC. Specifically, the view weight can measure the contribution of each view in clustering. The credal partition can provide a deeper understanding of the data structure by allowing samples to belong not only to singleton clusters, but also to a union of different singleton clusters, called meta-cluster. Experiment results demonstrate the effectiveness of the proposed WMVEC with respect to other state-of-the-art methods on real-world datasets. The code can be available at link.

Keywords: Evidential clustering · multi-view learning · belief functions · credal partition

1   Introduction

Clustering, as a core paradigm of unsupervised learning, has been widely used in various domains owing to its powerful capacity in data preprocessing [13]. Many representative clustering methods like k-means and fuzzy c-means (FCM) have been invented for single-view data [10]. With the prompt development of information technology, however, single-view data can no longer meet the actual needs, and we are often faced with multi-view data represented by different sources or features. For example, an image can be expressed by multiple feature descriptors, such as LBP, SIFT and HOG. In such situations, the aforementioned classical clustering methods will no longer be applicable.

To counter this flaw and explore multi-view data, a load of derivatives of multi-view clustering methods have been emerged. One of the most representative is hard partition-based [2,3,6,15]. For example, Cai *et al.* [2] suggested a $\ell_{2,1}$ norm-based multi-view k-means clustering, called RMVKM, to learn the weight of each view. Later, Chen *et al.* [3] developed a two-level variable-weighted k-means (TW-k-means) clustering for multi-view data by introducing two entropy regularizations to adjust the importance of the weights. Based on collaborative learning, Zhang *et al.* [15] presented a two-level weighted collaborative k-means (TW-Co-k-means) clustering to consider the weight of views and features. However, these multi-view clustering methods all consider that the relationship between objects and clusters is distinct, *i.e.*, an object can only belong to a cluster completely or not.

Unlike hard partition-based clustering methods, fuzzy clustering describes the uncertainty by assigning each object to various singleton clusters with different memberships [11]. Many multi-view methods based on the idea of fuzzy clustering have been developed [4,5,7,14]. Cleuziou *et al.* [4] proposed collaborative fuzzy clustering for multi-view data, named Co-FKM, whereas it equally considers the contribution of each view. Jiang *et al.* [7] proposed a weighted view collaborative fuzzy c-means clustering to identify the contributions of different views. Recently, Yang and Sinaga [14] suggested a feature-weighted multi-view FCM based on collaborative learning (Co-FW-MVFCM), which can automatically identify the contribution of each view and feature. Although these methods have achieved reasonable results to some extent, they cannot represent the imprecision in the result. In fact, uncertain and imprecise cluster structure is very ubiquitous in applications, that is, the cluster information of some objects is difficult to distinguish.

To address such issue, a new concept of credal partition based on the theory of belief functions is introduced [8,9]. Credal partition extends the concepts of hard, fuzzy and possibilistic partitions by assigning the objects, not only to singleton clusters, but also to the union of several singleton clusters (called meta-cluster) with different masses of belief. To date, some evidential clustering methods have been developed in the framework of belief functions [9,12]. Evidential c-means (ECM) clustering [9], as the evidential counterpart of FCM, can well characterize uncertain and imprecise cluster structures and derive effective clustering results. Unfortunately, ECM is built upon the assumption of access to single-view data. Hence, how to characterize the uncertainty and imprecision of multi-view data in cluster assignment and improve the clustering performance is an open issue.

In this paper, we propose a new adaptive weighted multi-view evidential clustering (WMVEC) method. Specifically, we introduce view weights to capture the different contributions of each view in clustering and employ entropy regularization to regulate the distribution of view weights.

The main contributions are summarized as follows:

1. Inspired by the theory of belief functions, we propose a novel WMVEC method, which can characterize the uncertainty and imprecision in clustering results.

2. We design the objective function of WMVEC and obtain the weight vector, cluster prototype matrix and credal partition matrix in an alternating optimization way.
3. We conduct extensive experiments on several real-world datasets to verify the superiority of WMVEC with respect to other state-of-the-art methods.

The remainder of this paper is organized as follows. Section 2 proposes the WMVEC method and illustrates its optimization process. In Sect. 3, the effectiveness and practicability of the proposed WMVEC method are analyzed using real-world datasets. Finally, we make a brief conclusion in Sect. 4.

2   Multi-view Evidential Clustering

As mentioned earlier, many real-world applications are currently crowded with multi-view data. However, how to characterize the uncertainty and imprecision in clustering results on such data is still a challenging task. Conventional evidential clustering is no longer adequate for the situations we encounter today.

In this section, we propose a novel adaptive weighted multi-view evidential clustering (WMVEC) method to simultaneously learn the consensus credal partition and the importance of each view.

2.1   Formulation

Given a dataset $\mathcal{X} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_\mathcal{H}\}$ and h-th view is expressed as $\mathbf{X}_h = \{\mathbf{x}_{1,h}, \mathbf{x}_{2,h}, \cdots, \mathbf{x}_{n,h}\} \in \mathbb{R}^{n \times \mathcal{P}_h}$. The objective function of WMVEC can be represented as follows:

$$\mathcal{J}_{WMVEC}(\mathbf{M}, \mathbf{V}, \mathbf{r}) = \sum_{h=1}^{\mathcal{H}} r_h \left(\sum_{i=1}^{n} \sum_{\substack{A_j \subseteq \Omega \\ A_j \neq \emptyset}} |A_j|^\alpha m_{ij}^\beta \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2 + \sum_{i=1}^{n} \delta_h^2 m_{i\emptyset}^\beta \right) \quad (1)$$

$$+ \lambda \sum_{h=1}^{\mathcal{H}} r_h \ln r_h$$

$$s.t. \sum_{A_j \subseteq \Omega, A_j \neq \emptyset} m_{ij} + m_{i\emptyset} = 1, \; i = 1, 2, \cdots, n; \; \sum_{h=1}^{\mathcal{H}} r_h = 1 \quad (2)$$

with

$$\bar{\mathbf{v}}_{j,h} = \frac{1}{|A_j|} \sum_{\omega_k \in A_j} \mathbf{v}_{k,h} \quad (3)$$

where r_h is the view weight of h-th view, m_{ij} is the (i,j)-th element of credal partition matrix $\mathbf{M} \in \mathbb{R}^{n \times 2^c}$ and denotes the mass of belief of i-th object associated to cluster A_j. $\bar{\mathbf{v}}_{j,h}$ is the j-th cluster prototype of h-th view. The parameter λ is used to adjust the distribution of view weights, and parameters α, β and δ

are the same as in ECM. The objective function $\mathcal{J}_{MVEC}$ consists of two parts, the first part is similar to ECM and computes the sum of intra-cluster weighted distances in each view by assigning weights to the views. The second part is a negative weight entropy, which regulates the effect of view weight.

2.2   Optimization

In this subsection, we provide an efficient iterative method to minimize (2). More specifically, we employ the Lagrange optimization method to alternately update one variable while leaving others fixed.

Updating. **M** With **V** and **r** fixed, **M** is computed by Theorem 1.

Theorem 1. *Suppose* $\mathbf{V} = \mathbf{V}^*$ *and* $\mathbf{r} = \mathbf{r}^*$ *are fixed,* $\mathcal{J}_{WMVEC}(\mathbf{M}, \mathbf{V}^*, \mathbf{r}^*)$ *is minimized iff:*

$$m_{ij} = \frac{\left(\sum_{h=1}^{\mathcal{H}} |A_j|^\alpha r_h \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2 \right)^{-\frac{1}{\beta-1}}}{\left(\sum_{A_z \neq \emptyset} \sum_{h=1}^{\mathcal{H}} |A_j|^\alpha r_h \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{z,h}\|^2 \right)^{-\frac{1}{\beta-1}} + \left(\sum_{h=1}^{\mathcal{H}} r_h \delta_h^2 \right)^{-\frac{1}{\beta-1}}} \tag{4}$$

$$m_{i\emptyset} = 1 - \sum_{A_j \neq \emptyset} m_{ij} \tag{5}$$

Proof. We employ Lagrange multipliers τ_i to solve the constrained minimization problem with respect to **M**. The objective function (2) can be rewritten as:

$$\mathcal{J}(\mathbf{M}, \tau_i) = \mathcal{J}_{WMVEC}(\mathbf{M}, \mathbf{V}, \mathbf{r}) - \sum_{i=1}^{n} \tau_i \left(\sum_{\substack{A_j \subseteq \Omega \\ A_j \neq \emptyset}} m_{ij} + m_{i\emptyset} - 1 \right) \tag{6}$$

By differentiating the Lagrangian with respect to the m_{ij}, $m_{i\emptyset}$, and τ_i and setting the derivatives to zero, we obtain:

$$\frac{\partial \mathcal{J}}{\partial m_{ij}} = \beta m_{ij}^{(\beta-1)} \sum_{h=1}^{\mathcal{H}} |A_j|^\alpha r_h \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2 - \tau_i = 0 \tag{7}$$

$$\frac{\partial \mathcal{J}}{\partial m_{i\emptyset}} = \sum_{h=1}^{\mathcal{H}} \beta r_h \delta_h^2 m_{i\emptyset}^{(\beta-1)} - \tau_i = 0 \tag{8}$$

$$\frac{\partial \mathcal{J}}{\partial \tau_i} = \sum_{A_j \subseteq \Omega, A_j \neq \emptyset} m_{ij} + m_{i\emptyset} - 1 = 0 \tag{9}$$

We thus have from (7):

$$m_{ij} = \left(\frac{\tau_i}{\beta}\right)^{\frac{1}{\beta-1}} \left(\frac{1}{\sum\limits_{h=1}^{\mathcal{H}} |A_j|^\alpha r_h \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2}\right)^{\frac{1}{\beta-1}} \tag{10}$$

$$m_{i\emptyset} = \left(\frac{\tau_i}{\beta}\right)^{\frac{1}{\beta-1}} \left(\frac{1}{\sum\limits_{h=1}^{\mathcal{H}} r_h \delta_h^2}\right)^{\frac{1}{\beta-1}} \tag{11}$$

Using (9)–(11), we can obtain the updating (4)–(5).

Updating. r With $\mathbf{M}$ and $\mathbf{V}$ fixed, $\mathbf{r}$ is computed by Theorem 2.

Theorem 2. *Suppose* $\mathbf{M} = \mathbf{M}^*$ *and* $\mathbf{V} = \mathbf{V}^*$ *are fixed,* $\mathcal{J}_{WMVEC}(\mathbf{M}^*, \mathbf{V}^*, \mathbf{r})$ *is minimized iff:*

$$r_h = \frac{\exp\left(\frac{-\sum\limits_{i=1}^{n}\sum\limits_{\substack{A_j \subseteq \Omega \\ A_j \neq \emptyset}} |A_j|^\alpha m_{ij}^\beta \|\mathbf{x}_{i,h}-\bar{\mathbf{v}}_{j,h}\|^2 + \sum\limits_{i=1}^{n}\delta_h^2 m_{i\emptyset}^\beta}{\lambda}\right)}{\sum\limits_{g=1}^{\mathcal{H}}\exp\left(\frac{-\sum\limits_{i=1}^{n}\sum\limits_{\substack{A_j \subseteq \Omega \\ A_j \neq \emptyset}} |A_j|^\alpha m_{ij}^\beta \|\mathbf{x}_{i,g}-\bar{\mathbf{v}}_{j,g}\|^2 + \sum\limits_{i=1}^{n}\delta_g^2 m_{i\emptyset}^\beta}{\lambda}\right)} \tag{12}$$

Proof. Lagrange multiplier σ is used to solve the constrained minimization problem with respect to $\mathbf{r}$. The objective function (2) becomes:

$$\mathcal{J}(\mathbf{r}, \sigma) = \mathcal{J}_{WMVEC}(\mathbf{M}, \mathbf{V}, \mathbf{r}) - \sigma\left(\sum_{h=1}^{\mathcal{H}} r_h - 1\right) \tag{13}$$

By differentiating the Lagrangian with respect to the r_h and σ setting the derivatives to zero, we obtain:

$$\frac{\partial \mathcal{J}}{\partial r_h} = \sum_{i=1}^{n}\sum_{\substack{A_j \subseteq \Omega \\ A_j \neq \emptyset}} |A_j|^\alpha m_{ij}^\beta \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2 + \sum_{i=1}^{n}\delta_h^2 m_{i\emptyset}^\beta + \lambda(1 + \ln r_h) - \sigma = 0$$

$$\tag{14}$$

$$\frac{\partial \mathcal{J}}{\partial \sigma} = \sum_{h=1}^{\mathcal{H}} r_h - 1 = 0 \tag{15}$$

We thus have from (14):

$$
r_h = \exp\left(\frac{\sigma - \lambda}{\lambda}\right) \exp\left(\frac{-\sum\limits_{i=1}^{n} \sum\limits_{\substack{A_j \subseteq \Omega \\ A_j \neq \emptyset}} |A_j|^{\alpha} m_{ij}^{\beta} \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2 + \sum\limits_{i=1}^{n} \delta_h^2 m_{i\emptyset}^{\beta}}{\lambda}\right)
$$
(16)

Using (15)–(16), we can obtain the updating (12).

***Updating.* V *while fixing* M *and* r** With $\mathbf{M}$ and $\mathbf{r}$ fixed, $\mathbf{V}$ is computed by Theorem 3.

Theorem 3. *Suppose* $\mathbf{M} = \mathbf{M}^*$ *and* $\mathbf{r} = \mathbf{r}^*$ *are fixed,* $\mathcal{J}_{WMVEC}(\mathbf{M}^*, \mathbf{V}, \mathbf{r}^*)$ *is minimized iff:*

$$
B_{lq,h} \triangleq \sum_{i=1}^{n} x_{iq,h} \sum_{A_j \ni \omega_l} |A_j|^{\alpha-1} m_{ij}^{\beta}, \quad \forall\, l = 1, c, \quad \forall\, q = 1, \mathcal{P}_h
$$
(17)

$$
H_{lk,h} \triangleq \sum_{i=1}^{n} \sum_{A_j \supseteq \{\omega_k, \omega_l\}} |A_j|^{\alpha-2} m_{ij}^{\beta}, \quad \forall\, k, l = 1, c
$$
(18)

$$
\mathbf{H}_h \mathbf{V}_h = \mathbf{B}_h
$$
(19)

where $B_{lq,h}$ *(resp.* $H_{lk,h}$*) is the* l, q-*th (resp.* l, k-*th) element of the matrix* $\mathbf{B}_h \in \mathbb{R}^{c \times \mathcal{P}_h}$ *(resp.* $\mathbf{H}_h \in \mathbb{R}^{c \times c}$*).*

Proof. The partial derivatives of $\mathcal{J}_{WMVEC}$ with respect to the cluster centers are given by:

$$
\frac{\partial \mathcal{J}}{\partial \mathbf{v}_{l,h}} = \sum_{i=1}^{n} \sum_{A_j \neq \emptyset} |A_j|^{\alpha} r_h m_{ij}^{\beta} \frac{\partial D_{ij,h}^2}{\partial \mathbf{v}_{l,h}}
$$
(20)

$$
\frac{\partial D_{ij,h}^2}{\partial \mathbf{v}_{l,h}} = -\frac{2}{|A_j|} \left(\mathbf{x}_{i,h} - \frac{1}{|A_j|} \sum_{A_j \ni \omega_l} \mathbf{v}_{l,h} \right)
$$
(21)

where

$$
D_{ij,h}^2 = \|\mathbf{x}_{i,h} - \bar{\mathbf{v}}_{j,h}\|^2
$$
(22)

Setting these derivatives to zero gives c linear equations that can be written as:

$$
\sum_{i=1}^{n} \mathbf{x}_{i,h} \sum_{A_j \ni \omega_l} |A_j|^{\alpha-1} r_h m_{ij}^{\beta} = \sum_{k=1}^{c} \mathbf{v}_{k,h} \sum_{i=1}^{N} \sum_{A_j \supseteq \{\omega_k, \omega_l\}} |A_j|^{\alpha-2} r_h m_{ij}^{\beta}
$$
(23)

The system of linear equations can be equally represented by (17) and (18).

For the convenience of implementation, the proposed WMVEC method is outlined in Algorithm 1.

Algorithm 1: Weighted Multi-View Evidential Clustering

Input: Multi-view data $\mathcal{X} = \{\mathcal{X}_1, \cdots, \mathcal{X}_{\mathcal{H}}\}$, $\mathcal{X}_h \in \mathbb{R}^{\mathcal{P}_h \times n}$, number of clusters c, and parameters α, β, λ, δ

Output: The optimal $\mathbf{M}$, $\mathbf{V}$, $\mathbf{r}$

1 Initialize $\mathbf{V}^0$, and $\mathbf{r}^0$;

2 while *not converge* **do**

3   Compute the prototypes of meta-clusters utilizing (3);

4   Update the credal partition matrix $\mathbf{M}$ utilizing (4)-(5);

5   Update the view weight vector $\mathbf{r}$ utilizing (12);

6   Compute matrices $\mathbf{B}$ and $\mathbf{H}$ utilizing (17)-(18);

7   Update the cluster prototype matrix $\mathbf{V}$ utilizing (19);

8 end

2.3 Computational Complexity

The computational complexity of ECM is $O(tn2^c)$, where t is the number of iterations, n and c represent the number of objects and clusters respectively. WMVEC is an extension of ECM in multi-view scenarios by introducing view weights. Therefore, the computational complexity of WMVEC is $O(t\mathcal{H}n2^c)$, and $\mathcal{H}$ is the number of views. It is worth noting that the number of clusters grows exponentially with the computational complexity, and the assignment of objects to high-cardinality meta-clusters is challenging to interpret in practice. Therefore, we can consider imposing constraints to limit the number of focal elements in the meta-cluster, for example, like in ECM, to keep only the meta-cluster containing two clusters. For higher cluster datasets, we can restrict the framework of discernment to consist of only singleton clusters, empty set, and Ω. By doing so, the computational complexity of WMVEC can be reduced to $O(t\mathcal{H}nc^2)$ and $O(t\mathcal{H}n(c+2))$.

3 Experiments

3.1 Experiment on Iris Dataset

In this experiment, we demonstrate the ability of credal partition in WMVEC on the Iris[1] dataset. The Iris dataset contains 150 objects with 3 clusters, where each object consists of four features (Sepal.Length, Sepal.Width, Petal.Length and Petal.Width). We split the Iris dataset into three views by different features, as shown in Fig. 1(a), (e), (i). We can see that $\{\omega_1\}$ is clearly distinguished from $\{\omega_2\}$ and $\{\omega_3\}$, whereas $\{\omega_2\}$ and $\{\omega_3\}$ are partially overlapped, and view 3 is much easier to distinguish between $\{\omega_2\}$ and $\{\omega_3\}$ than the other two views.

From Fig. 1(b)–(d), (f)–(h), (j)–(l), we can observe that WMVEC can clearly characterize the imprecision and uncertainty of cluster assignment. It carefully assigned indistinguishable objects to relevant meta-cluster (*i.e.* $\{\omega_2, \omega_3\}$). What

---

[1] https://archive.ics.uci.edu/ml/datasets/Iris/.

is more, the parameter α determines the size of the meta-cluster. A large α tends to correspond to lower imprecision and higher error, and vice versa. In applications, α should be adjusted according to the specific situation or user preferences. In the following experiments, we take $\alpha = 2$, as suggested in [9].
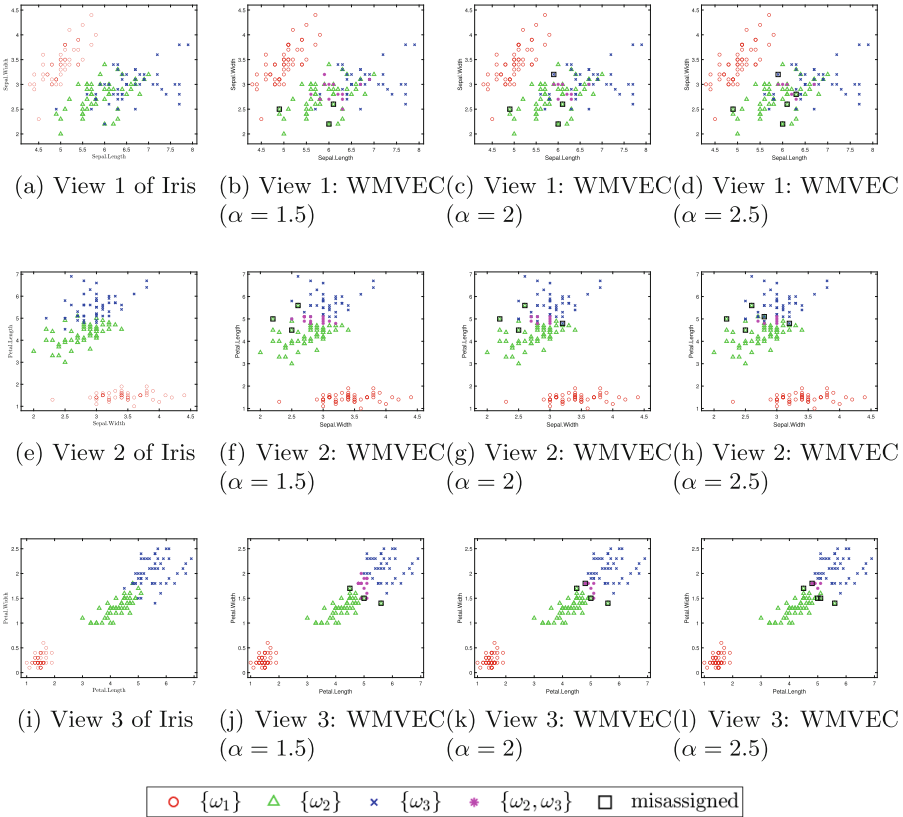

(a) View 1 of Iris   (b) View 1: WMVEC(c) View 1: WMVEC(d) View 1: WMVEC
                        $(\alpha = 1.5)$ $(\alpha = 2)$ $(\alpha = 2.5)$

(e) View 2 of Iris   (f) View 2: WMVEC(g) View 2: WMVEC(h) View 2: WMVEC
                        $(\alpha = 1.5)$ $(\alpha = 2)$ $(\alpha = 2.5)$

(i) View 3 of Iris   (j) View 3: WMVEC(k) View 3: WMVEC(l) View 3: WMVEC
                        $(\alpha = 1.5)$ $(\alpha = 2)$ $(\alpha = 2.5)$

$\bigcirc$ $\{\omega_1\}$ $\triangle$ $\{\omega_2\}$ $\times$ $\{\omega_3\}$ $*$ $\{\omega_2, \omega_3\}$ $\square$ misassigned

Fig. 1. Clustering results of Iris dataset.

3.2   Experiments on Real-World Multi-view Datasets

To verify the validity of the proposed method with respect to other methods, we constructed experiments on eight public multi-view datasets, including WebKB[2], Multiple features[3], Prokaryotic Phyla [1], SensIT Vehicle[4], HumanEva 3D Motion [1] and MSRCv1 [14]. Table 1 shows the basic information of the

used multi-view datasets, including the numbers of objects, clusters, views and features.

Table 1. The details of the used real-world datasets

Dataset	Objects	Clusters	Views	Views name	Features
WebKB (WebKB)	203	4	3	The text on web pages	1703
				The anchor text in hyperlinks	203
				The text in the title	203
Multiple features (MF)	2000	10	6	Fourier coefficients	76
				Profile correlations	216
				Karhunen-Love coefficients	64
				Pixel averages	240
				Zernike moments	47
				Morphological features	6
Prokaryotic Phyla (ProP)	551	4	3	Gabor	393
				Wavelet moments	3
				CENTRIST	438
SensIT Vehicle (SensIT)	300	3	2	Acoustic	50
				Seismic	50
HumanEva 3D Motion (Motion)	5000	5	2	X1_person	48
				X2_person	48
MSRCv1 (MSRCv1)	210	7	4	Color moment	24
				GIST	512
				LBP	256
				CENTRIST	254

We compare the performance of the proposed WMVEC method with respect to a number of related methods, which are shown as follows: FCM [11], ECM [9], Co-FKM [4], RMVKM [2], TW-k-means [3], SMVF [6], TW-Co-k-means [15], Co-FW-MVFCM [14] and MVASM [5]. Among them, FCM and ECM are two classical single-view clustering methods, RMVKM, TW-k-means, SMVF and TW-Co-k-means are representative multi-view clustering methods based on hard partition, Co-FKM, Co-FW-MVFCM and MVASM are state-of-art methods based on fuzzy partition. In order to fully evaluate the performance of the above mentioned methods, we applied two popular performance metrics[5], namely Accuracy (ACC) and Rand Index (RI). Note that the larger values of ACC and RI correspond to better clustering performance.

Table 2 and Table 3 show the clustering results of WMVEC and other methods on real-world datasets. Specifically, several points can be drawn from the observation.

---

[5] For credal partition, we can obtain a hard or fuzzy partition by Pignistic probability transformation [9].

1. Compared with single-view clustering methods (*i.e.* FCM and ECM), WMVEC generally has obvious advantages, which indicates that it is not appropriate to use single-view clustering methods to deal with multi-view data. Besides, we can see that on the second view of the Prop dataset, FCM and ECM both achieve relatively good results, however, these results are physically meaningless.
2. Compared with the multi-view clustering method, the proposed WMVEC achieves better performance compared to other methods. In particular, the clustering methods considering view weights can often obtain better results, which further demonstrates the rationality of introducing view weights.
3. More importantly, the advantage of WMVEC is not only that it can be transformed into hard or fuzzy partition to obtain outstanding performance, but WMVEC with credal partition can also offer deeper insight into the data. It can effectively characterize the uncertainty and imprecision in cluster assignment and avoid the risk of misassignment, which is very valuable in some cautious decision-making applications.

Table 2. Clustering results with various multi-view datasets on *ACC*

Method	Dataset					
	WebKB	MF	ProP	SensIT	Motion	MSRCv1
FCM(1)	0.5110	0.2005	0.4781	0.4699	0.6785	0.2877
FCM(2)	0.4923	0.1971	0.5448	0.3687	0.4246	0.2912
FCM(3)	0.7340	0.2859	0.3354	-	-	0.3798
FCM(4)	-	0.6643	-	-	-	0.3310
FCM(5)	-	0.2027	-	-	-	-
FCM(6)	-	0.2006	-	-	-	-
ECM(1)	0.6124	0.3402	0.5384	0.5097	0.4855	0.2861
ECM(2)	0.5312	0.2766	0.5783	0.3720	0.4832	0.4204
ECM(3)	0.6158	0.3429	0.5515	-	-	0.3692
ECM(4)	-	0.3564	-	-	-	0.3418
ECM(5)	-	0.3623	-	-	-	-
ECM(6)	-	0.3190	-	-	-	-
Co-FKM	0.4709	0.5240	0.4773	0.4567	0.5288	0.4440
RMVKM	0.5418	0.7926	0.5235	0.4769	0.7994	0.5437
TW-k-means	0.5387	0.7609	0.5576	0.4902	0.6959	0.6052
SWVF	0.5468	0.7745	0.5620	0.5075	0.7056	0.6121
TW-Co-k-means	0.6109	0.8188	0.5661	0.5139	0.7124	0.6335
Co-FW-MVFCM	0.6946	0.7410	0.5408	0.5133	0.6870	0.5857
MVASM	0.7161	0.6310	0.4993	0.4796	0.7347	0.5381
WMVEC	**0.7586**	**0.8300**	**0.5789**	**0.5667**	**0.8952**	**0.7619**

3.3   Parameter Study

Finally, we also tested the parameter sensitivity of WMVEC. The *ACC* and *RI* results with different parameters β and λ are shown in Fig. 2 and Fig. 3. We take β that varies in $\{1.1, 1.2, \cdots, 2.0\}$ and λ that varies in $\{e^0, e^1, \cdots, e^{10}\}$. It can be seen that the values of β and λ have a significant impact on the performance of the WMVEC. For each dataset, the WMVEC can gain better results if appropriate β and λ can be selected. In addition, we can clearly see that the WMVEC can often obtain better results with smaller β and larger λ.

Table 3. Clustering results with various multi-view datasets on *RI*

Method	Dataset					
	WebKB	MF	ProP	SensIT	Motion	MSRCv1
FCM(1)	0.6206	0.5733	0.5709	0.5279	0.8217	0.6566
FCM(2)	0.5110	0.5527	0.6360	0.5121	0.6301	0.5539
FCM(3)	0.6872	0.6921	0.5449	-	-	0.7573
FCM(4)	-	0.9144	-	-	-	0.7591
FCM(5)	-	0.5718	-	-	-	-
FCM(6)	-	0.5519	-	-	-	-
ECM(1)	0.5483	0.7265	0.4242	0.5924	0.6702	0.6327
ECM(2)	0.4042	0.5622	**0.6442**	0.5287	0.6639	0.6993
ECM(3)	0.5398	0.7271	0.4121	-	-	0.6961
ECM(4)	-	0.7584	-	-	-	0.6882
ECM(5)	-	0.7701	-	-	-	-
ECM(6)	-	0.7368	-	-	-	-
Co-FKM	0.6174	0.8714	0.6041	0.5781	0.7667	0.7879
RMVKM	0.6678	0.9499	0.6151	0.5879	0.8969	0.8305
TW-k-means	0.4131	0.9453	0.6356	0.5516	0.8453	0.8500
SWVF	0.4078	0.9347	0.5411	0.5776	0.8420	0.8490
TW-Co-k-means	0.5464	0.9590	0.6374	0.5689	0.8592	0.8608
Co-FW-MVFCM	0.6410	0.9231	0.6280	0.5873	0.8313	0.8535
MVASM	0.7097	0.9059	0.5832	0.5494	0.8482	0.8345
WMVEC	**0.7361**	**0.9560**	0.6414	**0.6029**	**0.9282**	**0.8929**

(a) WebKB   (b) SensIT   (c) MSRCv1
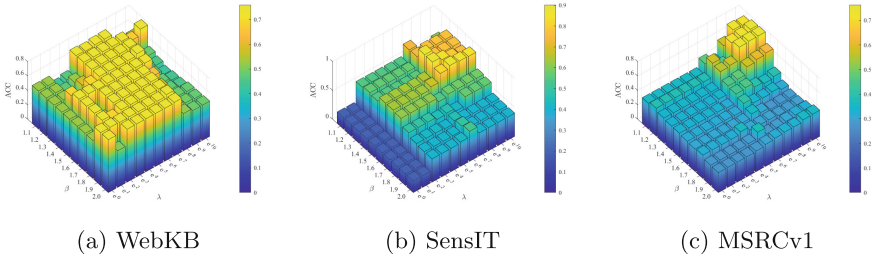
Fig. 2. Clustering results of multi-view datasets in terms of ACC.

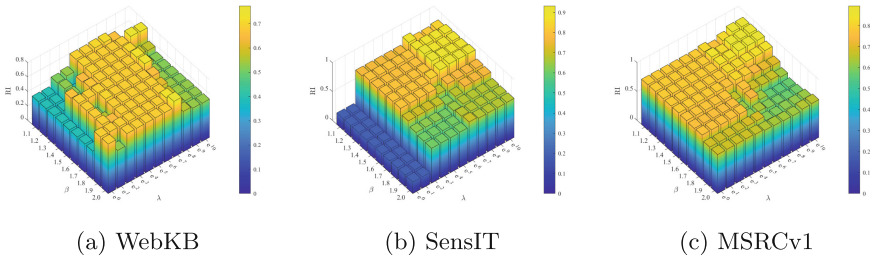(a) WebKB   (b) SensIT   (c) MSRCv1

Fig. 3. Clustering results of multi-view datasets in terms of RI.

4 Conclusion

In this paper, we propose an adaptive weighted multi-view evidential clustering (WMVEC) based on the framework of belief functions. As such, WMVEC is able to handle the uncertainty and imprecision of multi-view data clustering, and provides a credal partition, which extends the fuzzy, possibilistic and rough ones, allowing the representation of the imprecision and uncertainty about the assignment of objects to clusters. What is more, WMVEC can automatically learn the importance of views. Experimental results on several real-world datasets illustrate the potential and superiority of our proposed method. In future studies, we intend to extend the proposed method to deal with mixed types of data as well as incomplete data.

References

1. Benjamin, J.B.M., Yang, M.S.: Weighted multiview possibilistic c-means clustering with L2 regularization. IEEE Trans. Fuzzy Syst. **30**(5), 1357–1370 (2022)
2. Cai, X., Nie, F., Huang, H.: Multi-view k-means clustering on big data. In: IJCAI (2013)
3. Chen, X., Xu, X., Huang, J.Z., Ye, Y.: TW-k-means: automated two-level variable weighting clustering algorithm for multiview data. IEEE Trans. Knowl. Data Eng. **25**(4), 932–944 (2013)

4. Cleuziou, G., Exbrayat, M., Martin, L., Sublemontier, J.H.: CoFKM: a centralized method for multiple-view clustering. In: ICDM, pp. 752–757 (2009)

5. Han, J., Xu, J., Nie, F., Li, X.: Multi-view k-means clustering with adaptive sparse memberships and weight allocation. IEEE Trans. Knowl. Data Eng. **34**(2), 816–827 (2022)

6. Jiang, B., Qiu, F., Wang, L.: Multi-view clustering via simultaneous weighting on views and features. Appl. Soft Comput. **47**, 304–315 (2016)

7. Jiang, Y., Chung, F.L., Wang, S., Deng, Z., Wang, J., Qian, P.: Collaborative fuzzy clustering from multiple weighted views. IEEE Trans. Cybern. **45**(4), 688–701 (2015)

8. Liu, Z.: An effective conflict management method based on belief similarity measure and entropy for multi-sensor data fusion. Artif. Intell. Rev. 1–28 (2023)

9. Masson, M.H., Denœux, T.: ECM: an evidential version of the fuzzy C-means algorithm. Pattern Recognit. **41**(4), 1384–1397 (2008)

10. Oyewole, G.J., Thopil, G.A.: Data clustering: application and trends. Artif. Intell. Rev. **56**(7), 6439–6475 (2023)

11. Ruspini, E.H., Bezdek, J.C., Keller, J.M.: Fuzzy clustering: a historical perspective. IEEE Comput. Intell. Mag. **14**(1), 45–55 (2019)

12. Su, Z., Denœux, T.: BPEC: belief-peaks evidential clustering. IEEE Trans. Fuzzy Syst. **27**(1), 111–123 (2019)

13. Xu, R., Wunsch, D.: Survey of clustering algorithms. IEEE Trans. Neural Netw. **16**(3), 645–678 (2005)

14. Yang, M.S., Sinaga, K.P.: Collaborative feature-weighted multi-view fuzzy c-means clustering. Pattern Recognit. **119**, 108064 (2021)

15. Zhang, G., Wang, C., Huang, D., Zheng, W., Zhou, Y.: TW-co-k-means: two-level weighted collaborative k-means for multi-view clustering. Knowl.-Based Syst. **150**, 127–138 (2018)