DYNAMIC METRIC ACCELERATED METHOD FOR FUZZY CLUSTERING

WEI CAI¹, SHENGBING XU¹,²*, HONGXI XIA¹, BO LIU³, JIE XU¹
1 School of Applied Mathematics, Guangdong University of Technology, Guangzhou 510520, China
2 School of Computers, Guangdong University of Technology, Guangzhou, 510006, China
3 School of Automation, Guangdong University of Technology, Guangzhou, 510006, China
*Corresponding author: Shengbing Xu (e-mail: xushengbing@gdut.edu.cn).

ABSTRACT Features in data samples usually need a unified dimension by a standardization process before clustering. However, there still exists a nonstandardized metric in which the distance between samples is greater than 1 after features are standardized. It is difficult to find the optimal search path if the data sample metrics are not standardized.

To address this problem, we develop a dynamic-metric accelerated method for fuzzy clustering by introducing a metric matrix, whose diagonal elements consist of infinite norms of the metric matrix into the Fuzzy C-Means (FCM) clustering algorithm and its derived algorithms. More specifically, we focus on constructing a dynamic metric matrix that is used to unify the metric between data samples and updating cluster centers to optimize the search path of the cluster center. In addition, we propose a new evaluation index named the Coefficient of Variation Metric (CVM) to evaluate metric effectiveness.

The dynamic metric accelerated method, whose complexity remains unchanged, can effectively accelerate the iteration speed of fuzzy clustering. The comparisons between the algorithm using the dynamic metric accelerated method and the corresponding algorithm on UCI, business district and COVID-19 CT image datasets show the superiority of the dynamic metric accelerated method in accelerating effect and clustering performance.

INDEX TERMS Dynamic Metric, Fuzzy clustering, Iteration acceleration, Coefficient of variation metric

¹ This work was supported in part by the Natural Science Foundation of China under Grant 61876044 (Algorithm research and application of positive sample and unmarked sample learning).
1. INTRODUCTION

Generally, clustering depends on a metric that describes the similarity between data samples[1][2][3][4]. Metrics can express rich information, and metric-based clustering approaches in which the metrics of data samples are the main component of objective functions have been widely used in industrial applications[5]. For the above reasons, research on clustering algorithms focuses on algorithm design and metric research[6].

What's more, Metric plays an more important part in fuzzy clustering. There are many metrics in fuzzy clustering, such as Euclidean distance[7], Minkowski distance [8], Manhattan distance [9], Chebyshev distance [10], Mahalanobis distance [11] [12] [13], angular cosine [14] [15] [16], correlation coefficient [17], entropy [18] [19] [20], and Hamming distance [21]. In addition, in recent years, many studies have been done on metric learning of the kernel method [22] [23] [24] [25] [26]. The learning process of the fuzzy clustering algorithm depends on the metric that describes associations between different objects. Fuzzy c-means (FCM) uses Euclidean distance to measure relationship between samples and distance centers[27]; SFCM measures relationship between data samples, distance centers and guidance information by using Euclidean distance[28]; eSFCM measures data samples and cluster centers and entropy to measure guidance information by using Euclidean distance[29]; SMUC measures relationship between samples and distance centers and entropy to measure guidance information by using Mahalanobis distance[30]; In addition, there are some kernel methods [31][32]. However, the learning process of the relationship between data samples will be affected due to the limitations of metrics. The clustering algorithms based on Euclidean distance have the advantages of fast iteration convergence and stable results [33]. However, in the face of complex data in practical applications, these clustering algorithms are sensitive to data sample dimensions and cannot apply to nonlinear data [34]. The introduction of the Mahalanobis distance can solve these problems, but it will also lead to the exponential growth of algorithm computation requirements [30].

In the past, research mainly focuses on three kinds of metrics in fuzzy clustering: Entropy [35][36][37][38], Euclidean distance [39][40][41][42], and Mahalanobis distance [43][44][45]. Relative entropy measures the difference of data from the perspective of probability distribution and uncertainty [51][52][53]; it has the advantage of being insensitive to common noises [54][55]. Euclidean distance is the most popular metric in objective functions of fuzzy clustering because it can reflect the real distance in the sample space[46]. However, there exists a case that Euclidean distance ignores the differences among different features in the sample, which makes it difficult to reflect the associations between data samples[47]. The Mahalanobis distance can better reflect the correlation between data samples, so the process of fuzzy clustering is not affected by the feature dimensions, but the calculation of the inverse matrix of the covariance matrix in Mahalanobis distance greatly increases the complexity of the calculation[48][49][50]. In addition, none of the three metrics mentioned above will change with the iteration. We call this kind of metrics the static metrics. In our research, the standardization of static metrics can achieve normalization at the feature level, but can not guarantee the normalization of distance. This paper analyzes this phenomenon and proposes an algorithm acceleration method based on dynamic measurement.

Our work is summarized as follows:

1) We introduce distance standardization and how it affects the clustering process using static metric algorithms

2) We propose a dynamic metric accelerated method for fuzzy clustering that has better adaptability than the traditional static metric. In addition, we analyze the time complexity and optimality of metric

3) For the dynamic metric, we define a Coefficient of Variation Metric (CVM) to evaluate the effectiveness of the metric.

4) We use the dynamic metric accelerated method to improve the metric effect in clustering and set experiments on the classic UCI dataset, classic image dataset, business circle dataset [56] and COVID-19-CT dataset [57] to verify the effect of the dynamic metric accelerated method.

II. RELATED WORK

A. FUZZY C-MEANS CLUSTERING (FCM)

FCM is a widely used unsupervised fuzzy clustering algorithm. The Euclidean distance which is used in FCM is widely used in clustering algorithms.

Let us first assume that the sample set to be clustered is $X = \{x_1, x_2, \ldots, x_n\}$, where $x_j \in \mathbb{R}^d$ ($1 \leq j \leq n$) in the d-dimensional Euclidean space, and $c$ is the number of clusters. The objective function of FCM can be expressed as [27]:

$$J(U,V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m (x_j - v_i)^T I(x_j - v_i)$$  \hspace{1cm} (1)

For the convenience of introduction, the objective function is expanded as follows:

$$J(U,V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m (x_j - v_i)^T A(x_j - v_i)$$  \hspace{1cm} (2)

where $I$ is the identity matrix, $A$ is a metric matrix , only when $A = I$ , eq.2 is equivalent to eq.1 and the corresponding metric is Euclidean metric.
\( m \) is any real number \((m > 1)\) which denotes the degree of fuzziness, \( u_j \) is the membership degree of the \( j \)-th sample \( x_j \) belonging to the \( i \)-th cluster whose centroid is \( v_i \), \( U = \{u_j\} \), \( V = [v_1, v_2, \cdots, v_c] \), \( 1 \leq i \leq c \), \( 1 \leq j \leq n \), \( 2 \leq c < n \), \( u_j = (\mu_{1j}, \cdots, \mu_{cj})^T \) and \( u_j \) satisfies the following constraint condition:

\[
\sum_{j=1}^{c} u_{ij} = 1, \quad u_{ij} \geq 0
\]  

(3)

By minimizing (1) and using the Lagrange optimization, we obtain the following alternative update equations for the cluster center \( v_j \) and the membership degree \( u_{ij} \):

\[
v_j = \frac{\sum_{j=1}^{n} u_j^m x_j}{\sum_{j=1}^{n} u_j^m}
\]  

(4)

\[
u_{ij} = \left( \frac{(x_j - v_i)^T A(x_j - v_i)}{\sum_{h=1}^{c} (x_j - v_h)^T A(x_j - v_h)} \right)^{\frac{2}{m-1}}
\]  

(5)

**B. KERNEL-BASED FUZZY C-MEANS CLUSTERING OF METRIC ACCELERATED (KFCM)**

KFCM is an unsupervised fuzzy clustering algorithm that projects the original data into the kernel space and takes the inner product of the kernel space as the metric of the algorithm.

Let \( V = [v_1, v_2, \cdots, v_c] \) be the \( V \) cluster centers in the kernel space, \( c \) is the number of clusters. This minimizes the following objective function subject to conditions as considered in FCM:

\[
J_{KFCM}(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m (\phi(x_j) - \phi(v_i))^T I(\phi(x_j) - \phi(v_i))
\]  

(7)

\[
= \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m (2 - 2 \cdot K(x_j, v_i))
\]

where \( K(x_j, v_i) \) is the Gaussian radial basis function, and its form is as follows:

\[
K(x_j, v_i) = e^{-\frac{|x_j - v_i|^2}{2\sigma^2}}
\]  

(8)

In the iterative process of KFCM, the memberships and the cluster centers are updated as follows:

\[
u_{ij} = \frac{1}{\sum_{k=1}^{n} (1 - K(x_j, v_k))^{m-1}} \sum_{k=1}^{n} \mu_{ik}^m K(x_j, v_k) x_j
\]  

(9)

\[
v_i = \frac{\sum_{j=1}^{n} u_{ij}^m K(x_j, v_i) x_j}{\sum_{j=1}^{n} u_{ij}^m K(x_j, v_i)}
\]  

(10)

**C. SEMI-SUPERVISED FUZZY C-MEANS CLUSTERING (SFCM)**

SFCM is a semi-supervised fuzzy clustering algorithm that utilizes Euclidean distance to express the relationship of the membership matrix and the prior membership matrix. Its objective function is as follows:

\[
J_{SFCM}(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m (x_j - v_i)^T I(x_j - v_i)
\]

\[
+ \alpha \sum_{i=1}^{c} \sum_{j=1}^{n} (u_{ij} - \bar{u}_{ij}) (x_j - v_i)^T I(x_j - v_i)
\]

(11)

where \( m (m > 1) \) denotes the degree of fuzziness. As \( m \) tends toward 1, SFCM approaches HCM. \( u_{ij} (0 \leq u_{ij} \leq 1) \) is the membership degree of the \( j \)-th sample \( x_j \) belonging to the \( i \)-th cluster and \( v_i \) is its centroid. \( V = [v_1, v_2, \cdots, v_c] \), \( 1 \leq i \leq c \), \( 1 \leq j \leq n \), \( 2 \leq c \leq n \), \( u_j = (u_{1j}, \ldots, u_{cj}) \), and \( u_{ij} \) satisfies the following constraint condition:

\[
s.t. \quad u_{ij} \in [0,1], \sum_{i=1}^{c} u_{ij} = 1, (j = 1, 2, \ldots, n)
\]  

(12)

\( m \) is the weighted index, which is an empirical value, and the value is usually 2. Then, we obtain the iterative formula of the membership degree and clustering center:

\[
u_{ij} = \frac{1}{\sum_{k=1}^{n} (1 - K(x_j, v_k))^{m-1}} \sum_{k=1}^{n} \mu_{ik}^m K(x_j, v_k) x_j
\]

\[
v_i = \frac{\sum_{j=1}^{n} u_{ij}^m K(x_j, v_i) x_j}{\sum_{j=1}^{n} u_{ij}^m K(x_j, v_i)}
\]
\[
    u_{ij} = \frac{1}{1 + \alpha} \left[ 1 + \alpha \left(1 - \sum_{h=1}^{c} u_{ij}^h \right)^{m-1} \right] \frac{(x_j - v_i)^T I(x_j - v_i)}{\sum_{h=1}^{c} (x_j - v_h)^T I(x_j - v_h)} + \alpha \tilde{u}_{ij}
\]

(13)

\[
    v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^m x_j}{\sum_{j=1}^{n} \mu_{ij}^m}
\]

(14)

\[
    \tilde{u}_{ij} \text{ is an a priori membership matrix, which is transformed from label information; } \alpha \text{ is a predetermined suppression coefficient.}
\]

**D. ENTROPY SEMI-SUPERVISED FUZZY C-MEANS CLUSTERING (ESFCM)**

ESFCM is a semi-supervised fuzzy clustering algorithm that utilizes information entropy to express the relationship of the membership matrix and the prior membership matrix. Its objective function is as follows:

\[
    J(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}(x_j - v_i)^T I(x_j - v_i) + \lambda \sum_{i=1}^{c} \sum_{j=1}^{n} \left( u_{ij} - \tilde{u}_{ij} \right) \ln(u_{ij} - \tilde{u}_{ij})
\]

(18)

where is subject to conditions (2) and (6) and the following conditions:

\[
    M = \left[ \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij} u_{ij} x_j - \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^2 u_{ij}^2 x_j - \frac{1}{n} \sum_{j=1}^{n} \mu_{ij}^2 u_{ij}^2 v_i \right]^T
\]

(19)

There is the following optimal solution [30]:

\[
    v_i = \frac{\sum_{j=1}^{n} u_{ij} x_j}{\sum_{j=1}^{n} u_{ij}}
\]

(21)

\[
    u_{ij} = \tilde{u}_{ij} + \frac{e^{-\lambda (x_j - v_i)^T I(x_j - v_i)}}{\sum_{h=1}^{c} e^{-\lambda (x_j - v_h)^T I(x_j - v_h)}} (1 - \sum_{h=1}^{c} \tilde{u}_{ij})
\]

(22)

**III. ACCELERATED ITERATIVE METHOD FOR FUZZY CLUSTERING**

**A. MOTIVATION:**

At present, clustering algorithms usually need to standardize sample features to eliminate the influence of different feature dimensions before conducting iterative calculations. After data normalization, the value of each sample feature is limited to [0,1] for iterative calculations, but the value of the sample distance will exceed [0,1] after calculations, as shown in Figure 1:

![Diagram](image-url)
Figure 1. The numerical range is beyond [0,1] in the iterative process.

It can be seen from Figure 1 that the standardized feature does not mean that distance (i.e., metric) between samples is also standardized.

To verify the necessity of standardizing metrics, we built artificial datasets. The distance between samples in dataset (a) was less than 1, and some of the distances in the dataset (b) were more than 1 (as shown in Figures 2 and 3):

Figure 2. Synthetic dataset\(^2\) in which the distance between samples is less than 1.

Figures 2 and 3 represent the simulation dataset of dataset 1 and dataset 2, respectively, whose feature number is 2 (two-dimensional space) and sample number is 100.

Figure 3. Synthetic dataset in which the distance between samples exceeds 1.

Let us assume that the number of cluster centers is 2, the initial cluster center is \([0.3, 0.5; 7, 0.5]\), the maximum number of iterations of clustering in dataset (a) is 50, and the solution path-length of the two clustering centers is 0.0029. The maximum number of iterations of clustering in dataset (b) is 52, and the solution path length is 0.0097.

If the sample exhibits the phenomenon in Figure 3, then after the usual iterative process, the search path may become longer and indirect, as shown in Figure 4:

\[ J(v_i) \]

(a) \[ J(v_i) \]

(b)

Figure 4. Two possible cases caused by traditional methods (longer search distance).

We are committed to making the solving path closer to Figure 4 (a) instead of Figure 4 (b).

8. Metric Complexity

The Mahalanobis distance, which has a good effect on eliminating the dimensional effect, is commonly used to metric the continuous type of feature. At present, the Mahalanobis distance has a good influence on data of different dimensions, so the Mahalanobis distance is widely used for metric data of different dimensions. However, it is difficult for the Mahalanobis distance to apply to image recognition and other fields because it needs to calculate the covariance matrix and its inverse \([30]\), resulting in a large amount of calculations (as shown in Figure 5):

Figure 5. Comparison of computational complexity (time-consuming) between the Mahalanobis metric and Euclidean metric.

\(^2\)https://github.com/ChoiNgai/paper_DynamicMetricClustering/tree/main/data/4.18
In Figure 5, the number of randomly generated samples is fixed at 1000, and the number of features is from 1 to 1000. The time consumption of calculating the sample and distance and the sample data and the randomly generated $3 \times d$ size matrix is gradually calculated. Where the number of features is $d$. With the increase of features number, the amount of computation increases exponentially.

**C. DYNAMIC STANDARDIZATION OF EUCLIDEAN METRIC**

In the iterative process of the algorithm, there exists a case in which the distance between the sample and the cluster center is greater than 1, so it is necessary to compress the distances so that all the distances after the transformation are in $[0,1]$ (as shown in Figure 6).

![Figure 6. Distance standardization process (compress)](image)

When the distance between the sample and cluster center is small, the rounding error produced in the calculation process will have a great impact on the clustering process. Therefore, it is necessary to expand the distance by transformation. The effect of stretching is shown in Figure 7:

![Figure 7. Distance standardization process (expand)](image)

In Figures 6 and 7, the left side is a nonstandardized distance, and the right side is a standardized distance. To ensure that the mathematical relationship between distances is unchanged and that the algorithm is fast iterative, we satisfy the distance with the following relationship between the standardized distance and the nonstandardized distance:

$$\frac{d_{i,j}}{d_{i,j}^*} = \frac{d_{i,j}^*}{d_{i,j}^*} \cdot \frac{d_{i,j}^*}{d_{i,j}^*} \cdot \frac{d_{i,j}^*}{d_{i,j}^*} = \frac{d_{i,j}^*}{d_{i,j}^*} \cdot \frac{d_{i,j}^*}{d_{i,j}^*} \cdot \frac{d_{i,j}^*}{d_{i,j}^*}$$

(23)

where $d_{i,j}$ is the nonstandardized distance between $x_j$ and $v_i$; $d_{i,j}^*$ is the standardized distance between $x_j$ and $v_i$.

**D. DYNAMIC METRIC ACCELERATION**

The Euclidean distance between the sample and the cluster center expression in clustering expands from $\| x_j - v_i \|^2$ to $(x_j - v_i)^T A (x_j - v_i)$.

1) When $A = I$, where $I$ is the identity matrix, the corresponding metric is Euclidean distance.

2) When $A = M^{-1}$, where $M$ is from (19), the corresponding metric is the Mahalanobis distance.

3) When $A$ is a diagonal matrix whose elements consist of the positive infinite norm of the distance matrix, the corresponding metric is a dynamic metric whose matrix changes with each iteration. We call the method using dynamic metrics for fuzzy clustering as the dynamic metric accelerated method for fuzzy clustering.

Calculation steps of the metric matrix in the dynamic metric accelerated method:

1) Compute the Euclidean distance between the sample and the cluster center to get the Euclidean distance matrix;

2) Calculate the infinite norm according to the distance matrix;

3) Calculate the dynamic measure matrix $A$ based on the infinite norm.

The Euclidean distance between $x_j$ and $v_i$ is denoted as $D = \{d_{ij}\}$; the infinite norm of the matrix is:

$$\| D \|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} | d_{ij} |$$

(24)

Then, the metric matrix $A$ composed of the infinite norm of the metric matrix is as follows:

$$C = \begin{bmatrix} \| D \|_{\infty} & 0 & \cdots & 0 \\ 0 & \| D \|_{\infty} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \| D \|_{\infty} \end{bmatrix}$$

(25)

where $\| D \|_{\infty}$ is the infinite norm of the metric matrix:

$$A = C^{-1}$$

(26)
The metric matrix $A = \text{diag}(m_1, \ldots, m_d)$, and the expression of element $A$ is:

$$m_k = \frac{1}{\|D\|_{\infty}}$$ (27)

where $d$ is the number of features, $k = \{1, 2, \ldots, d\}$.

Therefore, the distance calculation expression of the dynamic metric is:

$$(x_j - v_i)^T A (x_j - v_i)$$ (28)

### IV. ANALYSIS OF WEIGHTING EXPONENT M IN FCM

The dynamic metric will inevitably affect the membership matrix when the iterative process changes, and the weighting exponent $M$ of the FCM model plays a role in determining the validity of FCM partitions. Therefore, we explore the impact of the weighting exponent $M$ in FCM when using the dynamic metric acceleration method. In this part, the effect of the value of $M$ is analyzed through experiments using the Wisconsin breast cancer dataset\(^3\). The internal evaluation index data reconstruction error rate \(^59\), the external evaluation index purity \(^60\) and the number of iterations are used to analyze the influence of the $m$ value on the dynamic metric and conventional Euclidean metric in FCM.

#### TABLE I

| M value/number of iterations | Euclidean metric | Dynamic metric |
|-----------------------------|------------------|----------------|
| m=2                         | 15               | 11             |
| m=3                         | 14               | 12             |
| m=4                         | 18               | 11             |
| m=5                         | 19               | 11             |
| m=6                         | 21               | 13             |
| m=7                         | 19               | 10             |

#### 2) INTERNAL EVALUATION INDEX:

$V_{RE}$ is the error rate of data reconstruction, which is defined as follows:

$$V_{RE} = \frac{1}{n} \sum_{i=1}^{n} \left\| I_g^j(i) - I_g(i) \right\|^2$$ (31)

It analyzes the difference between reconstructed data and the original data after clustering.

$I_g^j(i)$ is the gray level of the $i$-th sample of the reconstructed data:

$$I_g^j(i) = \frac{\sum_{k=1}^{c} u_{ki} I_g(i)}{\sum_{k=1}^{c} u_{ki}^2}$$ (32)

#### TABLE III

| M value/Vre | Traditional metric | Dynamic metric |
|-------------|--------------------|----------------|
| m=2         | $5.0695 \times 10^{-33}$ | $4.7624 \times 10^{-33}$ |
Tables 1 and 2 show the influence of the m value on the acceleration algorithm. In terms of cluster purity, dynamic metric does not affect the role of the m value in FCM, as shown in Table 1. Table 2 shows that with the increase in the m value, the clustering error rate of the FCM algorithm increases. At the same time, combined with Nikhil R. pal's conclusion[58], we can conclude that M = 2 is a better choice in the dynamic acceleration matrix.

E. EVALUATING INDICATOR AND VERIFICATION

To measure the effectiveness of the standardization of metrics, we analyzed the maximum and minimum distances between data sample and cluster centers.

FCM experiments were carried out on the Wisconsin breast cancer dataset\(^4\) with 2 clusters.

\(^4\)http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29

This process, we define some metric related concepts

- **Definition 1**: Iterative max/min distance
  The maximum / minimum distance from the sample's t-th iteration to \(v_i\) is:
  \[
  d_i^{(t)} = \|x_j - v_i^{(t)}\|^2
  \]
  where \(T\) is the maximum number of iterations, \(1 < t < T, 1 < i < c\).

- **Definition 2**: The extreme deviations of distance
  The extreme deviations of the distance from the sample to \(v_i^{(t)}\) are calculated as follows:
  \[
  r_i^{(t)} = \max_{1 < t < T} d_i^{(t)} - \min_{1 < t < T} d_i^{(t)}
  \]
  \[
  = \max_{1 < t < T} \|x_j - v_i^{(t)}\|^2 - \min_{1 < t < T} \|x_j - v_i^{(t)}\|^2
  \]
  However, due to \(\max d_i^{(t)} >> \min d_i^{(t)}\), the range of distance is mainly controlled by \(\max d_i^{(t)}\). To avoid this case and express the intensity of the transformation between distances better, we normalize it to obtain the improved range. To express the normalization of the distance range conveniently, we define the following function:
  \[
  \Phi(d_i^{(t)}) = \frac{d_i^{(t)} - \min |d_i^{(t)}|}{\max |d_i^{(t)}| - \min |d_i^{(t)}|}
  \]
  The improved extreme deviations of the distance are:
  \[
  R_i^{(t)} = \Phi(\max d_i^{(t)}) - \Phi(\min d_i^{(t)})
  \]
  The improved distance extreme deviations diagram is drawn, as shown in Figure 9:
**Figure 9.** Difference in the distance extreme deviations between different clusters

The difference between the maximum and the minimum of the distance between the sample and each cluster center is increasing. The change process of the distance is shown in Figure 10.

We call the above distance as the **Coefficient of Variation Metric** (CVM). The CVM is used to evaluate this case. Before defining $CVM$, we need to explain the **Variation Metric** $(VM)$.

- **Definition 3: Variation Metric**
  
  As Figure 11 shows, we define the D-value as $VM(t)$, and its mathematical expression is:
  
  $$VM(t) = \max |R(t)| - \min |R(t)|$$  \hspace{1cm} (37)
  
  where $t$ is the number of iterations. The smaller $VM(t)$ is, the smoother the iteration process and the better the normalization effect.

- **Definition 4: Coefficient of Variation Metric**

  $$CVM = \frac{1}{T} \sum_{i=1}^{T} \left( \frac{VM(t)}{VM(t) + VM(t+1)} \right) \cdot \frac{t}{T}$$  \hspace{1cm} (38)

  The coefficient of variation metric satisfies the following properties:

  1. If the mean value of $VM$ is 0.5 and the standard deviation is 0, then $CVM = 0.5$.
  2. If $VM$ and $VM^{(0)} > 0.5$ is monotonically increasing, then $0 < CVM < 0.5$.
  3. If $VM$ and $VM^{(0)} < 0.5$ is monotonically decreasing, then $0.5 < CVM < 1$.

  In addition, the value of CVM is independent of the initial value.

  One example of dynamic metric acceleration is the iteration process, and the variation of metric is shown in Figure 11.

  A larger CVM represents a more stable iteration for the metric.

  The CVM is calculated and shown in Figure 9 as: $CVM = 0.5013 \pm 0.0010$.

  After acceleration by the dynamic metric in Figure 11, $CVM = 0.8132 \pm 0.0110$.

  From the comparison of Figure 9 and Figure 11, we can see that the measured acceleration method effectively improves the optimization of the iteration.
V. OPTIMIZATION OF DYNAMIC METRICS

**Figure 10.** Schematic diagram of the change process of the sample distance difference

**Figure 11.** Schematic diagram of the change process of the sample distance difference
In order to explore the optimization effect of using dynamic indexes, we use FCM algorithm to experiment and try to verify its optimality. The dataset used in the experiment is shown on the left side of the table. The experiment reflects the optimal dynamic metric by using the change of internal and external evaluation indices before and after dynamic metric acceleration in multiple datasets. In parentheses, the left side is the algorithm without dynamic metric, and the right side is the algorithm with dynamic metric. We let $\alpha = 5$, $e = 10^{-5}$, the maximum number of iterations is 100, $\lambda = 10$, and randomly initialize cluster centers in experiments. Five labeled samples were used as supervisory information in all experiments.

We compare the traditional fuzzy clustering algorithm with the dynamic metric accelerated fuzzy clustering algorithm, FCM-M, for five datasets with the same $m$ value. The results are shown in Table 4. The results are the mean of 10 repeated experiments.

**TABLE IV**

| Datasets/evaluating indicator | Number of iterations | Purity | $V_{RE}$ | max distance |
|------------------------------|----------------------|--------|----------|--------------|
| Iris                         | (30,19)              | (0.8867,0.8867) | (1.0658×10^{-32}, 1.0658×10^{-32}) | (3.0816,0.0325) |
| Wine                         | (13,11)              | (0.9607,0.9607) | (2.8275×10^{-32}, 2.6003×10^{-32}) | (3.7917,0.0391) |
| Seed                         | (16,9)               | (0.8952,0.8952) | (1.4613×10^{-32}, 1.4455×10^{-32}) | (3.7378,0.0396) |
| Breast                       | (14,10)              | (0.9575,0.9575) | (4.9259×10^{-32}, 4.7587×10^{-32}) | (6.5557,0.0271) |
| Thyroid                      | (23,16)              | (0.9258,0.9258) | (5.5722×10^{-32}, 5.5566×10^{-32}) | (6.1550,0.0028) |

The experimental results show that the dynamic metric not only maintains the FCM clustering effect, but also achieves a better iterative speed for most data sets and ensures the standardization of the metric. Combined with the above results, we can conclude that the dynamic acceleration measurement method can achieve iterative acceleration without affecting the optimal solution, and can stably achieve the purpose of normalized distance.

VI. DISCUSSION ON TIME COMPLEXITY

**TABLE V**

| Euclidean distance | Mahalanobis distance | Dynamic metric accelerated method |
|--------------------|----------------------|----------------------------------|
| $O(n \cdot d \cdot c)$ | $O(d^4 \cdot n \cdot c)$ | $O(n \cdot d \cdot c)$ |

In the previous section, we present a dynamic acceleration method based on infinite norms. This section discusses the time complexity of various metric calculations.

Here, $n$ is the number of samples, $d$ is the number of features, and $c$ is the number of cluster centers.

The time complexity of calculating Euclidean distance is $O(n \cdot d \cdot c)$. The time complexity of computing the covariance matrix by the Mahalanobis distance is $O(n^2 \cdot d)$, and the time complexity of finding all elements in the Mahalanobis metric matrix is $O(d^4 \cdot n \cdot c)$. Thus, the time complexity of the Mahalanobis distance is $O(d^4 \cdot n \cdot c + n^2 \cdot d)$.

The time complexity of the three kinds of metrics are shown in Table 1.

VII. INFLUENCE ANALYSIS OF EDGE POINTS

The Euclidean distance of the sample points at the edge of the cluster is farther from other points, so change of the edge points will affect the infinite norm in the metric matrix, and our metric accelerated method depends on the norm in the dynamic metric. Therefore, this section discusses the impact of edge points in clusters on clustering speed and accuracy.
In this part, the sample pairs far away from the main points are analyzed by using iris experiment. Readers can obtain the preprocessed image datasets here\(^3\).

Figure 12. Add data points to simulate the change of edge points

Figure 12 (a) shows the iris dataset. Figures (b) and (c) add ten samples far away from the main point in one cluster of the iris dataset, and the sample points added in figure (c) are farther away from the main point.

Table 5 shows the influence of edge points on the acceleration algorithm. In terms of the maximum distance, the change of edge points changes the maximum distance, while in terms of clustering purity and iteration speed, the change of edge points do not affect the accuracy and speed of the algorithm.

VIII. EXPERIMENT

This section investigates the acceleration effect and performance of the dynamic-metric acceleration method applied to different algorithms in real datasets to verify its effectiveness.

1) EXPERIMENTAL SETUP

The experiment was conducted on a series of real datasets. These datasets include data with clusters that are not linearly separable, with noise, with numerous features and with partial redundancy [58]. In addition, image data classification [59] is also included. The statistical data of all datasets are summarized in Table 2. The image datasets have been preprocessed in the experiment, and readers can obtain the preprocessed image datasets here\(^6\).

To evaluate the effectiveness of the methods presented in this paper, the dynamic metric method is used to improve the fuzzy clustering algorithm based on Euclidean distance, Mahalanobis distance and entropy. In the experiment, we analyze the effect of the dynamic metric method by comparing the performance before and after the improvement.

2) ALGORITHM IMPROVEMENT

We choose the fuzzy clustering algorithm based on the Euclidean distance, Mahalanobis distance and entropy to verify the acceleration effect of the dynamic metric method, and we choose FCM as the unsupervised fuzzy clustering method based on the Euclidean distance, SFCM as the semi-supervised fuzzy clustering method based on the Euclidean distance, SMUC as the unsupervised fuzzy clustering method based on the Mahalanobis distance, and ESFCM as the semi-supervised clustering algorithm based on the Mahalanobis distance and entropy. We introduce the dynamic metric acceleration method into the FCM, KFCM, SFCM, eSFCM and SMUC algorithms to replace the static metric of the original algorithm. The algorithms that introduced dynamic metrics are called FCM-M, KFCM-M, SFCM-M, eSFCM-M and SMUC-M. The flow of the improved algorithm is as follows:

Table 6. The details of the datasets

| Dataset    | Number of samples | Number of samples | Cluster number |
|------------|-------------------|-------------------|----------------|
| Iris       | 150               | 4                 | 3              |
| Wine       | 178               | 13                | 3              |

\(^3\)https://github.com/ChoiNgai/paper_DynamicMetricClustering/tree/main/data/edge points/

\(^6\)https://github.com/ChoiNgai/paper_DynamicMetricClustering/tree/main/data/
A. FUZZY C-MEANS CLUSTERING OF METRIC ACCELERATED (FCM-M)

The membership matrix and the iterative formula of the clustering center of FCM are as follows (three iterative formulas):

\[ v_i = \frac{\sum_{j=1}^{n} (u_{ij})^m x_j}{\sum_{j=1}^{n} (u_{ij})^m} \]  
\[ A = \text{diag}(m_1, \ldots, m_d) \]  
\[ u_{ij} = \frac{[(x_j - v_i)^T A (x_j - v_i)]^{m-1}}{\sum_{k=1}^{c}[(x_j - v_i)^T A (x_j - v_i)]^{m-1}} \]

Algorithm 1. FCM-M

Input: \( n \): the number of samples; \( d \): samples dimensions; \( c \): cluster number; \( \varepsilon \): a small enough error; \( X \): dataset; \( T \): maximum iteration times. \( V \): random initialized clustering centroid. \( \bar{U} \): prior membership degree. \( \alpha \): trade-off parameter

Output: \( V \): clustering centroid matrix; \( U \): membership degree matrix

Initialize metric matrix \( A \) by (40); Initialize membership degree matrix \( U = [u_{ij}] \) by (41).

Repeat
Update the cluster prototype matrix \( V = [v_j] \) by (39); Update the Metric matrix \( A \) by (40); Update the membership degree \( U \) matrix by (40);

Calculate the objective function value \( J(t) \) by (2);
Iteration time \( t + 1 \);
Until \( J(t+1) - J(t) < \varepsilon \) or \( t = T \)

B. KERNEL-BASED FUZZY C-MEANS CLUSTERING OF METRIC-ACCELERATED (KFCM-M)

The membership matrix and the iterative formula of the clustering center of KFCM are as follows (three iterative formulas):

\[ v_i = \frac{\sum_{j=1}^{n} (u_{ij})^m \kappa(x_j, v_i) x_j}{\sum_{j=1}^{n} (u_{ij})^m \kappa(x_j, v_i)} \]  
\[ A = \text{diag}(m_1, \ldots, m_d) \]  
\[ u_{ij} = \frac{(1-e^{\|x_j - v_i\|^2 / -2\sigma^2})^{m-1}}{\sum_{k=1}^{c}(1-e^{\|x_j - v_i\|^2 / -2\sigma^2})^{m-1}} \]

Algorithm 2. KFCM-M

Input: \( n \): the number of samples; \( d \): samples dimensions; \( c \): cluster number; \( \varepsilon \): a small enough error; \( X \): dataset; \( T \): maximum iteration times. \( V \): random initialized clustering centroid. \( \bar{U} \): prior membership degree. \( \alpha \): trade-off parameter

Output: \( V \): clustering centroid matrix; \( U \): membership degree matrix

Initialize Metric matrix \( A \) by (43); Initialize membership degree matrix \( U = [u_{ij}] \) by (46).

Repeat
Update the cluster prototype matrix \( V = [v_j] \) by (44); Update the Metric matrix \( A \) by (43); Update the membership degree \( U \) matrix by (44); Calculate the objective function value \( J(t) \) by (11);
Iteration time \( t + 1 \);
The membership matrix and the iterative formula of the clustering center of SFCM are as follows (three iterative formulas):

\[
A = \text{diag}(m_1, \ldots, m_d)
\]

\[
\alpha_i = \frac{1}{1 + \alpha \left[ \frac{1 + \alpha \sum_{h=1}^{c} \bar{u}_{ij}}{\left( \sum_{h=1}^{c} \left( x_j - y_h \right)^T A (x_j - y_h) \right)^2} \right]}
\]

Algorithm 3. SFCM-M

Input: \( n \): the number of samples; \( d \): samples dimensions; \( c \): cluster number; \( \varepsilon \): a small enough error; \( X \): dataset; \( T \): maximum iteration times. \( V \): random initialized clustering centroid. \( \tilde{U} \): prior membership degree. \( \alpha \): trade-off parameter

Output: \( V \): clustering centroid matrix; \( U \): membership degree matrix

Initialize Metric matrix \( A \) by (46);
Initialize membership degree matrix \( U = [u_{ij}] \) by (47).

Repeat

Update the cluster prototype matrix \( V = [v_i] \) by (45);
Update the Metric matrix \( A \) by (46);
Update the membership degree \( U \) matrix by (47);
Calculate the objective function value \( J(t) \) by (15);

Iteration time \( t + 1 \);


E. SEMI-SUPERVISED METRIC-BASED FUZZY CLUSTERING OF METRIC-ACCELERATED (SMUC-M)

The membership matrix and the iterative formula of the clustering center of SMUC-M are as follows (three iterative formulas):

\[ v_i = \frac{\sum_{j=1}^{n} (u_{ij})^m x_j}{\sum_{j=1}^{n} (u_{ij})^m} \]  \hspace{1cm} (51)

\[ A = \text{diag}(m_1,\ldots,m_d) \]  \hspace{1cm} (52)

\[ u_{ij} = \tilde{u}_{ij} + \frac{e^{-\lambda(x_j-v_i)^TMA(x_j-v_i)}}{\sum_{h=1}^{c} e^{-\lambda(x_j-v_h)^TMA(x_j-v_h)}} (1 - \sum_{i=1}^{c} \tilde{u}_{ij}) \]  \hspace{1cm} (53)

Algorithm 5. SMUC-M

**Input:** \( n \): the number of samples; \( d \): samples dimensions; \( c \): cluster number; \( \epsilon \): a small enough error; \( X \): dataset; \( T \): maximum iteration times. \( V(0) \): random initialized clustering centroid. \( \tilde{U} \): prior membership degree. \( \alpha \): trade-off parameter

**Output:** \( V \): clustering centroid matrix; \( U \): membership degree matrix

Initialize Metric matrix \( A \) by (52);
Initialize membership degree matrix \( U = [u_{ij}] \) by (53).

**Repeat**

Update the cluster prototype matrix \( V = [v_i] \) by (51);
Update the Metric matrix \( A \) by (52);
Update the membership degree \( U \) matrix by (53);
Calculate the objective function value \( J(t) \) by (18);
Iteration time \( t + + \);

**Until**

\[ J(t+1) - J(t) < \epsilon \text{ or } t = T \]

3) EVALUATION INDEX

Experiments compare the clustering accuracy and running time before and after the improvement of the algorithm. The internal evaluation index data reconstruction error rate, the external evaluation index purity, and the number of iterations are used to evaluate the optimization effect of the iterative process.

In the following experimental Figures 11, 12 and 13, the dynamic metric acceleration method is represented by the L-\( \infty \) metric speedup. Moreover, we make \( \alpha = 5 \), \( e = 10^{-5} \), the maximum number of iterations is 100, \( \lambda=10 \), and randomly initialize cluster centers in experiments.

A. EXTERNAL EVALUATION INDEX:

Purity reflects the similarity between clustering results and reality. We propose an internal evaluation index \( \text{Purity}(\Omega,C) / \text{iter} \) to evaluate the improvement level of the clustering accuracy of the algorithm:

\[ \text{Purity}(\Omega,C) / \text{iter} = \frac{\sum_{k} \max_{j} | \omega_c \cap y_j |}{t} \]  \hspace{1cm} (54)

B. INTERNAL EVALUATION INDEX:

\( V_{\text{RE}} \) analyzes the difference between reconstructed data and the original data after clustering.

Equation (55) reconstructs data by using the membership matrix and clustering center obtained by clustering. The reconstructed data should be as similar to the original data as possible. A small \( V_{\text{RE}} \) value means that the algorithm created a well-clustered effect.

Fewer iterations mean a faster search speed. We found that the difference between the results of the two indices was small and stable over the corresponding intervals in the experiment. Therefore, we propose an internal evaluation index to \( V_{\text{RE} \cdot \text{iter}} \) to evaluate the improvement level of the iterative performance of the algorithm:

\[ V_{\text{RE} \cdot \text{iter}} = \frac{1}{n \sum_{i=1}^{n} \| I(i) - I(i) \|^2 \cdot t} \]  \hspace{1cm} (55)

where \( I \) is the number of iterations of the algorithm.
Figure 13. Comparison of the number of iterations required to run various algorithms.
From the experimental results in Figure 13, except for the KFCM algorithm, the algorithms have better effects after using the metric accelerated iteration method, which shows that this method is not suitable for the fuzzy clustering algorithm of the kernel function.
From the experimental results in Figure 14, the metric acceleration method still has a poor effect on the KFCM algorithm. Some datasets will not be improved on eSFCM and SMUC, and the effect is the same as that of the unimproved method, indicating that this method still has room for improvement in the nonlinear metrics.
From the experimental results in Figure 15, the metric acceleration method performs well in algorithms other than KFCM. From the comparison of SMUC algorithm with SFCM and eSFCM, it can be seen that our acceleration method is weaker than the algorithm based on Mahalanobis distance in the optimization effect of the algorithm based on Euclidean distance. The experiments of FCM and SFCM show that the optimization effect of unsupervised algorithm is lower than that of semi supervised algorithm.

**EXPERIMENTAL RESULTS ANALYSIS**

From the above experimental results, it can be seen that the metric acceleration method in this paper achieves high-quality results in the Euclidean distance and Mahalanobis distance metrics. However, for some nonlinear metric methods, the effect is not improved (such as KFCM and metrics with the kernel method), which indicates that this method is not suitable for nonlinear metrics, such as the kernel function. In the experiment, while combining the external evaluation index and iteration speed, the acceleration effect of dynamic metric is not substantial when dealing with linearly inseparable data. The optimization effect of our acceleration method based on Euclidean distance is weaker than that based on Markov distance, and the optimization effect of unsupervised algorithm is lower than that of semi supervised algorithm.

**IX. CONCLUSIONS**

While focusing on the problem of unstandardized distance in clustering algorithms, we develop a fuzzy clustering acceleration method based on dynamic metrics. This method standardizes the distance between instances to improve the search efficiency of the iteration process. We prove its stability in metric normalization, and verify that its optimization performance is not affected by the change of edge points in clusters. The method is applied to five common fuzzy clustering algorithms, and the experimental results on eight real datasets show that the method has an effective acceleration, while maintaining the stability of clustering accuracy. Furthermore, the algorithm uses the Chebyshev distance to standardize between features, which keeps the time complexity of the algorithm constant during metric iterations. In addition, the experimental results show that the dynamic metric acceleration method is less effective in the algorithm with the Gauss kernel function.

**ACKNOWLEDGMENT**

This work was supported in part by the Natural Science Foundation of China under Grant 61876044 (Algorithm research and application of positive sample and unmarked sample learning).
REFERENCES

[1] Ruspini E H . A new approach to clustering[J]. Information and Control, 1969, 15(1):22-32.

[2] Tamura S , Higuchi S , Tanaka K . Pattern Classification Based on Fuzzy Relations[J]. IEEE Transactions on Systems Man and Cybernetics, 1971, SMC1(1):61-66.

[3] Le K . Fuzzy relation compositions and pattern recognition[J]. Information Sciences, 1996, 89(1-2):107-130.

[4] Wu Z , Lejby R . An optimal graph theoretic approach to data clustering: theory and its application to image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(11):1101-1113.

[5] Gao Xinxiao, Xie Weixin. Research progress in the development and application of fuzzy clustering theory[J]. Chinese Science Bulletin, 1999, 44(021):2241-2250.

[6] Dai Tianchen. Research on metric learning and clustering algorithm based on transfer distance[D]. 2019.

[7] Dokmanic I , Parhizkar R , Ranieri J , et al. Euclidean Distance Matrices: A Short Walk Through Theory, Algorithms and Applications[J]. IEEE Signal Processing Magazine, 2015, 32(6):12-30.

[8] Groenen P J F , Kaymak U , Rosmalen J V . Fuzzy Clustering with Minkowski Distance Functions[M]// Advances in Fuzzy Clustering and Its Applications. Elsevier North-Holland, Inc. 2001.

[9] Kanzawa Y , Endo Y , Miyamoto S . KL-Divergence-Based and Manhattan Distance-Based Semi-supervised Entropy-Regularized Fuzzy c-Means[J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2011, 15(5):1057-1064.

[10] Loo Jiawei, YIN Zhiqiang, LIU Shuyan, Et al. Novel method for support vector clustering with cell growing[J]. Fuzzy Systems IEEE Transactions on, 2003, 11(4):518-527.

[11] Zhao X , Li Y , Zhao Q . Mahalanobis distance based fuzzy clustering algorithm for image segmentation[J]. Digital Signal Processing, 2015, 43:8-16.

[12] Huang S F , Lin Y H , Yih J M , et al. Fuzzy Clustering Algorithm Based on Mahalanobis Distances with Recursive Process[J]. International Journal of Intelligent Technologies & Applied Statistics, 2017, 10.

[13] Zhiwen Z U , Qin L I . Mahalanobis distance fuzzy clustering algorithm based on particle swarm optimization[J]. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition), 2019.

[14] Friedman M , Last M , Makover Y , et al. Anomaly detection in web documents using crisp and fuzzy-based cosine clustering methodology[J]. Information Sciences, 2007, 177(2):467-475.

[15] Jaferzadeh K , Kiani K , Mozaffari S . Acceleration of fractal image compression using fuzzy clustering and discrete-cosine-transform-based metric[J]. Iet Image Processing, 2012, 6(7):1024-1030.

[16] He H , Tan Y , Huang J . Unsupervised classification of smartphone activities signals using wavelet packet transform and half-cosine fuzzy clustering[C]// 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE, 2017.

[17] Thao N X, Ali M, Smarandache F. An intuitionistic fuzzy clustering algorithm based on a new correlation coefficient with application in medical diagnosis[J]. Journal of Intelligent & Fuzzy Systems, 2019, 36(1):189-198.

[18] Karayiannis N B , MECA: maximum entropy clustering algorithm[C]// Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence. Proceedings of the Third IEEE Conference on. IEEE, 1994.

[19] E. Yasunori, H. Yukihiro, Y. Makito. On semi-supervised fuzzy c-means clustering, in: Proceedings of IEEE International Conference on Fuzzy Systems, 2009, pp. 1119–1124.

[20] Zarinbal M, Zarandi M H F, Turksen I B. Relative entropy fuzzy c-means clustering[J]. Information sciences, 2014, 260: 74-97.

[21] Ionescu M, Ralescu A. Image clustering for a fuzzy hamming distance based ecb system[C]//Proceedings of the Sixteenth Midwest Artificial Intelligence and Cognitive Science Conference, MAICS-2005. 2005:102-108.

[22] Chiang J H , Hao P Y , A new kernel-based fuzzy clustering approach: support vector clustering with cell growing[J]. Fuzzy Systems IEEE Transactions on, 2003, 11(4):518-527.

[23] Li Y , Shen Y . Robust Image Segmentation Algorithm Using Fuzzy Clustering Based on Kernel-Induced Distance Measure[C]// International Conference on Computer Science & Software Engineering. IEEE Computer Society, 2008.

[24] Gong M , Liang Y , Shi J , et al. Fuzzy C-Means Clustering With Local Information and Kernel Metric for Image Segmentation[J]. IEEE Transactions on Image Processing, 2013, 22(2):573-584.

[25] Salehi F , Keyvanpour M R , Sharifi A . SMFKC-ER: Semi-supervised Multiple Kernel Fuzzy Clustering based on Entropy and Relative entropy[J]. Information Sciences, 2020.

[26] Zhao F , Zeng Z , Liu H , et al. Semisupervised Approach to Surrogate-Assisted Multiobjective Kernel Intuitionistic Fuzzy Clustering Algorithm for Color Image Segmentation[J]. IEEE Transactions on Fuzzy Systems, 2020, PP(99):1-1.

[27] Bezdik J C, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm[J]. Computers & geosciences, 1984, 10(2-3):191-203.

[28] Bui Q T, Vo B, Snasel V, et al. SFCM: A Fuzzy Clustering Algorithm for Extracting the Shape Information of Data[J]. IEEE Transactions on Fuzzy Systems, 2020.

[29] Giri N . Crucianu M , Boujemaa N . Semi-Supervised Fuzzy Clustering with Pairwise-Constrained Competitive Agglomeration[C]// The 14th IEEE International Conference on Fuzzy Systems, 2005. FUZZ ’05. IEEE, 2005.

[30] Yin X , Shu T , Huang Q . Semi-supervised fuzzy clustering with metric learning and entropy regularization [J]. Knowledge-Based Systems, 2012, 35(none):304-311.

[31] Wu C , Cao Z . Entropy-like Divergence Based Kernel Fuzzy Clustering for Robust Image Segmentation[J]. Expert Systems with Applications, 2020:114327.

[32] Zhang D Q, Chen S C. Clustering incomplete data using kernel-based fuzzy c-means algorithm[J]. Neural processing letters, 2003, 18(3):155-162.

[33] Kapil S, Chawla M. Performance evaluation of K-means clustering algorithm with various distance metrics[C]// 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICEPECES). IEEE, 2016: 1-4.

[34] Shen C, Kim J, Wang L. Scalable large-margin Mahalanobis distance metric learning[J]. IEEE transactions on neural networks, 2010, 21(9):1524-1530.
[35] Thao N X, Ali M, Smarandache F. An intuitionistic fuzzy clustering algorithm based on a new correlation coefficient with application in medical diagnosis [J]. Journal of Intelligent & Fuzzy Systems, 2019, 36(1): 189-198.

[36] Yao J, Dash M, Tan S T, et al. Entropy-based fuzzy clustering and fuzzy modeling[J]. Fuzzy sets and Systems, 2000, 113(3): 381-388.

[37] Yang M S, Nataliani Y. A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy[J]. IEEE Transactions on Fuzzy Systems, 2017, 26(2): 817-835.

[38] Salehi F, Keyvanpour M R, Sharif A. SMKFC-ER: Semi-supervised multiple kernel fuzzy clustering based on entropy and relative entropy[J]. Information Sciences, 2021, 547: 667-688.

[39] Hartigan J A, Wong M A. A K-Means Clustering Algorithm[J]. Applied Statistics, 1979, 28(1).

[40] Xing E P, Ng A Y, Jordan M I, et al. Distance Metric Learning with Application to Clustering with Side-Information.[C]// International Conference on Neural Information Processing Systems. MIT Press, 2002.

[41] Klwon F, Keller A. Fuzzy clustering based on modified distance measures[C]/International Symposium on Intelligent Data Analysis. Springer, Berlin, Heidelberg, 1999: 291-301.

[42] Flores-Sintas A, Cadenas J M, Martin F. A local geometrical properties application to fuzzy clustering[J]. Fuzzy Sets and Systems, 1998, 100(1-3): 245-256.

[43] Zhao X, Li Y, Zhao Q. Mahalanobis distance based on fuzzy clustering algorithm for image segmentation[J]. Digital Signal Processing, 2015, 43: 8-16.

[44] Li K, Zhou Y. An improved semi-supervised fuzzy clustering algorithm[J]. International Journal of Computer and Information Technology, 2013, 2: 115-120.

[45] Lai D T C, Garibaldi J M, Reps J. Investigating distance metric learning in semi-supervised fuzzy c-means clustering[C]//2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE, 2014: 1817-1824.

[46] Carter R L, Morris R, Blashfield R K. On the partitioning of squared Euclidean distance and its applications in cluster analysis[J]. Psychometrika, 1989, 54(1): 9-23.

[47] Malkauthekar, M.D. Analysis of euclidean distance and manhattan distance measure in face recognition [J] IET Conference Proceedings, 2013, p. 503-507.

[48] ICHIHASHI H, OHUE M, MIYOSHI T. Fuzzy C-Means Clustering Algorithm with Pseudo Mahalanobis Distances[C]/Proceedings of the Korean Institute of Intelligent Systems Conference. Korean Institute of Intelligent Systems, 1998: 148-152.

[49] Aziz A M. A new nearest-neighbor association approach based on fuzzy clustering[J]. Aerospace Science and Technology, 2013, 26(1): 87-97.

[50] Zeng S, Wang X, Duan X, et al. Kernelized mahalanobis distance for fuzzy clustering[J]. IEEE Transactions on Fuzzy Systems, 2020.

[51] Bezdek, James C. Pattern Recognition with Fuzzy Objective Function Algorithms [J]. Advanced Applications in Pattern Recognition, 1981, 22(1171):203-239.

[52] Pedrycz W, Waletzky J. Fuzzy clustering with partial supervision [J]. IEEE Transactions on Systems Man & Cybernetics Part B, 1997, 27(5):787-795.
