A multivariate grey incidence model for different scale data based on spatial pyramid pooling

ZHANG Ke*, CUI Le, and YIN Yao
Business School, Hohai University, Nanjing 211100, China

Abstract: In order to solve the problem that existing multivariate grey incidence models cannot be applied to time series on different scales, a new model is proposed based on spatial pyramid pooling. Firstly, local features of multivariate time series on different scales are pooled and aggregated by spatial pyramid pooling to construct n levels feature pooling matrices on the same scale. Secondly, Deng's multivariate grey incidence model is introduced to measure the degree of incidence between feature pooling matrices at each level. Thirdly, grey incidence degrees at each level are integrated into a global incidence degree. Finally, the performance of the proposed model is verified on two data sets compared with a variety of algorithms. The results illustrate that the proposed model is more effective and efficient than other similarity measure algorithms.

Keywords: grey system, spatial pyramid pooling, grey incidence, multivariate time series.

DOI: 10.23919/JSEE.2020.000052

1. Introduction

The grey incidence model is an important branch of the grey system theory. According to the model, the incidence degree between two factors is measured by the geometric similarity between their time series [1]. Since professor Deng proposed the grey incidence analysis model, it has become a well-known method for evaluation and decision because of its ease of applicability [2–4]. Many researchers have systematically studied grey incidence models from different perspectives, and developed a series of grey incidence models, such as absolute degree of grey incidence [5], type B [6], type C [7], type T [8] degrees of grey incidence, grey entropy incidence degree [9] and slope grey incidence [10]. However, most of these models are only applicable to incidence analysis between univariate time series. With the improvement of data storage and processing capabilities, multivariate time series (MTS) are widely used in a broad range of fields, including multimedia, economics, manufacturing and speech recognition [11].

Recently, some scholars extended the theoretical structure of grey incidence models to handle MTS. Zhang analyzed the geometric similarity of MTS in three-dimensional space, and put forward a matrix grey incidence model [12]. Wu proposed a grey convex incidence degree for MTS through introducing similar characteristics of convexity among factors [13]. Qian employed horizontal distance, incremental distance and variation distance to extract spatiotemporal characteristics of MTS, and constructed a grey matrix incidence model for MTS [14]. Zhang introduced the principal component analysis (PCA) to extract features of MTS and established a multivariate grey incidence model based on the PCA [15]. Wu developed a three-dimensional grey convex incidence model based on the two-dimensional model [16]. Liu applied the grid method to describing geometric characteristics of MTS in three-dimensional space and proposed a grey grid incidence model [17]. Li improved the multivariate grey incidence model with grey accumulating generate operation, and put forward a clustering method for MTS [18]. Cui introduced a development speed index and a growth rate coefficient as spatiotemporal characteristics of MTS and constructed a grey matrix similar incidence model [19]. Wu projected MTS into a vector sequence of space, and constructed a similarity and proximity incidence model based on angle and distance between space vectors [20]. Luo extracted growth and variation features of each index in MTS, and proposed a grey matrix incidence model based on root mean square distance [21]. Wang constructed a degree of grey trend incidence for both univariate time series and MTS [22]. Dai put forward a three-dimensional grey incidence degree based on multidimensional dynamic time warping distance [23]. Liu developed a new multivariate grey incidence model for relationship...
analysis between single index sequence and multiple index sequences [24]. In summary, these studies have expanded the application scope of grey incidence analysis and enriched models of multivariate grey incidence analysis.

However, grey incidence models have never been applied to MTS on different scales. In the real world, MTS on different scales is available everywhere. For example, different persons’ audio data of the same sentence are different scale time series, because of various speech rates. The incidence degree among MTS on different scales is a research hotspot in the grey incidence analysis theory. Moreover, multivariate time series analysis requires efficient similarity measure methods for clustering, classification and prediction of MTS on different scales. Although dynamic time warping (DTW) is the most robust similarity measure, it suffers from high computational cost.

Therefore, this paper constructs a new multivariate grey incidence model based on spatial pyramid pooling (MGIM-SPP) for MTS on different scales. Firstly, through spatial pyramid pooling, local features on the same scale are extracted from MTS on different scales, and feature pooling matrices on the same scale are constructed. Then, an incidence degree between feature pooling matrices is carried out by Deng’s multivariate grey incidence model to approximate the incidence degree between original different scale data. Finally, the performance of the proposed model is compared with a variety of similarity measure algorithms by two experiment data sets, to verify its effectiveness and efficiency.

As such, the primary contributions of this paper are: (i) modification of the multivariate grey incidence model for MTS on different scales; (ii) introduction of the proposed model as an efficient similarity measure for k-nearest neighbor (k-NN) classification of MTS on different scales.

The rest of this paper is ordered as follows. We present the architecture of the proposed model in Section 2. In Section 3, we discuss the dataset and experiment, evaluate the model on them, present results and analyze them. Section 4 concludes the paper.

2. Model construction

The MTS is a series of high-dimensional vectors, such as hydrological data [25], electroencephalogram (EEG) data [26], and flight data [27].

Definition 1[27] Given a time series
\[ X = \{T_1, T_2, \ldots, T_n\} \]  
(1)

where
\[ T_i = \{v_i(1), v_i(2), v_i(t), \ldots, v_i(m)\} \]  
(2)

and \(v_i(t)\) represents an observation value of the ith variable at time \(t\), then \(X\) is a multivariate time series.

An MTS \(X\) can be represented by an \(m \times n\) matrix.

\[
X_{m \times n} = \begin{bmatrix}
  x_{11} & x_{12} & \cdots & x_{1n} \\
  x_{21} & x_{22} & \cdots & x_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{m1} & x_{m2} & \cdots & x_{mn}
\end{bmatrix}
\]  
(3)

Definition 2 Suppose \(X_1\) and \(X_2\) are two MTS, and their sizes are \(m_1 \times n_1\) and \(m_2 \times n_2\) respectively. If \(m_1 \neq m_2\) or \(n_1 \neq n_2\), \(X_1\) and \(X_2\) are MTS on different scales.

Benefiting from the development of Internet of Things and sensor technologies, a large scale of MTS is being generated for economics and management research. For example, the Google search index is conducive to analyzing and predicting the GDP of a region or sales of a company. Due to the high frequency and numerous key words of Google search index, online information is of a different size from statistical GDP or sales. One challenge is identifying closely related key words for analysis and prediction through incidence analysis between MTS on different scales.

Because existing multivariate grey incidence cannot be applied to variable scale time series with different sizes of matrix, a feature pooling, namely spatial pyramid pooling, is introduced to solve the problem.

Feature pooling originated from Hubel and Wiese’s research on cat’s visual cortex [28]. They found the visual region consisted of simple cells and complex cells. Simple cells performed feature extraction operations, while complex cells combined local features from small spatial domains. This mechanism is feature pooling [29].

Because of translation invariance for feature aggregation, feature pooling is robust and fault-tolerant to noise [30]. Therefore, it is widely used in speech and image recognition algorithms, including neural perceptron [31], and convolutional neural network [32]. Moreover, a series of well-known feature extraction models, such as scale-invariant feature transform (SIFT) [33], histogram of oriented gradient (HOG) [34], generalized search trees (GIST) [35] and hierarchical maximum (HMAX) [36], adopt feature pooling.

This paper employs SPP, a new feature pooling method proposed in convolutional neural networks, and applies it to the feature pooling of MTS. SPP originated from spatial pyramid matching which applies spatial statistical properties to representing global information for recognizing scene categories. He firstly proposed a spatial pyramid pooling layer on the top of the last convolutional layer of deep convolutional networks in order to remove the fixed-size input constraint and generate fixed-length outputs [37]. SPP pools features and generates fixed-length
outputs which are then fed into the fully connected layers (or other classifiers). In other words, some information is aggregated to avoid the need for cropping or warping at the beginning [38].

SPP partitions a matrix into some divisions from the finest level to the coarsest level, and aggregates local features in them. There are three advantages of SPP. Firstly, SPP can generate a fixed length output regardless of the input size. In contrast, traditional sliding window pooling is sensitive to the length of time series. Secondly, SPP uses only a window size. Thirdly, SPP can pool features extracted on variable scales in virtue of the flexibility of input scales [39]. Therefore, it is feasible to introduce SPP into the grey incidence analysis for MTS on different scales.

Generally, SPP includes \( n \) levels of pooling from the finest to the coarsest. At a pyramid level, SPP is implemented as dynamic window pooling in which different scale time series are extracted by variable scale pooling windows. In a sliding pooling window, a pooling function extracts features of the pooling domain corresponding to the window. Then, a level feature pooling matrix of original MTS can be generated by aggregating features of different pooling domains. Finally, SPP matrices can be constructed by repeating the process mentioned above \( n \) times at each pyramid level. An example is shown in Fig. 1. In the figure an original MTS, an \( 8 \times 8 \) matrix, is pooled to three feature matrices by a 3-level SPP. Their sizes are \( 4 \times 4, 2 \times 2, \) and \( 1 \times 1 \) respectively.

![Fig. 1 An example of spatial pyramid pooling](image)

In order to process data in different sizes at each pooling level, SPP uses variable pooling windows according to sizes of the input data and the output feature. In addition, some zeros may be inserted into the original data in order to ensure the same output size. Suppose the input matrix size is \( m_{\text{input}} \times n_{\text{input}} \), the output matrix size is \( m_{\text{output}} \times n_{\text{output}} \), the pooling window size is \( l_{\text{kernel}} \times w_{\text{kernel}} \), the number of zero padding rows is \( m_{\text{zero}} \), the number of zero-padding columns is \( n_{\text{zero}} \), the horizontal movement step is \( l_{\text{stride}} \), and the vertical movement step is \( w_{\text{stride}} \). Among them,

\[
l_{\text{stride}} = l_{\text{kernel}}, \quad w_{\text{stride}} = w_{\text{kernel}},
\]

\[
l_{\text{kernel}} = \text{ceil}(m_{\text{input}}/m_{\text{output}}), \quad w_{\text{kernel}} = \text{ceil}(n_{\text{input}}/n_{\text{output}}),
\]

\[
m_{\text{zero}} = l_{\text{kernel}} \times m_{\text{output}} - m_{\text{input}}, \quad n_{\text{zero}} = w_{\text{kernel}} \times n_{\text{output}} - n_{\text{input}}\]

where \( \text{ceil}(\cdot) \) is the ceiling function.

Because of different pooling functions, SPP includes a bunch of pooling methods, such as maximum pooling, average pooling, and minimum pooling. The maximum pooling extracts the maximum value of elements in a pooled domain \( C \) as its characteristic, and the pooling function is denoted as \( \max(C) \). Conversely, the minimum pooling extracts the minimum value, and the pooling function is \( \min(C) \). Similarly, the mean of all elements in a pooled domain is extracted as the characteristic, and the pooling function is \( \text{ave}(C) \). Hence, the pooling function \( F(C) \) is determined according to the data characteristics of the MTS. If the positive feature is dominant among a time series, the maximum pooling function \( \max(C) \) is selected. On the contrary, the minimum pooling function \( \min(C) \) is selected. Otherwise, if the data feature distribution is relatively uniform, the average pooling function \( \text{ave}(C) \) is selected.

Taking a level-1 maximum SPP as an example, \( A \) and \( B \) are matrices of \( 4 \times 4 \) and \( 6 \times 4 \) respectively. In order to output feature pooling matrices of size \( 2 \times 2 \), the pooling kernel size of \( A \) needs to be \( 2 \times 2 \), while the pooling kernel size of \( B \) should be \( 3 \times 2 \), according to (6) and (7). After the feature pooling, \( A \) and \( B \) output two same size feature pooling matrices \( SPP_{\text{max}_A} \) and \( SPP_{\text{max}_B} \).

\[
A = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 4 \\ 3 & 7 & 4 & 3 \\ 5 & 8 & 6 & 8 \end{bmatrix}, \quad (10)
\]

\[
B = \begin{bmatrix} 4 & 7 & 7 & 2 \\ 2 & 1 & 5 & 9 \\ 5 & 8 & 6 & 2 \\ 2 & 4 & 5 & 7 \\ 1 & 2 & 3 & 7 \\ 2 & 4 & 8 & 9 \end{bmatrix}, \quad (11)
\]

\[
SPP_{\text{max}_A} = \begin{bmatrix} 4 & 5 \\ 8 & 8 \end{bmatrix}, \quad (12)
\]

\[
SPP_{\text{max}_B} = \begin{bmatrix} 8 & 9 \\ 4 & 9 \end{bmatrix}. \quad (13)
\]
Furthermore, Deng’s multivariate grey incidence model can be exploited to measure the similarity between feature pooling matrices.

**Definition 3** $X_0$ and $X_k$ ($k = 1, 2, \ldots, K$) are two MTS on different scales, and their SPP matrices are

$$SPP - X_0 = \{ SPP - X_{0l}, \ldots, SPP - X_{0l}, \ldots, SPP - X_{0L} \}$$

and

$$SPP - X_k = \{ SPP - X_{kl}, \ldots, SPP - X_{kl}, \ldots, SPP - X_{kL} \}$$

where

$$SPP - X_{0l} = \begin{bmatrix} a_{0l,1}^0 & \cdots & a_{0l,n_l}^0 \\ \vdots & \ddots & \vdots \\ a_{m_l,1}^0 & \cdots & a_{m_l,n_l}^0 \end{bmatrix}, \quad (16)$$

and

$$SPP - X_{kl} = \begin{bmatrix} a_{kl,1}^0 & \cdots & a_{kl,n_l}^0 \\ \vdots & \ddots & \vdots \\ a_{m_l,1}^0 & \cdots & a_{m_l,n_l}^0 \end{bmatrix}, \quad (17)$$

$SPP - X_{0l}$ and $SPP - X_{kl}$ are pooling feature matrices at the $l$th level of $X_0$ and $X_k$, respectively. The subscripts of $n_l$ and $n_l$ are the sizes of feature matrices.

Note that

$$\gamma_{0k}^l = \sum_{i=1}^{m_l} \sum_{j=1}^{n_l} \varepsilon_{ij}^{0kl}$$

where

$$\varepsilon_{ij}^{0kl} = \frac{d_{ij}^0 - d_{ij}^k}{d_{ij}^{0kl} + d_{ij}^{max}}$$

$\rho$ is the distinguishing coefficient between 0 and 1, and

$$d_{max} = \max_i \max_j \max_k d_{ij}^{0kl}, \quad (20)$$

$$d_{min} = \min_i \min_j \min_k d_{ij}^{0kl}, \quad (21)$$

$$d_{ij}^{0kl} = |a_{ij}^0 - a_{ij}^k|, \quad (22)$$

then $\gamma_{0k}^l$ is the grey incidence degree between $X_0$ and $X_k$ at the $l$th level of SPP.

In order to comprehensively measure the similarity between two MTS on variable scales, an incidence degree of all levels is integrated and noted as $\gamma_{0k}$.

**Definition 4** $\gamma_{0k}^l$ is the grey incidence degree between $X_0$ and $X_k$ based on the $l$th level, and $w_l$ is the weight of the $l$-level of spatial pyramid pooling, $w_l > 0$ and $\sum_{l=1}^{L} w_l = 1$. Then

$$\gamma_{0k} = \sum_{l=1}^{L} w_l \gamma_{0k}^l$$

is the grey incidence degree between $X_0$ and $X_k$ based on spatial pyramid pooling.

In the equation, the weights are recommended to be assigned according to the information of each level of SPP, by employing the variance of $\varepsilon_{ij}^{0kl}$ at the $l$th level of SPP, noted as $\sigma_l$, to measure the information.

$$w_l = \frac{\sigma_l}{\sum_{i=1}^{m_l} \sum_{j=1}^{n_l} (\varepsilon_{ij}^{0kl} - \sigma_l)^2}$$

$$\sigma_l = \frac{m_l \times n_l}{m_l \times n_l}$$

For convenience, they can be assigned the same value,

$$w_l = 1/L, \quad l = 1, 2, \ldots, L. \quad (27)$$

**Property 1** MGIM-SPP satisfies normality and pair symmetry of grey incidence axioms.

**Proof**

(i) The normality

If

$$d_{ij}^{0kl} = |a_{ij}^0 - a_{ij}^k| = d_{min},$$

then

$$\varepsilon_{ij}^{0kl} = 1.$$ 

If

$$d_{ij}^{0kl} = |a_{ij}^0 - a_{ij}^k| > d_{min},$$

therefore,

$$\varepsilon_{ij}^{0kl} < 1.$$ 

It is obvious that for any $i$, $j$ and $l$,

$$\varepsilon_{ij}^{0kl} \geq 0.$$ 

Hence,

$$0 \leq \gamma_{0k} \leq 1.$$ 

(ii) The pair symmetry

If $K = 1$, then there are only two MTS, $X_0$ and $X_1$.

$$\max_i \max_j d_{ij}^{01l} = \max_i \max_j d_{ij}^{10l},$$

$$\min_i \min_j d_{ij}^{01l} = \min_i \min_j d_{ij}^{10l},$$

Therefore,

$$\varepsilon_{ij}^{0kl} = \varepsilon_{ij}^{k0l},$$

and
Spatial pyramid pooling is primarily used for measuring global representation to form accurate feature representation matrix, and can remove the fixed-size input constraint of the multivariable grey incidence model. Nevertheless, a pooled feature matrix may lose some information compared with the original data. The information loss depends on the pooling window size and pooling function. For one thing, if a pooling window is undersized, its feature matrix will include some noises resulting in the bias of grey incidence analysis. For another, if it is oversized, the feature matrix will lose excessive information of original data causing the miscalculation of the grey incidence model. Furthermore, an average pooling function may weaken local features in a pooling window, and minimum and maximum pooling functions abandon some features. Therefore, the pooling window size and function should be determined through trial and error experiments.

In summary, the specific process of the multivariate grey incidence model based on spatial pyramid pooling is as follows:

(i) Determine the spatial pyramid pooling parameters according to the size of MTS, and choose appropriate pooling functions, according to (2) – (9).

(ii) The spatial pyramid pooling algorithm is used to pool different scales of time series data into feature pooling matrices on the same scale.

(iii) According to (14) – (27), Deng’s multivariate grey incidence model is deployed to calculate the grey incidence degree of the feature pooling matrices.

(iv) Analyze the ranking of the grey incidence degree, and measure the similarity between MTS.

3. Experiment and result analysis

Time series classification is one of the most important tasks in time series analysis, such as voiceprint recognition and sign language recognition. The k-NN with a similarity measure is a primary technique for time series classification. In voiceprint recognition, some respondents’ voiceprints are collected and stored in a multivariable time database. The technology can recognize the identity of an unknow sign language is determined by searching the most similar MTS in the corpus. Hence, the accuracy rate of a similarity measure for a sample X can be calculated according to (28).

\[ e_X = n_0/k \quad (28) \]

The above experiment is repeated to obtain the accuracy rate of n samples. According to the experimental method, \( e_X \) for any sample is included in a set \( \{0, 0.1, 0.2, \ldots, 1.0\} \), consisting of 11 possible values, recorded as \( \{e_1, e_2, e_3, \ldots, e_{11}\} \). The accuracy rate can be considered as a discrete random variable \( \varepsilon \). Consequently, the expected value of the accuracy rate can be calculated according to (29).

\[ p = \sum_{i=1}^{11} p(\varepsilon = e_i)e_i \quad (29) \]

where \( p(\varepsilon = e_i) \) is the frequency of the accuracy rate \( e_i \) appearing in experiments of all samples.

The matching performance of a similarity measure increases with a growing expected value \( p \). Thus, it is used as a comparative index between different similarity measurement methods. In the experiment, two sets of MTS datasets on different scales are selected, JV [42] dataset and ASL dataset [43]. Both datasets are obtained from the University of California Irvine (UCI) machine learning repository. The above experiment is conducted on the two datasets with the proposed model and some benchmark models such as point distribution (PD) [44], DTW [45], and PCA+DTW [46].

3.2 Comparison of different similarity measures

3.2.1 JV dataset experiment

The JV dataset is a dataset that records the pronunciation process of JVs with 12 variables. It contains nine tester data, namely nine classes. In the dataset, each tester pronounces 30 times, so there is a total of 270 samples. Be-
cause time spans of samples vary from 7 s to 29 s, it is a variable scale time series dataset.

In order to determine the parameters of the feature pooling matrix, different size feature pooling matrices and pooling functions are tested on JV datasets. The results are shown in Table 1. It can be seen from the table that the $\text{ave}(C)$ pooling function performs reasonably well than the others with the same size feature pooling matrix, due to fluctuation features of the dataset. For example, four samples of different classes in the JV dataset are shown in Fig. 2, where JV$_{X\_Y}$ represents the $X$th sample of the JV dataset, and its class is the tester $Y$.

### Table 1 Results of experiment on JV dataset with the MGIM-SPP method

| Output size of SPP | $F(c) = \text{min}(C)$ | $F(c) = \text{max}(C)$ | $F(c) = \text{ave}(C)$ |
|-------------------|-------------------------|-------------------------|-------------------------|
| $output_{m}=5$   | $k=1$ 0.79 0.79 0.7     | $k=5$ 0.94 0.89 0.83     | $k=10$ 0.79 0.73 0.69   |
| $output_{m}=6$   | $k=5$ 0.94 0.89 0.83     | $k=5$ 0.79 0.73 0.69     | $k=10$ 0.98 0.91 0.85   |
| $output_{m}=12$  | $k=1$ 0.94 0.87 0.8      | $k=5$ 0.85 0.81 0.76     | $k=10$ 0.93 0.9 0.84    |
| $output_{n}=6$   | $k=5$ 0.79 0.73 0.69     | $k=5$ 0.85 0.81 0.76     | $k=10$ 0.86 0.79 0.72   |
| $output_{n}=12$  | $k=5$ 0.79 0.73 0.69     | $k=5$ 0.85 0.81 0.76     | $k=10$ 0.94 0.89 0.85   |
| $output_{m}=5$   | $k=1$ 0.94 0.87 0.8      | $k=5$ 0.94 0.89 0.85     | $k=10$ 0.94 0.89 0.85   |
| $output_{n}=6$   | $k=5$ 0.94 0.87 0.8      | $k=5$ 0.94 0.89 0.85     | $k=10$ 0.94 0.88 0.82   |
| $output_{n}=12$  | $k=5$ 0.94 0.87 0.8      | $k=5$ 0.94 0.89 0.85     | $k=10$ 0.94 0.88 0.82   |

The values of four time series fluctuate around the mean value. Therefore, feature pooling matrix with $\text{ave}(C)$ can extract the long-term trend of a sample in the JV dataset. Table 1 also indicates that, for different size feature pooling matrices, the more features are extracted on the variable dimension, the higher the accuracy rate is.

Table 2 demonstrates the specific experiment results of the proposed model, comparing to PD, DTW and PCA+DTW, while Table 3 depicts the superiority of the proposed model over three existing models. According to the result, PD and PCA+DTW methods mismatch more times than the other methods, especially when the accuracy rate $e$ is lower than 0.5. The PCA+DTW method obtains an expected accuracy rate of 25% which cannot meet requirements of practical applications. Although the PD method outperforms PCA+DTW, its expected accuracy rate is still lower than a random guess.

On the contrary, the accuracy rate $e$ of the proposed MGIM-SPP and DTW methods are significantly higher than the other two algorithms, and especially the accuracy rate $e$ is greater than 0.5. Accordingly, it can be concluded that the proposed MGIM-SPP method is more suitable for the JV dataset, a classic small-scale MTS, than the DTW method.

In order to visually illustrate the experimental results of different methods, the 265th sample is selected randomly as an example noted as JV$_{265\_9}$. The four methods are employed to measure the similarity respectively for obtaining the most similar sample. The results are shown in Fig. 3. Since the PD method mainly focuses on the distribution characteristics of local important points, there is a high risk of misjudgment for JV$_{155\_6}$ without significant local important points. As a statistical method, PCA usually requires enough sample points to effectively solve the principal component vector, so the similarity measure error of PCA+DTW is great for small-scale datasets with fewer sample points. As Fig. 3 shows, the matching result...
obtained by the PCA+DTW method is observably different from the input sample in shape. As for the proposed MGIM-SPP method, it effectively extracts and aggregates the local features of the sample, and the geometric shape of the matching result is approximately identical with the input samples.

| Accuracy rate $e$ | MGIM-SPP | DTW | PD | PCA+DTW |
|-------------------|----------|-----|----|---------|
| $k=1$             | $k=5$    | $k=10$ | $k=1$ | $k=5$ | $k=10$ | $k=1$ | $k=5$ | $k=10$ |
| 0                 | 6        | 2    | 0  | 14      | 2      | 0  | 105 | 25 | 9  | 193 | 104 | 65 |
| 0.1               | —        | —    | 4  | —       | —      | 2  | —  | —  | 24 | —  | —   | 56 |
| 0.2               | —        | 5    | 4  | —       | 6      | 4  | —  | 36 | 16 | —  | 71  | 46 |
| 0.3               | —        | —    | 2  | —       | —      | 5  | —  | —  | 23 | —  | 50  | —  |
| 0.4               | —        | 9    | 7  | —       | 10     | 6  | —  | 41 | 28 | —  | 47  | 19 |
| 0.5               | —        | —    | 11 | —       | —      | 7  | —  | 36 | —  | —  | 17  | —  |
| 0.6               | —        | 11   | 15 | —       | 15     | 13 | —  | 72 | 49 | —  | 32  | 8  |
| 0.7               | —        | 16   | —  | —       | 20     | —  | —  | 37 | —  | —  | 6   | —  |
| 0.8               | —        | 48   | 35 | —       | 34     | 23 | —  | 58 | 28 | —  | 13  | 1  |
| 0.9               | —        | 38   | —  | —       | 46     | —  | —  | 13 | —  | —  | 2   | —  |
| 1.0               | 264      | 195  | 138 | 256    | 203   | 144 | 165 | 38 | 7  | 77  | 3   | 0  |

Table 3 Accuracy expectation values for JV dataset

| Parameter $k$ | MGIM-SPP | PD | DTW | PCA+DTW |
|---------------|----------|----|-----|---------|
| 1             | 0.966    | 0.611 | 0.948 | 0.285 |
| 5             | 0.912    | 0.560 | 0.905 | 0.243 |
| 10            | 0.869    | 0.511 | 0.866 | 0.213 |

3.2.2 ASL dataset experiment

The ASL is a sign language dataset containing 22 variables. The actions of respondents’ hands are described by 11 variables respectively. The first eight semantics (alive, all, answer, boy, building, buy, change-mind and cold), corresponding to 216 time series, are selected as the experiment dataset with a time range varying from 50 s to 79 s.

MGIM-SPP, PD, DTW, and PCA+DTW methods are also applied to the comparison of the similarity measure accuracy rate on the dataset. In the experiment, the segmentation form of the PD method is $2 \times 2$, the PCA+DTW...
method extracts the first 10 principal components of an MTS as its mode representation (the contribution rate of the first 10 principal components is greater than 95%), the feature pooling matrix size in the MGIM-SPP method is $25 \times 6$, and $F(C)$ is $\text{ave}(C)$. The test results are shown in Table 4.

| Correct rate $r$ | MGIM-SPP | DTW | PD | PCA+DTW |
|------------------|-----------|-----|----|---------|
| $k=1$ | $k=5$ | $k=10$ | $k=1$ | $k=5$ | $k=10$ | $k=1$ | $k=5$ | $k=10$ |
| 0 | 11 | 0 | 0 | 5 | 0 | 0 | 61 | 12 | 8 | 37 | 6 | 3 |
| 0.1 | — | — | 0 | — | — | 0 | — | — | 9 | — | — | 4 |
| 0.2 | — | 4 | 2 | — | 3 | 0 | — | 31 | 15 | — | 17 | 13 |
| 0.3 | — | — | 1 | — | — | 0 | — | — | 22 | — | — | 11 |
| 0.4 | — | 11 | 6 | — | 15 | 2 | — | 31 | 21 | — | 28 | 17 |
| 0.5 | — | — | 10 | — | — | 8 | — | — | 13 | — | — | 21 |
| 0.6 | — | 26 | 15 | — | 13 | 6 | — | 29 | 23 | — | 31 | 18 |
| 0.7 | — | — | 28 | — | — | 8 | — | — | 15 | — | — | 16 |
| 0.8 | — | 37 | 33 | — | 18 | 10 | — | 17 | 13 | — | 27 | 13 |
| 0.9 | — | — | 37 | — | — | 16 | — | — | 9 | — | — | 19 |
| 1.0 | 205 | 138 | 84 | 211 | 167 | 138 | 155 | 96 | 68 | 179 | 107 | 81 |

Table 4 indicates that the performance of the PD and PCA+DTW methods is greatly improved, but still inferior to the MGIM-SPP and DTW methods. For the PD method, because of the large sample size and complicated shape features of the ASL data, it is slightly insufficient that only nine quantile points are nominated to describe the point distribution characteristics of an MTS. As for the PCA+DTW method, because the sample points of the ASL dataset are more than the JV dataset, the principal component vectors of an MTS can be effectively extracted. Consequently, its performance is significantly boosted. Fig. 4 represents an example confirming the analysis.

Table 5 demonstrates that the MGIM-SPP and DTW methods are superior to the former two methods in terms of accuracy rate expectations. In virtue of feature pooling matrix, the proposed MGIM-SPP exhibits superiority through...
the effect of noise and redundancy reduction, and its recognition accuracy is comparable to DTW.

| Parameter \( k \) | MGIM-SPP | DTW | PD | PCA+DTW |
|-----------------|----------|-----|----|---------|
| 1               | 0.949    | 0.976 | 0.717 | 0.828 |
| 5               | 0.872    | 0.906 | 0.674 | 0.749 |
| 10              | 0.835    | 0.916 | 0.630 | 0.713 |

3.2.3 Time complexity comparison

Table 3 and Table 4 show that the performance of the proposed model is competitive to state-of-the-art DTW in terms of accuracy. However, the main advantage of our method is in accelerating the classification process rather than the accuracy. To illustrate the time cost of different methods, \( k \)-NN classification with four different similarity measures are run respectively on JV and ASL datasets. The experimental environment is Windows 7, CPU Intel(R) Core (TM) i5-3230 M 2.60 GHZ, Memory 4.00 GB, and Python 3.7. It can be seen from Table 6 that the running time of MGIM-SPP is identical to the PD method, and maintains the two or three order of magnitude performance gain compared to DTW and PCA+DTW. Note that there are many circumstances in which users would prefer to sacrifice some level of accuracy for considerable efficiency. Therefore, the advantage of the proposed model is its running time before its classification accuracy.

4. Conclusions

This paper constructs a new multivariate grey incidence analysis model for MTS on different scales. Based on the research, the following conclusions can be drawn:

(i) Combined with SPP, the application scope of the grey incidence analysis model has been extended to MTS on different scales. The new model also satisfies normality and pair symmetry of grey incidence axioms.

(ii) The proposed model attains state-of-the-art results in two \( k \)-NN classification experiments compared with three prominent algorithms. It can quickly be deployed in real-time systems and embedded systems because of its minimal preprocessing and efficient performance.

Developing tuning method for parameters of SPP according to target time series could further improve the accuracy of the proposed model, and it is left for future work.

References

[1] LIU S F, FORREST J Y L. Grey systems: theory and applications. Berlin: Springer Science & Business Media, 2010.
[2] THANKACHAN T, PRAKASH K S, MALINI R, et al. Prediction of surface roughness and material removal rate in wire electrical discharge machining on aluminum-based alloys composites using taguchi coupled grey relational analysis and artificial neural networks. Applied Surface Science, 2019, 472(SI): 22 – 35.
[3] UZUN G. Analysis of grey relational method of the effects on machinability performance on au tempered vermicular graphite cast irons. Measurement, 2019, 142(12): 122 – 130.
[4] HAERI S A S, REZAIEJ A. A grey-based green supplier selection model for uncertain environments. Journal of Cleaner Production, 2019, 221(16): 768 – 784.
[5] LIU Y, LIU S F, FORREST J. A new grey absolute degree of grey incidence model and application. Chinese Journal of Management Science, 2012, 20(5): 173 – 177. (in Chinese)
[6] WANG Q. The grey relational analysis of B-mode. Journal of Huazhong University of Science and Technology, 1989, 6(6): 77 – 82. (in Chinese)
[7] WANG Q Y, ZHANG Y Y. The relational analysis of C-mode. Journal of Huazhong University of Science and Technology, 1999, 27(3): 75 – 77. (in Chinese)
[8] TANG W X. The concept and the computation method of left correlation degree. Journal of Applied Statistics and Management, 1995, 15(1): 34 – 37. (in Chinese)
[9] ZHANG Y S, GAO X J, DENG J L. Grey relation entropy method of grey relation analysis. Systems Engineering-Theory & Practice, 1996, 16(8): 7 – 11. (in Chinese)
[10] DANG Y G, LIU S F, LIU B, et al. Improvement on degree of grey slope incidence. Engineering Science, 2004, 6(3): 41 – 44. (in Chinese)
[11] SINGH N, LAKI H, VEJVAR M. Big data technology: challenges, prospects, and realities. IEEE Engineering Management Review, 2019, 47(1): 58 – 66.
[12] ZHANG K, LIU S F. Extended clusters grey incidences for panel data and its applications. Systems Engineering-Theory & Practice, 2010, 30(7): 1253 – 1259. (in Chinese)
[13] WU L F, LIU S F, YAO L G, et al. Grey convex relational degree and its application to evaluate regional economic sustainability. Scientia Iranica, 2013, 20(1): 44 – 49.
[14] QIAN W Y, WANG Y H, DANG Y G, et al. Grey matrix relational modeling and its application based on multivariate panel data. Systems Engineering, 2013, 31(10): 70 – 74. (in Chinese)
[15] ZHANG K, ZHANG Y Y, QU P. Comprehensive multivariate grey incidence degree based on principal component analysis. Journal of Systems Engineering and Electronics, 2014, 25(5): 840 – 847.
[16] WU L F, LIU S F. Panel data clustering method based on grey convex relation and its application. Control and Decision, 2013, 28(7): 1033 – 1036, 1045. (in Chinese)
[17] ZHANG K, ZHANG Y Y, QIAN W Y, et al. Grey grid incidence model based on panel data. Systems Engineering-Theory & Practice, 2014, 34(4): 991 – 996. (in Chinese)
[18] ZHANG K, ZHANG Y Y, QIAN W Y, et al. Grey grid incidence model based on panel data. Systems Engineering-Theory & Practice, 2014, 34(4): 991 – 996.
[19] ZHANG K, ZHANG Y Y, QIAN W Y, et al. Grey grid incidence model based on panel data. Systems Engineering-Theory & Practice, 2014, 34(4): 991 – 996. (in Chinese)
[20] ZHANG K, ZHANG Y Y, QIAN W Y, et al. Grey grid incidence model based on panel data. Systems Engineering-Theory & Practice, 2014, 34(4): 991 – 996. (in Chinese)
ZHANG Ke et al.: A multivariate grey incidence model for different scale data based on spatial pyramid pooling

[21] LUO D, YE L L, WEI B L, et al. Grey matrix incidence model for panel data and its application in drought vulnerability risk. Control and Decision, 2018, 33(11): 2051 – 2056. (in Chinese)

[22] WANG J J, HIPEL K W, DANG Y G. An improved grey dynamic trend incidence model with application to factors causing smog weather. Expert Systems with Applications, 2017, 87: 240 – 251.

[23] DAI J, YAN Y, HE Y H. Grey incidence clustering method based on multidimensional dynamic time warping distance. Journal of Systems Engineering and Electronics, 2017, 28(5): 946 – 954.

[24] LIU Z, DANG Y G, XIAO Y. A new multivariate grey incidence model and its application to economic growth driving analysis. Journal of Grey System, 2018, 30(4): 116 – 133.

[25] PRASAD R, DEO R C, LI Y, et al. Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and Boruta-random forest hybridizer algorithm approach. Catena, 2019, 177: 149 – 166.

[26] KHOJANDI A, SHYLO O, ZOKAEINIKOO M. Automatic EEG classification: a path to smart and connected sleep interventions. Annals of Operations Research, 2019, 276(1): 169 – 190.

[27] WANG B K, LIU D T, WANG W J, et al. An improved grey dynamic trend incidence model with application to factors causing smog weather. Expert Systems with Applications, 2018, 84(5): 253 – 262.

[28] RANZATO M, BOUREAU Y, CUN Y L. Sparse feature learning for deep belief networks. Advances in Neural Information Processing Systems, 2007, 20: 1185 – 1192.

[29] SCHERER D, MULER A, BEHNKE S. Evaluation of pooling operations in convolutional architectures for object recognition. Berlin: Springer, 2010.

[30] BOUREAU Y, PONCE J, LECUN Y. A theoretical analysis of feature pooling in visual recognition. Proc. of the International Conference on Machine Learning, 2010: 125 – 133.

[31] FUKUSHIMA K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 1980, 36(4): 193 – 202.

[32] NIOKA H, ASATANI S, YOSHIMURA A, et al. Classification of C2C12 cells at differentiation by convolutional neural network and its application to cell recognition. Cell, 2018, 31(1): 87 – 93.

[33] GUO L H, GUO C G, LI L, et al. Two-stage local constrained sparse coding for fine-grained visual categorization. Science China Information Sciences, 2017, 61(1): 18104.

[34] SHARIF S, MOHAMED N, MOMEN S, et al. Classification of Bangla compound characters using a HOG-CNN hybrid model. Proc. of the International Conference on Computing and Communication Systems, 2018: 403 – 411.

[35] PAUL A, MALATHY C. An innovative approach for automatic genre-based fine art painting classification. Proc. of the Advanced Computational and Communication Paradigms, 2018: 19 – 27.

[36] LIU D Q, YUE S G. Event-driven continuous stdp learning with deep structure for visual pattern recognition. IEEE Trans. on Cybernetics, 2018, 49(4): 1 – 14.

[37] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 – 1916.

[38] BASBUG A M, SERT M. Acoustic scene classification using spatial pyramid pooling with convolutional neural networks. Proc. of the IEEE 13th International Conference on Semantic Computing, 2019: 128 – 131.

[39] YANG Z, HE X Y, YANG F, et al. Combining partial parameter transfer strategy and spatial pyramid pooling for biological-cell classification. Proc. of the 10th International Conference on Graphics and Image Processing, 2019: 11069.

[40] KARIM F, MAJUMDAR S, DARABI H, et al. Multivariate LSTMs for time series classification. Neural Networks, 2019, 116(8): 237 – 245.

[41] YU C Y, LUO L T, CHAN L L, et al. A fast LSH-based similarity search method for multivariate time series. Information Sciences, 2019, 476(7): 337 – 356.

[42] KUDO M, TOYAMA I, SHIMBO M. Multidimensional curve classification using passing-through regions. Pattern Recognition Letters, 1999, 20(11/12/13): 1103 – 1111.

[43] KADOUS M W. Temporal classification: extending the classification paradigm to multivariate time series. Sydney: University of New South Wales, 2002.

[44] GUAN H S, JIANG Q S, WANG S R. Pattern matching method based on point distribution for multivariate time series. Journal of Software, 2009, 20(1): 67 – 79. (in Chinese)

[45] SHEN J Y, HUANG W P, ZHU D Y, et al. A novel similarity measure model for multivariate time series based on LMNN and DTW. Neural Processing Letters, 2017, 45(3/4): 925 – 937.

[46] BANKO Z, ABONYI J. Correlation based dynamic time warping of multivariate time series. Expert Systems with Applications, 2012, 39(17): 12814 – 12823.

Biographies

ZHANG Ke was born in 1983. He received both his B.S. and M.S. degrees in electronic information engineering from Nanchang Hangkong University in 2004 and 2007 respectively, and Ph.D. degree in system engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China. Currently, he is an associate professor at the Business School, Hohai University. His research interests include grey system theory and uncertainty system modeling. E-mail: kezhang@hhu.edu.cn

CUI Le was born in 1997. She received her B.S. degree in information management and information system from Hohai University in 2018. Currently, she is a postgraduate in Hohai University. Her research interests are data mining and machine learning. E-mail: cui_le19@foxmail.com

YIN Yao was born in 1995. He received his B.S. degree in information management and information system from Hohai University in 2017. He is a postgraduate in Hohai University. His research interest is machine learning. E-mail: 18260061373@163.com