2017 4th International Conference on Information Science and Control Engineering

A Shape Based Similarity Measure for Time Series Classification with Weighted Dynamic Time Warping Algorithm

Yanqing Ye, Caiyun Niu, Jiang Jiang*, Bingfeng Ge, Kewei Yang
College of Information System and Management
National University of Defense Technology
Changsha, China
Email: yeyanqing@nudt.edu.cn, caiy_niu@yahoo.com,
jiangjiangnudt@163.com, bfge.nudt@gmail.com,kayyang27@163.com

***Abstract*—Time series similarity measure is an essential issue in time series data mining, which can be widely used in various applications. With an eye to the fact that most current measures neglect the shape characteristic of time series, this paper proposes a shape based similarity measure. By introducing a shape coefficient into the traditional weighted dynamic time warping algorithm, an improved version, shape based weighted dynamic time warping (SWDTW) algorithm is proposed. Specifically, the ways to measure univariate and multivariate time series similarity with SWDTW are presented. Finally, in order to verify the effectiveness of the proposed similarity measure, both 1NN classification and similarity search experiments are carried out using datasets derived from UCR Time Series Classification Homepage. By comparing the SWDTW similarity measure with other measures, the results show that the proposed SWDTW measure is more of accuracy and robust.**

***Keywords*-time series; similarity measure; shape based weighted dynamic time warping**

I. Introduction

Time series data mining (TSDM) is the process of extracting hidden information from a large amount of time series data, one core issue of which is similarity measure. Almost all of the tasks in TSDM, such as retrieval, clustering and classification, need to find a suitable distance measure to compare the similarity/dissimilarity between pairwise time series [1]. Due to the characteristics of high dimensionality, autocorrelation, shifting in time and amplitude, and noisy, it's not straightforward to derive a method that can measure the similarity of time series efficiently and effectively[2].

Common methods to measure the similarity of time series are Euclidean distance(ED)[3], [4] and dynamic time warping(DTW)[5], [6]. Compared with the previous one, DTW is a dynamic programming method that can measure the similarity for time series of different lengths, stretching and bending. As a result, it can achieve a better accuracy and good robustness, but high computational complexity, and sometimes can lead to abnormal matching problem. Therefore, various improvement methods for DTW have been proposed, including LB_Keogh lower bound [7], differential dynamic time warping distance[8], [9], weighted dynamic time warping distance(WDTW)[10], [11], etc.. Compared with DTW algorithm, WDTW adds weight to the base distance of the points. In other words, it tends to match points with narrow time interval, which can avoid abnormal matching problem as possible. However, WDTW algorithm only sees two basic characteristics of time series, time and numerical characteristics, but ignores their shape characteristic. It can be seen as the shape of a curve or several curves(multivariate time series), which is constituted by the numerical point sorted in time. For example, two points have a large time interval, but their shapes are similar, i.e., both points are peak or bottom points in curves. In this condition, WDTW algorithm will avoid their matching, however, it ignores the importance of the shape matching.

Therefore, this paper proposes a shape based similarity measure for time series. By introducing a shape coefficient into the traditional WDTW proposed in [9], a shape based weighted dynamic time warping algorithm (SWDTW) is presented to measure the similarity for both univariate and multivariate time series. Finally, by comparing SWDTW algorithm with other methods in the nearest neighbor classification and similarity search experiments, it has showed that the proposed SWDTW similarity measure is more of accuracy and robustness.

II. Shape Based Weighted Dynamic Time Warping Algorithm

In this section, we propose a shape based weighted dynamic time warping algorithm to measure the similarity of time series. Similar to DTW, SWDTW is to find the minimum matching distance by dynamic programming, allowing the offset of time series in the time and numerical axis. Their differences lie in the calculation of the base distance between local points. In DTW, each point in time series is equally dealt with in the calculation of the base distance. Normally, the square of Euclidean distance between points are used. In order to solve the abnormal matching problem of DTW, WDTW algorithm gives the base distance a weight. The weight will be larger when two points have a larger interval in time, which can possibly avoid abnormal

978-1-5386-3013-6/17 $31.00 © 2017 IEEE
DOI 10.1109/ICISCE.2017.32

matching between points. While in SWDTW, the weight is relevant to both the time interval and shape characteristics of the matching points. When two points have a narrow time interval and similar shape, the weight is little. In other words, when two points in the matched time series are in the same corresponding place, and the shape of the curve here is similar, then SWDTW algorithm will give the base distance a small weight and support their matching. In the following, a detailed description of the SWDTW distance calculation process is presented.

For a given time series $X_{K\times m} = \{x_k, k = 1, 2, ..., K\}^T$ and $Y_{K\times n} = \{y_k, k = 1, 2, ..., K\}^T$, in which, $x_k = \{x_{ki}, i = 1, 2, ..., m\}$ and $y_k = \{y_{kj}, j = 1, 2, ..., n\}$. When $K = 1$, X and Y are univariate time series, and when $K \geq 2$, X and Y are multivariate time series. $A_{m\times n} = (a_{ij})_{m\times n}$ represents the distance matrix between X and Y, in which a_{ij} is calculated by the base distance measure. In traditional DTW method, multivariate time series in one dimensional space is seen as univariate time series in multi-dimensional space and the square of Euclidean distance is chosen as the base distance measure, i.e.,$a_{ij} = \sum_{k=1}^{K} (x_{ki} - y_{kj})^2$. While in SWDTW measure, the weighted square of Euclidean distance is chosen as the base distance, i.e. $a_{ij} = w_{i-j} \sum_{k=1}^{K} (x_{ki} - y_{kj})^2$, that is to say, it is necessary to give the weight coefficient w_{i-j} between two corresponding points. $S = \{s_p, p = 1, 2, ..., P\}$ represents the dynamic warping path in the distance matrix, which needs to satisfy the following conditions:

(1)$\max\{m, n\} \leq P \leq m + n - 1$;

(2)$s_1 = a_{11}$,$s_P = a_{mn}$;

(3) if $s_p = a_{ij}$ and $s_{p+1} = a_{i'j'}$, then $0 \leq j' - j \leq 1$.

If $L_w(i, j)$ represents the weighted cumulative distance between the point i of X and j of Y, then

$$L_w(i,j) = \min[L_w(i-1,j-1), L_w(i-1,j), L_w(i,j-1)] + a_{ij}, \tag{1}$$

where $L_w(1, 1) = a_{11}$.

There are more than one dynamic path. In all paths, $L_w(m, n)$ represents the best weighted dynamic time warping path with the minimum value of the weighted cumulative distance. Figure 1 shows an example of the matching process in SWDTW. There are three time series, sample 1, sample 2 and sample 3, and the matrix represents the cumulative distance matrix. The dashed line is the dynamic warping path between sample 1 and sample 2, and the dashed one represents the path between sample 1 and sample 3. As shown, both lines are curves, which means that time series are dynamically matched. The solid line is more curved than the dashed line, which shows the dynamic matching process for time series of the different category is more complicated and of higher complexity. Thus, the length of the warping path can be used for representing the dissimilarity of time series. The length of warping path between sample 1 and sample 3 is longer than that between sample 1 and sample 2. Thus, compared with sample 3, sample 2 is more similar to sample 1. $SWDTW(X, Y)$ represents the similarity of time series X and Y measured by SWDTW algorithm, thus $SWDTW(X, Y)$ can be obtained by

$$SWDTW(X,Y) = \min(\sqrt{L_w(m,n)}). \tag{2}$$

In summary, the SWDTW algorithm can be represented by

$$\begin{cases} a_{ij} = w_{i-j} \sum_{k=1}^{K} (x_{ki} - y_{kj})^2 \\ L_w(i,j) = \min[L_w(i-1,j-1), L_w(i-1,j), L_w(i,j-1)] + a_{ij} \\ SWDTW(X,Y) = \min(\sqrt{L_w(i,j)}). \end{cases} \tag{3}$$

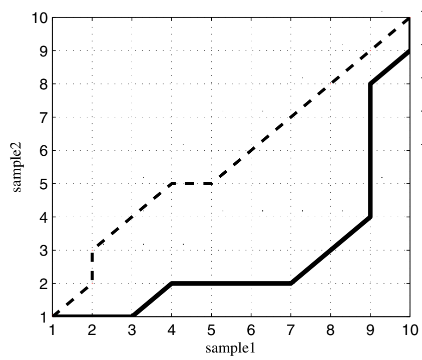

Figure 1: An example of shape based weighted dynamic time warping algorithm. In figure, three samples have been extracted from the synthetic_control dataset (detailed information can be seen in experiments section), in which, sample 1 and sample 2 belong to the same category and sample 3 belong to another category. As seen, the dashed line represents the dynamic warping path between sample 1 and sample 2 while the solid line represents the warping path between sample 1 and sample 3.

In the following, the detailed calculating process of the weight factors w_{i-j} between two corresponding points i and j is presented. First, the importance vectors $E_X = \{e_{ki}\}_{K\times m}$ and $E_Y = \{e_{kj}\}_{K\times n}$ of the points in two time series X and Y are obtained by the local extreme point detection algorithm, in which both e_{ki} and e_{kj} are determined by:

$$e_{ki} = \begin{cases} 1, \text{when } X_i \text{ is the local maximum} \\ -1, \text{when } X_i \text{ is the local minimum} \\ 0, \text{when } X_i \text{ is none of local maximum and minimum} \end{cases} \tag{4}$$

As mentioned before, as a numerical point sequence sorted by time, time series not only have numerical and time characteristic, but the shape feature. In this paper, we draw attention on the local point distribution of the time series. In sequence, points can be classified into three categories, which are local maximum(peak), local minimum(bottom) and normal. Therefore, when matching two time series, we tend to match points with the same shape, e.g., "peak-peak" or "bottom-bottom", and avoid "peak-bottom" matching. Thus, the weight of two corresponding points is given as follows:

$$w_{i-j} = w_{max}/(1 + exp(-g \times (c \times R - m_c \times K))) \quad (5)$$

where $c = |i - j|$ is the distance coefficient, obviously, $0 \leq c \leq max(m, n)$; R is the shape coefficient, which can be determined by $R = \sum_{k=1}^{K} r_k, r_k = |e_{ki}e_{kj} - 1|, k = 1, 2, ..., K$; m_c is the middle point of X or Y, i.e., $m_c = [m/2]$ or $[n/2]$; w_{max} is the upper limit of the weight; g is a constant that can effectively control the level of penalty on the time interval between the point and the shape difference. When $g = 0$, the weight values of all points are the same, that is, the traditional DTW method. With the increase of g, the penalty for both the time interval and the unmatched shape of two points increases. In [9], it is pointed out that a good effect can be obtained when the value of g is between 0.01 and 0.6. Unlike WDTW, the influence of g on the results depends not only on the time interval, but also on the shape differences of the sequences. According to the definition of the weight value w_{i-j}, it is an increasing function of the shape coefficient R and the distance coefficient c.

According to the definition of shape coefficient, there are three values for r_k as follows:

$$r_k = \begin{cases} 0, \text{when } e_{ki} = 1, e_{kj} = 1 \text{ or } e_{ki} = -1, e_{kj} = -1 \\ 1, \text{when } e_{ki} = 0 \text{ or } e_{kj} = 0 \\ 2, \text{when } e_{ki} = 1, e_{kj} = -1 \text{ or } e_{ki} = -1, e_{kj} = 1 \end{cases} \quad (6)$$

Therefore, the range of shape coefficient $R \in \{0\} \bigcup \{1, 2, ..., K\} \bigcup \{2 \times i, i = 1, 2, ..., K\}$, i.e. when all dimensions are "peak-peak", or "bottom-bottom", the weight function is only dominated by shape coefficient, and the value reaches its minimum value $w_{min} = w_{max}/(1 + exp(g \times m_c \times K))$. In other words, it encourages "peak-peak" or "bottom-bottom" points to match. When all dimensions are "peak-bottom" matching, shape coefficient reaches its maximum value $R = 2K$, which can expand the influence of the shape coefficient and avoid the matching of the peak point and bottom point. When there are all none extreme points matching, the shape coefficient $R = K$, at the point the weight value is not dominated by the shape coefficient R but the distance coefficient c.

As described before, the calculating process of the SWDTW algorithm is based on the univariate time series. But it is not to say that this method can only be used to calculate the similarity for one-dimensional time series. It can also be used for measuring the similarity for multi-dimensional time series. Similar to other univariate similarity measures, SWDTW can be used to calculate each corresponding dimension of multivariate time series, respectively. Then the sum of the result from the measures of all variables can be used as the similarity measure for multivariate time series. It should be noted that in order to avoid the numerical value of each dimension affecting other dimensions, before measuring the similarity, normalization needs to be carried out to standardize the whole time series. Normally, the "min-max" normalization is used. For the MTS X with K dimensions, for each dimension $x_k = \{x_{ki}, i = 1, 2, ..., n_k\}, k = 1, 2, ..., K$,

$$\hat{x}_{ki} = (x_{ki} - min(x_k))/(max(x_k) - min(x_k)). \quad (7)$$

However, in real application, it can only be used for small scale of few variables or short length as a result of high computational complexity. For multivariate time series with large scale, we can represent them by dimensionality reduction method, such as principal component analysis or linear segmentation. Afterwards, we can use SWDTW to measure the transformed time series. The detailed information will be discussed in our later work.

III. Experiments

In order to verify the effectiveness of the proposed SWDTW similarity measure, in this section, we adopts the measure in the nearest neighbor (1NN) classification and similarity search experiments in comparison with other common measures, i.e. Euclidean distance(ED), dynamic time warping(DTW), and weighted dynamic time warping (WDTW) algorithms. As mentioned before, the proposed SWDTW measure can be used to measure the similarity for both univariate and multivariate time series. Thus, in this section, we carried out experiments on both univariate and multivariate time series datasets. The detailed process and the results are shown as follows.

A. Data sets

The univariate time series data sets used derive from the UCR Time Series Classification Homepage [12], which possesses a multitude of the time series data sets that are publicly available and labeled. The lengths of samples in the datasets vary from 96 to 570, and the number of classes per dataset varies from 2 to 9. The size of the training dataset (testing dataset) varies from 28 (28) to 1000 (4000). The datasets contain data from various domains, such as robotics, handwriting recognition and medicine. In the experiment, 10 univariate time series datasets are selected. In the multivariate time series classification experiment, the data sets used derive from the UCI homepage, in which EEG [13], Japanese Vowel (JV) [14] and Robot (Execution) Failure (REF) [15] dataset are selected. The EEG data set was derived from two types of experimental objects. The number of variables is 64, and the length is 256. Here we randomly select 60 samples from two different types of samples, named co2a0000364 and co2c0000337. The JV data set is a set of records of different individual pronunciation, which contains 9 categories of samples with 12 variables and the various length of 7-29. The REF data set records the failure of the robot, which includes LP1-LP5 five data sets. The first subset LP1 data set is used in this paper.

The software of MATLAB (version R2015a) was used in the computational process. A PC with Intel Core i7-4710MQ 2.05 GHZ processor (8GB RAM) and Windows 7 Professional SP1 (64-bit) operating system served as the hardware platform.

B. The nearest neighbor classification experiment

1) Experiment settings: In this part, the proposed similarity measure is validated by the 1NN classification method. Following Wang et al. [16] and Serra et al. [17], the nearest-neighbor classifier has several advantages for use as the validation method. First of all, the performance of the 1NN classifier is decided by the similarity measure it selects. Second, the 1NN classifier is parameter-free, thus easily implemented and compared with other measures. Third, it has been suggested that the simple 1NN classifier can obtain the best results in time series classification. Furthermore, in most of the existing research, the 1NN classification method has been used as the evaluation criterion of the similarity measure. The error rates are calculated as the results. In both WDTW and SWDTW, the upper limit of weight $w_{\max}$ is set as 2, g is taken as 0.08.

2) Result and analysis: 1NN classification results are shown in Table I. All the results obtained from non-random measures are better than that from the random baseline similarity measure. Our proposed SWDTW outperforms other measures in most datasets except Olive Oil, EEG and REF(LP1) datasets. Additionally, it can achieve zero-error classification in some datasets, i.e., CBF, Trace, Two Patterns and Coffee datasets. Both DTW and WDTW also win in some of datasets. Furthermore, ED obtains 2.83 for average rank, only being superior to the random baseline. To show the comparison of the similarity measures more clearly, the box plot of the performance ranks of each measure is presented in Figure 2. The measures are SMDTW, WDTW, DTW, ED and random baseline from left to right in horizontal axis. As shown, from a global basis, our proposed SWDTW measure can obtain the best average performance in all datasets, followed by WDTW and DTW. In addition, ED and random measures are inferior to other measures.

In order to further assess the statistical difference between the error ratio of our proposed measure and other similarity measures, the well-known Wilcoxon signed-rank test is employed. The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test. It usually serves to

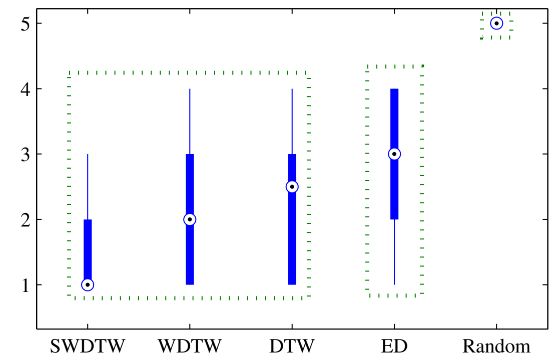

Figure 2: Box plot for the distribution of performance ranks of each measure across datasets. Dashed lines denote statistically significantly equivalent groups of measures ($p < 0.05$)

compare two repeated measurements (or related samples) by assessing whether their mean ranks in population differ [17]. In order to comprehensively compare similarity measures, we employed the average error ratio of all datasets as input. According to common practice, the threshold of the significance level is set as 5%. The statistical analysis results are shown in Figure 2. The green dashed line denotes the statistically equivalent group measures. As seen, the SDTW, WDTW and DTW are statistically equivalent. Although our proposed SWDTW do not outperform WDTW and DTW measures, our main contribution is proposing an idea to extract the shape characteristic of the time series in similarity measure, which can also outperform other methods in accuracy or efficiency in some cases.

C. Similarity search experiments

1) Experiments settings: In order to further verify the proposed SWDTW similarity measure, in this part, we carry out the K-nearest similarity search experiments on both univariate time series and multivariate time series datasets of three different scales, i.e., small, medium and large. In univariate time series datasets. In univariate datasets, Gunpoint, Trace and Lighting-2 datasets, lengthes of which are 150, 275 and 637, respectively. While in multivariate datasets, EEG, Japanese Vowel(JV) and REF(LP1) datasets are chosen. Similar to the classification experiment, the SWDTW similarity is compared with ED, DTW and WDTW similarity measures.

For the time series datasets with n samples, select i^{th}, $i = 1, 2, 3..., n$ sample from testing datasets to be used as matching samples first. Second, find out k samples in training dataset that are the most similar with the matched samples by using the similarity measure methods listed above, where $k = 1, 5$ or 10. Finally, find out samples that belong to the same category with the matched sample and

Table I: Error rates in 1NN classification for different data sets. The data sets include 10 univariate and 3 multivariate data sets. In the front 5 columns, the detailed information about datasets has been given. In the last 5 columns, the error rates of the classification experiments for each datasets based on random, Euclidean distance, dynamic time warping, weighted dynamic time warping and shape based weighted dynamic time warping algorithms are presented, respectively. As ED algorithm can only measure time series with the same lengthes, thus, it can't be applicable to JV dataset.

Datasets	**Length**	**Size of train sets**	**Size of test sets**	**Number of classes**	**Number of variables**	**Random**	**ED**	**DTW**	**WDTW**	**SWDTW**
Gun-point	150	50	150	2	1	0.506	0.087	0.093	0.087	**0.067**
CBF	128	30	900	3	1	0.655	0.148	0.003	0.004	**0**
Trace	275	100	100	4	1	0.757	0.240	**0**	**0**	**0**
Two patterns	128	1000	4000	4	1	0.743	0.093	**0**	**0**	**0**
Lighting-2	637	60	61	2	1	0.488	0.246	0.131	0.115	**0.098**
ECG200	96	100	100	2	1	0.515	**0.120**	0.230	0.200	0.160
Fish	463	175	175	7	1	0.871	0.217	0.177	**0.160**	**0.160**
Beef	470	30	30	5	1	0.763	0.467	0.500	0.500	**0.400**
Coffee	286	28	28	2	1	0.394	0.250	0.179	0.179	**0**
Olive Oil	570	30	30	4	1	0.644	0.133	0.167	**0.133**	0.167
EEG	256	30	30	2	**64**	0.767	0.067	**0.033**	0.067	0.067
Japanese vowel	7-29	135	135	9	**12**	0.948	-	**0.037**	**0.037**	**0.037**
REF(LP1)	15	43	45	4	**6**	0.667	0.422	**0.400**	0.578	0.489
			Average error rate			0.648	0.207	0.159	0.169	**0.134**
			Average rank			5.00	2.83	2.33	2.17	**1.50**

the number is denoted as h_i, and calculate the matching error rate $e_i = 1 - h_i/k, i = 1, 2, ..., n$. According to the definition, $e_i \in [0, 0.1, 0.2, ..., 1]$, which has 11 values. After repeating n times, ε represents the random variable of experiment error rate, then the mathematical expectation value $e*$ of ε can be calculated as follows.

$$e* = \sum_{i=1}^{n} p(\varepsilon = e_i) \times e_i$$

Table II: Error rates in similarity search experiment on different datasets including 3 univariate and 3 multivariate data sets. Results with the minimum values are bold.

Datasets	**k**	**Random**	**ED**	**DTW**	**WDTW**	**SWDTW**
Gun-point	**1**	0.507	**0.087**	0.093	**0.087**	**0.087**
	5	0.505	0.248	**0.259**	0.241	**0.239**
	10	0.479	**0.351**	0.335	0.327	**0.326**
Trace	**1**	0.740	0.240	**0**	**0**	**0**
	5	0.740	0.448	**0.014**	**0.014**	**0.014**
	10	0.757	0.507	**0.065**	0.072	0.075
Lighting-2	**1**	0.541	0.246	0.131	0.115	**0.098**
	5	0.449	0.348	**0.223**	0.246	0.226
	10	0.482	0.380	**0.321**	0.336	0.323
EEG	**1**	0.400	0.067	**0.033**	0.067	0.067
	5	0.800	0.267	0.300	**0.260**	**0.260**
	10	0.900	**0.350**	0.383	0.357	0.357
Japanese Vowel	**1**	0.733	-	**0.037**	**0.037**	**0.037**
	5	0.800	-	0.129	**0.118**	0.120
	10	0.900	-	0.209	0.202	**0.201**
REF(LP1)	**1**	0.859	0.378	**0.333**	0.444	0.400
	5	0.800	0.436	**0.382**	0.538	0.484
	10	0.900	0.493	**0.436**	0.560	0.613

2) *Result and analysis:* As shown in Table II, the results of similarity search using five measures for six datasets when

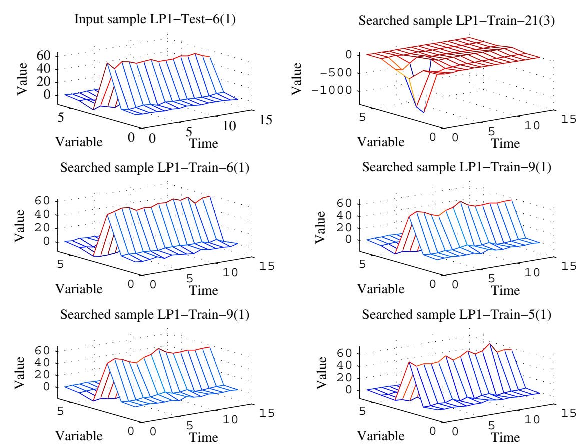

Figure 3: The example result of similarity search for LP1 dataset. The first subfigure is the 3D image of the searched sample, which is the 6^{th} sample in testing dataset of LP1, and the remaining subfigures are the searched samples found by random baseline, ED, DTW, WDTW and SWDTW measures, respectively. They are the 21^{th},6^{th},9^{th},9^{th},5^{th} samples in training dataset of LP1, respectively.

k equals to 1, 5 and 10 are given. From a global view, with the increase of k, the error rates increase, which shows that the similarity measures used are more accurate in the nearest neighbor search experiment. Similar to classification experiment, all the none-random measures can outperform random measures and none measures can outperform all other measures in similarity search experiment. SWDTW algorithm can achieve a best performance in Gun-point and

Trace datasets, while DTW algorithm wins in both Lighting-2 and REF(LP1) datasets. It can show that compared with multivariate time series datasets, SWDTW algorithm can perform better in univariate time series dataset. In order to compare the matching results of different methods more intuitively, this paper takes the sixth sequence of LP1 testing dataset as input sample, and the sample category is 1, which is represented by LP1-Test-6(1), and finds the most similar samples in the training datasets by the above five methods, for example, the result obtained by ED is represented by LP1-ED-6(1). In the similarity matching experiment, the most similar samples found by Random, ED, DTW, WDTW and SWDTW algorithm are samples numbered 21, 6, 9, 9 and 5. In those samples, except for sample 21 belonging to category 3, other samples belong to the category 1 that is in accordance with testing sample. Because LP1 is a multivariate time series, the samples searched by different methods are displayed with 3D images in Figure 3. As shown, from the shape view, the fluctuation trend of the sample found by SWDTW method is the most similar with the testing sample, so as to further demonstrate the validity of the SWDTW algorithm.

IV. Conclusion

In the light of the condition that most existing methods ignore the shape characteristic of time series, this paper proposes a shape based similarity measure for time series. By introducing a shape coefficient into the weight function, a shape based weighted dynamic time warping algorithm is proposed that can effectively capture the numerical, time and the shape characteristics. In experiments, 1NN classification and similarity search for both univariate and multivariate time series are carried out and verify the effectiveness of the proposed SWDTW method.

However, the method proposed in this paper still has some shortcomings. Similar to DTW algorithm, due to the complexity of the dynamic programming, the computational complexity is very high and may restrict its real applications, especially for multivariate time series of long length. Therefore, in our future research, we will draw our attention on improving the computational efficiency of the SWDTW algorithm. We will first consider those methods that have successfully improve the efficiency of the DTW algorithm.

Acknowledgment

The authors would like to thank the supports by Program of National Natural Science Foundation of China(Nos. 71571185, 71671186 and 71501182) and the UCR Time Series for providing online datasets.

References

[1] Lehnertz K. Non-linear time series analysis of intracranial EEG recordings in patients with epilepsyan overview. International Journal of Psychophysiology, 1999, 34: 45-52.

[2] He Xiaoxu. Research on key issues in time series data mining. Hefei: University of Science and Technology of China, 2014: 25-29.

[3] Yi B K, Faloutsos C. Fast Time Sequence Indexing for Arbitrary Lp Norms. Proceedings of the 26th International Conference on Very Large Data Bases, Cairo, Egypt: Morgan Kaufmann Publishers, 2001: 385-394.

[4] Helman S, Merav S, Sigal B, Most Discriminating Segment-longest Common Subsequence (MDSLCS) Algorithm for Dynamic Hand Gesture Classification. Pattern Recognition Letters, 2013, 34: 1980-1989.

[5] Keogh E, Ratanamahatana A. Exact Indexing of Dynamic Time Warping. Knowledge and Information Systems, 2005, 7(3): 358-386.

[6] Zhang Z, Tavenard R, Bailly A, et al. Dynamic Time Warping under limited warping path length[J]. Information Sciences, 2017, 393:91-107.

[7] Kim S W, Sanghyun P, Chu W W. An Index-based Approach for Similarity Search Supporting Time Warping in Large Sequence Databases. Proceedings of the 17th International Conference on Data Engineering, 2001: 607-614.

[8] Tomasz G, Maciej L. Non-isometric Transforms in Time Series Classification Using DTW. Knowledge Based Systems, 2014, 61: 98-108.

[9] Tomasa G, Maciej L. Multivariate Time Series Classification with Parametric Derivative Dynamic Time Warping. Expert Systems with Applications, 2015, 42: 2305-2312.

[10] Jeong Y S, Jeong M K, Omitaomu O A. Weighted Dynamic Time Warping for Time Series Classification. Pattern Recognition, 2011, 44: 2231-2240.

[11] Rodriguez J J, Alonso C J, Maestro J A. Support vector machines of interval based features for time series classification. Knowledge-Based Systems, 2005, 18: 171-178.

[12] Hettich, S. and Bay, S. D. (1999). The UCI KDD Archive [http://kdd.ics.uci.edu]. Irvine, CA: University of California, Department of Information and Computer Science.

[13] Begleiter H. EEG Database [EB/OL]. [2009-11-6]. http://kdd.ics.uci.edu/databases/eeg/eeg.html.

[14] Mineichi Kudo, Jun Toyama, Masaru Shimbo. Japanese Vowels [EB/OL]. [2009-11-6]. http://kdd.ics.uci.edu/databases/Japanese Vowels/JapaneseVowels.html.

[15] Lopes L S, Camarinha-Matos L M. Robot Execution Failures [EB/OL]. [2009-11-6]. http://kdd.ics.uci.edu/databases/robotfailure/robotfailure.html.

[16] X. Wang, A. Mueen, H. Ding, G. Trajcevski, P. Scheuermann, E. J. Keogh, Experimental comparison of representation methods and distance measures for time series data, Data Mining & Knowledge Discovery, 2013, 26: 275-309.

[17] J. Serra, J. L. Arcos, An empirical evaluation of similarity measures for time series classification. Knowledge Based Systems, 2014, 67: 305-314.